ABSTRACT

Title of Dissertation: STRUCTURE OF CONTINENTAL LITHOSPHERE FROM TRANSDIMENSIONAL BAYESIAN INVERSION OF SURFACE WAVES AND RECEIVER FUNCTIONS

Chao Gao, Doctor of Philosophy, 2019

Dissertation directed by: Associate Professor Vedran Lekić, Department of Geology

Continental crust and the underlying lithospheric mantle make up the continental lithosphere of the Earth. Our understanding of its structure and composition is limited by the inaccessibility of Earth's deep interior. Seismic imaging utilizing complementary seismic data provides unique constraints on the present-day structure of continental lithosphere. However, while recent efforts have improved the resolution of seismic images, the quantification of uncertainties remains challenging due to the non-linearity and the non-uniqueness of the geophysical inverse problem. To gain insights into the composition, formation, and evolution of the continental lithosphere, an interdisciplinary approach that incorporates seismological, geodynamical, and geochemical contributions is needed. In this dissertation, I

implement a model-space search approach – transdimensional Bayesian inversion – to explore seismological constraints of continental lithosphere. I utilize seismic observables including Rayleigh and Love wave dispersion, Rayleigh wave ZH ratio, and Ps receiver functions to invert for shear velocity (Vs), compressional velocity (Vp), density (ρ), and radial anisotropy (ξ) profiles of lithospheric structure. I begin by systematically investigating the effects of parameterization choices on inversion results using synthetic tests. Then, I proceed to tackle several technical challenges regarding the accurate retrieval of multi-parameter velocity structures from large seismic arrays in the presence of sediment layers. Finally, I apply these techniques to create a shear velocity model (TBI-NGP) of the lithosphere across the Northern Great Plains of the United States using data from the *EarthScope* Transportable Array. This probabilistic seismic model enables statistical assessment of the elastic properties in an Archean craton and Paleoproterozoic orogen. Subsequently, I incorporate seismic constraints with geophysical and geochemical measurements to infer the composition of continental crust in the region.

STRUCTURE OF CONTINENTAL LITHOSPHERE FROM TRANSDIMENSIONAL BAYESIAN INVERSION OF SURFACE WAVES AND RECEIVER FUNCTIONS

by

Chao Gao

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2019

Advisory Committee:
Associate Professor Vedran Lekić, Chair
Associate Professor Maria K. Cameron
Associate Professor Michael N. Evans
Assistant Professor Nicholas C. Schmerr
Staff Scientist Lara S. Wagner

Dedication

I dedicate this dissertation to my families: my parents, my grandparents, my brother Fei, and my fiancée Xizheng. Thank you for encouraging me to pursue my dream and supporting me through this long journey.

Acknowledgements

There is an old saying in China: “Swift horses are usually found but not the same as the person who has good judgments to spot them”. Six years ago, when I started reaching out to potential graduate study advisors, I was not confident about my application. Ved responded to my email with a detailed list of potential projects and strongly encouraged me to apply to Maryland. Later I was offered an opportunity to work with Ved, for which I am always grateful. Ved tried his best to make sure that I have everything needed to be successful in my PhD. Aside from the patient guidance and insightful discussions, Ved gave me the freedom to work on the projects that I am most excited about. He set high standard for my research, and he is always available and encouraging when I got stuck in research. Ved did not just “teach” me how to tackle challenging problems, he showed me how to think and work like a scientist. I enjoyed the time when we sit together discussing papers and debugging codes. PhD study is not always fun and easy as I imagined, I was lucky to work with Ved. Thank you, Ved, for being my mentor.

I want to thank all the members of my dissertation defense committee: Lara Wagner, Michael Evans, Nicholas Schmerr, and Maria Cameron for their constructive reviews of my dissertation and participation in my defense. Special thanks to Lara Wagner for traveling from DC to College Park to attend my defense. Thank you for your suggestions and reviews of the low-velocity zone paper as well. I want to thank Nick for his support and efforts through my PhD. Nick’s door was always open when I needed external expertise in seismology and structural geology. I was able to work on a quite different project of studying meltwater beneath a Greenland glacier thanks to

Nick. I took three classes with Mike during my graduate study, from which I learned a lot about data analysis, statistics, and numerical modeling in geoscience. Mike met with me on a weekly basis for one semester to discuss my research, which really helped me with my first publication. Many thanks to you, Mike.

I would like to acknowledge my colleagues and friends in the Seismology Group: Raj Moulik, Scott Burdick, Tolu Olugboji, Ross Maguire, Zac Reeves, Kristel Izquierdo, Erin Cunningham, Karen Pearson, Phillip Goodling, and Anthony Mautino, for all the paper readings, practice talks, and discussions on science. My thanks to Zac for practicing English with me and introducing American culture to me. I also appreciate all the conversations with Raj about science and career choices.

A huge thanks to my fiancée, Xizheng. We came to Maryland together for graduate school six years ago, and I could not have done this without her. Her determination, critical thinking, and positive attitude always inspire me. Thank you for all the encouragement, and for all the fun and good times we shared.

Thanks to all my friends at Maryland. Thanks to Yu Huang, Hailong Bai, Huan Cui, Ming Tang, Kang Chen, Jiangyi Hou, Manyi Wang, Israel Martínez, Songjie Wang, Tiange Xing, Quancheng Huang, Goeun Ha, Grace Ni, and fellow graduate students in the Department of Geology, for making my graduate school life unforgettable.

My research was financially supported by the National Science Foundation grant EAR 1352214, 1361325, the teaching assistantship from Department of Geology, Summer Research Fellowship from UMD graduate school, Packard

Fellowship for Science and Technology, and the ESSIC Travel Award from the Earth System Science Interdisciplinary Center.

Table of Contents

List of Figures

Chapter 1: Introduction

The lithosphere is Earth's outermost rigid shell part that consists of the crust and the underlying lithospheric mantle. The lithosphere can be classified into oceanic and continental due to remarkable distinctions that have been observed between the two, especially in the crust. Earth's oceanic crust is relatively thin (<8km) and young (<200 Myr) with a mafic composition. In contrast, the continental crust is thick (25 – 70 km), billions of years old, and is more differentiated. The continental crust preserves rocks with diverse lithologies, offering a unique opportunity to study the formation and evolution of the Earth. The continental crust evolved from a predominantly mafic composition to today's andesitic composition with a vertically stratified layered structure (e.g. Christensen and Mooney, 1995; Rudnick and Gao, 2003; Huang et al., 2013; Hacker et al., 2015). Unlike the relatively well-constrained upper crust, the nature of the mid and lower continental crust is still unclear as direct observations are limited. Seismic imaging serves as a unique and robust approach to pose constraints on the elastic properties of Earth's inaccessible interior.

Recent development in seismic imaging techniques (Shapiro et al., 2005; Langston and Liang, 2008; Lin et al., 2012) and installation of massive seismic arrays (e.g. *EarthScope* USArray Transportable Array) provide an opportunity for obtaining high-resolution, comprehensive seismic constraints on the structure and properties of Earth's continental lithosphere. Successfully leveraging these techniques and multiple datasets requires the simultaneous application of complementary seismic observables, as well as the development of joint inversion techniques that can handle flexible

parameterizations and disparate noise characteristics of the datasets. Moreover, to further incorporate the seismic constraints with geochemistry, geodynamics, and petrology studies, a single best-fit model from classical seismic inversion scheme without accurate quantification of uncertainties is insufficient.

I hypothesize that by applying Bayesian inference into the seismic inversion, the representative sample of the set of acceptable models given the data – which we call the ensemble solution – will allow us to quantify the uncertainties of the inversion result. Recent studies have also suggested that transdimensional sampling methods, in which the number of parameters describing the interior itself is treated as unknown, are advantageous in problems with multiple parameter types, which are common in seismology.

I will briefly introduce the transdimensional Bayesian (TB) method for seismic inversion at the beginning of this dissertation. I will then demonstrate the application of TB inversion two seismic observables: surface wave dispersion and Rayleigh wave ZH ratio. Surface wave dispersion is the measurement of propagation velocity for seismic surface waves. The frequency dependence of their sensitivity to Earth's structure allows them to be used to constrain Earth's subsurface structure. Rayleigh wave ZH ratio is a measurement of the ellipticity of its particle motion, being the ratio of the vertical to horizontal amplitude of particle motion for a Rayleigh wave. It is shown to be sensitive to elastic properties and density in Earth's crust. While inversion of these two data types has been widely used for constraining the lithosphere structure, particularly since the emergence of ambient seismic noise measurements, systematic

studies of the model parameter uncertainties and trade-offs are limited. The TB approach is well suited for such quantification due to its less restrictive parameterization and probabilistic framework.

In Chapter 2 of this dissertation, I focus on three aspects of seismic inversion: data sensitivities, assumed scaling among parameters (compressional wave speed, V_P, shear wave speed, V_S, density and radial anisotropy) and parameterization choices. In an ideal case, the inversion results should be dominated by data sensitivity without introducing potential bias. We show that under realistic circumstances, while surface wave data provide relatively strong constraints on the posterior distribution of V_S and, to a lesser extent, V_P, common parameterization choices can potentially bias structure estimates. This is particularly the case for radial anisotropy (ξ) – which refers to transverse isotropy with a vertical axis of symmetry and no azimuthal dependence – due to the inability to distinguish variations of V_P and density from those of ξ. Inferred results therefore depend substantially on the parameterization and scaling choices. Two types of model parameterization are proposed for TB inversion involving multiple types of parameters. I demonstrate that by implementing an independent parameterization for different physical quantities, the inversion can avoid imposing identical model geometry across multiple types of model parameters, and obtain better model estimates with reduced trade-offs. I advocate for such a parameterization in TB inversion of radial anisotropy using surface wave data, and when targeting disparate V_P and V_S structures such as those associated with the α-β quartz transition.

As a short note on material in Chapter 2, I document the technical details of my implementation for predicting surface wave dispersion and ZH ratio for a layered model based on the work by Lai and Rix (1998). The MATLAB-based program improves the prediction compared to standard techniques by taking into account the sphericity and radial anisotropy of the Earth.

To further constrain the sharp change of seismic velocities within the continental lithosphere, receiver functions are included for a joint inversion together with surface wave dispersion. Receiver functions are time series computed from three-component seismograms that are related to the seismic response of Earth structure near the receiver. Joint inversion including receiver function data has recently proliferated in crustal seismology due to the development of model-space search approaches. There are several challenges to the application of joint inversion to data from large seismic arrays that span across various tectonic settings. A key complication to the interpretation of seismic data arises due to the heterogeneous near-surface sedimentary layers. This issue has been addressed in the literature by imposing stronger prior information of the inversion (Shen et al., 2013, 2016) or by degradation of receiver function data (Chai et al., 2015). However, there remain substantial published studies that do not address or discuss the issue directly, despite being situated in areas where sediment-related contamination of receiver functions is to be expected.

In Chapter 4 of this dissertation, I investigate the reliability of crustal structure inferences using an implementation of transdimensional Bayesian joint inversions. I

show that reverberations from shallow layers such as sedimentary basins produce spurious low-velocity zones when inverted together with surface wave data of insufficiently high frequency. Therefore, reports of such layers in the literature based on inversions using receiver function data should be re-evaluated. I demonstrate that a simple resonance-removal filter can suppress these effects and yield reliable estimates of crustal structure, and I document the details of its implementation in receiver-function based inversions.

Chapter five of this dissertation is an application of transdimensional Bayesian joint inversion of a large seismic array based on the theoretical development of the preceding chapters. I focus on the lithospheric seismic imaging of the Northern Great Plains (NGP) of the United States using Transportable Array data. I present a shear velocity model – THB-NGP – for the top 100 km of the Northern Great Plains following accurate characterization of the sedimentary basin. Seismic velocity structures of different tectonic settings are quantitatively analyzed and then incorporated with lab measured wave speed to infer the composition of the NGP mid-lower crust. Together with previous geochemical, petrological studies of the Northern Great Plains crust, I discuss implications of our work for the formation and evolution of the Archean crust and the Paleoproterozoic orogenic belts.

Chapter 2: Consequences of Parameterization Choices in Surface Wave Inversion: Insights from Transdimensional Bayesian Methods

Gao, C., & Lekić, V. (2018). Consequences of parameterization choices in surface wave inversion: insights from transdimensional Bayesian methods. Geophysical Journal International, 215(2), 1037-1063. https://doi.org/10.1093/gji/ggy310

Abstract

Inversion of surface wave data for crustal and upper mantle structure is a staple of passive seismology, particularly since the advent of techniques enabling surface wave dispersion and Rayleigh wave ellipticity measurements from ambient noise. Recent development and application of transdimensional Bayesian (TB) seismic inversion offers an approach to quantify model parameter uncertainties and trade-offs with fewer assumptions than traditional methods. Using synthetic tests, we investigate choices in the implementation of TB for the inversion of surface wave dispersion and Rayleigh wave ellipticity to constrain the structure of Earth's continental lithosphere. We focus on three aspects of the inversion: limitation of data sensitivity, assumed scaling among parameters (compressional wavespeed, Vp, shear wavespeed, Vs, density, radial anisotropy), and parameterization choices. We show that while surface wave data provide relatively strong constraints on the posterior distribution of Vs and, to a lesser extent, Vp, common parameterization choices can potentially bias structure estimates. This is particularly the case for radial anisotropy (ξ), due to the inability to distinguish variations of Vp and density from those of ξ. Inferred results therefore depend substantially on the parameterization and scaling choices. We illustrate how layered parameterizations can, in the TB framework, recover smoothly-varying

profiles, and quantify the number of layers recoverable at different levels of measurement uncertainty. Finally, we propose two types of model parameterization for TB inversion involving multiple types of parameters. We demonstrate that by implementing an independent parameterization for different physical quantities, we can avoid imposing identical model geometry across multiple types of model parameters, and obtain better model estimates with reduced trade-offs. We advocate for such a parameterization in TB inversion of radial anisotropy using surface wave data, and when targeting disparate Vp and Vs structures such as those associated with α-β quartz transtion.

2.1 Introduction

Seismic surface waves are strongly excited by shallow earthquakes, and more easily recorded at large epicentral distances compared to body waves due to lesser geometrical spreading. Because their sensitivity to structure depends on frequency, their propagation velocity does as well. Therefore, measurements of surface wave dispersion (SWD) provide constraints on crustal and upper mantle structure with excellent global coverage and high lateral resolution (e.g. Romanowicz, 2002). Indeed, seismic tomography based on surface wave dispersion has been widely used to constrain the 3-D shear wave velocity in crust and upper mantle both on a global scale (Montagner and Tanimoto, 1991; Trampert and Woodhouse, 1996; Shapiro and Ritzwoller, 2002; Ekstrom et al., 2014; Pasyanos et al., 2014) and regional scale (e.g. Simons et al., 2002; Huang et al., 2003; Yao et al., 2006, 2008; Lin et al., 2008; Wagner et al., 2012). Differences between the dispersion of Rayleigh and Love waves led to the discovery of radial anisotropy in the upper mantle (Anderson, 1961; Aki, 1968), and are now routinely used to constrain profiles and lateral variations of radial anisotropy of Earth's crust and upper mantle (Panning and Romanowicz, 2004; Kustowski et al., 2008; Ferreira et al., 2010). The particle motion of fundamental mode Rayleigh wave is elliptical. The measurement of the ratio of the vertical to horizontal amplitude of particle motion (ZH ratio) can be shown to be sensitive to elastic properties and density in the crust, and has also been used to constrain shallow Earth structure (Boore and Toksoz, 1969; Tanimoto and Rivera, 2008; Yano et al., 2009; Lin et al., 2014).

Inferring Earth properties from seismic data is a non-unique inverse problem because seismic observables provide only limited constraints (Franklin, 1970). Love and Rayleigh waves depend on density and 13 independent elastic parameters (e.g. Chen and Tromp, 2007), which can vary laterally and with depth. Therefore, inversions of SWD and ZH ratio data inherently involve significant non-uniqueness due to trade-offs among model parameters; when linearized and posed in terms of matrix inversion, the large number of model parameters yields singular matrices requiring regularization for a solution to be obtained. To best represent the Earth's structure given the available seismic observation, certain assumptions are often made to simplify the model. For example, the structure may be assumed to be layered or to vary smoothly with depth (e.g. Constable et al., 1987). Within each layer, the elastic properties might be assumed to be isotropic, so that they can be described with just three parameters: density (ρ), shear (Vs) and compressional (Vp) wavespeed. Another common assumption is that of radial anisotropy (transverse isotropy), which involves three additional parameters: the squares of the ratios of wavespeeds of horizontally and vertically-polarized waves, $\xi = (\mathrm{V}_{SH}/\mathrm{V}_{SV})^2$ and, $\varphi = (\mathrm{V}_{\mathrm{PH}}/\mathrm{V}_{PV})^2$ respectively, as well as a parameter, η, describing wavespeeds at intermediate directions.

Even with these simplifying assumptions, constraints provided by SWD data are insufficient to reliably infer all the model parameters, particularly those to which the seismic observables are weakly sensitive – such as Vp, φ, η, and ρ. Based on empirical trends, variations in these parameters are often scaled to variations in better-resolved parameters such as Vs (e.g. Brocher, 2005) and ξ (e.g. Montagner and Anderson, 1989). Indeed, φ, η, are assumed and remain a workhorse of structural

seismology. Recently, global tomographic models in which the scaling relationships are allowed to vary with depth or laterally have also been performed (e.g. Simmons et al., 2009; Moulik and Ekstrom, 2016).

Standard inversion approaches are ill-suited for studying how parameterization choices and scaling assumptions affect the accuracy of seismic structures inferred from surface wave data. To start with, a certain amount of regularization – in the form of smoothing, damping, or a priori covariance among parameters – must be imposed to obtain a solution in the first place. Uncertainty analysis developed for linear(izable) problems can be applied (e.g. Backus and Gilbert, 1967; Tarantola and Valette, 1982) to study the tradeoffs between inferences of Vs, Vp, and density. However, SWD and ZH ratios depend on elastic properties in a non-linear way; their sensitivity to a parameter of interest can depend on the value of that and other parameters. Therefore, these linear approaches may not be appropriate. Even if they were appropriate, the analysis may depend on the parameterization – for example, for different choices of number and thickness of layers – limiting their generalizability. On the other hand, model space search methods do not require an inversion to be performed, and offer the potential to quantify the uncertainty of inferences even in highly non-linear problems (Mosegaard, 1998). Yet, application of these approaches to over-parameterized problems is stymied by the curse of dimensionality (e.g. Tarantola. 2005), and has, until recently, required the parameterization to be chosen prior to inversion.

To better represent the uncertainties from seismic imaging results and to incorporate complementary seismic observables with increasingly available measurements, seismic transdimensional Bayesian (TB) inversion has been developed

(Malinverno, 2002; Bodin et al., 2009; Agostinetti and Malinverno, 2010; Bodin et al., 2012). Under a Bayesian framework, all information is described in terms of probabilities. This allows for data uncertainties as well as prior assumptions about model parameters to be explicitly accounted for. Furthermore, since Bayesian inversion seeks an ensemble solution instead of a single best-fitting model, quantification of uncertainties of inferred model parameters and correlations between them is relatively straightforward. In contrast to traditional inversion methods, which treat the number of model parameters as a constant chosen prior to inversion, transdimensional inversion includes it as an unknown determined by the data (Sambridge et al., 2013). With a more flexible model parameterization, transdimensional inversion also more easily accommodates multiple data types with different, and therefore complementary, sensitivities to the seismic structure.

The TB method offers a new opportunity to quantify effects of parameterization choices and assumptions of scaling among parameters, enabling a reassessment of uncertainties in surface wave dispersion and insight into outstanding questions, such as the origin of the relatively poor agreement among radially anisotropic global shear velocity models (Chang et al., 2015). Under a TB framework, we can eliminate scaling assumptions as well as assumptions concerning the number and thickness of structural layers, while simultaneously constraining multiple model parameters to various degrees.

In this paper, we use TB inversion to systematically explore the ability of surface wave dispersion and ZH ratios to constrain profiles of Vs, Vp, ρ, and ξ beneath a seismic station, under various model parameterization choices. Synthetic data of

these two observables are inverted individually and jointly to investigate the complementarity of data sensitivity, the consequences of parameterization choices, and the influence of assumptions about scaling relationships between physical quantities.

2.2 Method

2.2.1 Seismic Bayesian Inversion

Bayes' theorem (Bayes and Price, 1763) relates the probability (p) of an model (m) conditional on a dataset (d), written as $p(m|d)$, to the probability of observing the dataset given a model, i.e. $p(d|m)$:

$$p(m|d) = \frac{p(d|m) \times p(m)}{p(d)} \quad (2.1)$$

Here the model is represented by a vector quantity that includes all the model parameters of interest. Similarly, all observed data comprise a vector d. The aim of Bayesian inference is to quantify the posterior probability density $p(m|d)$, which is the probability density of the model parameter given the observed data (Smith 1991). The term $p(d)$ is called the evidence. Notice that $p(d)$ is not a function of m, and should remain constant as we vary the model parameter under the same setting, allowing us to write:

$$p(m|d) \propto p(d|m) \times p(m) \quad (2.2)$$

Seismic inversion is the procedure of using the measurements made on seismic records (i.e. data, d) to infer a model (m) that quantitatively describes the Earth's, typically inaccessible, interior. In such case, the posterior is the probability of certain

seismic structure given the observation and the prior. It is therefore proportional to the product of likelihood -- the probability of observing the data given a seismic structure – and the prior probability on the model parameters. While the forward problem of predicting the outcome of some measurements given a complete description of the physical system has a unique solution, the inverse problem does not. This non-uniqueness arises both from data measurement errors and the insufficiency of information contained in the data. Unlike many common approaches to solving such inverse problems, which seek to reduce the non-uniqueness by introducing prior information in the form of smoothing or damping operators (e.g. Constable et al., 1987; Menke, 2012), the Bayesian approach embraces the non-uniqueness and represents it in probabilistic terms.

The prior

In the Bayesian framework, the prior information $p(m)$ is used to describe our knowledge about the parameters that describe the model prior to introducing data (Sivia, 2006). If the parameters that we are interested in inferring correspond to an unknown number *n* of physically non-overlapping regions, the prior can be separated into two terms:

$$p(m) = p(m_e, n) = p(m_e|n) \times p(n) \quad (2.3)$$

Here m_e stands for the parameters describing the seismic structure (Vp, Vs, density and the physical location of the regions). We use a uniform distribution for *n* over the interval $I = \{n \in N \mid n_{min} < n \leq n_{max}\}$. Hence,

$$p(n) = \begin{cases} \frac{1}{\Delta n} & if\ n \in I \\ 0 & otherwise \end{cases} \quad (2.4)$$

where $\Delta n = (n_{max} - n_{min})$.

In this study, the target model is parameterized in depth using Voronoi nuclei (Aurenhammer, 1991); the region nearest to a given Voronoi nucleus is defines a layer of constant elastic parameters specified for that Voronoi nucleus. The boundary between adjacent layers is defined as the midpoint between two Voronoi nuclei (See Figure 2.1). Since we are interested in the profiles of multiple types of parameters, we propose two different parameterization schemes, illustrated in Figure 2.1.

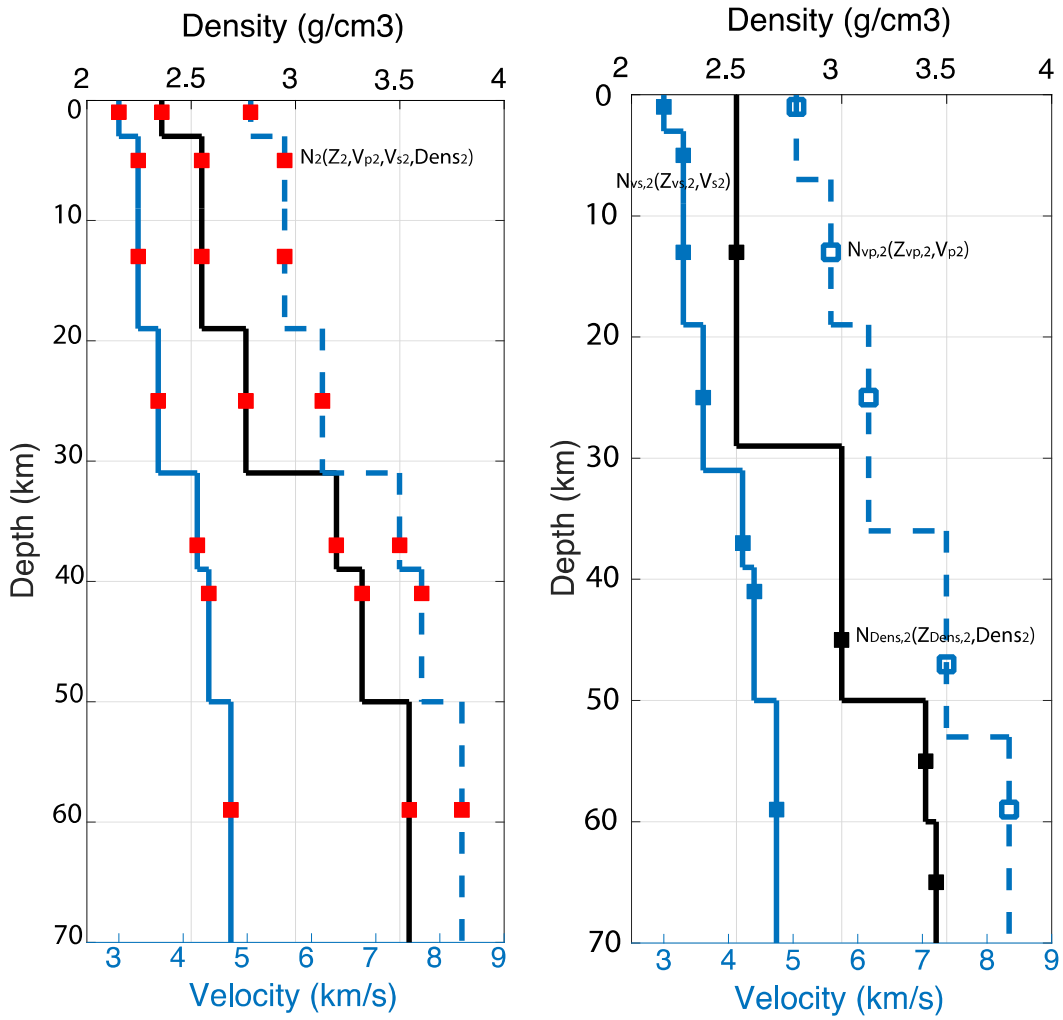


Figure 2.1 Two schemes for using Voronoi nuclei (squares) to characterize a 1D isotropic seismic velocity model (Vs – blue, Vp – dashed, ρ – black). In the "attached" scheme (left panel), each Voronoi nucleus carries all three parameters (Vs, Vp and density). In the "independent" scheme (right panel), the three types of parameters are carried by independent sets of Voronoi nuclei. The 2nd Voronoi nucleus is labeled in both panels to illustrate differences between the two parameterization schemes. The model shown in left panel is the target model used to generate synthetic data inverted in later sections.

For the first type of parameterization, we allow each Voronoi nucleus to specify all types of parameters. We call this type of parameterization "attached". Given a number of cells n, the probability distributions for the $4 \times n$ parameters, 1D

Voronoi nucleus position (z), shear velocity (Vs), compressional velocity (Vp) and density (ρ) are assumed to be independent from each other, and so can be written as:

$$p(m_e|n) = p(z|n) \times p(v_s|n) \times p(v_p|n) \times p(\rho|n) \quad (2.5)$$

To minimize the amount of prior information introduced, we assume uniform distributions over specific intervals. For example, if we define $J_s = \{v_{s,i} \in \mathcal{R} \,|V_{min} < v_i < V_{max}\}$ we have:

$$p(v_{s,i}|n) = \begin{cases} \frac{1}{\Delta v_s} & if\ v_{s,i} \in J_s \\ 0 & otherwise \end{cases} \quad (2.6)$$

where $\Delta v = (V_{max} - V_{min})$. Since the shear velocity in each Voronoi nucleus is assumed to be independent (i.e. no smoothing is imposed),

$$p(v_s|n) = \prod_{i=1}^{n} p(v_{s,i}|n) \quad (2.7)$$

Similarly, we can write:

$$p(v_p|n) = \prod_{i=1}^{n} p(v_{p,i}|n) \quad (2.8)$$

$$p(\rho|n) = \prod_{i=1}^{n} p(\rho_i|n) \quad (2.9)$$

For a 1D-layered model, the possible positions of the Voronoi nuclei are distributed along depth. If we assume that there are *N* possible positions for *n* Voronoi nuclei, there are then $\frac{N!}{n!(N-n)!}$ possible configurations. Again, we assign an equal probability to each of the configurations, and can then write:

$$p(z|n) = \left[\frac{N!}{n!\,(N-n)!}\right]^{-1} \quad (2.10)$$

Combining together equations (2.3) (2.4) (2.5) (2.6), the full prior probability density function can be written as:

$$p(m) = \begin{cases} \dfrac{n!\,(N-n)!}{N!\,(\Delta v_s)^n(\Delta v_p)^n(\Delta\rho)^n\Delta n} & if(n \in I \; and \; \forall i \in [1,n], v_{si} \in J_s, v_{pi} \in J_p, \rho_i \in J_\rho) \\ 0 & otherwise \end{cases} \quad (2.11)$$

For the second type of parameterization, we assign independent sets of Voronoi nuclei to each type of parameter. Hence we call this type of parameterization "independent". In this case, we have

$$p(m_e|n) = \prod_{j=1}^{3} p\left(z|n_j\right) \times p\left(m_{e,j}|n_j\right) \quad (2.12)$$

Here $m_{e,j}$ stands for the elastic parameters (Vs, Vp, and ρ) the Voronoi nuclei carry. Unlike the first scheme, the number and the position of the Voronoi nuclei are independent from each other for different elastic parameters. In this way, we do not force all type of elastic parameters to be attached to a single Voronoi nucleus, which ideally will allow a more flexible parameterization. Similarly, we have

$$p\left(m_{e,j}\right) = \frac{n_j!\,(N-n_j)!}{N!\left(\Delta m_{e,j}\right)^{n_j}\Delta n_j} \quad if\ (n_j \in I \; and \; \forall i \in [1,n_j], m_{e,j} \in J) \quad (2.13)$$

Each of the two types of parameterization has certain advantages for particular problems; we will further explore this in this paper in the joint inversion of SWD and ZH ratios in the discussion section. Bodin et al., (2016) described an alternative type of parameterization where additional parameters constraining anisotropy are proposed on existing isotropic shear velocity structures, and such proposed model is accepted

based on the constraints from data only. Since all the anisotropic parameters are proposed attaching to the velocity layers, the geometry of the anisotropic structure will depend on the velocity structure to some degree. This kind of parameterization appears to lie between the two we proposed in terms of the dependence among different types of parameters.

Likelihood function

The likelihood $p(d|m)$ quantifies how likely we would be to observe the data if the actual structure were described by the set of parameters in vector m. We use a least squares misfit function to describe the consistency between the predicted and observed data:

$$\Phi(m) = \left\| \frac{g(m) - d}{\sigma_d} \right\|^2 \quad (2.14)$$

where $g(m)$ is the predicted data and σ_d^2 is the estimated variance describing the data uncertainties. This misfit function is appropriate for data with normally-distributed errors, and yields the following likelihood:

$$p(d|m) \propto \exp\left(\frac{-\Phi(m)}{2}\right) \quad (2.15)$$

In the rest of this paper, we do not explicitly contaminate our synthetic data with noise because our likelihood function takes into account the effect of noise if we assume it to be uncorrelated across different periods, as is commonly done in the literature.

2.2.2 Transdimensional Sampling

Bayes' theorem quantifies how the posterior distribution is affected by the choice of the prior. The assumptions we make in formulating the inversion influence the outcome. In seismic inversion, assumptions about number of parameters are often made to fit the linearized inverse problem and to reduce non-uniqueness. These assumptions are often motivated by previous knowledge about the studied region. The risk in making these assumptions is that they could be biased or incorrect. The geophysical inversion literature abounds in examples in which the choice of the parameterization affects the inversion to different extents due to different degrees of correlation among model parameters. As an example, Trampert and Snieder (1996) showed how truncated expansions of basis functions could bias seismic tomography models. The motivation for applying a transdimensional sampling method into the inversion is to allow flexibility that does not require, but can nevertheless accommodate, strong prior assumptions about the model parameterization.

Allowing a flexible parameterization without any regulation may lead to another problem, where the model will contain complexities arising from attempting to fit details of the data as closely as possible. Since the data we measure contains error due to both instrumental and environmental noise sources, fitting the detailed data is ill-advised. The Bayesian formulation of model selection is naturally parsimonious (Malinverno, 2000; Sivia et al., 2006); Malinverno (2002) showed that this is also the case with transdimensional Bayesian inversion. This means that if we have two competing models with different numbers of parameters that both fit the data equally well, the Bayesian formulation will favor the simpler model. Combining the

transdimensional sampling method and the Bayesian framework, transdimensional Bayesian inversion therefore allows a more flexible parameterization with fewer assumptions made.

2.2.3 Reversible Jump Markov Chain Monte Carlo For Multi-Parameter Seismic Structure

We apply a reversible-jump Markov chain Monte Carlo algorithm to carry out the transdimensional Bayesian inversion. The Markov chain Monte Carlo (McMC) is an iterative algorithm that draws random steps from a desired distribution; with sufficient number of iterations, the models are sampled proportional to their posterior probability, p(m|d). The rjMcMC algorithm consists of two stages, proposing a new model (m') by perturbing the current model (m) and deciding accepting or rejecting it.

In a case of transdimensional sampling, the acceptance probability is:

$$\alpha = \min\left[1, \frac{p(m')p(d|m')q(n,m|n',m')}{p(m)p(d|m)q(n',m'|n,m)}|\boldsymbol{J}|\right] \quad (2.16)$$

Note that here the proposal ratio is different than in the fixed parameterization case. $\boldsymbol{J}$ is the Jacobian matrix of the transformation from *m* to *m'*. It is needed to account for the scale change only when there is a dimension change during the sampling process (Green 2003). In our case of discrete Voronoi positions, $|\boldsymbol{J}|$ equals to one (Bodin et al., 2009). Therefore, the Jacobian is unity for each case of the rjMcMC sampling process and can be ignored.

An important part of designing an rjMcMC is choosing how to perturb the current model m into m' with some randomness, i.e. how to efficiently sample the parameter space. A schematic representation of our rjMcMC algorithm is shown in

Figure 2.2. Following the approach of Bodin et al., (2009), we perturb the current model by randomly choosing one the four options with equal probability. However, since we propose two parameterization schemes for dealing with the multi-parameter seismic structure, the Markov chain could behave differently, especially when dimension changes are involved. We derive the relevant expressions in the Appendix.

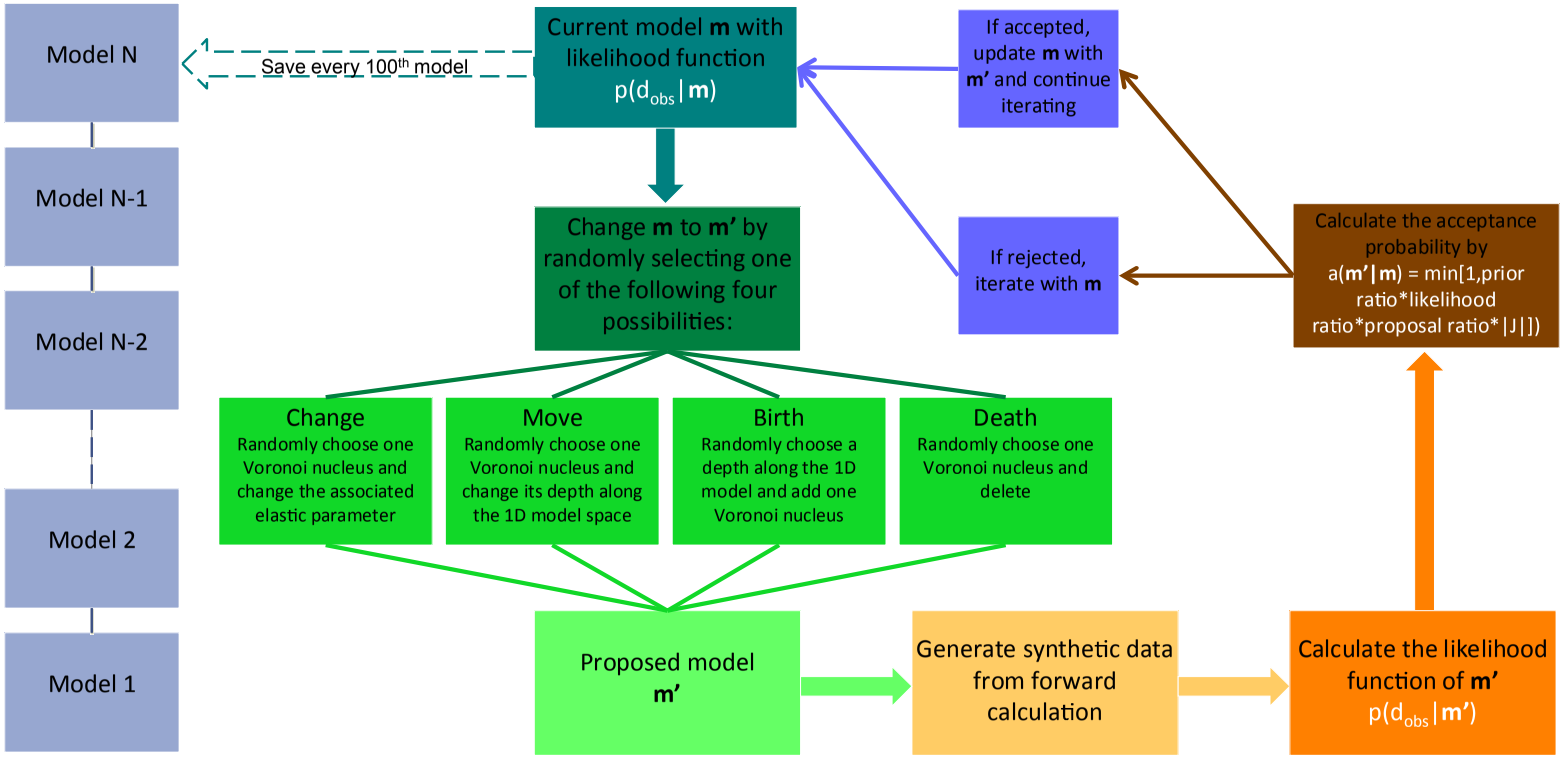


Figure 2.2 A schematic overview of the reversible-jump Markov chain Monte Carlo algorithm used for inversion. Every 100th model is saved to the ensemble solution to increase the independence of sampled model.

2.2.4 Target Model and Forward Problem

To test the performance of the transdimensional Bayesian joint inversion, we choose a realistic isotropic, layered target model (left panel of Figure 2.1) for the synthetic tests. The target model has a 3 km thick sedimentary layer, underlain by a two-layer crystalline crust with Moho at 31 km depth. The upper mantle shallower than 70 km is represented by three layers with increasing velocity. The velocities remain constant below 70 km. In the target model, Vs, Vp and density follow the empirical relations from Brocher (2005). The target model is designed in this way for the convenience of later discussion of scaling relationship effects. To predict SWD and ZH ratios, we use the reflectivity method (Aki and Richard, 2002; Hisada, 1994) to solve the eigenvalue problem for both Rayleigh wave and Love wave in an elastic, vertically heterogeneous medium, based on the implementation of Lai and Rix (1998). Later in the discussion about TB inversion of radial anisotropy, we modify the forward code to compute the surface wave dispersion given elastic parameter A, C, N, L, and F according to Harkrider and Anderson (1962) and Bhattacharya and Arora (1997). The code also takes into account the sphericity of the Earth based on the formulation of Bhattacharya (1996). The technical details of the implementation can be found in chapter 3 of this dissertation. We validate our implementation by comparing our predictions to those from MINEOS (Masters et al., 2011) for the upper 200 km of the PREM model (Dziewonski and Anderson, 1981). Given the same period range, our implementation costs around 0.1 second to predict SWD, which is much faster than MINEOS. All of the software is written in MATLAB.

2.3 Results

2.3.1 TB Inversion of SWD

We perform transdimensional Bayesian inversion of both Rayleigh and Love wave dispersion data computed from synthetic input structures in the 5 to 100 sec period range. This period range covers both the ambient noise data range and part of the teleseismic data range. For teleseismic earthquake data, the surface wave dispersion data are usually measured between approximately 30 s and 250 s (Laske and Masters, 1996; Ekstrom et al., 1997; van Heijst and Woodhouse, 1999; Boschi and Ekstrom, 2002; Trampert and Woodhouse, 2003; Ekstrom 2011; Ma et al., 2014). The dispersion data below 25 second period are relatively difficult to measure from teleseismic data due to scattering and potential for cycle skipping. Dispersion measurements made on ambient noise correlations are typically in the ~5 to 40 s range (e.g. Ekstrom, 2014). The combination of these two period ranges comprehensively constrains Vs in the crust and upper mantle. We choose to not include dispersion data at periods larger than 100 s, because their primary sensitivity is below the depth range of interest in this manuscript (upper 70 km). The uncertainty of surface wave dispersion measurements can be affected by data quality (e.g. signal-to-noise ratio), data coverage (e.g. distribution of earthquakes and stations), and measurement method (e.g. whether or not the smoothness of the dispersion curves is exploited). We assign a realistic 3% uncertainty to the dispersion measurements at each period, and assume that measurements at different frequencies are uncorrelated. It should be noted that the

assumption of uncorrelated data uncertainty, while ubiquitous in the literature, warrants further investigation.

In this test we only invert for Vs due to the limited constraints surface wave dispersion data have for Vp and density. We assign an uniform prior between 2.5 km/s and 5.5 km/s for Vs. Bodin et al., (2009) suggested that when the data constraints are strong enough, the choice of the broad prior has little effect on the posterior. Our synthetic tests with different priors for Vs support such conclusion. Meanwhile, the choice of prior range for weakly constrained model parameters such as density during SWD inversion could significantly affect the posterior. For example, when inverting for Vs and density using SWD with independent parameterization, we find that broadening the prior range twice for density would resulting the preferred number of layer from the posterior to change from 5 to 3, while broadening the Vs prior range does not change the preferred number of layer from the posterior. During the inversion, the Vp and density are calculated using the empirical relations according to Brocher (2005), as used in the synthetic model (see Figure 2.1). We want to point out that assuming a scaling relationship imposes additional prior information, requiring fewer model parameters to be inverted for. When the scaling relation is properly chosen, it will help reduce the variation in the posterior yielding tighter constraints on parameters of interest. Theoretically, an incorrect scaling relation on Vp and Vs will bias the estimates of both parameters. However, our synthetic tests suggest that for typical Vp/Vs ranges for crustal studies, imposing incorrect Vp/Vs only biases Vp estimates while the Vs estimates remain indistinguishable from the posterior obtained with the correct Vp/Vs. We find this to be the case even when only Rayleigh wave

dispersion is used in the inversion. We attribute this behavior to the much greater sensitivity of Vs compared to Vp for surface wave dispersion data. On the other hand, not assuming a scaling relation corresponds to a less informative prior; due to the naturally parsimonious nature of Bayesian inference, given the same observation, this will lead to a more simplified posterior. In later sections, we introduce more data types in joint inversions, allowing us to release Vp and density from the empirical scaling relations and allow them to vary independently.

The rjMcMC starts with a random initial structure. After a burn-in period during which the convergence is achieved, we save the accepted models into the ensemble solution, for a total of 5 million iterations. Due to the nature of the Markov chain, each time we perturb the current model, only a small part of the proposed model is different from the current model. Therefore, consecutive models are highly correlated, even when the acceptance rate is optimal. To increase the independence of the model ensemble, we choose to save every 100^{th} sampled model. We primarily rely on two approaches to assess the progress of the rjMcMC chain and to estimate the number of iterations needed to achieve convergence. First, we monitor how misfit evolves with iterations, making sure that it remains low. Second, we run several chains with different starting models and compare the statistical properties of the ensemble solutions obtained from each. For each chain's ensemble solution, we calculate the root-mean-square deviation to ensure that they are indistinguishable from one another after the burn-in period (See Figure 2.A1 and Figure 2.A2).

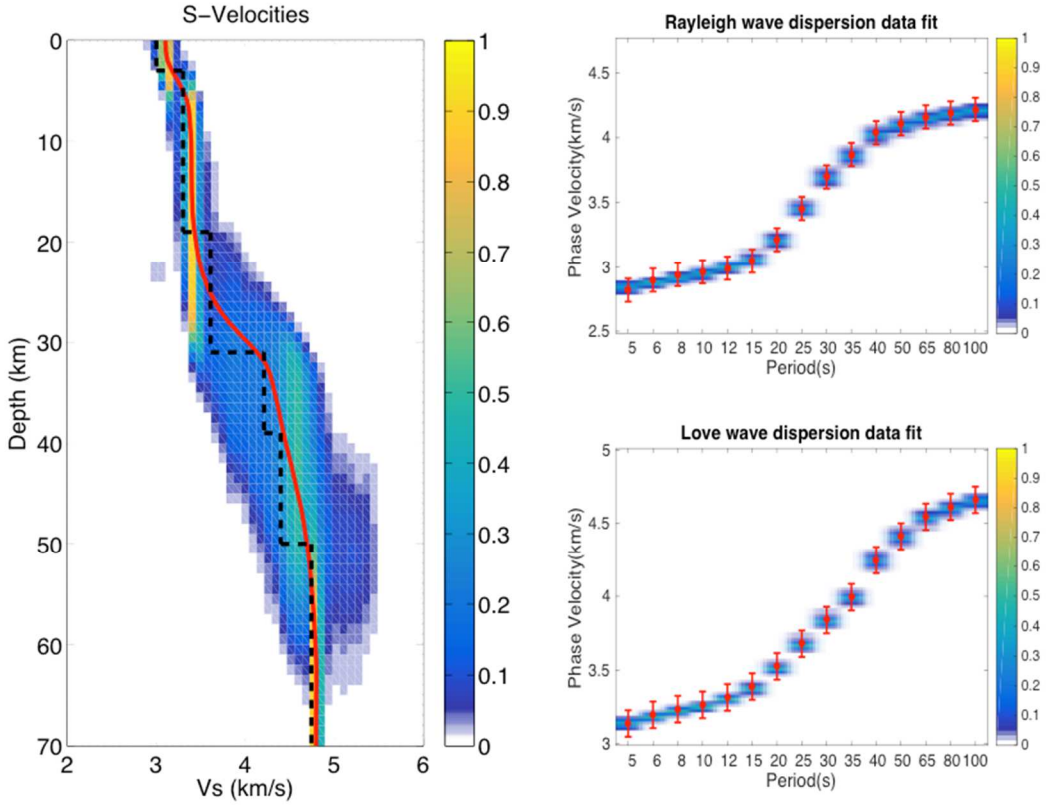


Figure 2.3 (left) Vs depth distributions retrieved using transdimensional Bayesian inversion of synthetic surface wave dispersion data. The ensemble solutions are displayed as probability density functions at each depth, with warmer colors corresponding to higher posterior probabilities, and the solid red line denoting the 5% trimmed mean of the posterior. The target model (black dashed lines) is used to generate synthetic data. Synthetic data (red) with specified 1σ uncertainties for Rayleigh (right top) and Love (right bottom) waves; data predicted by the ensemble solution plotted as probability density functions.

The posterior model density plot is shown in Figure 2.3. At every 1 km, we evaluate all the seismic velocities from the ensemble and normalize them to compute the probability density function (PDF). The PDF is represented so that warm colors correspond to higher probability and cool colors indicate lower probability. We want to point out that in such PDF plots, the absolute value of the probability is a function of the bin size used in plotting. Therefore, in some cases low absolute probability

density does not necessarily represent poor resolution. The target model is plotted as a black dashed line for reference. It can be seen that we are able to resolve absolute Vs at different levels along the depth range allowed in the inversion. We find that Vs is constrained better at shallow compared to deeper depths in terms of posterior variance, mainly due to the sharper sensitivity kernel for shallow Vs structure from short period surface wave dispersion. It can also been seen that while the SWD is able to constrain the absolute value of Vs, it tends to retrieve the sharp velocity jump in the target model as a smooth transition. This is expected from the fact that SWD measurements depend on the integral of elastic properties across a range of depths. The uncertainties of Vs below 50km seem to decrease with depth; we interpret this as a result of the fact that the period range of SWD we use here (up to 100 second) still has substantial sensitivity between 50km and 70km and only one single layer is preferred by the TBI at this depth range.

2.3.2 TB Inversion of ZH Ratios

Having explored the ability of surface wave dispersion constraining shear velocity structures, we turn our attention to ZH ratios, and perform a Transdimensional Bayesian inversion of ZH ratios computed for the same synthetic model shown in Figure 2.1. Due to difficulty of reliably measuring ZH ratios at long periods (Ferreira and Woodhouse, 2007), we restrict our attention to the 5-40 s period range, and assign 1% uncorrelated uncertainty to the observation at each period. In reality, depending on whether standard deviation of the repeat measurements or the standard deviation of the mean of the repeated measurement is used, the measurement error for ZH ratios could be as large as 3% to 10% (See Lin et al., 2012 and Lin et al.,

2014); however, to illustrate the ability of ZH ratios to constrain elastic properties, we choose a relatively small value of uncertainty that might be achieved under ideal circumstances.

Our initial tests show that, compared to SWD, ZH ratios have limited potential for constraining elastic properties deeper than 20 km. When we invert for Vs, Vp and density using an attached-type parameterization, without assuming scaling relationship among parameters, the retrieved seismic structure shows large variations along depth and absolute Vp, Vs and density values are systematically biased at most depths. Even when we restrict the parameter space to the upper 25 km, the ensemble solutions show that we are unable to resolve the profiles of Vs, Vp, and density simultaneously. Motivated by this finding, we invert only for Vs, and use empirical relations of Brocher (2005) to scale to Vp and density.

As with the ZH ratio inversion, the total number of rjMcMC iterations is 5 million, of which the first 2.5 million are the burn-in period, in all ZH ratio individual inversions in this section. The convergence rate of ZH ratios individual inversion is similar to the SWD inversion. With different starting models, the Markov chain is considered to have converged after about ~2×10^5 iterations. Therefore, we consider 2.5million iteration to be a safe choice for burn-in period. As is shown in Figure 2.4, the retrieved seismic structure is well constrained in terms of Vs, and fit to data is excellent.

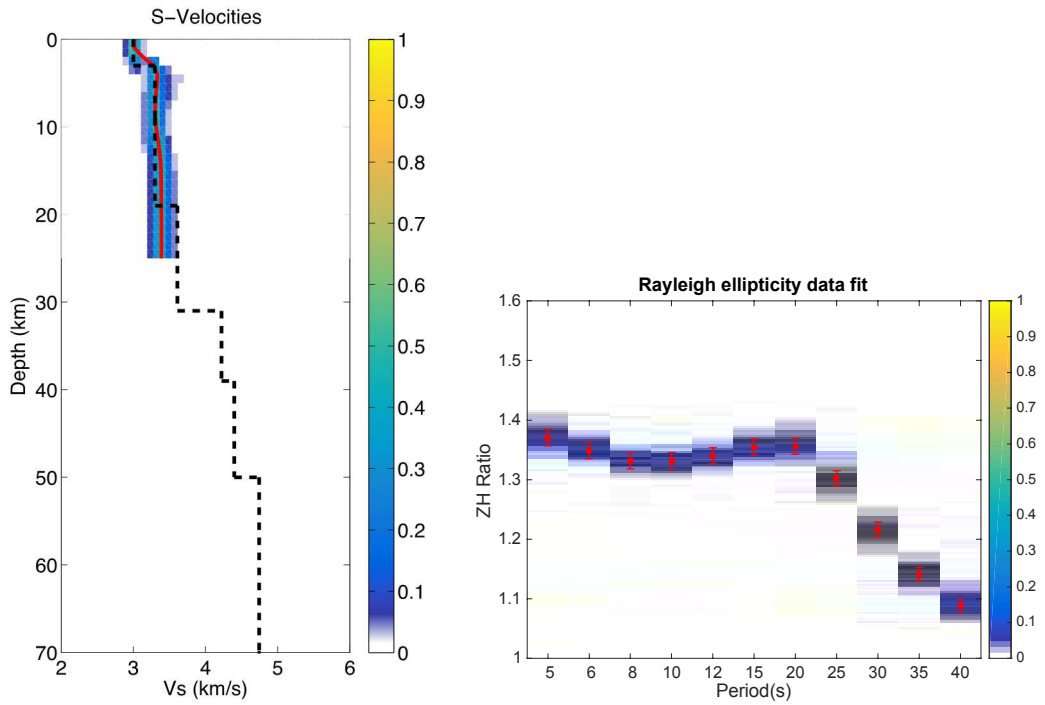


Figure 2.4 (left) Vs profile retrieved from transdimensional Bayesian inversion using synthetic ZH ratio data. The ensemble solutions are displayed as probability density functions at each depth, with warmer colors corresponding to higher posterior probabilities, and the solid red line denoting the 5% trimmed mean of the posterior. The target model (black dashed lines) is used to generate synthetic data. (right) Data fit of ZH ratio inversion. The red stars with their associated uncertainties (1σ error bars) are the synthetic data used as an input of the TBI, while the color tracks the probability density of the data predicted by the ensemble.

The TBI tests using ZH ratios are very informative. First, they show that compared to SWD, ZH ratios have limited ability to resolve structure below the crust. The shallow sensitivity of ZH ratios is well documented in the literature, with investigators usually using this data to constrain structure in the uppermost crust (e.g. Lin et al., 2012). Second, even though the strength of ZH ratios is their sensitivity to Vp, Vs and density, allowing a unconstrained inversion with all three types of parameters perturbed achieves very little resolution of structure due to nearly-total trade-offs among parameters. When we reduce the number of parameters by fixing

scaling relationships among them, we achieve better outcomes. Therefore, we should keep in mind that without additional constraints, it might not be practical to resolve an accurate and precise seismic structure from ZH ratios alone. These additional constraints could come either from the prior, such as by imposing scaling relationships among parameters appropriate for the geological setting of the inversion, or from the inclusion of other seismic observables to perform a joint inversion, such as surface wave dispersion, which is the next topic we turn our attention to.

2.3.3 TB Joint Inversion of SWD and ZH Ratios

In the previous sections, we showed that TB inversion is able to retrieve seismic structures with an adaptive parameterization using seismic observables one at a time. Here we conduct a transdimensional Bayesian joint inversion by combining the Rayleigh wave dispersion, Love wave dispersion, and ZH ratios. The motivation for doing joint inversion is to combine the strengths of different seismic data types to invert for a more comprehensive structure. The expected improvement in the ability to retrieve Vp and density structure also makes joint inversion a good example to illustrating the differences between the two types of parameterization proposed in this study.

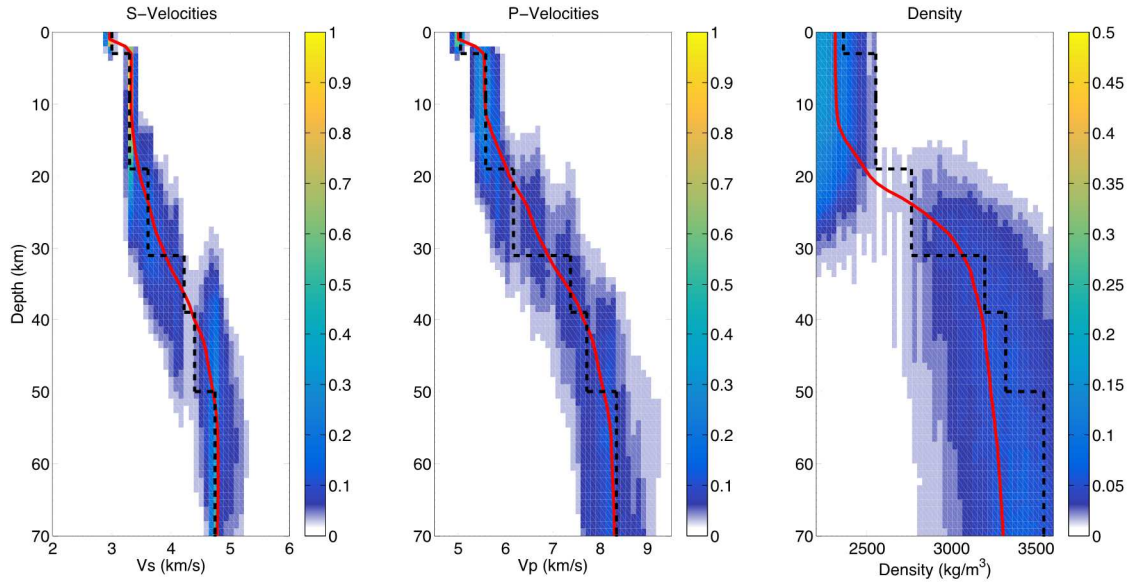


Figure 2.5 Vs (left), Vp (middle), density (right) retrieved from transdimensional Bayesian joint inversion using synthetic SWD and ZH ratio data. In this test all three types of parameters have independent parameterization. The ensemble solutions are displayed as probability density functions at each depth, with warmer colors corresponding to higher posterior probabilities, and the solid red line denoting the 5% trimmed mean of the posterior. The target model (black dash lines) is used to generate synthetic data.

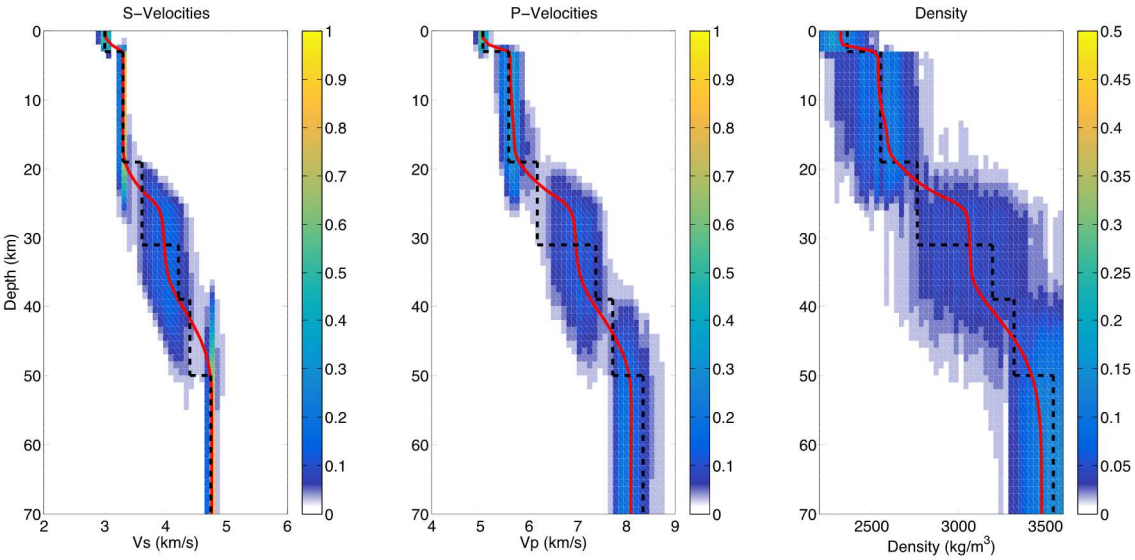


Figure 2.6 Vs (left), Vp (middle) and density (right) retrieved from transdimensional Bayesian joint inversion using synthetic SWD and ZH ratio data. In this test all three types of parameters share the same geometry. The ensemble solutions are displayed as probability density functions at each depth, with warmer colors corresponding to higher posterior probabilities, and the solid red line denoting the 5% trimmed mean of the posterior. The target model (black dash lines) is used to generate synthetic data.

We first invert for Vp, Vs and density using the independent type parameterization. To better illustrate the data sensitivity, we do not impose any scaling factors between Vs, Vp and density. We want to point out that throughout this study, when Vp is inverted, it is actually parameterized as Vp/Vs ratio. Inverting Vs and Vp/Vs ratio is equivalent to inverting Vs and Vp if given the same prior. As is shown in Figure 2.5, both Vs and Vp are well constrained above 20 km in terms of absolute value and variance compared to either of the individual inversion. When it comes to deeper structure, the variance of Vp increases significantly. Our calculation of normalized root-mean-square error for Vp and Vs ensemble (Figure 2.A3) also shows that Vs is better constrained than Vp at most of the depths. The behavior of the joint inversion is consistent with our expectations: SWD is able to constrain Vs in crust and upper mantle (with some sensitivity to Vp), while ZH ratios are able to constrain Vp and Vs in the crust. The preferred model from the density ensemble has two layers, while the Vs and Vp structures favor a five-layer model. This is because the much weaker data constraints on density yield density structures that are simpler than the actual target model or the retrieved Vp and Vs structures when inverted using an independent parameterization, in accordance with the lesser ability to resolve this parameter.

The attached type parameterization is much more common in the seismic literature on the inversion of surface wave data than is the independent type parameterization discussed above (e.g. Shapiro and Ritzwoller, 2002; Yao et al., 2008; Chai et al., 2015; Shen & Ritzwoller 2016). In Figure 2.6, we show the retrieved

structures using an attached type parameterization. The density ensemble shows a better fit to the true model, and contains more detailed structure. We should keep in mind that by using an attached type parameterization, we impose a more informative prior that all types of parameters share the same geometry. When such prior information is valid, we could expect a better-resolved structure. This also suggests that prior assumptions about co-variance of parameters should be justified before being applied to actual inversion because their effects are significant, particularly for ill-resolved parameters like density. To illustrate the potential pitfalls of using attached-type parameterization, we show an example where synthetic data in computed for a structure in which the density does not share the same geometry with Vs and Vp. When we use the attached type parameterization to perform a TB inversion of this data, the retrieved density structure exhibits artifacts that reflect the major features of Vs structure. This leads to a biased and misleading estimate of density (See Figure 2.A4).

2.4 Discussion

2.4.1 Transdimensional vs. Fixed-parameterization Inversion

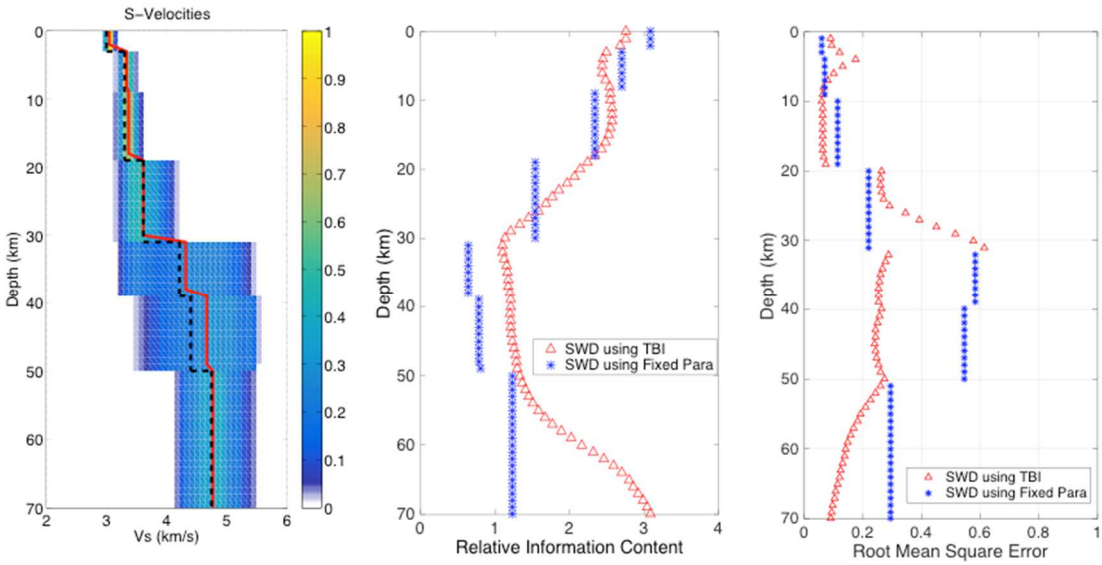


Figure 2.7 (left) Vs depth distributions retrieved using Bayesian inversion of synthetic surface wave dispersion data when parameterization is fixed to the same parameterization used to compute the synthetic dispersion data. The ensemble solutions are displayed as probability density functions at each depth, with warmer colors corresponding to higher posterior probabilities, and the solid red line denoting the 5% trimmed mean of the posterior. The target model (black dashed lines) is used to generate synthetic data. (middle) Kullback-Leibler divergence for TBI of surface wave dispersion (red triangle) and fixed-parameterization inversion of surface wave dispersion (blue star). (right) Root-mean-square error for TBI of surface wave dispersion (red triangle) and fixed-parameterization inversion of surface wave dispersion (blue star).

In Section 2.3.1, we showed that transdimensional Bayesian inversion can recover a Vs profile from surface wave dispersion measurements while treating the number of model parameters as an unknown. However, inversion of surface wave dispersion data is most frequently done with a fixed parameterization (e.g. Hermann, 2013). To gain insight into the relative advantages and disadvantages of a transdimensional inversion, we perform a Bayesian inversion with a fixed parameterization and compare our results to those obtained in Section 2.3.1. Using a

starting model with the correct geometry, we only perturb the shear velocity during the Markov chain Monte Carlo. The retrieved ensemble structure is showed in Figure 2.7. Like the transdimensional inversion, the fixed-parameterization inversion is able to recover the Vs. However, while both the absolute value and variance of Vs are well constrained at shallow depths (<31km), the variances at deeper depths increase significantly. We calculate the Kullback-Leibler divergence (KLD) for the posterior of both inversions with respect to their prior probability density function. The KLD for discrete probability distributions is defined as:

$$D_{KL}(P||Q) = \sum_i P(i)\, log \frac{P(i)}{Q(i)} \qquad (2.17)$$

The KLD from Q to P, denoted $D_{KL}(P\|Q)$, can be interpreted as the information gained when one revises one's beliefs from the prior probability distribution Q to the posterior probability distribution P (Kullback & Leibler, 1951). In Bayesian statistics, when P is the posterior and Q is the prior, KLD can be interpreted as a measure of the information gained from the data that transformed the prior distribution into the posterior distribution. The calculated KLD (Figure 2.7, middle panel) shows that the fixed-parameterization inversion produced posterior solutions containing less information at deeper depths compared to the transdimensional inversion. A similar conclusion can also be drawn based on the comparison of root-mean-square error of these two tests (see Figure 2.7 right panel), which shows that, at deeper depths, TBI yields lower errors than the fixed parameterization inversion.

The apparent superiority of TBI may be counterintuitive, since the fixed-parameterization inversion imposed a stronger and perfectly accurate prior (since the parameterization was fixed to that of the target model). Generally, the more restrictive the prior is, the less uncertain the posterior should be. However, in this case, due to the trade-offs between shear velocities at different depths, the variance of the posterior is larger than the one using TBI. We stress that the relative advantage of TBI over traditional fixed-parameterization inversion would be even greater in the more realistic scenario using actual, rather than synthetic, surface wave dispersion measurements, because in that scenario, the parameterization would not be known a priori. Moreover, the ensemble result from TBI allows statistical inferences of potential discontinuities along depths because of the less restrictive assumptions made on parameterization (Bodin et al., 2012).

2.4.2 Resolving Gradual Changing Seismic Structures Using Layered Parameterizations

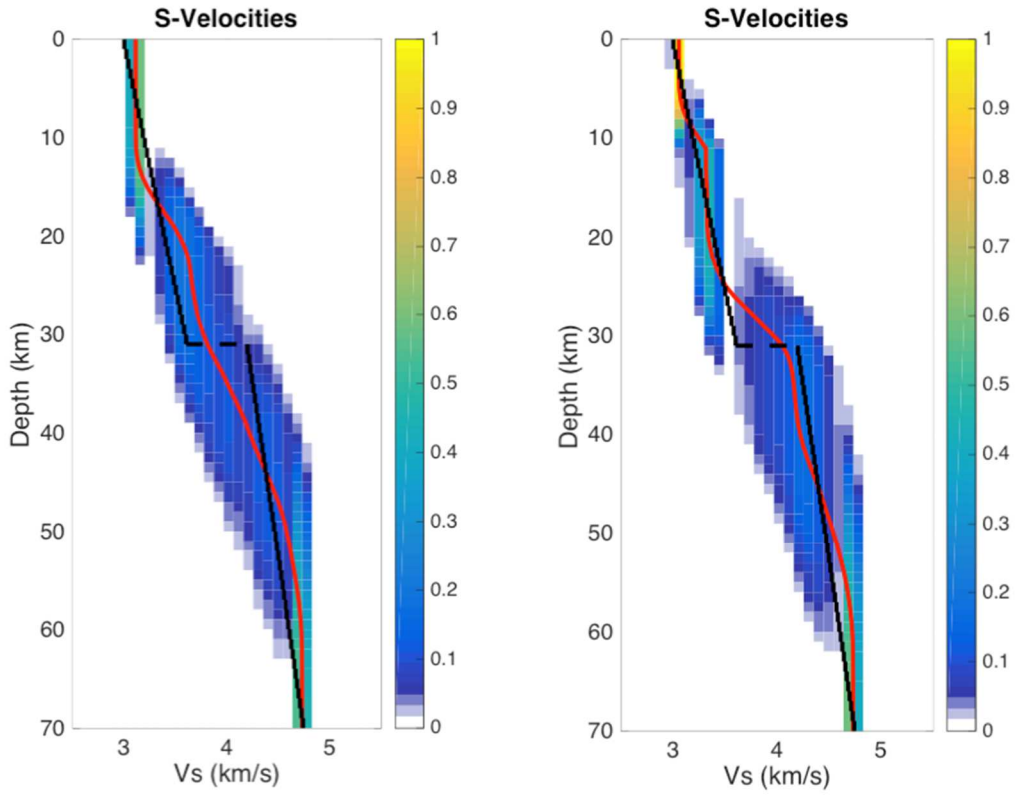


Figure 2.8 Vs depth distributions retrieved using transdimensional Bayesian inversion of synthetic surface wave dispersion data. The period range of SWD used is 5 to 100 s in the left panel and 2 to 100 s in the right panel. The ensemble solutions are displayed as probability density functions at each depth, with warmer colors corresponding to higher posterior probabilities, and the solid red line denoting the 5% trimmed mean of the posterior. The target model (black dashed lines) is used to generate synthetic data.

In our inversion for 1D layered seismic structure, we assumed that elastic properties remain constant within each layer. While such assumption is widely made in SWD inversion, gradient-based model parameterization has also been employed in the literature (Gosselin et al., 2017). Since our inversions are inherently parsimonious, this assumption sets up an inconsistency between layered profiles preferred by the prior information (via parameterization), and gradual ones potentially preferred by

realistic data. To test the implications of this assumption for resolving gradually changing velocity structures, we perform the following test. We compute synthetic Love and Rayleigh wave dispersion data in a target model with a 31 km thick crust, within which the Vs increases linearly from 3.0 km/s to 3.6 km/s. The shear velocity jumps from 3.6 km/s to 4.2 km/s at Moho, and then increases linearly to 4.74 km/s at the depth of 70 km. Vp and density are scaled to Vs, following the empirical relations from Brocher (2005) in both the target model and the later synthetic tests. We use both Rayleigh and Love wave dispersion data from 5 s to 100 s with uncertainties of 3% to invert for shear velocity, and obtain the Vs profiles shown in Figure 2.8 (left panel). We find that despite parameterizing the inversion with layers of constant properties, the ensemble solution partially resolves the gradually increasing velocity. The greatest exception occurs in the upper crust (above around 12 km), where the model ensemble shows a velocity jump overlaid by constant velocities. When surface wave dispersion data between 2 s and 5 s period are included, the model ensemble better resolves the gradually increasing velocity structure at the top of the crust (Figure 2.8, right panel). This test shows that the lack of data constraints in the uppermost crust is the main reason for the oversimplified structure. We can see a similar tendency of the inversion toward constant-velocity layers at the bottom of the model, where constraints from the data decrease once more. These tests suggest that throughout most of the crust, velocity gradients can be retrieved using TBI even with a layered parameterization; nevertheless, interpretations of gradients from inversions parameterized with constant-property layers should be cautious in areas where data constraints are less strong.

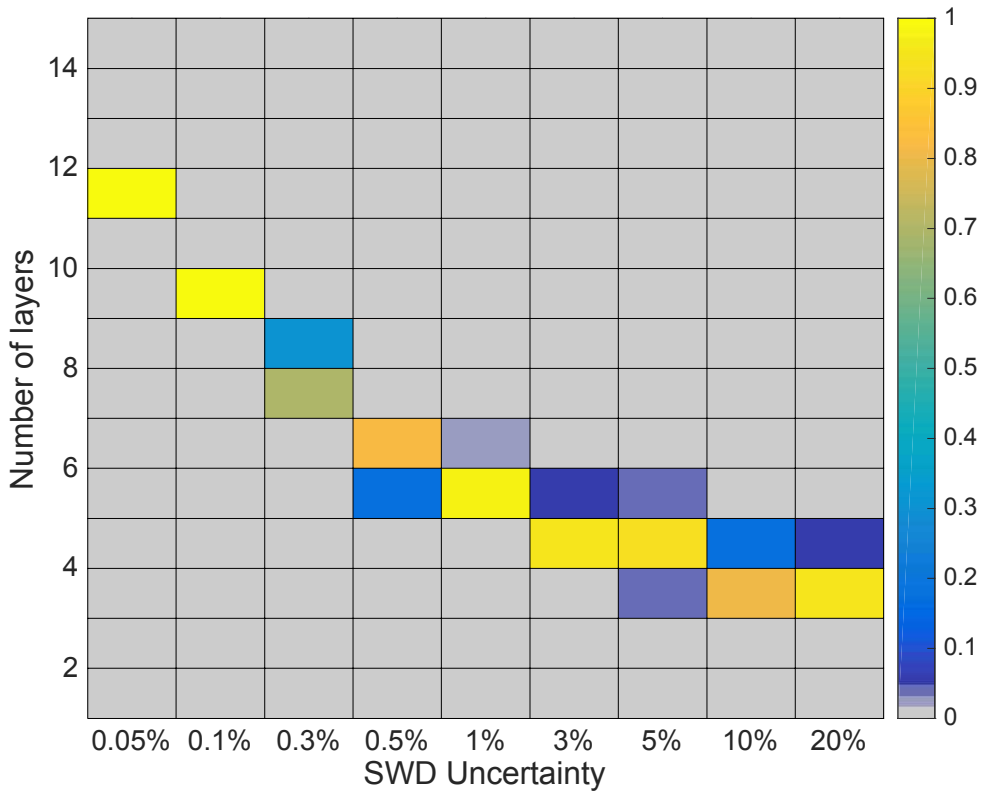


Figure 2.9 Posterior probability distribution of the number of layers in the retrieved ensemble for transdimensional inversion of surface wave dispersion measurements with different measurement uncertainties. The period range of both Rayleigh and Love wave dispersion used is 2 to 100 s.

Aside from the period range of SWD, a large measurement error could be another cause for the limited data constraint, because greater measurement error degrades the amount of information contained in the data. Here we test the effect of different measurement errors on the model complexity of the retrieved solution. We perform transdimensional inversion of 2 to 100 s Rayleigh and Love wave dispersion data with data measurement uncertainties ranging from 0.05% to 20%. The target model used to generate synthetic data is the same gradually changing model shown in Figure 8. For each test, we validate the convergence of the rjMcMC with the

procedure described in Section 2.3.1, and then use the distribution of number of layers for the retrieved ensemble as a measure of model complexity. As is shown in Figure 2.9**,** as data measurement uncertainty increases, the preferred number of layers for the retrieved ensemble decreases. Since the target model has a gradually changing velocity with depth, a decrease in the preferred number of layers would eventually lead to an over-simplified structure that does not accurately reflect velocity gradients with depth.

2.4.3 Attached vs. Independent Parameterizations

Despite the popularity of TB approaches, little discussion has concerned the parameterization of inversions involving multiple types of physical parameters. For model space search approaches – such as TB inversion – the number of model parameters is not limited by the number of measurements. This enables us to employ different parameterization schemes in the inversions, and quantify the effect of these choices on the posterior solution. Bodin et al., (2016) proposed a parameterization scheme for anisotropy inversion in which anisotropic parameters are proposed based on the existing isotropic structure. In this paper, we discussed two types of parameterization for dealing with multi-parameter problems: attached and independent.

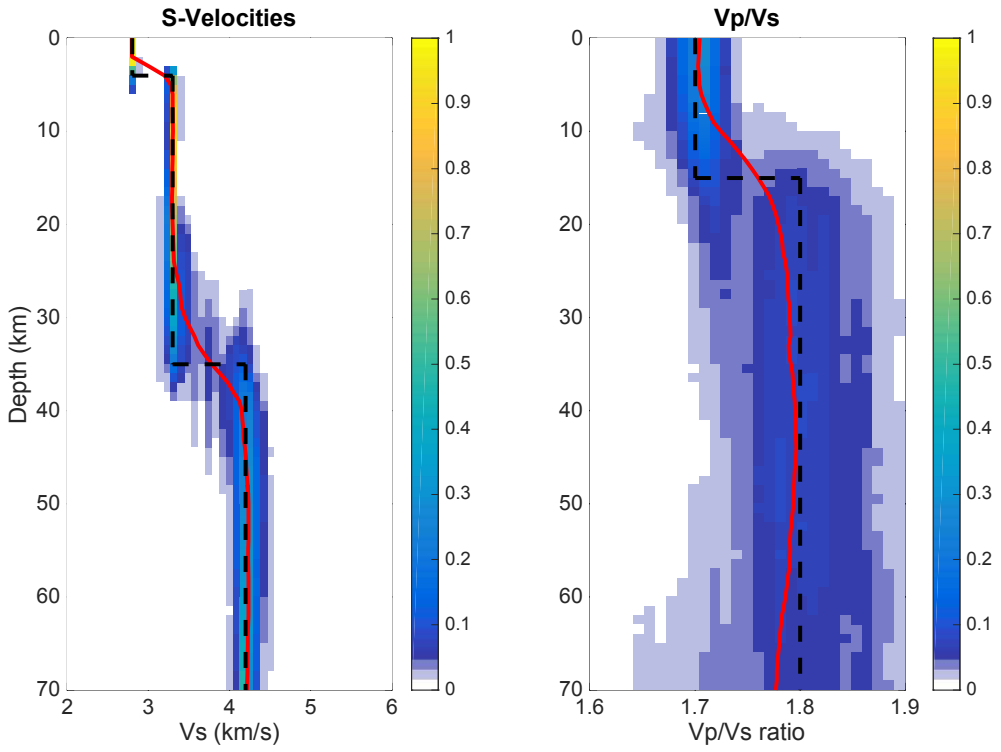


Figure 2.10 Vs (left) and Vp/Vs ratio (right) depth distribution retrieved using transdimensional Bayesian inversion of synthetic surface wave dispersion data. Independent parameterization is used for Vs and Vp/Vs ratio in this test. The ensemble solutions are displayed as probability density functions at each depth, with warmer colors corresponding to higher posterior probabilities, and the solid red line denoting the 5% trimmed mean of the posterior. The target model (black dashed lines) contains a jump in Vp/Vs due to the α-β quartz transition that is not accompanied by a change in Vs [Diaferia and Cammarano, 2017].

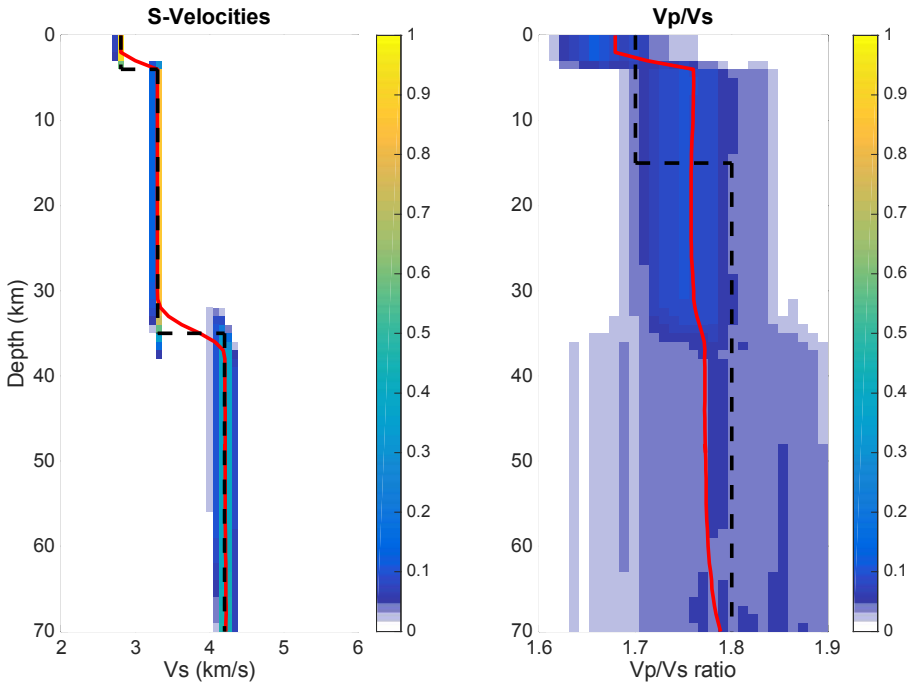


Figure 2.11 Vs (left) and Vp/Vs ratio (right) depth distribution retrieved using transdimensional Bayesian inversion of synthetic surface wave dispersion data. Attached parameterization is used for Vs and Vp/Vs ratio in this test. The ensemble solutions are displayed as probability density functions at each depth, with warmer colors corresponding to higher posterior probabilities, and the solid red line denoting the 5% trimmed mean of the posterior. The target model (black dashed lines) contains a jump in Vp/Vs due to the α-β quartz transition that is not accompanied by a change in Vs [Diaferia and Cammarano, 2017].

In regions where the geotherm is sufficiently high, the α-β quartz transition is expected to occur in the middle-lower crust, resulting in a sharp Vp/Vs ratio increase that is not accompanied by a significant change in Vs (Kuo-Chen et al., 2012; Diaferia and Cammarano, 2017). We show a synthetic test where the target model has a simplified one-layer crust on top of mantle Vs structure, and impose a mid-crustal Vp/Vs ratio increase from 1.7 to 1.8 representing the effects of the α-β quartz transition. This model is motivated by Figure 4 of Diaferia and Cammarano (2017). We perform TBI of surface wave dispersion data (5-100 second range) using

independent and attached parameterization of Vs and Vp/Vs ratio. The ensemble results show that when independent parameterization is applied (Figure 2.10), the inversion resolves both the Vs and Vp/Vs structure accurately, despite the distinct geometries. However, when attached parameterization is applied (Figure 2.11), the retrieved structure of Vp/Vs is strongly affected by the resolved geometry of Vs, to which the data is primarily sensitive.

This test illustrates how parameterization choices can be crucial to detecting complex structures with distinct geometries for different seismic parameters. Specifically, it shows that surface wave studies aiming to detect the α-β quartz transition in the middle-lower crust should employ a parameterization flexible enough to not preclude its detection.

2.4.4 Constraining Radial Anisotropy Using TB Inversion

In our previous synthetic tests, we assumed isotropic, layered structure, which may not always be an appropriate assumption to make, depending on the geological setting of the seismic station being analyzed. Due to lattice-preferred orientation (LPO) of anisotropic minerals or shape-preferred orientation (SPO) of different rock layers or fractures, seismic wave velocities will depend on polarization and propagation directions (Crampin et al., 1984). Studying seismic anisotropy in the crust and upper mantle can provide us insights into crust and mantle deformation (Kendall, 2000; Becker et al., 2003), mantle composition (Montagner and Anderson, 1989), lithosphere and asthenosphere coupling (Silver and Holt, 2002; Becker et al., 2006), and the net rotation of the lithosphere (Becker, 2008). Here, we restrict our attention to

radial anisotropy, in which the elastic properties of the medium can be described by 5 independent elastic coefficients (A, C, F, L, N; Love, 1911) at each location. Seismic observations including surface waves and normal mode data are often used to constrain the radial anisotropy of the Earth (e.g. Ekstrom and Dziewonski, 1998; Lekic and Romanowicz, 2011; Chang et al., 2014; Moulik and Ekstrom, 2014). Radial anisotropy in the Earth is often due to layering. However, recently studies also suggest that a large portion of anisotropy presented in the tomographic models may be due to unmapped discontinuities (Bodin et al., 2015).

To study shear wave radial anisotropy, we use the Voigt average isotropic shear wave velocity Vs and radial anisotropy parameter $\xi = (V_{SH}/V_{SV})^2$ instead of isotropic shear velocity alone to represent the seismic structure. With limited studies discussing the relationship between isotropic Vs and ξ geometries, we propose to use independent parameterization to represent such ignorance.

For the synthetic test, we set the vertical shear wave velocity V_{SV} for the target model to be same as the Vs value from the isotropic target model we used in previous sections. We set the horizontal shear wave velocity V_{SH} to be different from the V_{SV} structure so that the radial anisotropy parameter $\xi = (V_{SH}/V_{SV})^2$ has a value of 1.149 between 19 km and 50 km and a value of 1.000 at other depths (Figure 2.12, black dashed line). Vp and density in the target model are derived from isotropic Vs using the empirical relationship from Brocker et al., (2005). To systematically investigate the effects of parameter trade-offs and data uncertainties on the retrieved structures, we perform a series of synthetic tests with different combinations of parameter types and data uncertainties.

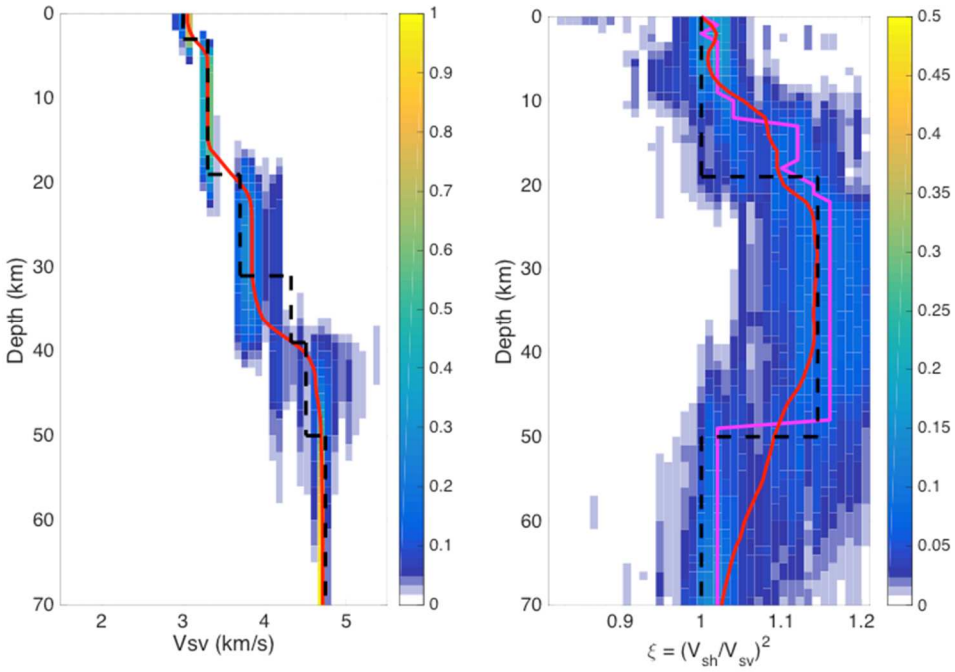


Figure 2.12 Anisotropy inversion with independent parameterization and no scaling on Vp. The ensemble solutions (Vs on the left, ξ on the right) are displayed as probability density functions at each depth, with warmer colors corresponding to higher posterior probabilities, the solid red line denoting the 5% trimmed mean of the posterior, and the magenta line denoting its mode. The target model (black dashed lines) is used to generate synthetic data.

We start by inverting surface wave dispersion data, assuming a 2% measurement error uncorrelated between periods. Here we invert for Vs, ξ, and Vp without assuming scaling relationships between any of them. In order to study the effect of trade-offs between different model parameters in a systemic way, we scale density to Vs, using expressions in Brocher et al., (2005). The prior on ξ is set to be a uniform distribution between 0.81 and 1.21. The retrieved structure is shown in Figure 2.12. Compared to the velocity structures, the ensemble of ξ spreads widely across the prior space. Despite the large variance, the inversion is able to resolve an anisotropic layer between 19 and 50 km. In Figure 2.12, we plot the trimmed mean and the mode of the posterior PDF to better illustrate the inversion result. Both the trimmed mean

value and the mode are overestimated between 10 and 19 km, as can be seen by comparing them against the target structure shown in black.

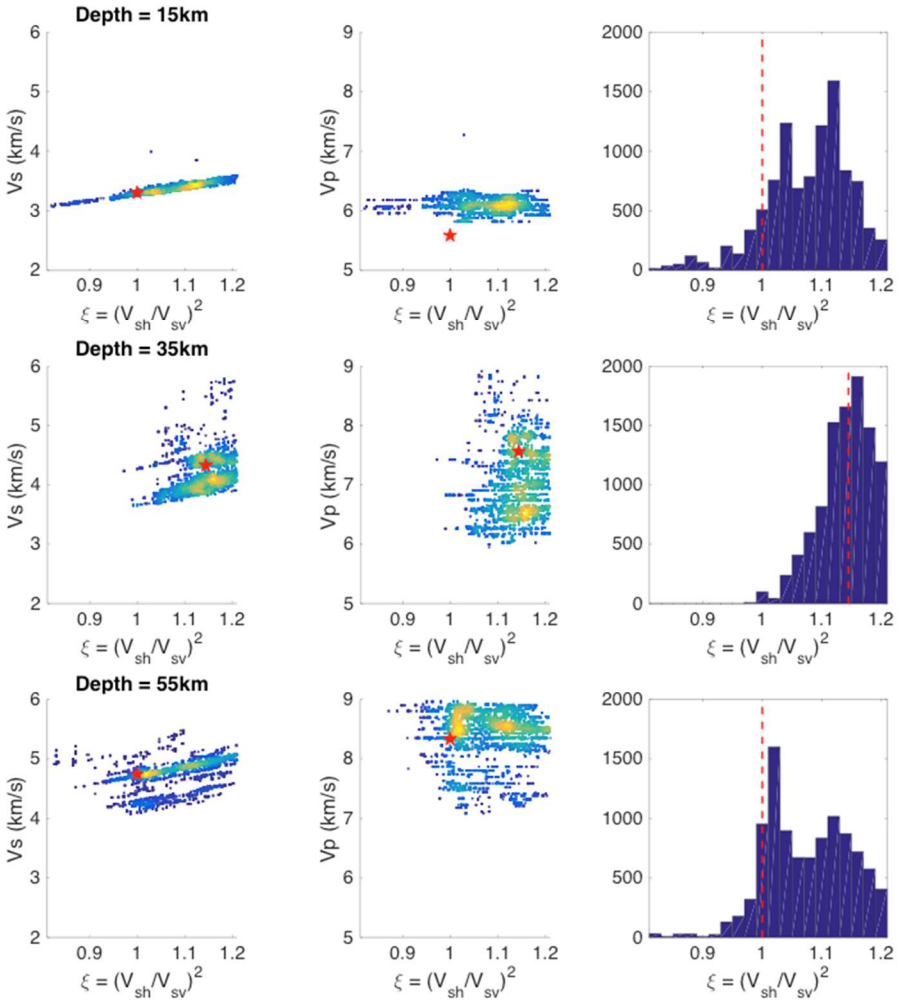


Figure 2.13 Parameter trade-offs and marginal posterior PDF at 15 km (upper), 35 km (middle), and 55 km (bottom) from the radial anisotropy inversion with no scaling relation assumed between Vs and Vp. Left panels show the scatter plot of Vs and ξ values from the 10,000 models in the ensemble solution. The scatter plot is colored by the density of points to better reveal the parameter trade-off. Warm colors denote higher probabilities and cool colors denote lower probabilities. Middle panels show colored scatter plots but for Vp and ξ values in the ensemble. The right panels are the marginal posterior PDF of ξ. The true value at that depth is plotted as red dashed line.

In Figure 2.13, we show the scatter plots colored by their density from the posterior PDF of Vs, Vp and ξ at three different depths. The scatter plots between Vp and ξ show that there are trade-offs between these two parameters. These tradeoffs result from the fact that in a radially anisotropic medium, Rayleigh waves, whose sensitivity is primarily to V_{SV}, are also sensitive to Vp, while Love waves, whose primary sensitivity is to V_{SH}, are not sensitive to Vp. This tradeoff between Vp and ξ limits our ability to retrieve radial anisotropy given the limited constraints provided by surface wave dispersion data. Similarly, we test the tradeoff between density and ξ by performing an inversion for Vs, ξ and density without assuming a fixed scaling relationship between them, while fixing the Vp scaling to Vs, using the expressions in Brocher et al., (2005). The target model is same as the one in Figure 2.12. The retrieved structure of ξ is shown in Figure 2.14 (left). While the ensemble results resolve an anisotropic structure approximately between 19 km and 50 km, both the trimmed mean and the mode are underestimated within this depths range. Moreover, the thickness of the anisotropic layer is not well constrained. We interpret this as a result of tradeoff between density and ξ limiting our ability of resolving radial anisotropy giving the limited constraints provided by surface wave dispersion data.

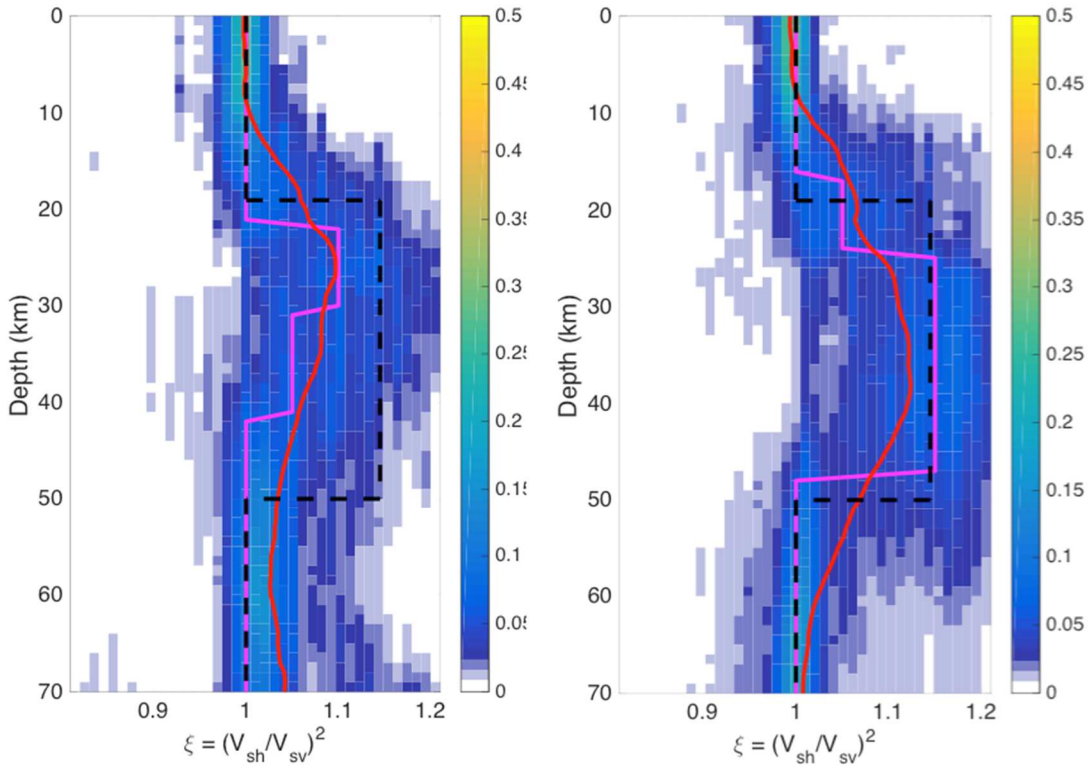


Figure 2.14 Anisotropy inversions with independent parameterization and no scaling on density (left) / correct scaling on Vp and density (right). The ensemble solutions are displayed as probability density functions at each depth, with warmer colors corresponding to higher posterior probabilities, the solid red line denoting the 5% trimmed mean of the posterior, and the magenta line denoting its mode. The target model (black dashed lines) is used to generate synthetic data.

To validate the effect of the trade-off on the inference of ξ, we apply the same empirical relationship from Brocher et al., (2005) used in the target model to derive both Vp and density from Vs in our synthetic test to reduce the trade-off between Vp, density and ξ. We keep the rest of the inversion set-up the same as in the previous two tests. The retrieved ξ structure is shown in Figure 2.14 (right). Compared to the tests with no scaling applied to Vp or density, the radial anisotropy structure is better-constrained at several depths. The estimated ξ from the mode of posterior PDF is closer to the true value between 24 and 48 km. The mode of the ensemble in this test

also better tracks the thickness of the anisotropic layer. We plot the marginal posterior PDF in Figure 2.15 at several depths to better illustrate the retrieved ξ against the actual structure. In Figure 2.15 we also show the scatter plot between Vp and ξ, which illustrated the reduced Vp- ξ trade off compared to the results shown in Figure 2.13.

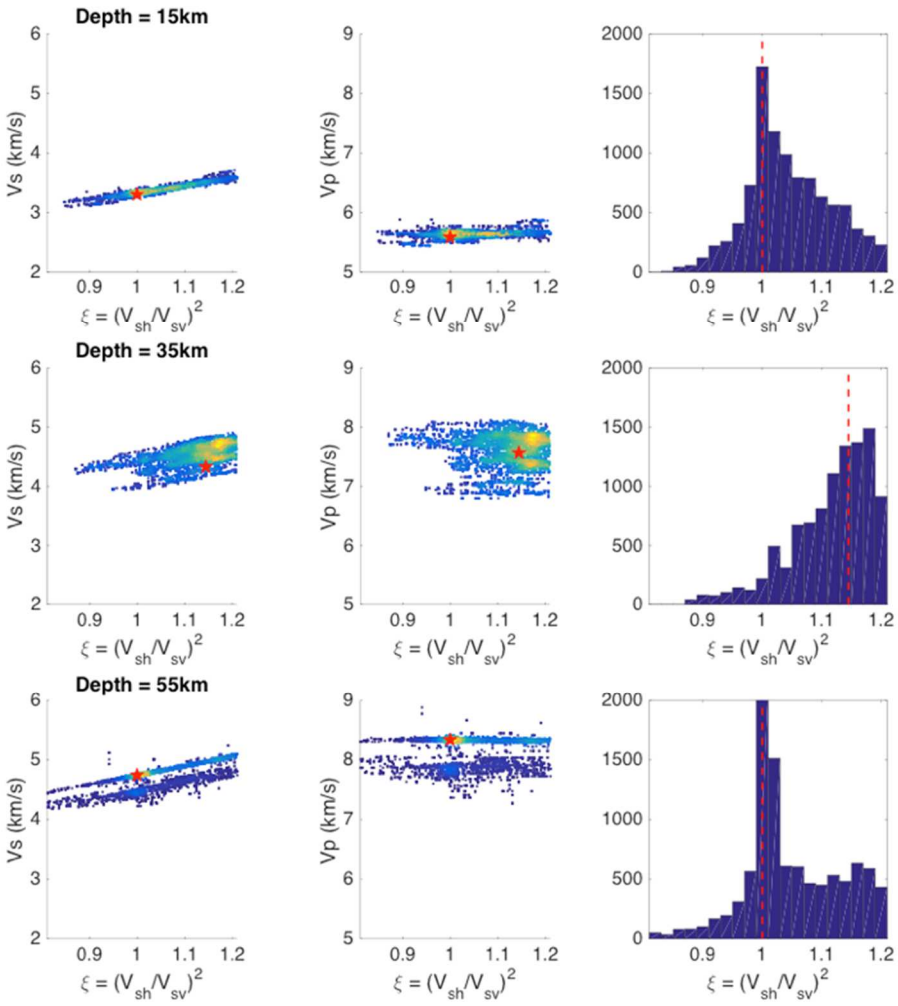


Figure 2.15 Parameter trade-offs and marginal posterior PDF from the radial anisotropy inversion with correct scaling relation between Vp and Vs at 15 km (upper), 35 km (middle), and 55 km (bottom). Left panels show the scatter plot of Vs and $\boldsymbol{\xi}$ values from the 10,000 models in the ensemble solution.. Middle panels are the colored scatter plots but for Vp and $\boldsymbol{\xi}$. The right panels are the marginal posterior PDF of $\boldsymbol{\xi}$. The true value at that depth is plotted as red dashed line.

These three tests together show that the tradeoffs between Vp and ξ as well as between density and ξ affect the inversion result for ξ. This means that without introducing additional data, applying a proper scaling between Vp, density and ξ helps resolve the radial anisotropy; however, on the other hand, Vp-Vs-density scaling assumptions need to be carefully made since an unsuitable scaling could bias the estimate of ξ. As a comparison, in Figure 2.16 we show the retrieved anisotropy structure where we apply a constant Vp/Vs ratio of 1.68 as the scaling during the inversion. As expected, the incorrect scaling biases our estimate of Vp. The retrieved anisotropy structure also deviates from the true value at several depths. We estimate an anisotropic layer between 24 km and 48 km for the correct scaling case, while for the fixed Vp/Vs case, the anisotropic layer is only recovered between 23 km and 38 km. In Figure 2.16, we calculated the root-mean-square error (RMSE) of the ensembles for the three radial anisotropy tests against the true value from the input model. It can be seen that when assumptions about Vp-Vs scaling are incorrect, the disagreement between the posterior and the true value is the largest between 20 km and 50 km, which covers most of the anisotropic layer. The RMSE in the incorrect scaling case can be as large as 0.16, while the largest RMSE in the correct scaling is 0.12.

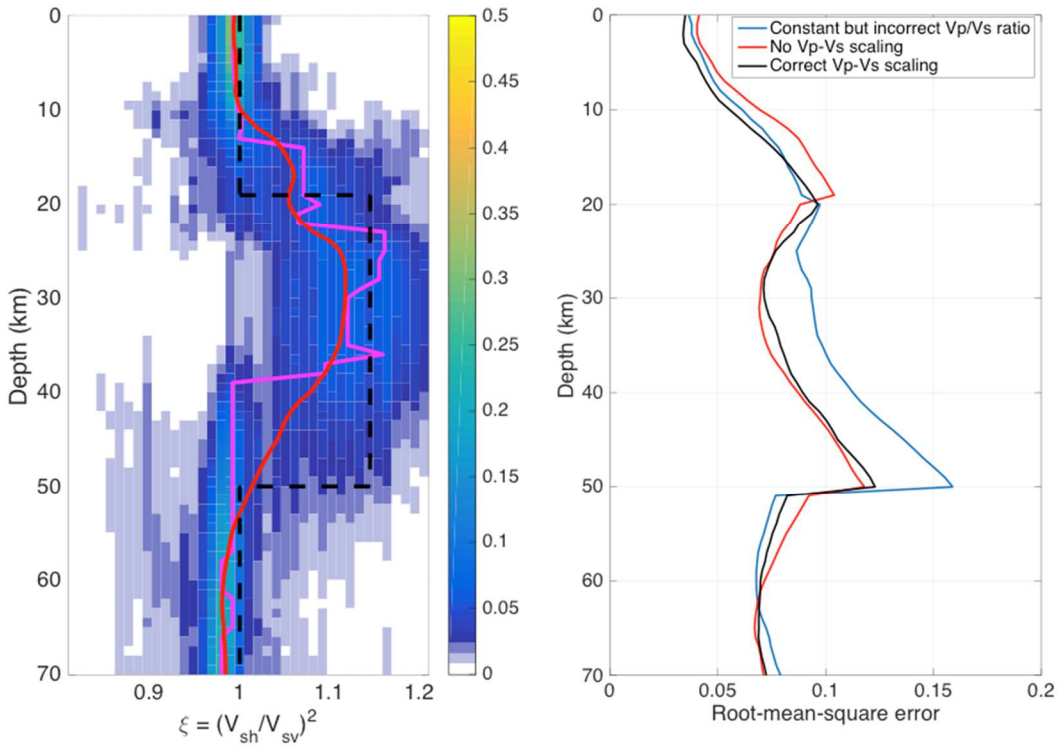


Figure 2.16 (left) Anisotropy inversion with independent parameterization and a fixed Vp/Vs ratio of 1.68. The ensemble solutions are displayed as probability density functions at each depth, with warmer colors corresponding to higher posterior probabilities, the solid red line denoting the 5% trimmed mean of the posterior, and the magenta line denoting its mode. The target model (black dashed lines) is used to generate synthetic data. (right) Root-mean-square error of radial anisotropy inversion ensemble against the input model. The blue curve is the TB inversion with fixed Vp/Vs = 1.68 using Rayleigh and Love wave dispersion. The red curve is the same inversion except using no Vp-Vs scaling. The black curve uses the correct Vp-Vs scaling as the input model while the rest of the setting is same as the previous two.

Roy and Romanowicz (2017) investigated the effect of assuming a fixed Vp/Vs on the inversion of surface wave dispersion and converted body wave (P-to-S) data for Vs radial anisotropy. By comparing results obtained fixing Vp/Vs to those obtained treating Vp/Vs as unknown, they concluded that the slight difference in the choice of Vp/Vs would not affect the retrieved structure. While the inclusion of body wave data does not provide direct constraint on radial anisotropy, it is expected to

improve the constraints on the depths and impedance contrasts across discontinuities in the velocity structures, and indirectly lead to better resolution of radial anisotropy. Therefore, the work of Roy and Romanowicz (2017) complements the analysis presented here, in which we explore the effects of scaling assumptions on inversion of surface wave dispersion alone, without including observables such as receiver function. Furthermore, unlike this study, Roy and Romanowicz (2017) assume the geometry of variations in isotropic Vs and radial anisotropy to be the same. As is discussed in Section 2.3.3, when such assumption is justified, it helps resolve the posterior; our study provides a different, complementary perspective on how different parameterization choices could affect the resolution of radial anisotropy.

2.5 Conclusion

Transdimensional Bayesian inversion has recently gained increasing attention in the area of geophysics. Its applications to various topics including seismic tomography (Young et al., 2013; Petrescu et al., 2016; Calo et al., 2016; Burdick and Lekic, 2017; Olugboji et al., 2017), earthquake source inversion (Dettmer et al., 2014), receiver function estimation (Kolb and Lekic, 2014), controlled source exploration geophysics (Ray et al., 2014; Gehrmann et al., 2015), and geoacoustic inversion (Dettmer et al., 2010), and viscosity inversion (Rudolph et al., 2015), show the utility of flexible yet naturally parsimonious parameterization in geophysical inversion as well as the capability of uncertainty quantification of inversion results that better represent the data sensitivity.

In this paper, we first show that we are able to retrieve seismic structures using surface wave dispersion and Rayleigh wave ZH ratio individually and together with less restrictive assumptions. Our synthetic tests suggest that although SWD and the ZH ratios are sensitive to Vp, Vs and density to various degrees, neither dataset is individually able to resolve a comprehensive structure. However, because of the flexibility of the transdimensional inversion, we can easily combine the ZH ratio data with surface wave dispersion in a joint inversion. We show that TB inversion can take advantage of the complementary sensitivity of the two data types to simultaneously constrain the continental lithosphere Vs structure, as well as the crustal Vp structure.

By comparing the results from the transdimensional inversion and inversion with fixed but correct parameterization, we illustrate that a fixed parameterization with strong prior information could bias the estimate of model parameters. One might argue that given a fixed but correct parameterization, the inversion results should represent the true uncertainties of the model parameters. However, the estimates of model parameters include not just the value of elastic parameters but also their distribution along depth (i.e. layering). A fixed parameterization is equivalent to using a prior that assumes no uncertainty for the latter, which affects the estimates of the former due to model parameter trade-offs.

While 1D layered model with constant elastic value within the layer is assumed in our transdimensional Bayesian inversion, our synthetic tests suggest that given robust data constraints, the layered structure parameterization is still able to resolve structures in which elastic properties change gradually with depth. We show that the limited SWD resolution and large measurement error could both result in

gradient structures appearing oversimplified when using parameterizations based on constant-velocity layers.

We then explore the effects of choices of parameterization on the retrieval of isotropic and anisotropic structure from surface wave inversion. Specifically, we propose and contrast two distinct parameterization choices: attached, in which all parameters of interest share the same geometry; and, independent, in which the geometry of different physical parameters can vary.

Using synthetic tests, we show that the attached type scheme tends to yield results whose geometry is mainly determined by the parameter that is best constrained by the data at hand. When other parameters share the same geometry with the best-resolved parameter, the use of attached-type parameterization is advised (See Figure 2.7). On the other hand, when parameters do not share the same geometry, the estimate of weakly constrained parameter could be biased due to trade-offs (See Figure 2.A4). The attached-type parameterization we discuss in this study is similar to the scheme proposed by Bodin et al. (2016), since parameters of different type share the same geometry. It differs from the Bodin et al. scheme, in that since the anisotropic structure is sampled by adding/removing anisotropic parameters from an existing layer, our attached type parameterization assumes uniform prior on the additional parameters and samples them together.

The independent type parameterization we introduce offers a more flexible parameterization containing less prior information. By assuming no correlation between the geometry of different parameters, we are able to detect potentially complex structures with distinct geometries of different parameters while using an

optimally parsimonious number of parameters. This situation would accompany the α-β quartz transition as shown in Figure 2.10 and discussed in section 2.4.3. Additionally, it would be expected in the presence of a layer of partial melt where Vs drops dramatically but Vp does not, which might be associated with volcanic regions, regions with elevated temperatures in the lower crust, or even glacial firn aquifers recently seismically characterized in Greenland (Montgomery et al., 2017).

Seismic Bayesian inversion has been used to investigate the radial anisotropy of the Earth (Shapiro and Ritzwoller, 2002; Beghein and Trampert, 2004; Beghein et al., 2014; Calo et al., 2016). In particular, Calo et al. (2016) applied a transdimensional inversion using surface wave dispersion as well as receiver function data, and relied on empirical scaling laws between Vs and Vp to reduce parameter trade-offs. Here we show that inferences of radial anisotropy from surface wave dispersion are affected by prior information imposed on the inversion process. Specifically, applying proper scaling relations between Vp, Vs and density helps improve the constraints on radial anisotropy, while inaccurate assumptions about Vp-Vs-density scaling can bias estimates of radial anisotropy.

Previous studies have highlighted the potential of unmodeled crustal structure to bias inferences of upper mantle radial anisotropy (e.g. Bozdag and Trampert, 2008; Lekic et al., 2010; Ferreira et al., 2010), but the effect of Vp-Vs-density scaling assumptions has not garnered equal attention. Therefore, we stress that careful choices must be made to prevent the estimates of radial anisotropy from being biased due to unmodeled Vp and density structure. We find that trade-offs between Vp and radial

anisotropy can increase root-mean-square error by 33% in estimates of radial anisotropy.

Our inversion for radial anisotropy adopts an independent parameterization in which the geometries of isotropic Vs and ξ are not assumed to be the same. With fewer assumptions made to avoid potential bias in the inversion results, we are able to resolve the main anisotropic feature in the synthetic model. The independent-type parameterization scheme is particularly appropriate since radial anisotropy need not share the same geometry as isotropic wavespeeds, as pointed out by Montagner (2002).

In this study, we considered three sources of uncertainty: limitation of data sensitivity; assumed scaling among parameters; and the choice of parameterization, including both the number of parameters and attached vs. independent parameterizations for multi-parameter problems. The discussion of different sources of uncertainties presented herein should help inform choices for inversions on surface wave measurements on their own and in combination with other, complementary data types (e.g. receiver functions). While better constraints on seismic velocity profiles are expected when combining multiple seismic observables, it is necessary to attribute the influence of certain parameterization choices. This becomes particularly important for inferences of parameters such as radial anisotropy and density – which are less well constrained by available data – since our findings suggests that model parameterization can significantly bias them.

2.6 Appendix

2.6.1 Acceptance Probabilities for Attached and Independent Parameterization

Attached Type Parameterization

For the "attached" type parameterization, all three elastic parameters are assigned to one Voronoi nucleus. Below, we specify the proposal probabilities corresponding to each of the four possible steps in the Markov chain: change, move, birth, and death.

1) Change

Randomly choose one of the Voronoi nuclei and randomly change all three parameters (Vp, Vs and ρ) based on a normal distribution with specified standard deviation. The proposal ratio is given by:

$$q_{vs}(v_s'|v_s) = \frac{1}{\sigma_{vs}\sqrt{2\pi}} \exp\left\{-\frac{(v_s' - v_s)^2}{2\sigma_{vs}^2}\right\} \quad (2.A1)$$

The proposed value of Vs is denoted by a prime, and is obtained by:

$$v_s' = v_s + u\sigma_{vs} \quad (2.A2)$$

where σ_{vs} is the standard deviation for perturbations to Vs, and u is a random number generated from the standard normal distribution. Expressions for Vp and density are analogous:

$$q_{vp}(v_p'|v_p) = \frac{1}{\sigma_{vp}\sqrt{2\pi}} \exp\left\{-\frac{(v_p' - v_p)^2}{2\sigma_{vp}^2}\right\} \quad (2.A3)$$

$$v_p' = v_p + u\sigma_{vp} \quad (2.A4)$$

$$q_\rho(\rho'|\rho) = \frac{1}{\sigma_\rho\sqrt{2\pi}}\exp\left\{-\frac{(\rho'-\rho)^2}{2\sigma_\rho^2}\right\} \quad (2.A5)$$

$$\rho' = \rho + u\sigma_\rho \quad (2.A6)$$

In this type of model perturbation, the expressions for the proposal ratios can be used to compute the final acceptance probability. Equations 2.A1, 2.A3, and 2.A5 clearly satisfy detailed balance conditions:

$$q_{vs}(v'_s|v_s) = q_{vs}(v_s|v'_s) \quad (2.A7)$$

$$q_{vp}(v'_p|v_p) = q_{vp}(v_p|v'_p) \quad (2.A8)$$

$$q_\rho(\rho'|\rho) = q_\rho(\rho|\rho') \quad (2.A9)$$

Since the three parameters are perturbed at the same time, we have

$$q_{vs}(v'_s|v_s)q_{vp}(v'_p|v_p)q_\rho(\rho'|\rho) = q_{vs}(v_s|v'_s)q_{vp}(v_p|v'_p)q_\rho(\rho|\rho') \quad (2.A10)$$

Therefore,

$$\frac{q(m|m')}{q(m'|m)} = 1 \quad (2.A11)$$

2) Move

In a move step, the depth of a randomly chosen Voronoi nucleus is perturbed based on a normal distribution with specified standard deviation.

$$q_c(c_i'|c_i) = \frac{1}{\sigma_c\sqrt{2\pi}}\exp\left\{-\frac{(c_i'-c_i)^2}{2\sigma_c^2}\right\} \quad (2.A12)$$

Even though the model parameterization changes, the number of Voronoi nuclei does not change, so no additional parameter is added in this step. When considering the proposal ratio of changing from m to m', equation 2.A12 clearly satisfies detailed balance:

$$q_c(c'_i|c_i) = q_c(c_i|c'_i) \quad (2.A13)$$

Therefore,

$$\frac{q(m|m')}{q(m'|m)} = 1 \quad (2.A14)$$

3) Birth

In the birth step, we randomly choose a depth defined by the uniform prior, and create a new Voronoi nucleus. The velocity and density values corresponding to the newly created nucleus are drawn from normal distributions centered on the current values of Vp, Vs and density at the chosen depth. The standard deviations of the normal distributions ($\sigma_{vp2}, \sigma_{vs2}, \sigma_{\rho 2}$) can differ from, but in this study are fixed to be the same as, ones used in the Change step.

The proposal probability for assigning a velocity v_s' to the new Voronoi nucleus (denoted by subscript $n+1$) is given by:

$$q(v_s'|m) = \frac{1}{\sigma_{vs2}\sqrt{2\pi}} \exp\left\{-\frac{\left(v_{s,n+1}' - v_{s,i}\right)^2}{2\sigma_{vs2}^2}\right\} \quad (2.A15)$$

where $v_{s,i}$ is the velocity of the Voronoi nucleus closest to the depth of the newly-born nucleus. Analogous expressions can be written for Vp and density:

$$q(v_p'|m) = \frac{1}{\sigma_{vp2}\sqrt{2\pi}} \exp\left\{-\frac{\left(v_{p,n+1}' - v_{p,i}\right)^2}{2\sigma_{vp2}^2}\right\} \quad (2.A16)$$

$$q(\rho'|m) = \frac{1}{\sigma_{\rho 2}\sqrt{2\pi}} \exp\left\{-\frac{(\rho_{n+1}' - \rho_i)^2}{2\sigma_{\rho 2}^2}\right\} \quad (2.A17)$$

We now turn our attention to defining the probabilities related to the depth of the Voronoi nucleus. Assume that we have n Voronoi nuclei in the current model m, and there are N possible positions in total for placing a Voronoi nucleus. The probability of having a $(n+1)th$ Voronoi nucleus in the rest of the available positions will be:

$$q(c'|m) = \frac{1}{N-n} \quad (2.A18)$$

The reverse process is to delete the added Voronoi nucleus from m'. The probability of deleting one Voronoi nucleus from $(n+1)$ nuclei is

$$q(c|m') = \frac{1}{n+1} \quad (2.A19)$$

The associated probabilities of removing the elastic parameters when their associated Voronoi nucleus is deleted are:

$$q(v_s|m') = 1 \quad (2.A20)$$

$$q(v_p|m') = 1 \quad (2.A21)$$

$$q(\rho|m') = 1 \quad (2.A22)$$

We write the proposal ratio as:

$$\frac{q(m|m')}{q(m'|m)} = \frac{q(c|m')}{q(c'|m)} \frac{q(v_s|m')}{q(v_s'|m)} \frac{q(v_p|m')}{q(v_p'|m)} \frac{q(\rho|m')}{q(\rho'|m)} \quad (2.A23)$$

Substituting expressions A.15-A.22 into equation A.23, we obtain:

$$\left(\frac{q(m|m')}{q(m'|m)}\right)_{birth} = \frac{(2\pi)^{3/2}\sigma_{vs2}\sigma_{vp2}\sigma_{\rho 2}(N-n)}{(n+1)} \exp\left[\frac{(v'_{s,n+1}-v_{s,i})^2}{2\sigma_{vs2}^2} + \frac{(v'_{p,n+1}-v_{p,i})^2}{2\sigma_{vp2}^2} + \frac{(\rho'_{n+1}-\rho_i)^2}{2\sigma_{\rho 2}^2}\right] \quad (2.A24)$$

4) Death

The death step is the exact reverse of the birth step, since one of the existing *n* Voronoi nuclei, denoted by index *j*, is randomly chosen and deleted to create a model m' with $(n-1)$ nuclei. For the death step, since it is supposed to be the exact reverse

step of birth, here we consider the situation of changing from n to $(n-1)$ Voronoi nuclei, we have

$$\left(\frac{q(m|m')}{q(m'|m)}\right)_{death} = \frac{n}{\sigma_{vs2}\sigma_{vp2}\sigma_{\rho 2}(2\pi)^{3/2}(N-n+1)} \exp -\left[\frac{\left(v'_{s,j}-v_{s,i}\right)^2}{2\sigma_{vs2}^2} - \frac{\left(v'_{p,j}-v_{p,i}\right)^2}{2\sigma_{vp2}^2} - \frac{(\rho'_j-\rho_i)^2}{2\sigma_{\rho 2}^2}\right] \tag{2.A25}$$

where index *i* denotes the Voronoi nucleus closest to the deleted nucleus.

5) Acceptance Probability

With the expression of proposal probabilities, we can now specify the acceptance probabilities for the attached parameterization. In the move and change step, the number of model parameters does not change, and the proposal ratios are unity. The acceptance probability can be written as:

$$\alpha(m'|m) = \min\left[1, \frac{p(m')}{p(m)} \cdot \frac{p(d|m')}{p(d|m)}\right] \quad (2.A26)$$

Since the dimension of the model does not change, and the priors on all the parameters are uniform, the prior ratio will be either zero or unity, and the acceptance probability can be simplified to:

$$\alpha(m'|m) = \begin{cases} \min\left[1, \dfrac{p(d|m')}{p(d|m)}\right] & if\ \forall i \in [1,n], v_{si} \in J_s, v_{pi} \in J_p, \rho_i \in J_\rho \\ 0 & otherwise \end{cases} \quad (2.A27)$$

For a birth step, according to equation (2.11), the prior ratio takes the form

$$\left(\frac{p(m')}{p(m)}\right)_{birth}$$

$$= \begin{cases} \dfrac{n+1}{(N-n)\Delta v_s \Delta v_p \Delta\rho} & if\ (n+1) \in I, and\ v'_{s,n+1} \epsilon J_s\ , v'_{p,n+1} \in J_p, \rho'_{n+1} \in J_\rho \\ 0 & otherwise \end{cases} \quad (2.A28)$$

Substituting (2.15), (2.A24), and (2.A28) into (2.16), the acceptance probability becomes

$$\alpha(m'|m)_{birth}$$

$$= \begin{cases} \dfrac{(\sigma_{vs2}\sigma_{vp2}\sigma_{\rho 2})(2\pi)^{3/2}}{\Delta v_s \Delta v_p \Delta\rho} \exp\left\{\dfrac{\left(v'_{s,n+1} - v_{s,i}\right)^2}{2\sigma_{vs2}^2} + \dfrac{\left(v'_{p,n+1} - v_{p,i}\right)^2}{2\sigma_{vp2}^2} + \dfrac{(\rho'_{n+1} - \rho_i)^2}{2\sigma_{\rho 2}^2} - \dfrac{\Phi(m') - \Phi(m)}{2}\right\} \\ \qquad if\ (n+1) \in I, and\ v'_{s,n+1} \epsilon J_s\ , v'_{p,n+1} \in J_p, \rho'_{n+1} \in J_\rho \\ 0 \qquad otherwise \end{cases} \quad (2.A29)$$

For the death step, the prior ratio should be the inverse of equation (2.A28), except we change from n nuclei to (n-1) nuclei. Therefore, substituting (2.15), (2.A25) into (2.16), we have:

$$\alpha(m'|m)_{death}$$

$$= \begin{cases} \dfrac{\Delta v_s \Delta v_p \Delta\rho}{\sigma_{vs2}\sigma_{vp2}\sigma_{\rho 2}(2\pi)^{3/2}} \exp\left\{-\dfrac{\left(v'_{s,j} - v_{s,i}\right)^2}{2\sigma_{vs2}^2} - \dfrac{\left(v'_{p,j} - v_{p,i}\right)^2}{2\sigma_{vp2}^2} - \dfrac{\left(\rho'_j - \rho_i\right)^2}{2\sigma_{\rho 2}^2} - \dfrac{\Phi(m') - \Phi(m)}{2}\right\} if\ (n-1) \in I \\ 0 \qquad otherwise \end{cases} \quad (2.A30)$$

Independent Type Parameterization

In the "independent" type parameterization, we have three independent sets of Voronoi nuclei, each specifying Vs, Vp or density. When changing the current model in the Markov chain, we first randomly choose one out of the three types of parameters (Vs, Vp, density) with equal probability. Once a specific type of parameter

is chosen, the rest of the process and the proposal probabilities are identical to that described in Bodin et al., (2009).

Once the type of parameter to perturb is chosen, it is straightforward to derive the acceptance probability. We take perturbing density as an example. When choosing to change the density value or to move one of the Voronoi nuclei (denoted by index *i*), the dimension of the model does not change. Therefore, we have:

$$\alpha(m'|m) = \begin{cases} \min\left[1, \dfrac{p(d|m')}{p(d|m)}\right] & if\ \forall i \in [1,n], \rho_i \in J_\rho \\ 0 & otherwise \end{cases} \quad (2.A31)$$

For a birth step:

$$\alpha(m'|m)_{birth} = \begin{cases} \dfrac{(\sigma_{\rho 2})(2\pi)^{1/2}}{\Delta\rho} \exp\left\{\dfrac{(\rho'_{n+1} - \rho_i)^2}{2\sigma_{\rho 2}^2} - \dfrac{\Phi(m') - \Phi(m)}{2}\right\} & if\ (n+1) \in I, and\ \rho'_{n+1} \in J_\rho \\ 0 & otherwise \end{cases} \quad (2.A32)$$

where *n* is the number of Voronoi nuclei defining the density structure. Finally, for the death step in which Voronoi nucleus *j* is deleted, and Voronoi nucleus *i* is the remaining nucleus closest to the deleted one, we have:

$$\alpha(m'|m)_{death} = \begin{cases} \dfrac{\Delta\rho}{\sigma_{\rho 2}(2\pi)^{1/2}} \exp\left\{-\dfrac{(\rho'_j - \rho_i)^2}{2\sigma_{\rho 2}^2} - \dfrac{\Phi(m') - \Phi(m)}{2}\right\} & if\ (n-1) \in I \\ 0 & otherwise \end{cases} \quad (2.A33)$$

2.6.2 Convergence Analysis for TB Inversion of SWD

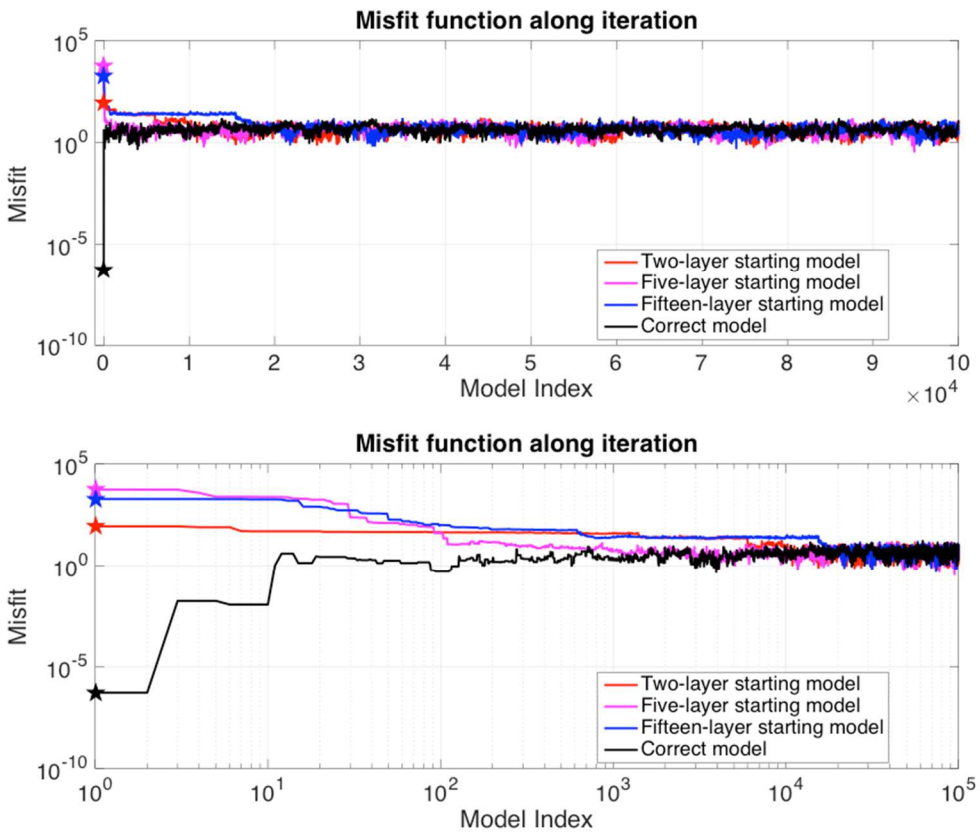


Figure 2.A1 The evolution of misfit along the rj-McMC with different starting models. Four starting models with different initial numbers of layers are used here to invert surface wave dispersion data between 5 s and 100 s. Every 100th model from a total of 10 million iterations is plotted on a linear (top) and logarithmic (bottom) x-axis. After about $\mathbf{2 \times 10^5}$ iterations, all four rj-McMC remain at a low misfit. We choose a burn-in period of $\mathbf{2.5 \times 10^6}$ iterations to ensure the rj-McMC have converge before sampling the ensemble solution.

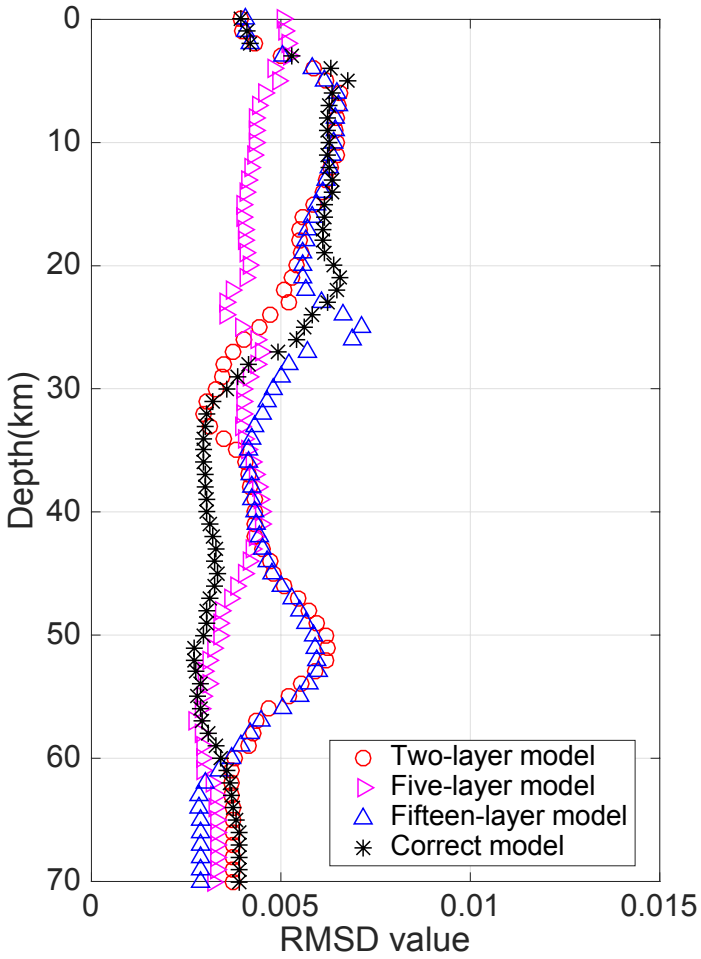


Figure 2.A2 Root-mean-square deviation of Vs posterior probability density functions from TB inversion of surface wave dispersion using four different starting models. We use the following equation to calculate to root-mean-square deviation (RMSD) for the four ensembles.

$$RMSD_{vs}(z) = \sqrt{\frac{\sum_{i=1}^{n(z)}(y_i(z) - \overline{y_i(z)})^2}{n(z)}} \quad (2.A34)$$

Here z stands for the depth; n stands for the number of bins used to discretize the PDF; $y_i(z)$ is the value of the PDF at the depth z; $\overline{y_i(z)}$ is the mean value of the PDF across four ensembles. The RMSDs of the four ensembles are close to each other and have low absolute values across all depths. This confirms that after the chosen burn-in period, the ensemble results from different starting models are indistinguishable from each other.

2.6.3 Normalized root-mean-square error of Vs and Vp posterior from Figure 2.5

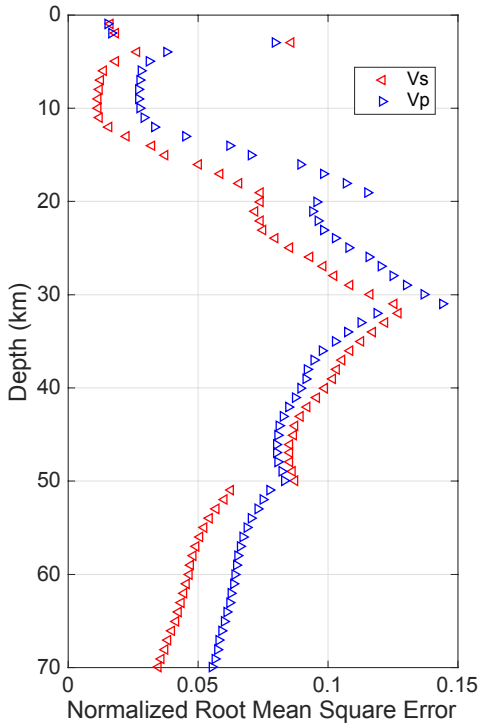


Figure 2.A3 Normalized root-mean-square error of Vs (red) and Vp (blue) inversion ensemble against the input model. The ensemble is taken from test shown in Fig. 5 where Vs, Vp and density are inverted using SWD and ZH ratio. RMSE of Vs and Vp are normalized by their mean at that given depth.

Chapter 3: Short Note on Surface Wave Dispersion Calculation of a Radial Anisotropic Spherical Earth

To predict SWD and ZH ratios, we use the reflectivity method (Kennet and Kerry, 1979; Aki and Richard, 2002; Hisada, 1994) to solve the eigenvalue problem for both Rayleigh wave and Love wave in an elastic, vertically heterogeneous medium, based on the implementation of Lai and Rix (1998).

Consider an isotropic and multilayered half-space medium, the free elastodynamic equation without source term can be written, in frequency domain, as (see Aki and Richard, 2002)

$$-\rho^{(j)}\omega^2\boldsymbol{u}^{(j)}(\boldsymbol{x},\omega) = \left[\lambda^{(j)} + 2\mu^{(j)}\right]\boldsymbol{\nabla}\boldsymbol{\nabla}\cdot\boldsymbol{u}^{(j)}(\boldsymbol{x},\omega) - \mu^{(j)}\boldsymbol{\nabla}\times\boldsymbol{\nabla}\times\boldsymbol{u}^{(j)}(\boldsymbol{x},\omega) \quad (3.1)$$

for $z^{(j-1)} < z < z^{(j)}$, j = 1, 2, 3, … , N, N+1. Where $0 = z^{(0)} < z^{(1)} < \cdots < z^{(j)} < \cdots < z^{(N)} < z^{(N+1)} = +\infty$; $\rho^{(j)}$, $\lambda^{(j)}$, $\mu^{(j)}$ are the density and Lame parameters for the *j*th layer, respectively. For horizontal layered media, the seismic wavefield can be represented and studied in terms of an equation of the type (Aki and Richard, 2002):

$$\frac{d}{dz}\boldsymbol{f}^{(j)}(z) = \boldsymbol{A}^{(j)}\boldsymbol{f}^{(j)}(z) \quad (3.2)$$

$\boldsymbol{f}^{(j)}(z)$ is the motion-stress vector for the *j*th layer, for P-SV waves it has dimensions of 4x1, for SH waves it has dimensions of 2x1. The constant matrix $\boldsymbol{A}^{(j)}$ has dimensions of 4x4 and 2x2 for P-SV and SH waves, respectively. This becomes a linear differential eigenvalue problem with displacement eigenfunctions (first two terms of $\boldsymbol{f}^{(j)}(z)$) and stress eigenfunctions (last two terms of $\boldsymbol{f}^{(j)}(z)$). It has been

shown (Aki and Richard, 2002) that inside each layer, the analytic solution for equation (3.2) has the following form:

$$\boldsymbol{f} = \mathbf{E} \cdot \boldsymbol{\Lambda}(z) \cdot \mathbf{R} \quad (3.3)$$

Several techniques have been proposed and developed for solving this linear eigenproblem (Thomson, 1950; Haskell, 1953; Kennett, 1983). For the multi-layered media in our case, we use the method proposed by Kennett and Kerry (1979) and developed by other researchers (Luco and Apsel, 1983; Chen, 1993; Hisada, 1994). It is based on the use of reflection and transmission coefficients to construct reflection and transmission matrices for a stratified media. The readers are referred to the studies cited above for the details of the method and implementation as they are well documented in the literature.

The flat-layered Earth model needs to be modified, especially for propagation of long-period surface wave, to take into account the sphericity of the Earth. The Earth-flattening transformation is introduced in seismology (Alterman et al., 1961; Biswas and Knopoff, 1970; Bhattacharya, 1996) to simulate the wave propagation in a spherical layered Earth using a formulation in isotropic flat-layered Earth. A new depth variable *z* is introduced as:

$$r = a \exp(^{-z}/_{a}) \, , or \; z = -a \cdot \ln(r/a) \quad (3.4)$$

Here, *a* is the radius of the Earth, r is the distance from the center of the Earth. The transformation for velocities can be written as:

$$\alpha = (^{a}/_{r})\alpha_0 \quad (3.5)$$

$$\beta = (^{a}/_{r})\beta_0 \quad (3.6)$$

α_0 and β_0 are the shear and compressional velocity in the flat-layered model, respectively.

Arora et al., (1996) showed that assuming the elastic parameter λ, μ is proportional to r^p, where r is the radial distance from the center of the Earth and p is an arbitrary constant, the solutions of the equations of P-SV motion can be written in form of exponential functions. We choose $p = -2$ in our study as Bhattacharya (1996) shows that the Earth-flattening transformation becomes further simplified for P-SV motion when $p = -2$. The expression for the displacement-stress matrix is shown to be similar as the one obtained from a flat Earth. Bhattacharya and Arora (1997) further extended the study to solve for the equations of P-SV motion for radial-anisotropic medium. As for propagation of Love wave in radial-anisotropic layered model, Anderson (1962) showed that the equations of SH motion can be solved by substituting two pseudo-parameters into the isotropic displacement-stress matrix. We discuss the details of this work in the later section.

To take into account the sphericity and radial anisotropy of the Earth, we modify the reflectivity method to compute the surface wave dispersion given elastic parameter of the transverse isotropic medium A, C, N, L, and F according to Anderson (1962) and Bhattacharya and Arora (1997). We then write

$$[A, C, N, L, F, \rho] = [A_0, C_0, N_0, L_0, F_0, \rho_0 ({}^{r}/_{a})^{-2}] \times ({}^{r}/_{a})^{p} \quad (3.7)$$

Here, ρ is the density. A_0, C_0, N_0, L_0, F_0 are the elastic parameter in the flat-layered model. It has been shown that the wave propagation in a transverse isotropic spherical-layered Earth, the displacement field of P-SV motion can be written as fourth-order differential equation:

$$\frac{d^4x}{dz^4} - [M_1 + M_2 - M_3]\frac{d^2x}{dz^2} + [M_1M_2 - M_4]x = 0 \quad (3.8)$$

x here can be either the vertical displacement ν or radial displacement μ. The expression for M_1, M_2, M_3, M_4 are

$$M_1 = \frac{q^2}{a^2} + \frac{1}{C_0}\left[\frac{n(n+1)}{a^2}L_0 + \frac{4(A_0 - N_0 + qF_0)}{a^2} - \omega^2\rho\right] \quad (3.9)$$

$$M_2 = \frac{(1-q)^2}{a^2} + \frac{1}{L_0}\left[\frac{n(n+1)}{a^2}A_0 - \frac{2N_0}{a^2} - \omega^2\rho\right] \quad (3.10)$$

$$M_3 = \frac{n(n+1)}{a^2}\frac{(L_0 + F_0)^2}{C_0L_0} \quad (3.11)$$

$$M_4 = \frac{n(n+1)}{a^2}\frac{[2(A_0 - N_0) + qF_0 + (1-q)L_0]^2}{a^2C_0L_0} \quad (3.12)$$

$$q = -(p+1)/2 \quad (3.13)$$

The roots (eigenvalues) of equation (8) can then be written as:

$$\nu_{\alpha,\beta}^2 = \left[(M_1 + M_2 - M_3) \pm \sqrt{(M_1 + M_2 - M_3)^2 - 4(M_1M_2 - M_4)}\right]/2 \quad (3.14)$$

Write the displacement-stress matrix as:

$$\boldsymbol{f} = [\sqrt{n(n+1)}\nu, \mu, -\sqrt{n(n+1)}\tau, -\sigma]^T \quad (3.15)$$

Recall that the displacement-stress matrix can be written the in form $\boldsymbol{f} = \mathbf{E} \cdot \boldsymbol{\Lambda}(z) \cdot \mathbf{R}$, where elements of **E** are given by

$$E_{11} = -1 \qquad E_{21} = -\gamma_\alpha$$

$$E_{31} = L_0[\nu_\alpha + (q-1)/a + \sqrt{n(n+1)}\gamma_\alpha/a]$$

$$E_{41} = C_0\gamma_\alpha\nu_\alpha + (2F_0 + C_0 q)\gamma_\alpha/a - \sqrt{n(n+1)}F_0/a$$

$$E_{12} = \gamma_\beta \qquad E_{22} = 1$$

$$E_{32} = -L_0[\gamma_\beta\nu_\beta + (q-1)\gamma_\beta/a + \sqrt{n(n+1)}/a]$$

$$E_{42} = -C_0\nu_\beta - (2F_0 + C_0 q)/a + \sqrt{n(n+1)}F_0\gamma_\beta/a \qquad (3.16)$$

E_{i3} is obtained from E_{i1} by replacing ν_α by $-\nu_\alpha$ and E_{i4} is obtained from E_{i2} by replacing ν_β by $-\nu_\beta$ (i = 1, … , 4). $\sqrt{n(n+1)}/a$ can be replaced by $\sqrt{[k^2 - 1/(4a^2)]}$, where k is wave number. γ_a, γ_b are given as below:

$$\gamma_a = \frac{[\nu_a(L_0 + F_0) - \{2(A_0 - N_0) + qF_0 + (1-q)L_0\}/a]\sqrt{n(n+1)}/a}{C_0\nu_a^2 + \omega^2\rho_0 - [4(A_0 - N_0 + qF_0) + C_0 q^2]/a^2 - L_0 n(n+1)/a^2} \qquad (3.17)$$

$$\gamma_\beta = \frac{\left[-\nu_\beta(L_0 + F_0) - \{2(A_0 - N_0) + qF_0 + (1-q)L_0\}/a\right]\sqrt{n(n+1)}/a}{L_0\nu_\beta^2 + \omega^2\rho_0 + [2N_0 - L_0(1-q^2)]/a^2 - A_0 n(n+1)/a^2} \qquad (3.18)$$

Therefore, the normal modes and Green's function can be obtained by using **E** given in equation (3.16) and eigenvalues given in equation (3.14) in the implementation of isotropic flat-layered Earth in Rix and Lai (1998).

The propagation of Love wave in a transverse isotropic spherical-layered Earth is much simpler to modify from the isotopic flat-layered case. We first perform the Earth-flattening transformation using equation (3.4), (3.5), (3.6). Anderson (1962) showed that upon the substitution of the following pseudo-parameters

$$d'_{(j)} = \left(\frac{N_{(j)}}{L_{(j)}}\right)^{1/2} d_{(j)} \quad (3.19)$$

$$\mu'_{(j)} = (L_{(j)} \cdot N_{(j)})^{1/2} \quad (3.20)$$

into the displacement-stress matrix for Love waves (Haskell, 1953), it is equivalent to the isotropic case. Here $d_{(j)}$, the thickness of the ***j***th layer is replaced by $d'_{(j)}$; $\mu'_{(j)}$ is the psudo-rigidity of the ***j***th layer.

We validate our implementation by comparing our predictions to those from MINEOS (Masters et al., 2011) for the upper 1000km of the anisotropic PREM model (Dziewonski and Anderson, 1981). Given the same period range, our implementation costs around 0.1 second to predict SWD, which is much faster than MINEOS. All of the software is written in MATLAB. Figure 3.1 shows the comparison between the predictions of MINEOS and the reflectivity method assuming isotropic flat PREM and anisotropic spherical PREM. In Figure 3.2, we plot the different between the reflectivity method and the MINEOS prediction, our radial anisotropic spherical PREM prediction shows only 0.3% relative difference for Rayleigh wave and 1% relative difference for Love wave.

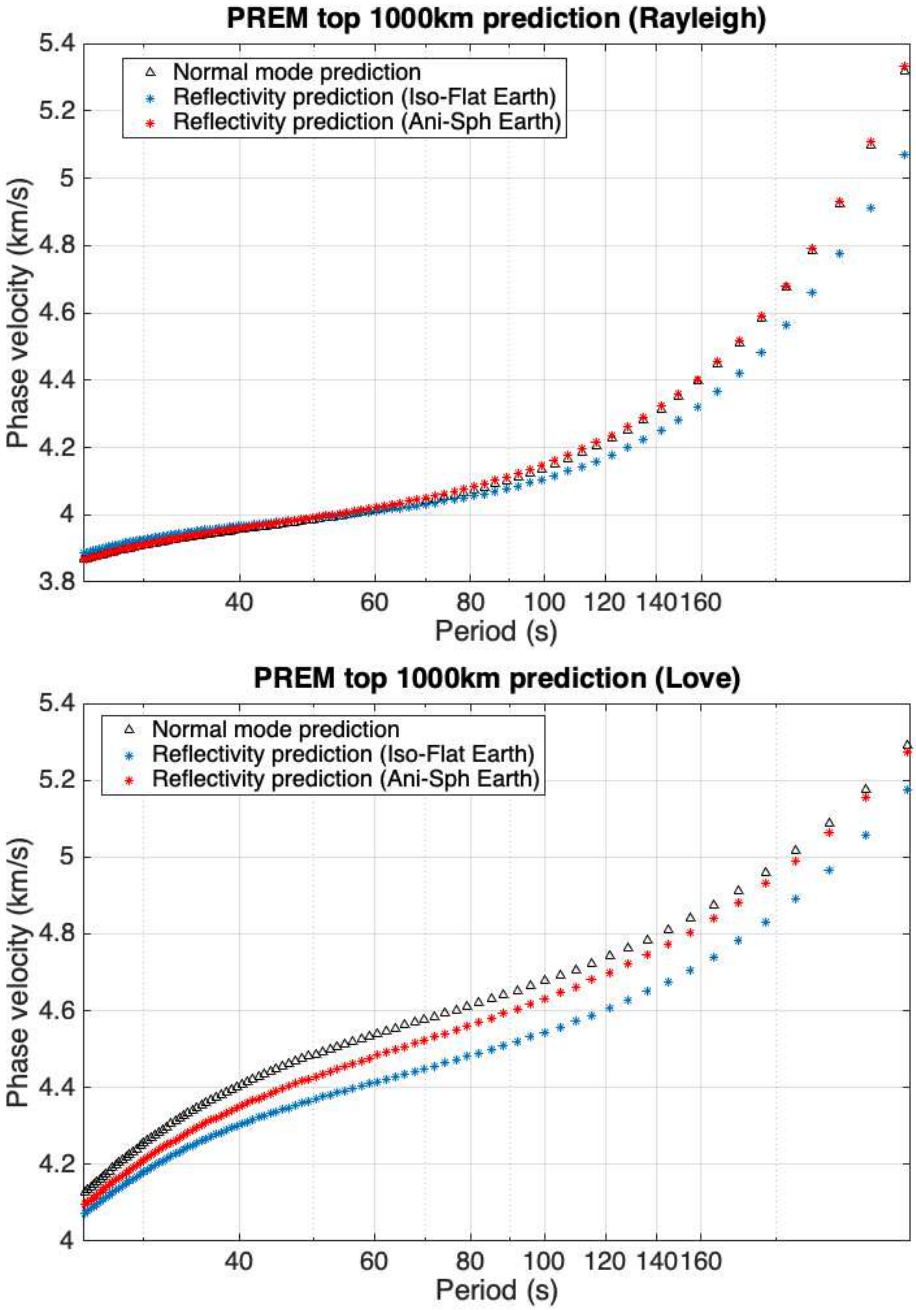


Figure 3.1 Rayleigh (top) and Love (bottom) wave dispersion prediction of PREM model top 1000 km structure. Black triangles represent the normal mode prediction from MINEOS (Masters et al., 2011); blue stars represent the reflectivity method prediction using isotropic flat-layered PREM; red stars represent the reflectivity method prediction using radial anisotropic spherical PREM.

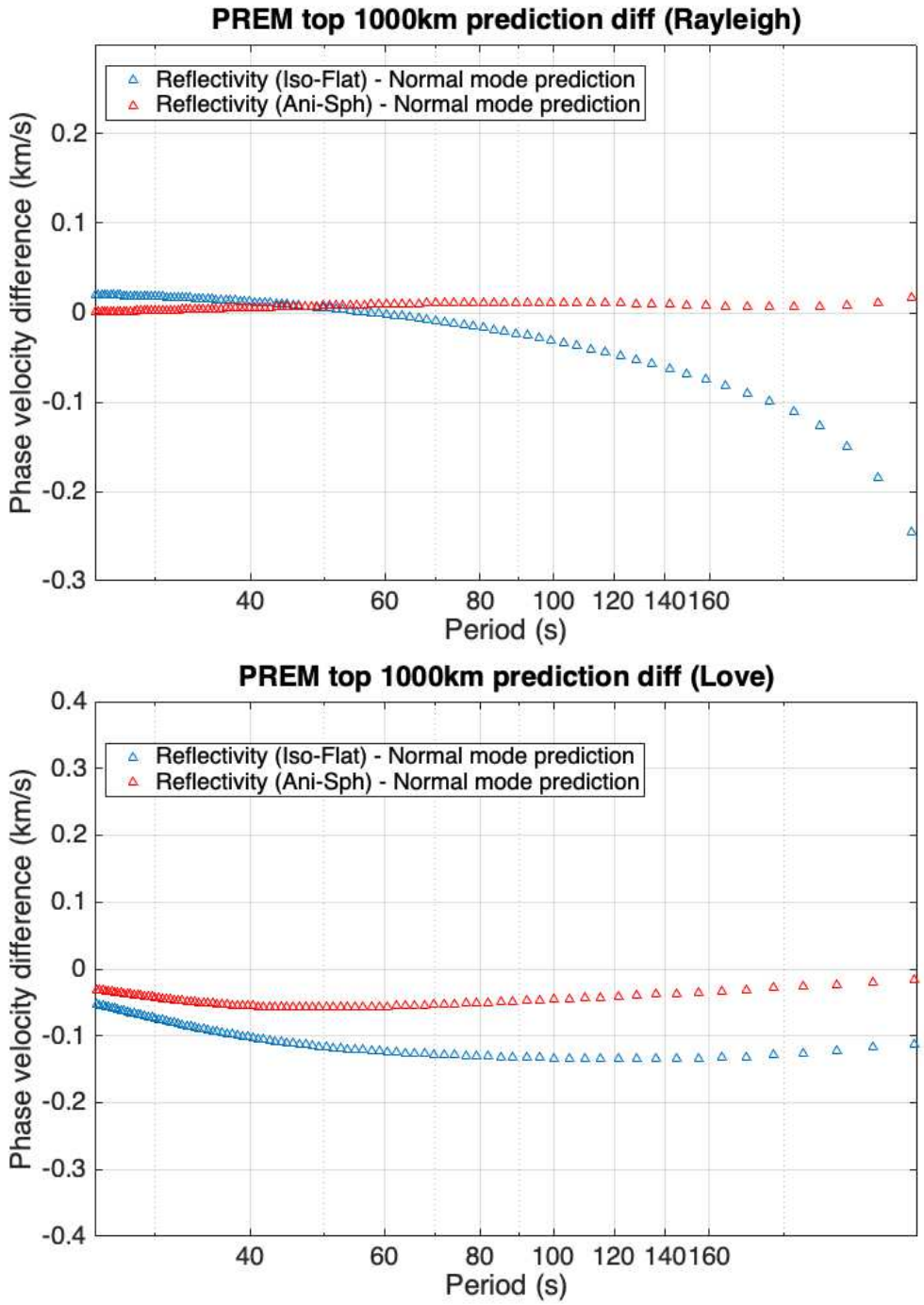


Figure 3.2 Differences between Rayleigh (top) and Love wave dispersion prediction of PREM model top 1000 km structure between MINEOS and reflectivity method. Blue triangles represent the value of reflectivity method prediction using isotropic flat-layered PREM subtract from MINEOS prediction; red triangles represent the reflectivity method prediction using radial anisotropic spherical PREM subtract from MINEOS prediction.

Chapter 4: Spurious Low Velocity Zones in Joint Inversion of Surface Waves and Receiver Functions

Gao, C., Cunningham, E., & Lekic, V., Spurious low velocity zones in joint inversion of surface waves and receiver functions. *In revision with Geophysical Journal International.*

Abstract

Low-velocity layers within the crust can indicate the presence of melt and lithologic differences with implications for crustal composition and formation. Seismic wave conversions and reverberations across the base of the crust or intra-crustal discontinuities, analyzed using the receiver function method, can be used to constrain crustal layering. This is commonly accomplished by inverting receiver functions jointly with surface wave dispersion. Recently, the proliferation of model-space search approaches has made this technique a workhorse of crustal seismology. We show that reverberations from shallow layers such as sedimentary basins produce spurious low-velocity zones when inverted for crustal structure with surface wave data of insufficiently high frequency. Therefore, reports of such layers in the literature based on inversions using receiver function data should be re-evaluated. We demonstrate that a simple resonance-removal filter can suppress these effects and yield reliable estimates of crustal structure, and advocate for its use in receiver-function based inversions.

4.1 Introduction

Constraints on the layering and seismic velocities of Earth's continental crust provide key insights on its composition and formation (e.g. Christensen and Mooney, 1995; Rudnick and Fountain, 1995). However, accurate quantification of crustal velocity structures from seismological observations is challenging, with primary constraints often coming from active-source experiments (for a recent review, see Mooney, 2015). Recently, passive source seismic investigations of continental crust structure have benefited from the installation of massive seismic arrays (e.g. USArray Transportable Array; Superior Province Rifting Earthscope Experiment [Zhang et al., 2016]) and the development in improved imaging techniques (Shapiro et al., 2005; Langston and Liang, 2008; Kumar & Bostock, 2008; Rychert and Harmon, 2016).

Consequently, efforts at combining different seismic observables to improve the constraints on crustal models have grown. Among those, joint inversions of surface wave dispersion (SWD) and receiver functions (RFs) (Özalaybey et al., 1997; Julia et al., 2000; Tkalčić et al., 2006; Bodin et al., 2012; Shen et al., 2013; Chai et al., 2015) have gained in popularity due to their complementary constraints: absolute seismic velocities from SWD and sharp impedance contrasts from RFs. The resulting crustal models revealed features in greater detail and with smaller uncertainties, including sharp discontinuities in the lithospheric mantle (Calo et a., 2016; Bodin et al., 2016) and crustal low velocity zones in various tectonic settings (Ward et al., 2014; Li et al., 2017; Li et al., 2018).

While the complementarity of SWD and RF sensitivity helps reduce the non-uniqueness of the seismic inverse problem (e.g. Özalaybey et al., 1997), differences in their inherent resolving power can introduce inconsistencies between the two data types (e.g. Chai et al., 2015). The consequences of these inconsistencies on the inversion results have not been adequately quantified, especially across diverse geologic settings. Additional complications arise in regions where the low-velocity sedimentary layers overlay crystalline continental crust; the strong impedance contrast at the bottom of the sediment layer can produce high amplitude, long duration reverberations in the RFs that could bias estimates of crustal thickness (Yeck et al., 2013; Yu et al. 2015).

In this study, we assess the potential for artifacts in joint seismic inversions for crustal structure using a novel implementation of transdimensional Bayesian inversion of SWD and RFs measured across 49 stations of the EarthScope Transportable Array. We focus on the Trans-Hudson Orogen and Superior Craton, part of which is overlain by the Williston Basin. We compare our results with analyses using synthetic data and discuss approaches for mitigating these effects.

4.2 Data and Method

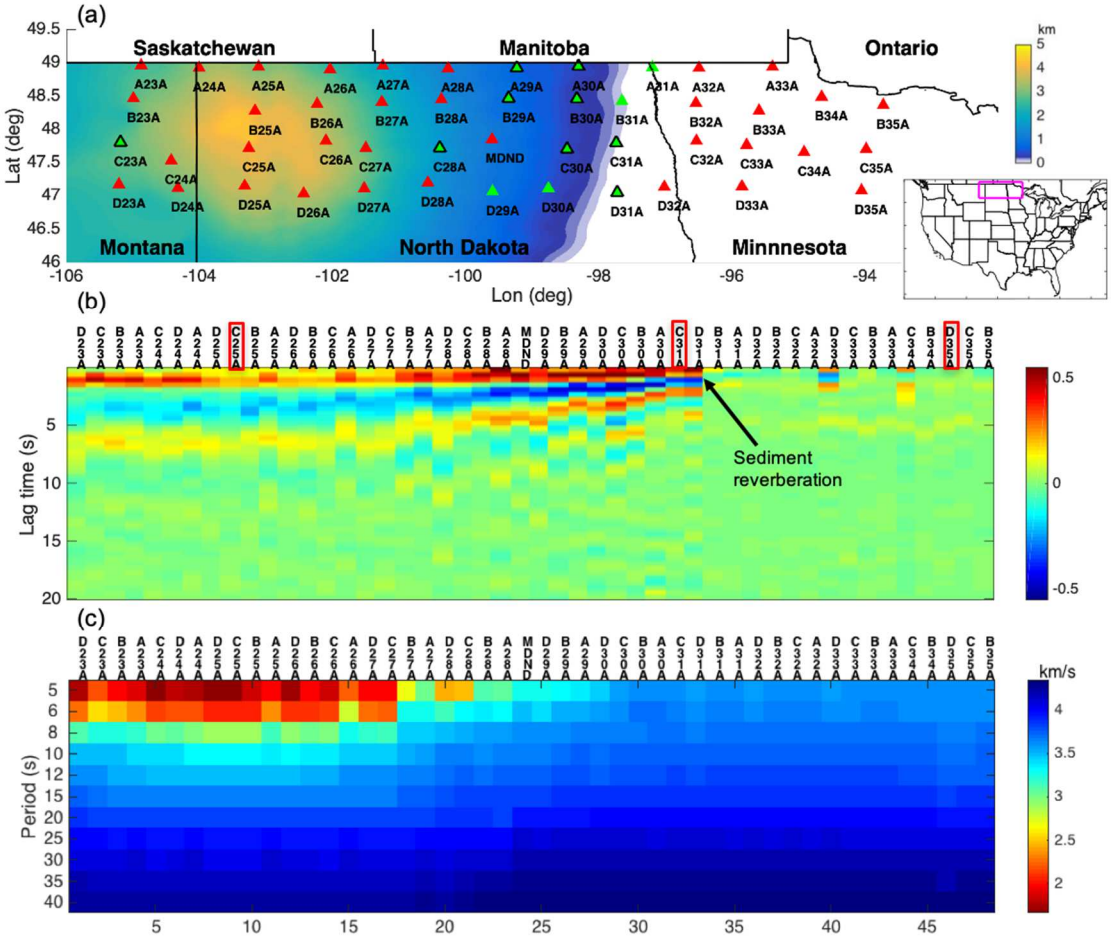


Figure 4.1 (a) Sediment thickness (Marshak et al., 2017) across the study area (magenta box in inset); Transportable Array (TA) stations (triangles). Red triangles show the stations where sediment correction is not needed, green triangles show the stations where the sediment correction is applied in order to retrieve reliable posterior. The green triangles with black contour are the stations where spurious low velocity zones are resolved when sediment correction is not applied to RFs for the joint inversion. (b) Ps RFs and (c) Love wave phase velocities at the TA stations shown in Fig.1a, sorted by longitude from west to east. Three stations are highlighted in (b) as they will be discussed in detail in the paper.

4.2.1 Ps Receiver Functions from Transportable Array

We focus on the east side of the Williston Basin where the thickness of the sedimentary layer varies from 0 to 5 km (See Figure 4.1(a)) from east to west (e.g.

Marshak et al., 2017; Nelson et al., 1993), and utilize data from EarthScope Transportable Array.

At each station, we compute P-to-s receiver functions using 300 second three-component waveforms around the P arrival time from Mw>5.5 events at 30-90° epicentral distance, which are transformed into the P-SV system using the free-surface transform matrix (Kennett, 1991) by minimizing the SV energy on the P component at the time of the P arrival. We follow the Abt et al. (2010) procedure and cull the dataset to include data with a minimum Z-to-R cross correlation of 0.3, and a maximum 25s discrepancy between the automatically determined arrival time and prediction for ak135. We window the P and SV waveforms prior to deconvolution and apply a fourth order Butterworth band-pass filter to waveforms of 0.03-1 Hz. We then use the iterative time domain deconvolution with Gaussian half-amplitude half-width of ~0.5 s to calculate the receiver functions (Ligorria and Ammon, 1999).

This yields between 76 and 338 RFs at each station, which we divide into 2-6 equal-sized bins based on their ray parameters to explicitly account for ray-parameter dependence of RF waveforms. In each RF bin, we apply bootstrap sampling to compute the average Ps RF and estimate its uncertainty. Figures 4.S1 – 4.S3 shows the individual Ps RFs as well as the average Ps RFs after binning for the three stations discussed in this paper. Due to the limited deployment time (18 – 24 months) of the Transportable Array, RF analyses using TA data generally suffer from incomplete back-azimuthal coverage. In this study, we pre-check the RFs with respect to back azimuth to make sure they do not exhibit strong signs of azimuthal anisotropy (Figure 4.S1 – 4.S3). Since we already binned the receiver functions based on their ray-

parameter, and the limited number of events available for the Transportable Array, it is impossible to further divide the dataset by backazimuth to account for anisotropy while maintaining robustness of the RF estimates. In Figure 4.S1 – 4.S3, individual RFs are colored according to their bin number based on the ray parameter, as described in the main text, to illustrate that this binning allows us to account for potential amplitude variations within the RF.

The average Ps RF for each station are arranged by longitude and shown in Figure 4.1(b). Strong reverberations at the beginning of the Ps RFs are observed to the west, suggesting a layer on top of the crust bounded by large impedance contrast; such signal is not seen to the east. Both features are consistent with the trend of basement depth from Marshak et al. (2017). To quantify uncertainty of binned RFs, we could use RFs within bins as samples for estimating the data covariance matrix (C_D). However, the relatively small number of samples available for this computation reduces its reliability, and yields singular and ill-conditioned C_D (e.g. Ledoit and Wolf, 2004). Instead, we represent C_D with 3 parameters (Kolb and Lekic, 2014) obtained by minimizing the L_1 norm to the average covariance vs. lag time estimated across RFs (See Figure 4.2). The data covariance matrix is estimated at each station following the procedure illustrated in Figure 4.2; therefore, our noise parameterization should be able to capture different noise characteristics at different locations.

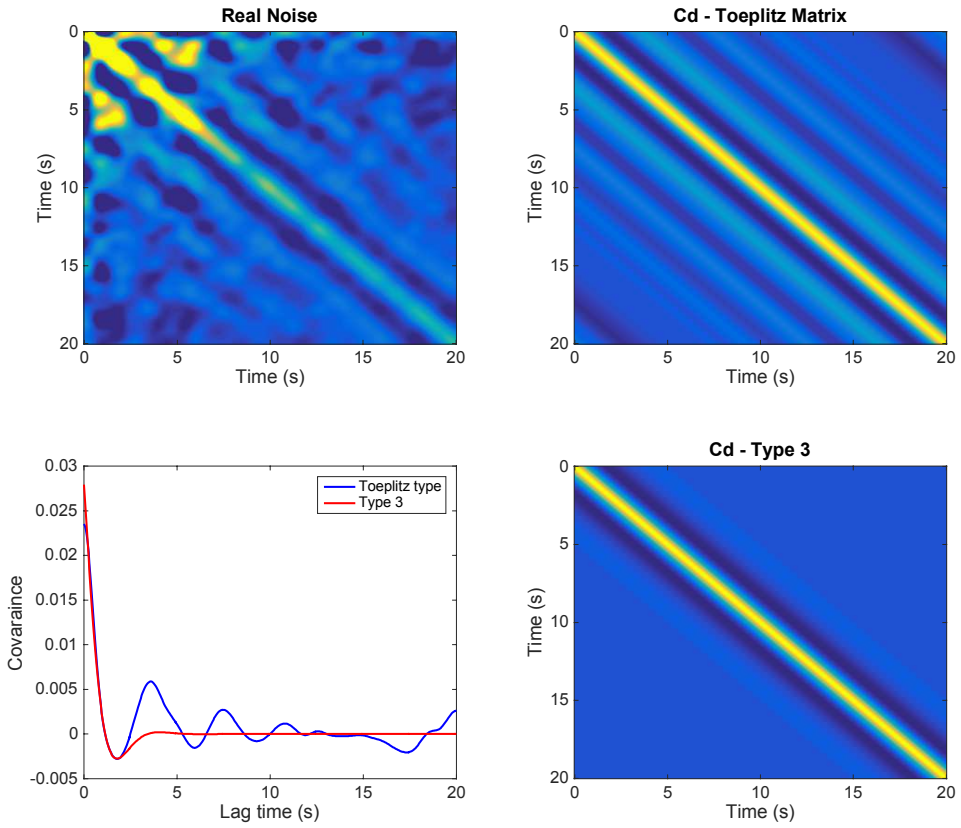


Figure 4.2 Noise characterization of receiver functions. (top left) Data covariance matrix estimated from 338 Ps receiver functions calculated at station MDND-TA. (top right) A Toeplitz-type data covariance matrix constructed from the average covariance as a function of lag-time determined from the data covariance matrix in the top left. (bottom left) Data covariance from Toeplitz type data covariance matrix versus type 3 parameterization proposed by Kolb and Lekic (2014) as a function of lag time. (bottom right) Covariance matrix corresponding to best-fit type 3 parameterization to the average covariance as a function of lag-time.

4.2.2 Surface Wave Dispersion From Transportable Array

At the station coordinates, we also extract Rayleigh and Love wave dispersion curves between 5 and 40s period from the transdimensional hierarchical Bayesian (THB) phase velocity maps (Olugboji et al., 2017) based on ambient noise measurements of Ekström (2017). The THB approach used to create the phase velocity maps yields ensembles of solutions, enabling uncertainties to be quantified. Figure 1c

shows the Love wave dispersion data projected onto the TA stations in the study area; the dispersion curves arranged by station longitude reflect the trends of the sedimentary thickness map, confirming that sediment thickness variations manifest clearly in the SWD data. However, unlike the Ps RF data, SWD data on the edge of the Williston basin do not bear the signature of the sediment layer even at the shortest period measured (See B29A – D31A from Figure 1b and Figure 1c). This difference will be further discussed in the results section.

4.2.3 Transdimensional Bayesian Joint Inversion with a Progressive Inclusion Scheme

Joint inversions of SWD and RFs are highly non-linear and non-unique (e.g. Ammon et al., 1990). Though linearized inversions have been performed in the past (Owens et al., 1984; Kosarev et al., 1993), they are easily trapped by local minima, making the final model strongly dependent on the starting model. Furthermore, linearized inversions often seek a single model that minimizes the misfit function and approximate uncertainty of the model parameters in ways that may not adequately represent the full uncertainty. Here, we opt for a Bayesian inversion using a model-space sampling framework that embraces the non-linearity and non-uniqueness of seismic inversion. Seismic Bayesian joint inversions of SWD and RFs with fixed parameterizations (Shen et al., 2013; 2016) have been used to retrieve an ensemble of lithospheric shear velocity structures compatible with the data, enabling uncertainty quantification on a dataset of continental scale. To allow a more flexible parameterization that introduces less prior information, a transdimensional sampling method was proposed and applied to the joint inversion of SWD and RFs (Agostinetti

and Malinverno, 2010; Bodin et al., 2012). Unlike most of the seismic inverse problems, transdimensional inversions treat the number of model parameters as an unknown. It has been shown that the posterior distribution acquired through transdimensional Bayesian (TB) inversions is naturally parsimonious (Malinverno, 2002) due to the Bayesian formulation of model selection (Malinverno 2000; Sivia et al., 2006). This property of TB inversions punishes more complicated models and therefore restricts the inverse problem from over-fitting the data. Instead, TB inversions produce an ensemble solution where the model complexity is primarily determined by the data itself.

Implementations of TB inversions of SWD and RFs have previously been described in the literature (e.g. Bodin et al., 2012). We adopt a similar approach by applying a reversible jump Markov chain Monte Carlo (rj-McMC) algorithm, but differ from Bodin et al. (2012) in some respects: 1. We explicitly use covariance matrices representing data uncertainty, ensuring optimal information extraction from the two data types; 2. We do not impose fixed V_p-V_s and density-V_s scaling relationships based on prior knowledge, because inverting for V_p/V_s and density along with V_s minimizes potential bias (Dettmer et al., 2015; Kim et al., 2016; Gao and Lekic, 2018). We assign uniform distributions between 0.5-5.5 km/s, 1.6-1.9, and 2.6-3.1 g/cm^3 for V_s, V_p/V_s and density, respectively; 3. We implement a progressive inclusion scheme for the RF data, which is described below.

A major challenge for probabilistic sampling approaches is achieving convergence quickly and efficiently exploring the model space, especially when involving multi-modal density functions such as those associated with RF inversion.

We use a progressive-inclusion scheme to optimize convergence. We start by inverting SWD data for $5 \cdot 10^5$ iterations using the approach of Gao and Lekic (2018), who found that convergence is approached in around 10^5 iterations regardless of starting model. Using the latter $2.5 \cdot 10^5$ steps, we construct an estimate of a velocity structure that we then use as a starting model for the joint inversion of SWD and RF data. The joint inversion also proceeds in a progressive fashion, initially including the first 3s of the Ps RFs to compute the likelihood using the Mahalanobis distance (for $1 \cdot 10^5$ iterations), before proceeding to include the first 5s for additional $1 \cdot 10^5$ iterations, and finally inverting the 15s-long Ps RFs for additional $3 \cdot 10^5$ iterations to complete the inversion. $1 \cdot 10^4$ samples are drawn from this final chain to create the ensemble solution, which can be used to compute the uncertainty on the V_p and V_s inferences. During final segment of the progressive inclusion, we ensure the misfit remains low and stable along the iteration after the burn-in period. Additionally, we verify that the statistical properties of the models do not change between the first and second half of the ensemble. Finally, we also show the comparison of the model prediction from the posterior and the observation (right panel of Figures 4.S4 – 4.S6) to assess the convergence of the rj-McMC.

4.3 Results

4.3.1 Joint Inversion of SWD and Ps RFs

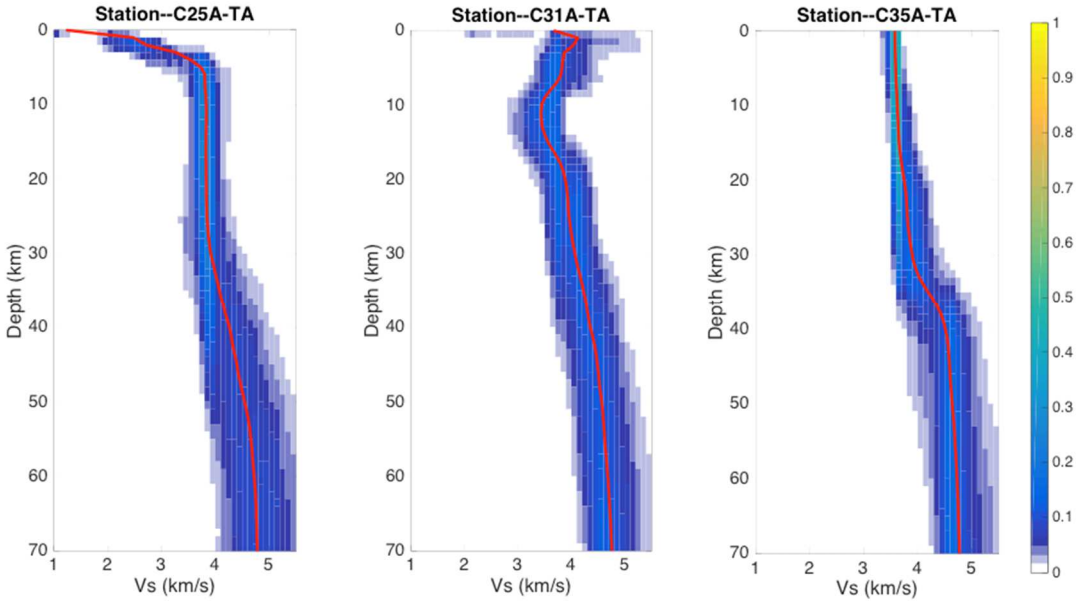


Figure 4.3 V_S depth distributions from transdimensional Bayesian inversion of surface wave dispersion and Ps RFs for stations C25A-TA (left), C31A-TA (middle), C35A-TA (right). The ensemble solutions are displayed as probability density functions at each depth, with warmer colors corresponding to higher posterior probabilities, and the solid red line denoting the 5% trimmed mean of the posterior. NB: The low-velocity layer at 8 -17 km depth at C31A is likely an artifact (see discussion in Section 4.4).

We perform a TB joint inversion of SWD and Ps RFs at three TA stations within the study area, selected as archetypes of different scenarios controlled by the thickness of the sedimentary layer (Figure 4.3). The sedimentary layer thicknesses for C25A, C31A and C35A are approximately 4, 1 and 0 km, respectively (Marshak et al., 2017). Though we invert for structure down to 100 km depth, here we only present the top 70 km in order to highlight the crustal structure. To aid in the identification of

layering in the ensemble models, we compute and plot the probability of a V_S change – called a transition probability – at every 1 km depth (see Figures 4.S4 – S6).

The posterior distributions of V_S retrieved from the joint inversion are intriguing. For station C35A, we see a 36 km thick, two-layer crustal structure. According to the transition probability, the most abrupt V_S jump within the crust occurs around 20 km. The average shear velocity above 20km is 3.62 km/s, while the average shear velocity between 21 km and 36 km is 3.86 km/s. No sedimentary layer is resolved at this location. The SWD and RF predicted by the ensemble solution reasonably fit the observations (Figure 4.S6), confirming that the models are consistent with the data.

Unlike C35A, both C25A and C31A resolve a thin, low velocity layer at shallowest depths, which we interpret as the sedimentary layer. The posterior distribution of V_S at C25A resolves two layers within the 2.9 km thick sediment package (See Figure 4.S7 for a zoomed in version): The top layer has an average V_S of 1.12 km/s and a thickness of 0.8 km; the bottom layer has an average shear velocity of 2.47 km/s and a thickness of 2.1 km. No clear crustal layering is resolved for this station and the Moho depth is around 50 km where the trimmed mean of V_S changes from 3.8 to 4.6 km/s. The V_S structure retrieved from C31A station shows a one-layer, 0.8 km thick sedimentary layer (See Figure 4.S7). Based on the estimate of trimmed mean, the shear velocity changes drastically from 1.90 to 3.74 km/s right below the sedimentary layer. While the crust beneath C31A appears to be relatively fast in V_S in the upper crust, V_S drops significantly at around 8 km depth to 3.57 km/s, and an intra-crustal LVZ appears to span the 8-17 km depth range. To demonstrate that such LVZ

feature is not due to the choice of the summary statistics of the trimmed mean, we analyze the posterior distribution obtained by the Bayesian inversion, which is shown in Figure 4.S7. A multi-modal velocity distribution below the C31A sediment layer is observed across the ensemble, demonstrating the non-uniqueness of the inversion due to lack of data constraints. The posterior distribution of C31A between the depth of 5 and 15 km appears to be non-Gaussian, with a switch of skewness from positive to negative around 8 km. In contrast, the posterior at the same depth for C25A is more Gaussian-like with a tighter variation, suggesting a relatively well-constrained velocity structure. In Figure 4.S8, we show the individual models from the ensemble solution – representing samples from the posterior – sorted by their fit to the data in an ascending order. This arrangement makes it very clear not only that a significant portion of the posterior bears a clear LVZ signature in the mid-crust of C31A, but that LVZs are more prominent on the left, meaning that models with LVZ are preferred by the data. The Moho depth is inferred at 40 km. For both stations, the SWD and RF predicted by these structures agree well with observations (Figure 4.S4, S5).

4.3.2 Joint Inversion of SWD and Ps RFs, Correcting for Sediment Layer

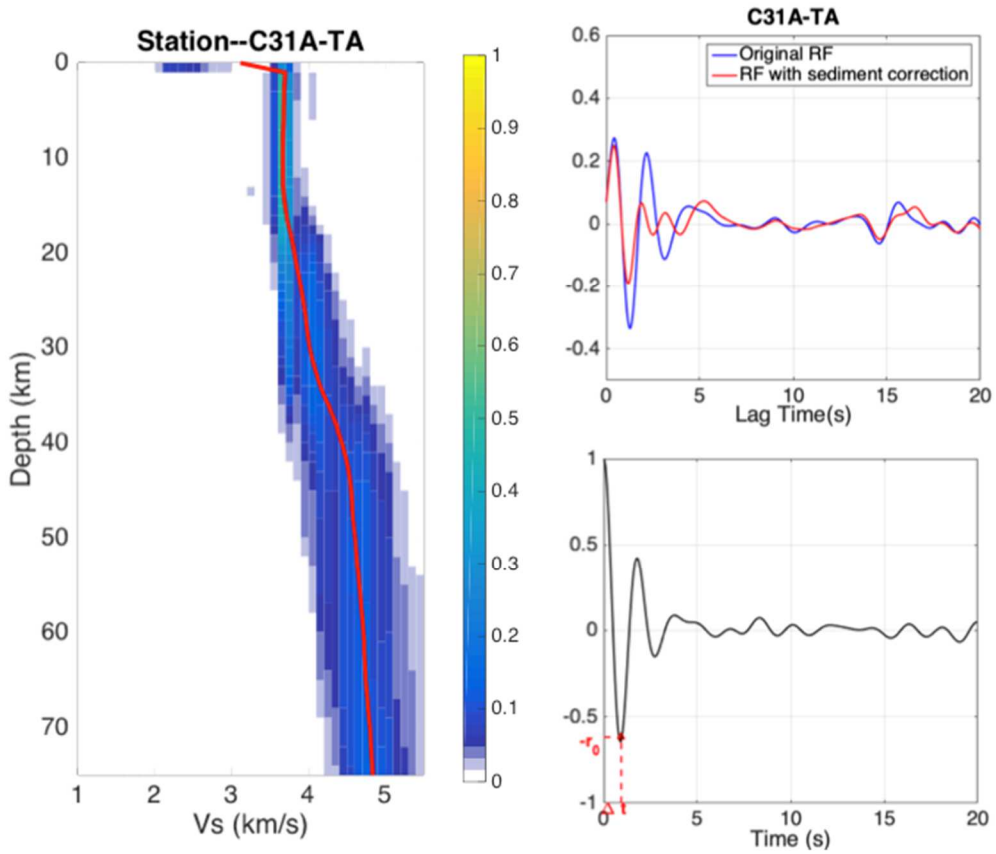


Figure 4.4 (left) V_S depth distributions from TB inversion of SWD and resonance-removed Ps RF data at station C31A-TA. The ensemble solutions and the 5% trimmed mean of the posterior (red) are shown as in Figure 4.3. (right top) Comparison of RFs before (blue) and after (red) applying the resonance removal filter. (right bottom) Autocorrelation of original RF showing how Δt and r_0 are measured.

The significant impedance contrast between the sedimentary layer and the underlying crystalline crust can produce large amplitude reverberations that appear as a decaying, oscillatory signal in RFs. These shallow layer reverberations can overprint signals from direct conversions across the Moho and other intra- and sub-crustal interfaces, making it difficult to resolve accurate crustal structure (Zelt and Ellis, 1999). Yu et al. (2015) proposed an approach to effectively remove such near surface reverberation by applying a resonance removal filter. To assess the significance of

these reverberations to joint SWD and RF inversion for crustal structure, we use this filter to obtain resonance-removed RFs that we invert jointly with SWD data using our TB approach.

To construct the resonance removal filter, the travel time of the S reverberation in sediment, Δt, and the relative strength of the Ss reverberation, r_0 are needed. The effect of the reverberation removal filter on the RFs can be dramatic. Figure 4.4 shows a comparison between the original mean RF and the mean resonance-removed RF for station C31A. It can be seen that the potential sediment reverberation right after the direct sediment conversion has been significantly reduced after applying the filter, clarifying signals from later lag-times, corresponding to greater depths.

After computing the resonance removal filter for each station, we repeat our joint TB inversion using the resonance-removed RFs but introduce a fixed one-layered "sediment" at the top of the model. Since the travel time of the S reverberation in sediment, Δt, is used to construct the removal filter, our fixed "sediment" layer needs to have the correct combination of thickness and V_S to be consistent with the Δt. During the inversion, we first propose a V_S value for the "sediment" layer, V_{s0}, and then calculate the thickness, H, using the equation:

$$H = (\Delta t/2)\sqrt{V_{s_0}^{-2} - p^2}, \quad (4.1)$$

where p stands for the median ray parameter of the binned receiver function. At each step of the rj-McMC chain, we introduce a 1-in-5 chance of perturbing V_{s_0} in the range of $0 - V_{s_1}$ km/s; where V_{s_1} is the shear velocity value of the layer right beneath the fixed "sediment" layer. This allows H to vary during the inversion while remaining consistent with the resonance removal filter.

The V_S posterior distribution of the inversion is show in Figure 4.4. The retrieved sedimentary layer has a thickness of 1 km and an average V_S of 2 km/s, which is similar as the sedimentary layer resolved in the inversion with original RFs. The most prominent difference between the two inversion results is the crustal structure. In this inversion, no crustal LVZ is resolved; the V_S increases from 3.59 km/s to 4.04 km/s along the depth, with one abrupt V_S jump around 18 km. The Moho depth is inferred at 37 km.

Apart from the inversion with station C31A, we also perform TBI with SWD and resonance-removed RFs for C25A, where the rj-McMC struggles to converge. Given the same number of iterations, the retrieved ensemble for C25A (Figure 4.S9) resolves multiple sharp V_S drops within a small depth range. Meanwhile, the predicted RFs and Love wave dispersion at 5, 6 second from the ensemble fits poor with the observation, suggesting the inversion is unable to converge, especially for the shallow structure.

4.4 Discussion

The TB inversion results from the three selected stations show intriguing and contrasting effects of sediment on the retrieval of structure. The three selected seismic stations are located in the Superior Craton and Trans-Hudson Orogen (Whitmeyer and Karlstrom 2007). Crustal LVZs are not expected in this region, and active source seismological studies find no evidence for them (Nelson et al., 1993). At C35A, where no sediment layer is present, the TB inversion resolves a seismic structure where the V_S increases with depth. In that case, our implementation of rj-McMC with a

progressive inclusion scheme enables a fast, stable convergence. The shear velocity structure is well constrained in terms of both absolute value and sharp changes such as crustal layering and Moho.

At station C31A, lying on 1 km of sediment, we resolve the sedimentary layer in both inversions, whether using the original RFs or the resonance-removed RFs. However, the inversion using the original RFs also contains an unexpected LVZ in the mid-crust; no LVZ is present in the inversion of RFs where sedimentary reverberations have been removed. We interpret the LVZ retrieved as an artifact resulting from the non-uniqueness of inverting RFs containing large amplitude oscillatory reverberations together with SWD measurements that lack constraints on the sedimentary layer, either because it is too thin for the period range of SWD or because it is a local feature not resolved in the construction of the phase velocity maps. Another possibility is that although both RFs and SWD contain sufficient constraints at the topmost depth, they introduce conflicting information leading to the spurious LVZ. To rule out this possibility, we conduct a TB joint inversion with the SWD starting from 15 to 40s and uncorrected RFs (Figure 4.S10). In this way, we further reduce the strength of SWD at shallow depths. The retrieved ensemble still resolves a crustal LVZ, consistent with the interpretation that the presence of LVZ is not due to the conflicting constraints but is instead due to lack of constrains on absolute V_S. This implies that crustal velocity structure should be carefully examined when SWD measurements in short period are unavailable, or when only RFs are used.

We confirm the paucity of SWD constraint on the sediment at C31A, by performing a TB inversion using only SWD. The retrieved V_S structure (Figure 4.S11) does not resolve any low-velocity layer on top of the crust, strongly suggesting that the SWD data do not see a low-velocity layer. Meanwhile, the RF data show large amplitude oscillations at early lag times (Figure 4.4), indicating the existence of a large impedance contrast at shallow depths. When only the Ps RF is used in the TB inversion (Figure 4.S12), the retrieved Vs structure resolves as spurious mid-crustal low-velocity layer, as expected. We stress that absolute velocities are not well constrained in the RF only inversion due to lack of data sensitivity. Once we apply the resonance removal filter, the signals after the first peak are significantly reduced, confirming that the signal comes from the sediment reverberation, rather than from abrupt, intra-crustal V_s drops. Therefore, in joint inversions of SWD and uncorrected RFs, constraints on V_S in the shallowest few kilometers only come from the first few seconds of the RFs, which, on their own, only very weakly constrain V_S (e.g. Bodin et al., 2012). However, since this part of the RFs is dominated by both the direct conversion and sedimentary multiples, the joint inversion readily obtains structures with redundant sharp velocity changes due to the lack to constraints on absolute V_S.

To further investigate the origin of the intra-crustal LVZs, we carry out synthetic tests that reproduce them and illustrate why applying resonance removal to the RFs helps prevent their appearance. Figure 4.5(a) shows two simple velocity models constructed to allow differences in predicted SWD and RFs to be straightforwardly attributed to the presence of a sediment layer. One model (red) has 1 km of sediment (V_S=1.5 km/s), on top of a two-layer crust with a Moho depth of

41km. The V_S is 4.3 km/s below the Moho. The second model is identical to the first, except that it lacks a sediment layer. Differences in predicted SWD due to the sediment are limited to only the shortest periods (Figure 4.5(b)); given that typical measurement uncertainties in this period range are ~0.1 km/s, they would be difficult to detect in realistic data. However, the differences between RFs due to the sediment are much more prominent; S wave reverberations within the sediment produce large amplitude oscillations that overprint the direct Moho conversion.

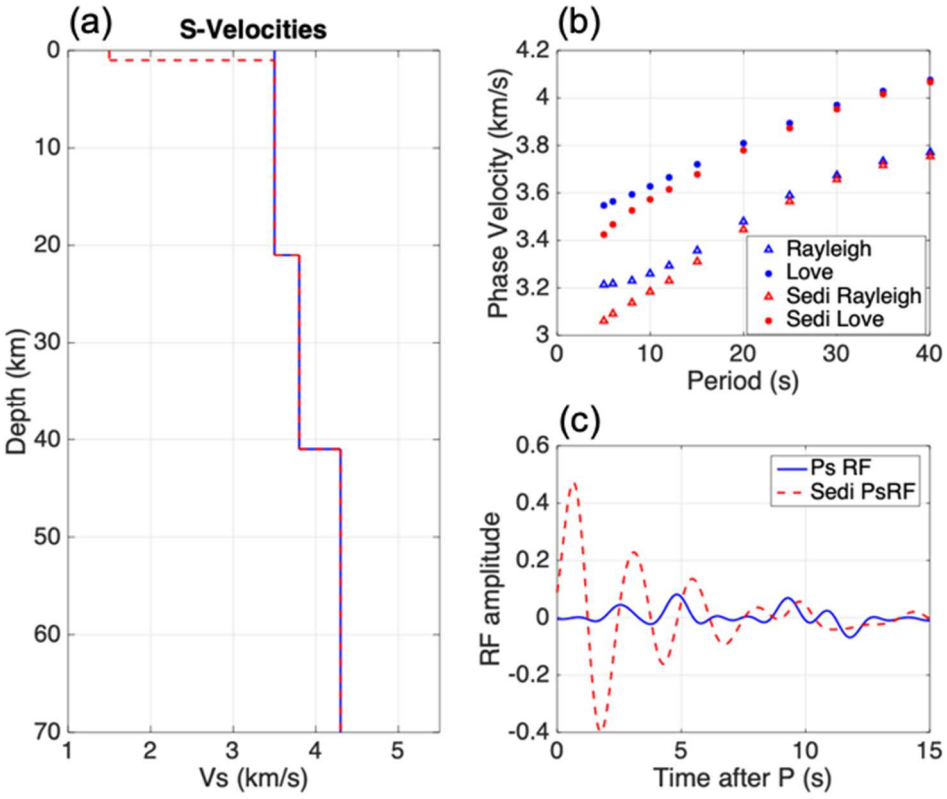


Figure 4.5 (a) 1D synthetic V_S model with (red) and without (blue) a sedimentary layer. (b) Associated Rayleigh and Love phase velocities and (c) Ps RFs from the model with (red) and without (blue) a sedimentary layer.

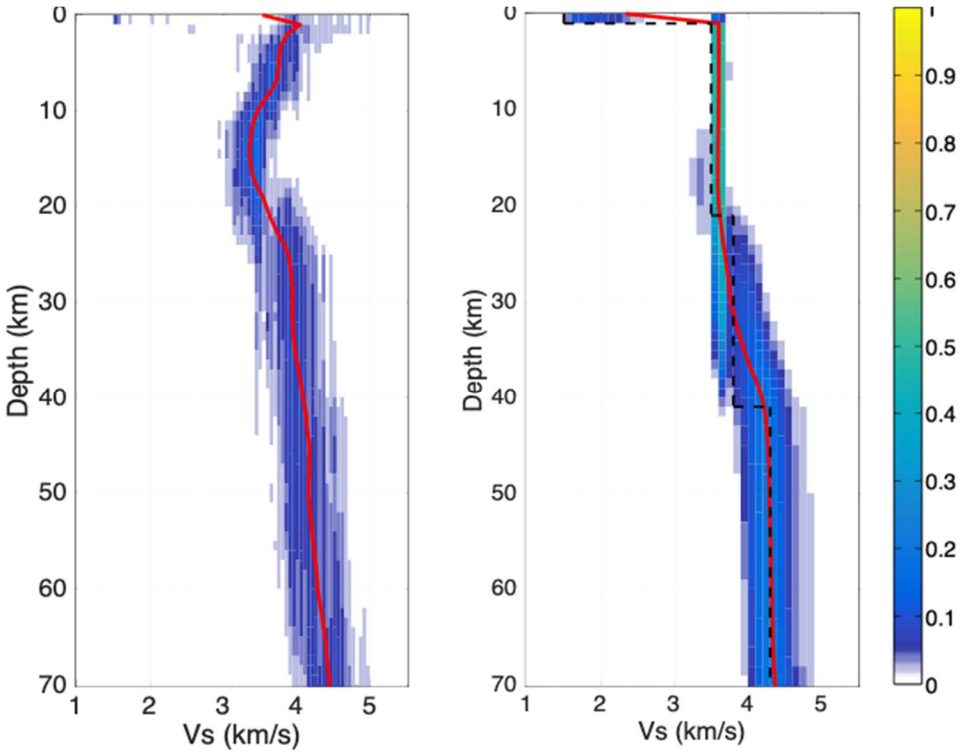


Figure 4.6 (left) V_S depth distributions from TB inversion of synthetic SWD and Ps RF data. SWD data are generated for model without a sedimentary layer, while the Ps RF data is generated for the model with a sedimentary layer. The ensemble solutions are displayed as probability density functions at each depth, and the solid red line denoting the 5% trimmed mean of the posterior. (Right) Same as the left panel, except that the resonance removal filter has been applied to the Ps RF. The model used to generate Ps receiver function is showed in black dashed lines.

In order to re-create a scenario where SWD measurement is incapable of constraining the sediment layer, we perform joint TB inversion using calculated SWD data from the model without sediment together with the RFs from the model with sediment. The retrieved V_S structure (Figure 4.6 (left)) resolves a mid-crustal LVZ that does not exist in either synthetic model. After applying a reverberation removal filter to the RFs and repeating the inversion, the LVZ no longer appears in the retrieved V_S structure (Figure 4.6 (right)).

One might argue that the spurious LVZ results from using SWD predicted from a model without a sediment layer. We agree that given sufficiently accurate and precise SWD measurements, the appearance of the spurious LVZ could be mitigated (Figure 4.S13). However, the situation illustrated by our synthetic test is common in practice mainly due to the way SWD is measured. Unlike RFs, which are single-station-based observables, SWD curves are extracted from phase velocity maps whose resolution is imperfect due to data coverage and modeling assumptions. Therefore, while RF data are sensitive to the subsurface structure directly beneath the seismic station, surface wave dispersion curves might not be, particularly when dealing with structures of limited spatial extent and/or data from less dense seismic deployments.

Previously, Chai et al. (2015) suggested that this potential discrepancy in sensitivity between RFs and SWD should be addressed by spatially smoothing the RFs to make their resolution more comparable to that of SWD maps. Here, we show that this degradation of RF data is not always needed. By applying a resonance-removal filter to our RFs, we are able to reduce the non-uniqueness of the inverse problem. Additionally, by requiring the two-way travel time of S wave in the sediment (Δt) to fit the value estimated from the autocorrelation of the RFs, we pose a stronger prior constraint on the TB inversion to help better constrain the sediment structure. However, crustal V_S structures constrained from RFs only or RFs together with long period SWD data should be interpreted with extreme caution due the possibility of crustal artifacts.

The joint inversion with SWD and resonance-removed RFs for station C25A does not converge in our tests. As is suggested by Yu et al., (2015), the resonance

removal filter may perform poorly when the sediment layer is too thick due to several reasons: the rapid decay of reverberations with time; weak impedance contrast at the bottom of the sediment layer that is unable to generate significant reverberations; or complexity within the sediment layer requiring more accurate modeling of the reverberations. At C25A, we believe that the complex sediment structure is the main reason for the failed convergence of the joint inversion. During the inversion, the resonance removal filter that is applied to all proposed RFs assumes a single-layer sediment, which is unlikely. On the contrary, the inversion with SWD and original RFs resolved a more realistic shear velocity structure, with no crustal artifacts. Unlike station C31A, the SWD data used for station C25A show a strong low velocity signal at shallow depths. Our inversion with only SWD for C25A confirms this observation (Figure 4.S14), the retrieved V_S structure resolves a clear low-velocity sediment layer. We believe that the strong signal of the sediment layer in the SWD data, as well as the relatively low measurement error associated with them, helps better constrain the shallow part of the structure beneath C25A even without applying the resonance removal filter. Therefore, due to better constraints from SWD and the unsuitability of a single-layer resonance removal filter we advise that inversions use SWD and the original RFs in locations where the thick sediment layers are expected. We want to point out that although it is clear that the sediment removal filter to suitable for the thin, single layer sediment setting, sediment thickness should not be used as the sole criteria to determine whether or not to apply sediment correction. Both the layering and the frequency range of SWD used should also be considered.

With limited information of the sediment structure available as prior, it is challenging to determine the complexity of the sediment layering as a guidance for application of sediment removal filter before performing join inversion. Cunningham and Lekic (in review) proposed approaches for evaluating the effectiveness of the sediment removal filter by comparing the change of RF waveform as well as the fit of the RF auto-correlation to a decaying sinusoid. In Figure 4.1 (a), we highlighted the seismic stations where the sediment removal filter is needed to retrieve accurate velocity structures using green color. The sediment thicknesses beneath these stations are also shown in color. In addition, we show the comparisons between the joint inversion of original RFs and of the sediment-corrected RFs for five seismic stations in Figure 4.S15. Crustal LVZs are observed at all five stations, and are effectively removed by the sediment-correction.

We also want to point out that applying a sediment resonance removal filter does not prohibit the retrieval of a real crustal LVZ. In Figure 4.S16, we show a synthetic test where a mid-crust LVZ is imposed in the true model. By performing TB joint inversion of SWD and sediment-corrected Ps RFs, the posterior is able to resolve the LVZ without significantly losing constraints on other features of the true model, such as the sediment layer and Moho.

4.5 Conclusion

In this chapter, we investigated the reliability of crustal structure inferences using an implementation of transdimensional Bayesian (TB) method capable of simultaneously inverting surface wave dispersion and receiver functions. We

documented that shallow sedimentary layers can produce spurious intra-crustal LVZs in joint seismic inversions and proposed an approach for removing these artifacts. We found that in locations where surface wave data and RFs show consistent signal of the shallowest layers – no sediment is expected by both data types or the sedimentary layer is sufficiently thick to be seen by both data types – TB joint inversions are able to retrieve realistic crustal structures. Finally, we showed that applying a sediment resonance removal filter to our RFs can effectively remove the sediment reverberation, enabling recovery of a more realistic crustal model without LVZs.

Seismological studies have identified crustal LVZs in various geological settings (Kind et al., 1996; Beck & Zandt, 2002; Li et al., 2003; Zorin et al., 2002). While most such LVZs are reported in active orogens, some studies have suggested their presence in cratonic regions (Chen et al., 2015). Our results indicate that the inferences of LVZs from joint inversions should be scrutinized in sedimented regions. Furthermore, they provide guidance on best practices for avoiding spurious intra-crustal LVZs and performing reliable joint inversion of SWD and RFs under different shallow-layer scenarios.

4.6 Supplementary Figures

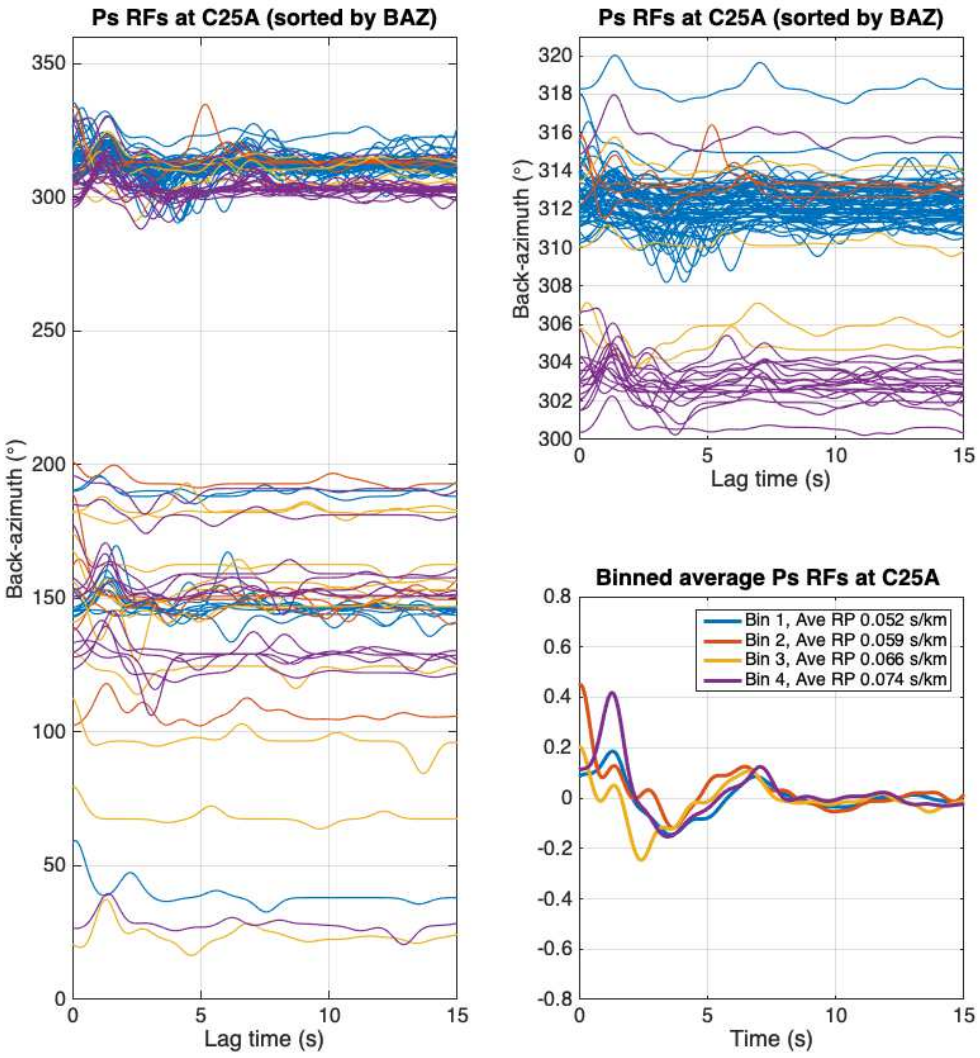


Figure 4.S1 Individual receiver functions at station C25A sorted by the back azimuth (left, top right for zoomed-in view between 300 and 321 degree). Bottom right panel shows the average RF at different bins based on their ray parameters. Individual RFs are colored based on their bin number in the left and top right panel using the same color scheme as bottom right panel.

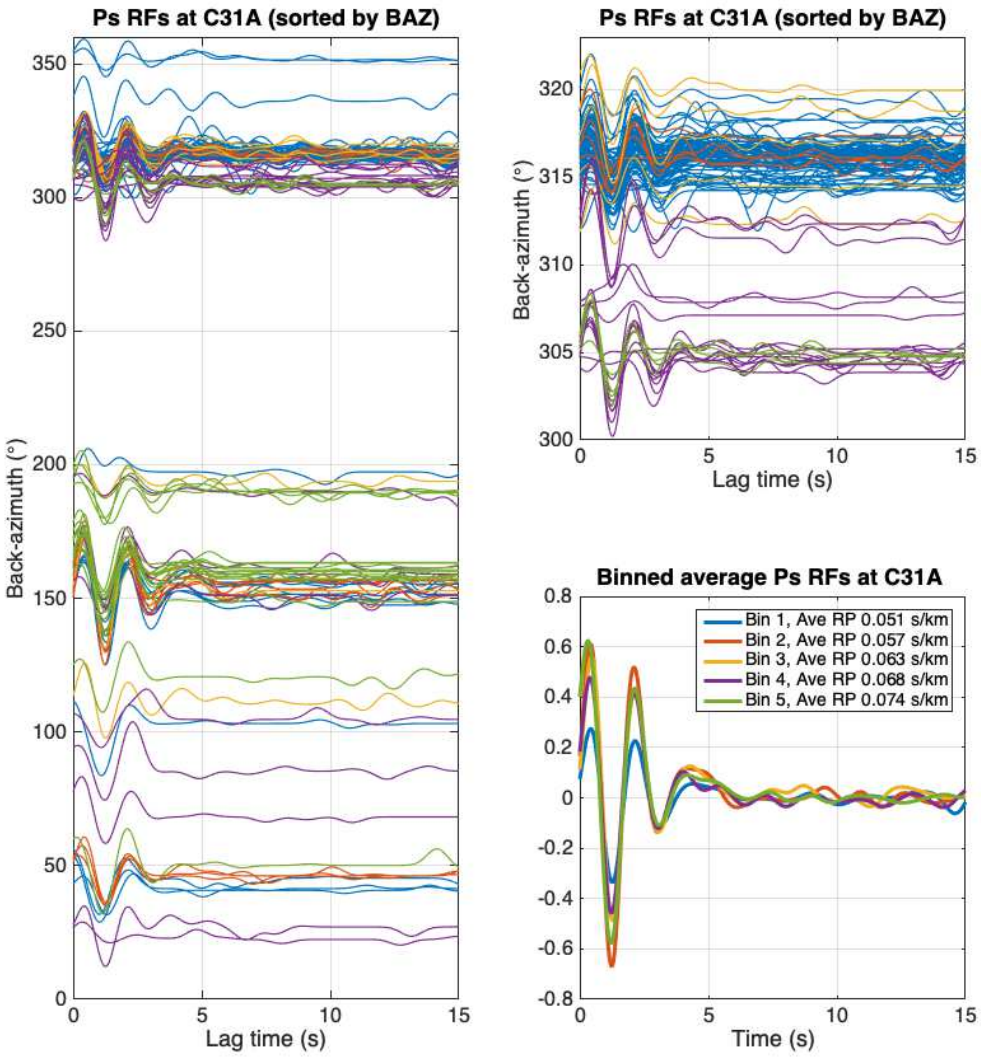


Figure 4.S2 Individual receiver functions at station C31A sorted by the back azimuth (left, top right for zoomed-in view between 300 and 323 degree). Bottom right panel shows the average RF at different bins based on their ray parameters. Individual RFs are colored based on their bin number in the left and top right panel using the same color scheme as bottom right panel.

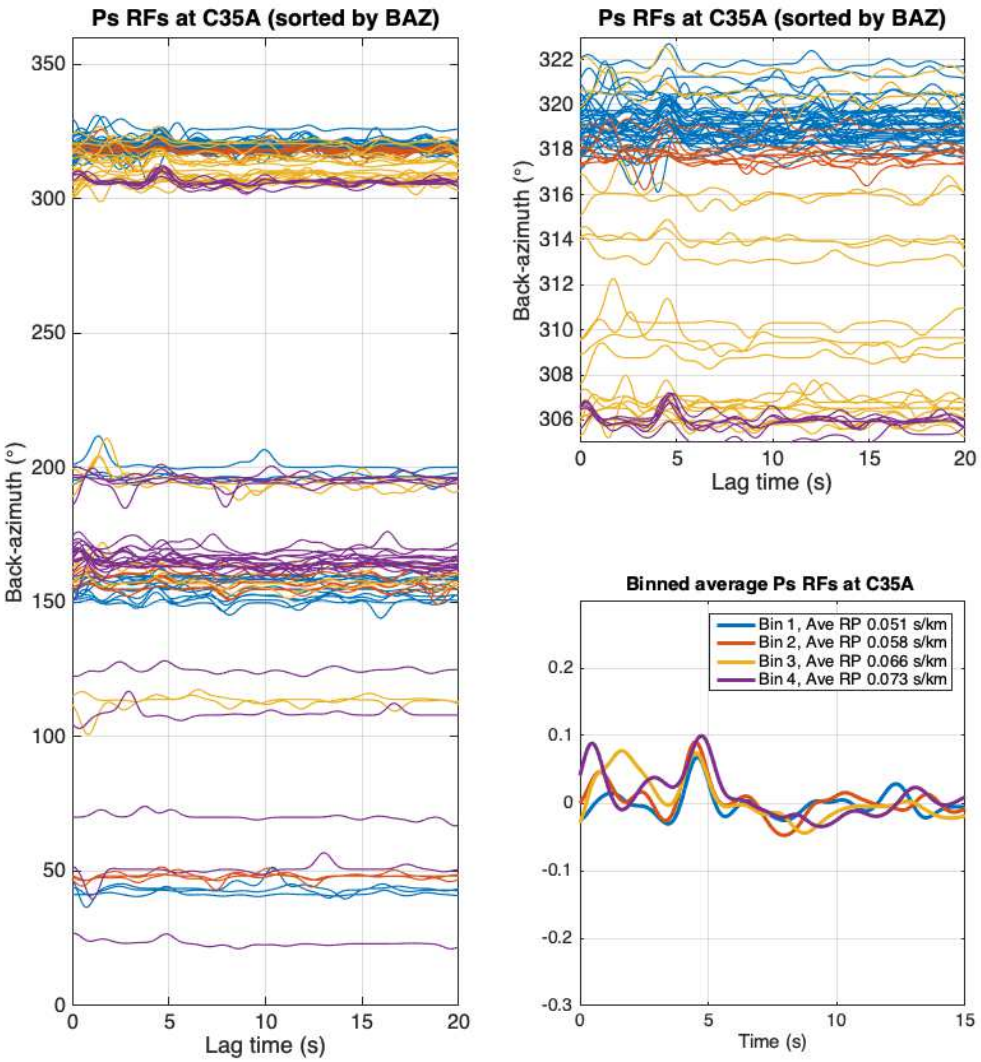


Figure 4.S3 Individual receiver functions at station C31A sorted by the back azimuth (left, top right for zoomed-in view between 300 and 323 degree). Bottom right panel shows the average RF at different bins based on their ray parameters. Individual RFs are colored based on their bin number in the left and top right panel using the same color scheme as bottom right panel.

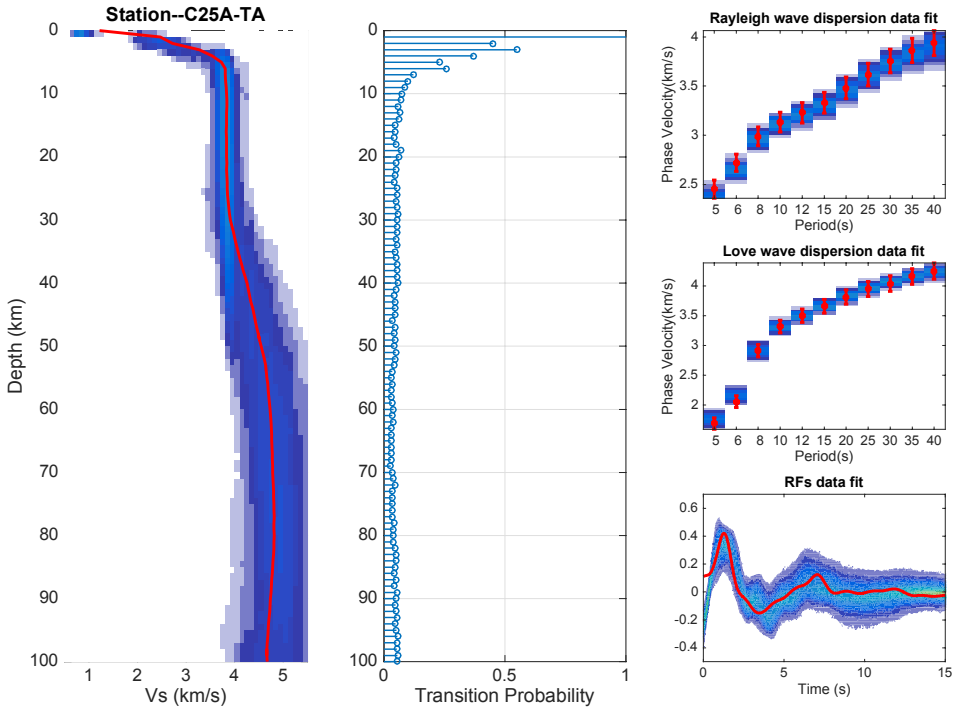


Figure 4.S4 (left) V_s depth distributions retrieved using transdimensional Bayesian inversion of SWD and Ps RF data from station C25A-TA (right: red). The ensemble solutions are displayed as probability density functions at each depth, with warmer colors corresponding to higher posterior probabilities, and the solid red line denoting the 5% trimmed mean of the posterior. (middle) Transition probability of the Vs ensemble calculated at every kilometer. Note that the transition probability only takes into account the occurrences of the velocity change, not the change of value associated with them. Therefore, when large amplitude velocity changes that have a low probability show up as simply low probability ones in the transition probability plots. Conversely, high probability but low amplitude velocity changes show up as high probabilities of transition. (right) From top to bottom: Rayleigh wave dispersion, Love wave dispersion, and original Ps receiver function data fit. Data are denoted in red while values predicted by the ensemble solution are plotted as probability density functions. The inversion scheme is described in the detailed method description section in the supporting information.

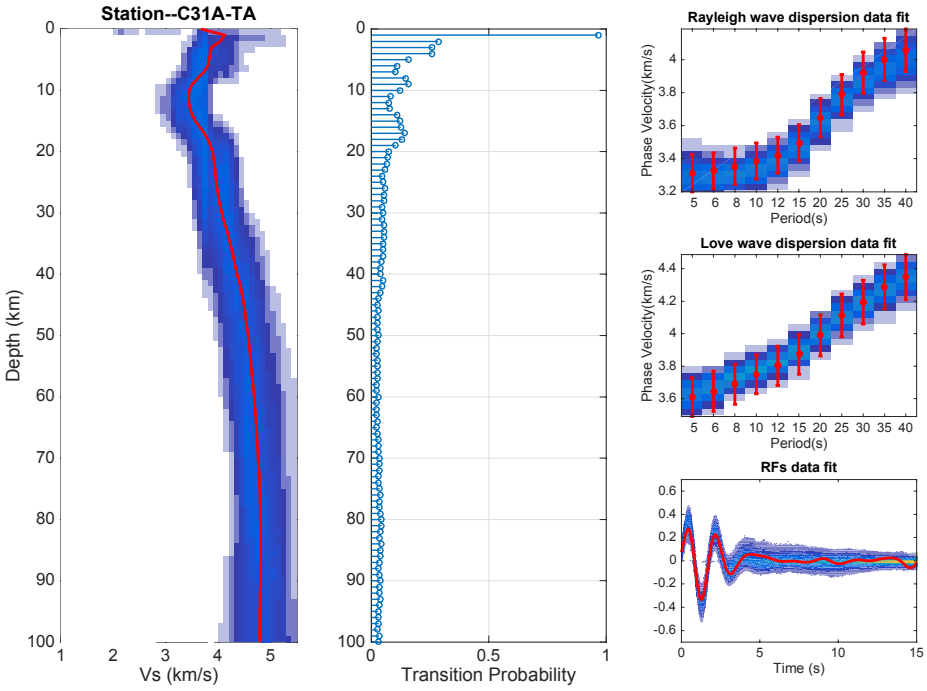


Figure 4.S5 Same as Figure 4.S4, but for station C31A-TA.

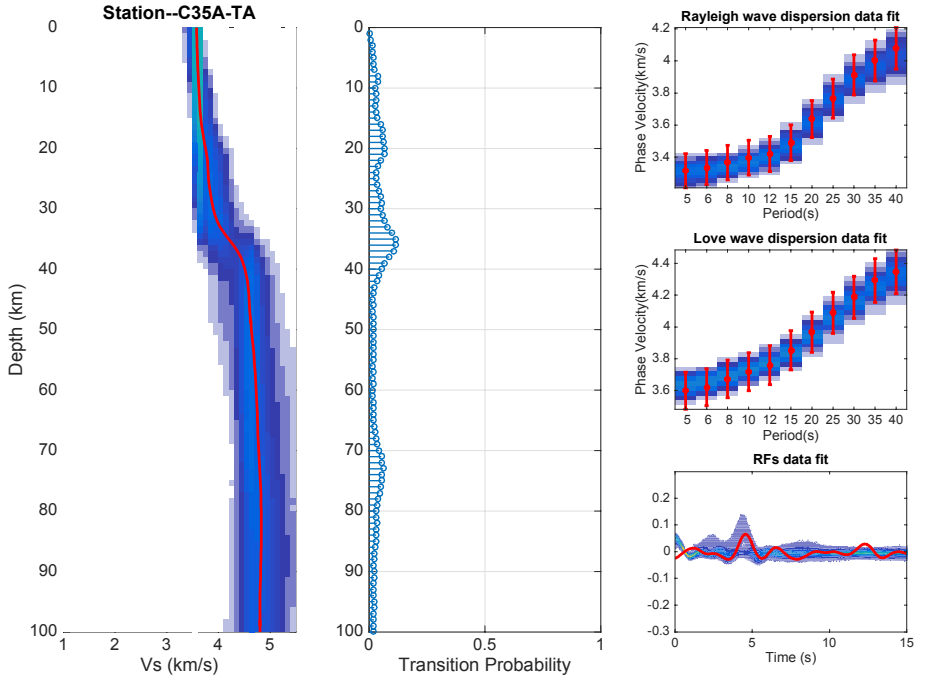


Figure 4.S6 Same as Figure 4.S4, but for station C35A-TA.

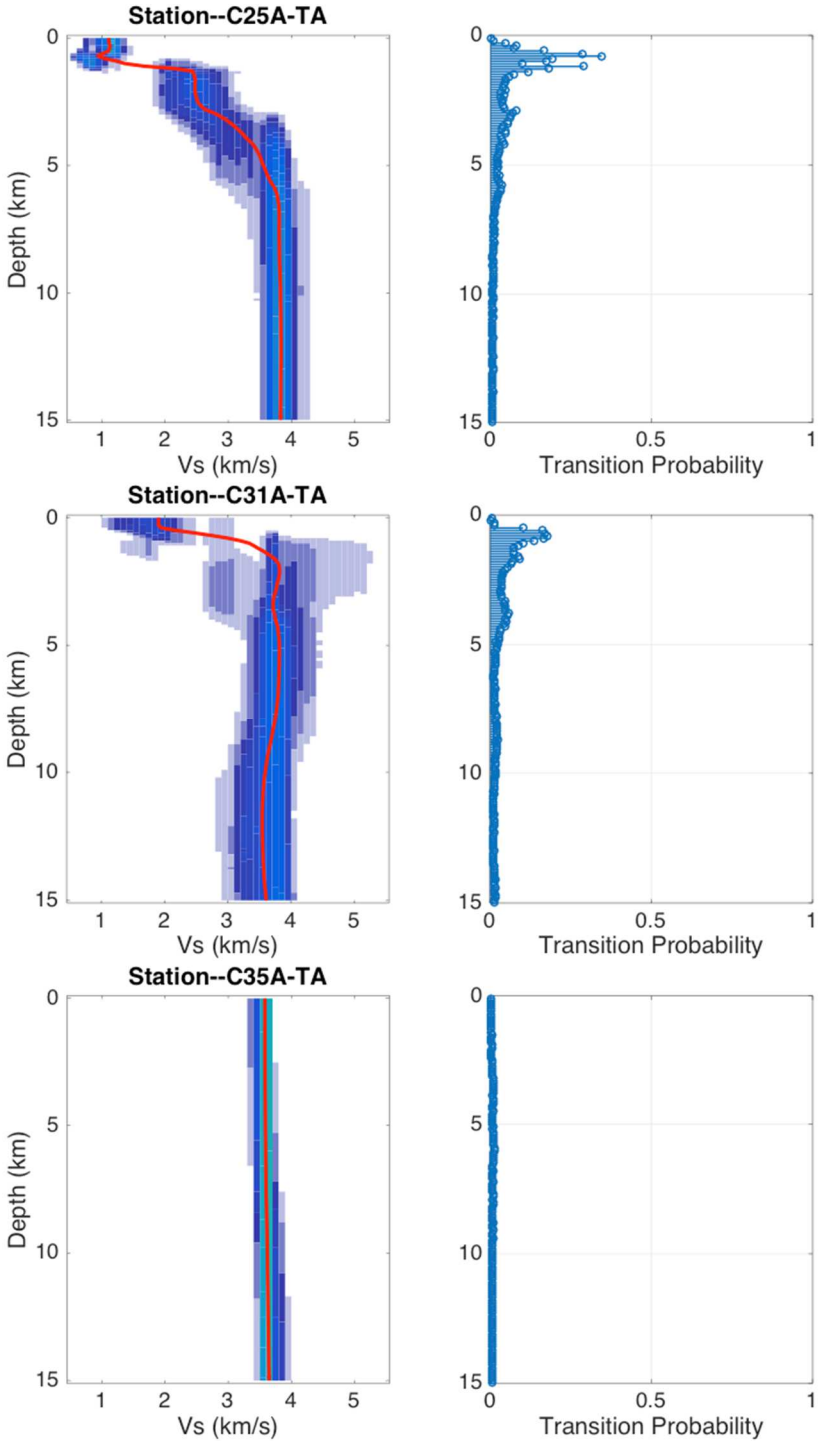
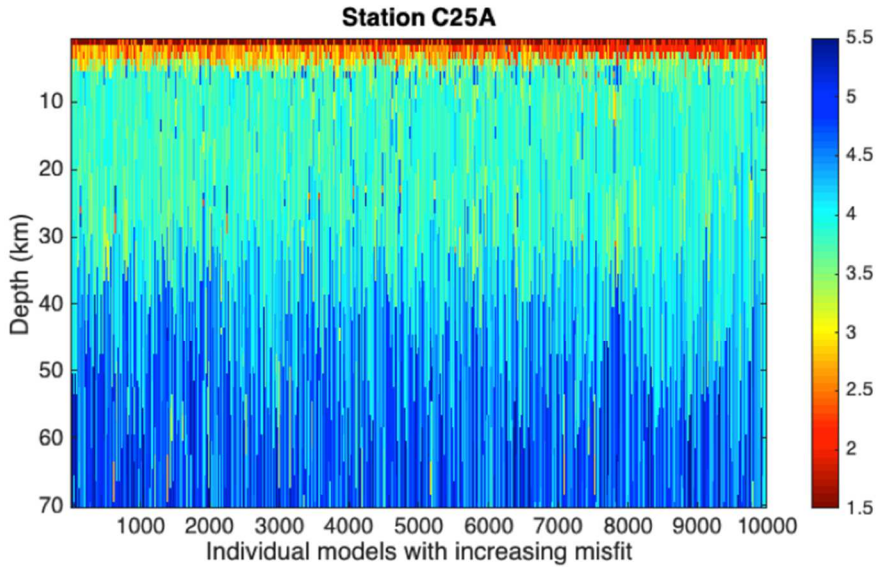
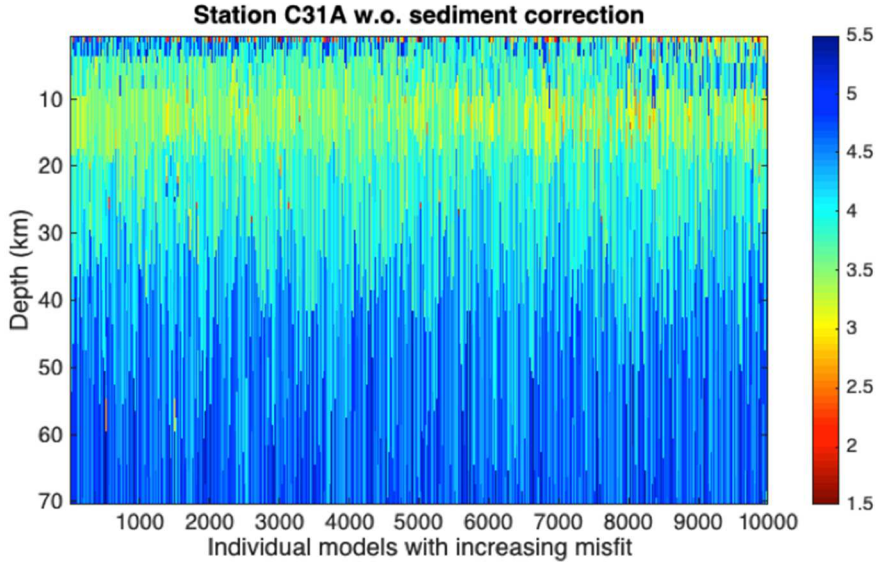


Figure 4.S7 (left) Vs depth distributions at the top 15 km retrieved using transdimensional Bayesian inversion of surface wave dispersion and Ps receiver function data. The ensemble solutions are displayed as probability density functions at each depth, with warmer colors corresponding to higher posterior probabilities, and the solid red line denoting the 5% trimmed mean of the posterior. (right) Transition probability of the Vs ensemble calculated at every 100 meters.


Station C25A
Depth (km)
10
20
30
40
50
60
70
1000 2000 3000 4000 5000 6000 7000 8000 9000 10000
Individual models with increasing misfit
5.5
5
4.5
4
3.5
3
2.5
2
1.5


Station C31A w.o. sediment correction
Depth (km)
10
20
30
40
50
60
70
1000 2000 3000 4000 5000 6000 7000 8000 9000 10000
Individual models with increasing misfit
5.5
5
4.5
4
3.5
3
2.5
2
1.5

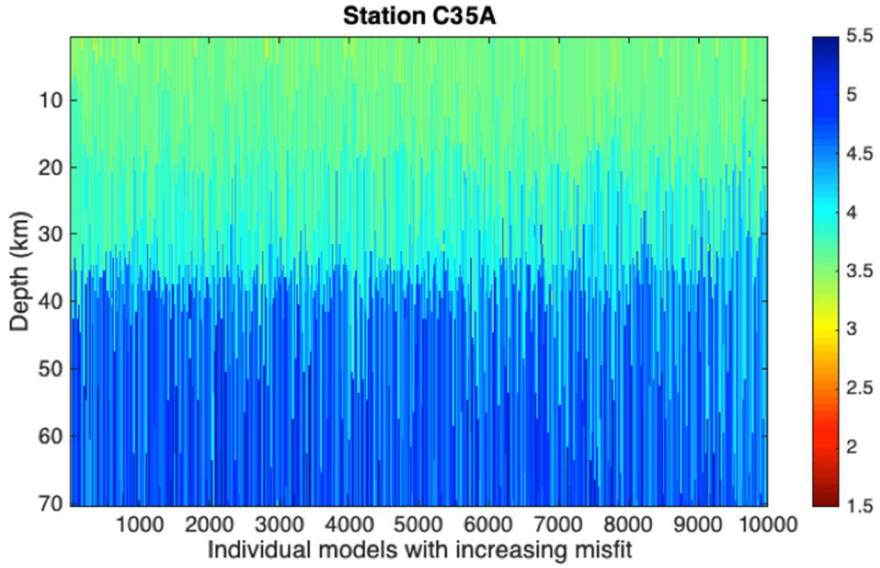

Station C35A
Depth (km)
10
20
30
40
50
60
70
1000 2000 3000 4000 5000 6000 7000 8000 9000 10000
Individual models with increasing misfit
5.5
5
4.5
4
3.5
3
2.5
2
1.5

Figure 4.S8 Individual 1D velocity models from the ensemble solution of C25A, C31A and C35A joint inversion (Figure 4.2) sorted by their misfit in an ascending order.

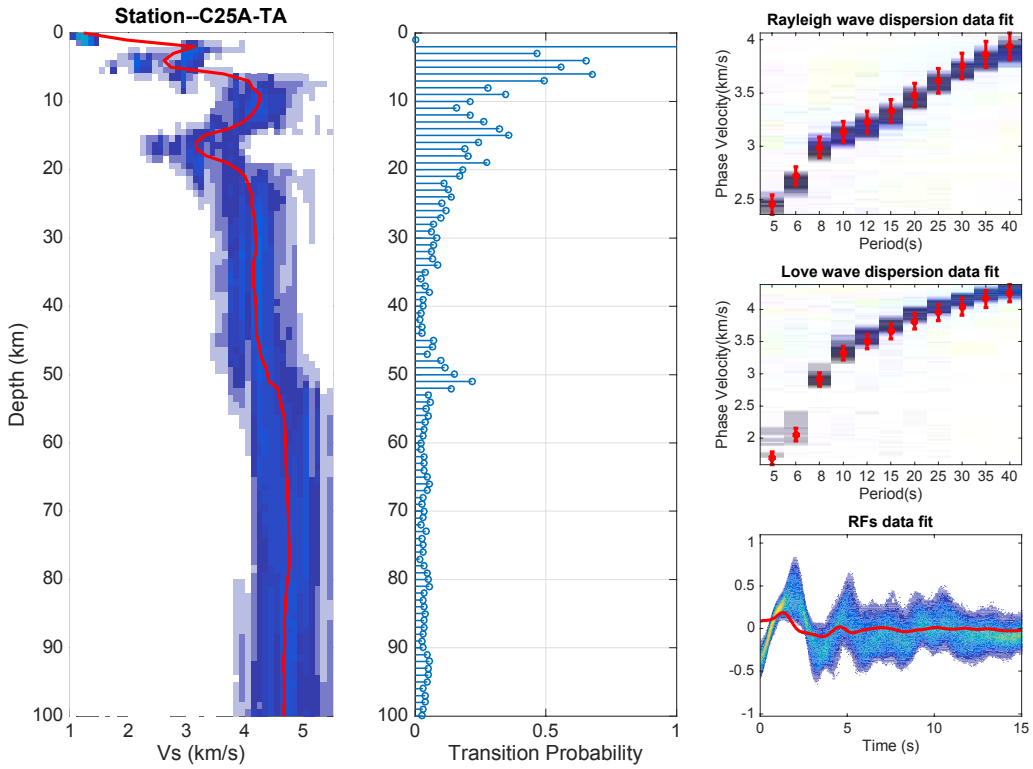


Figure 4.S9 Same as Figure S1, except that the resonance removal filter has been applied to Ps receiver function for C25A-TA.

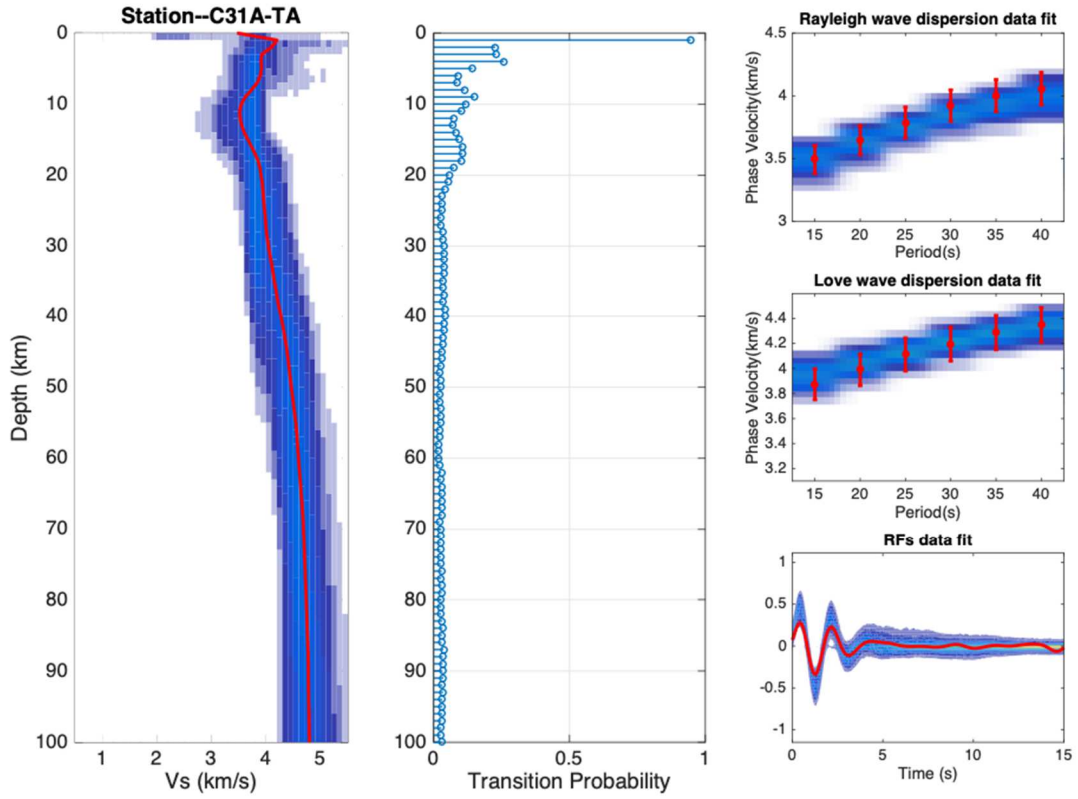


Figure 4.S10 Same as Figure 4.S2, but SWD in the 15 to 40 s range is used here instead of the 5 to 40 s period range used elsewhere.

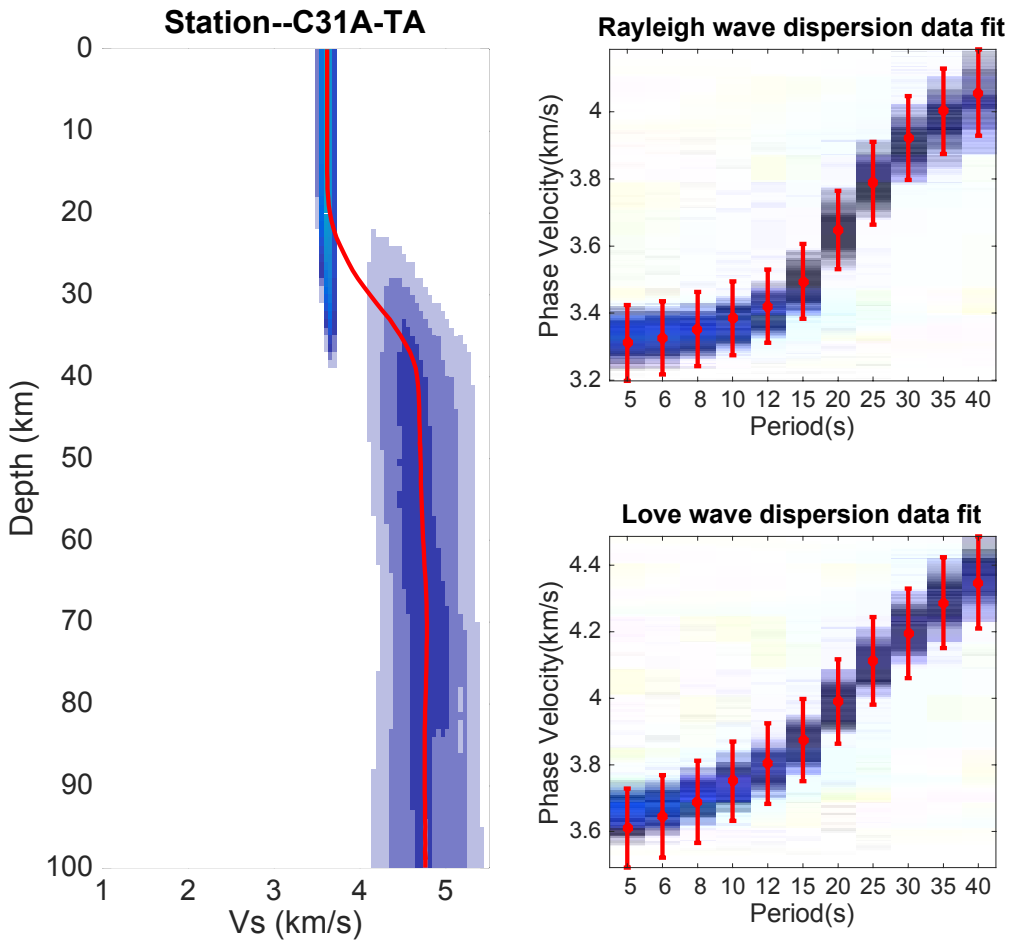


Figure 4.S11 Vs (left) depth distributions retrieved using transdimensional Bayesian inversion of surface wave dispersion data from station C31A-TA. The ensemble solutions are displayed as probability density functions at each depth, with warmer colors corresponding to higher posterior probabilities, and the solid red line denoting the 5% trimmed mean of the posterior. (right) From top to bottom: Rayleigh wave dispersion, Love wave dispersion data fit. Data are denoted in red color; while values predicted by the ensemble solution are plotted as probability density functions.

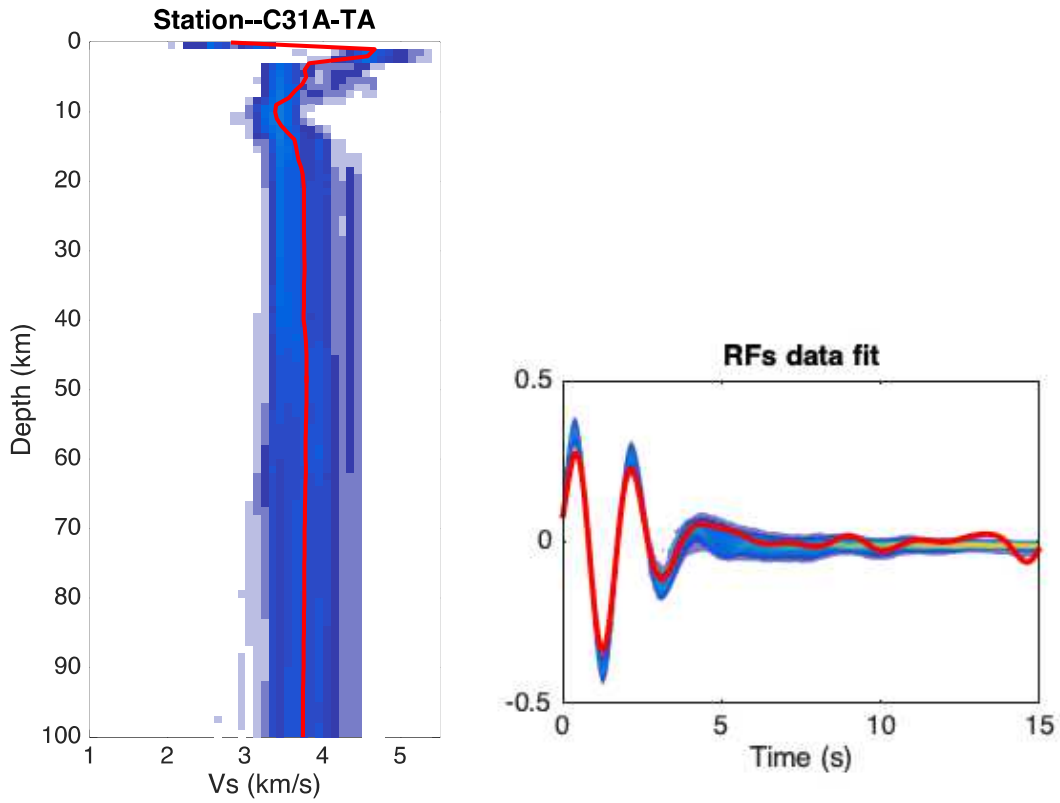


Figure 4.S12 (left) Vs depth distributions retrieved using transdimensional Bayesian inversion of Ps RF data from station C31A-TA (right: red). The ensemble solutions are displayed as probability density functions at each depth, with warmer colors corresponding to higher posterior probabilities, and the solid red line denoting the 5% trimmed mean of the posterior. (right) Original Ps receiver function data fit. Data are denoted in red while values predicted by the ensemble solution are plotted as probability density functions.

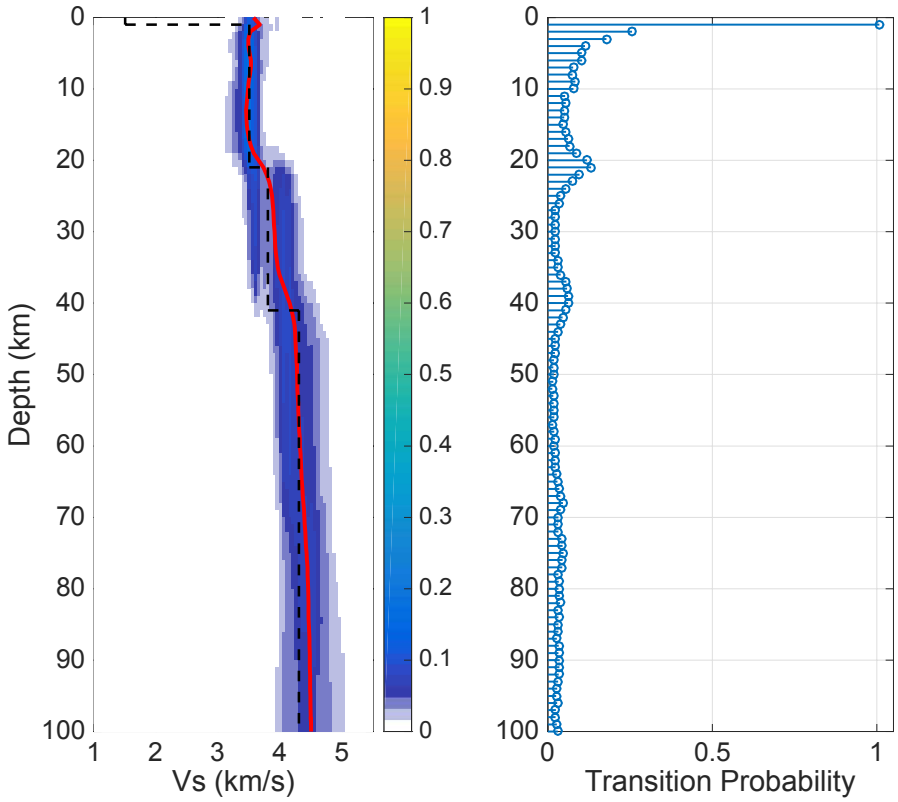


Figure 4.S13 (left) Vs depth distributions retrieved using transdimensional Bayesian inversion of synthetic surface wave dispersion and Ps receiver function data. The ensemble solutions are displayed as probability density functions at each depth, with warmer colors corresponding to higher posterior probabilities, and the solid red line denoting the 5% trimmed mean of the posterior. The target model (black dashed lines) is used to generate synthetic data. (right) Transition probability of the Vs ensemble calculated at every kilometer.

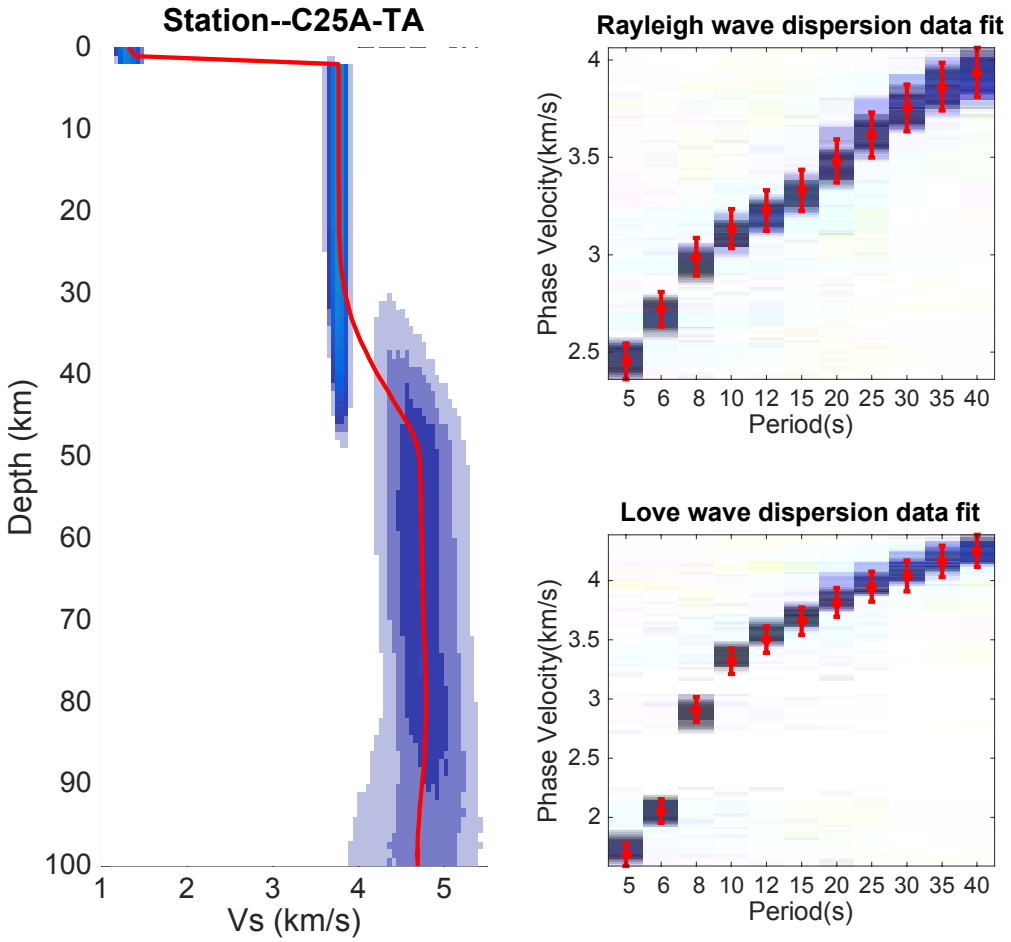


Figure 4.S14 Same as Figure 4.S11, but for C25A-TA.

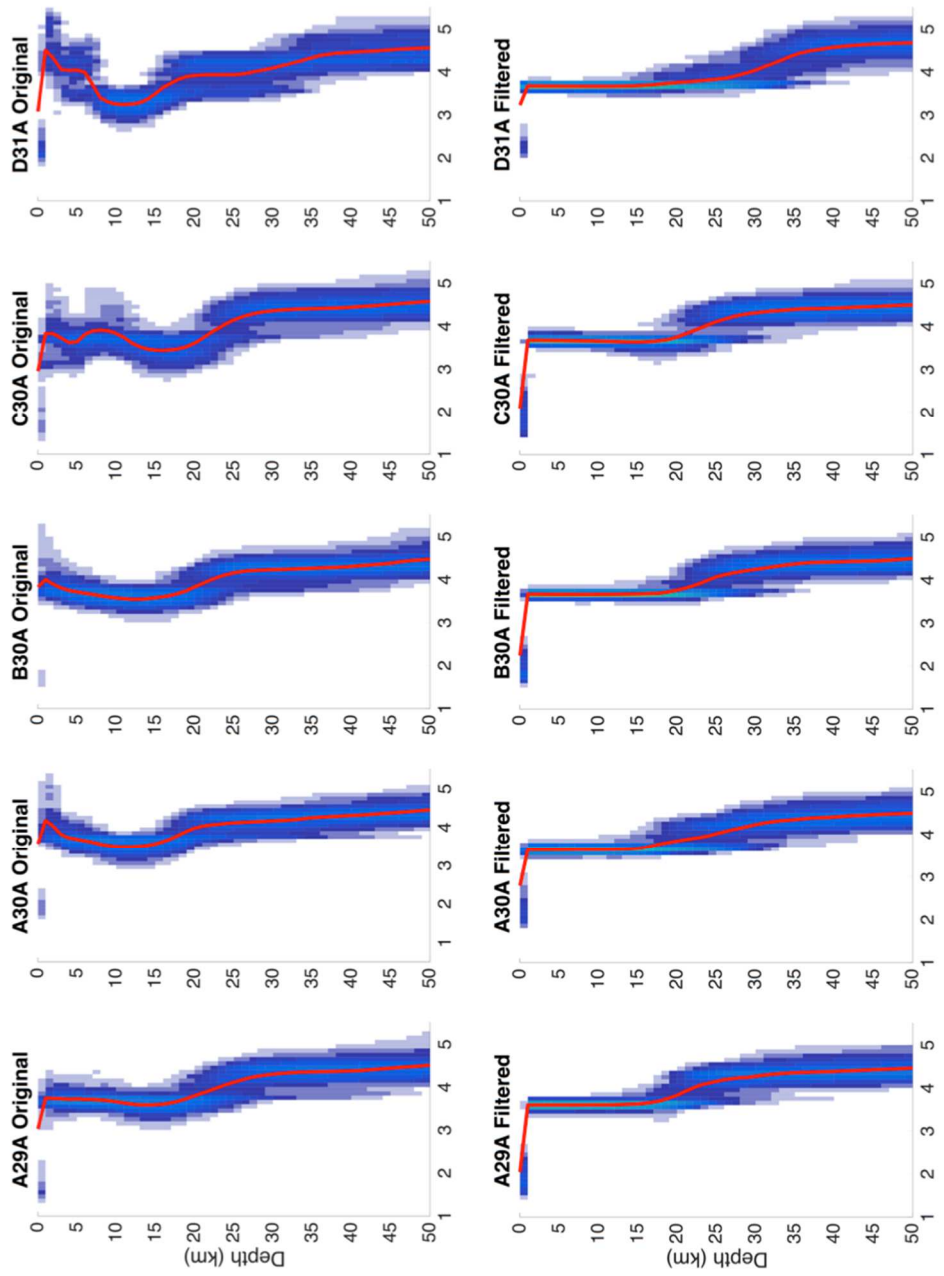


Figure 4.S15 Posterior distribution from joint inversion with SWD and original RFs (upper panel) and from joint inversion with SWD and sediment-corrected RFs (lower panel) for five selected stations.

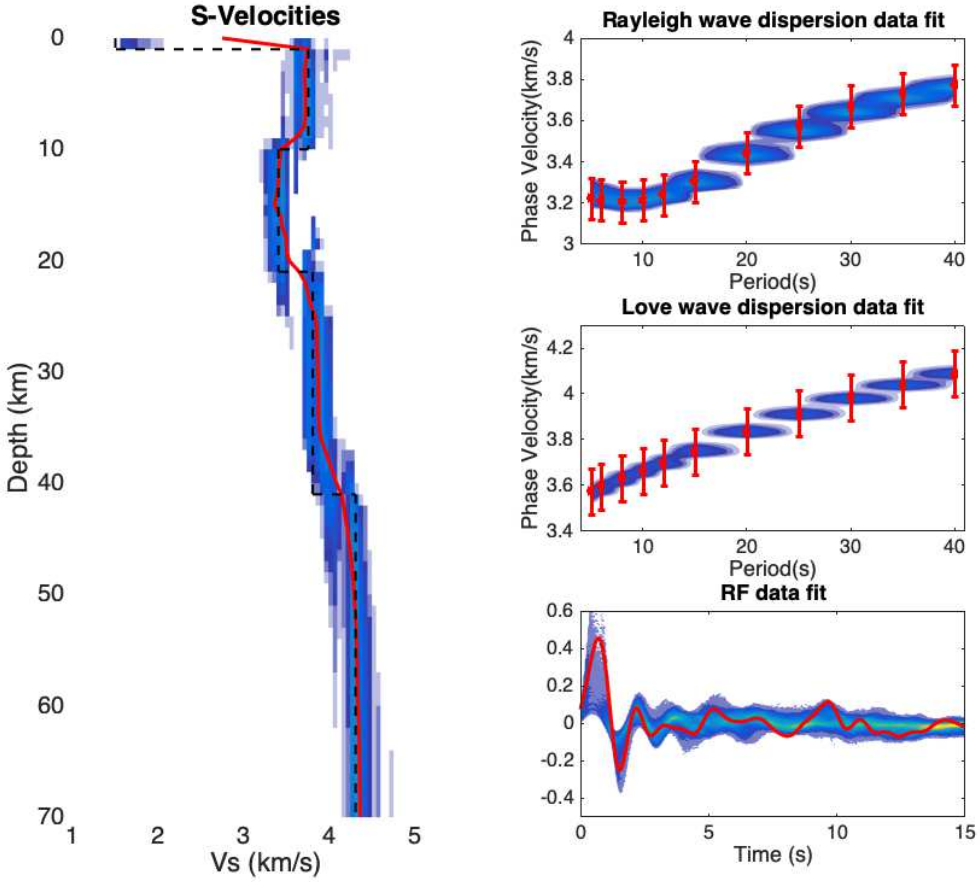


Figure 4.S16 Posterior of transdimensional Bayesian inversion of surface wave dispersion and sediment-corrected receiver functions for synthetic model (shown in black dashed line). The ensemble solutions are displayed as probability density functions at each depth, with warmer colors corresponding to higher posterior probabilities, and the solid red line denoting the 5% trimmed mean of the posterior. (right) From top to bottom: Rayleigh wave dispersion, Love wave dispersion, and original Ps receiver function data fit. Data are denoted in red while values predicted by the ensemble solution are plotted as probability density functions.

Chapter 5: Seismic Imaging of the Northern Great Plains Crust

Abstract

The Northern Great Plains of the United States consists of geological provinces that have been created and modified over 2.5 Ga. The relatively well-preserved history makes it an ideal area to investigate the formation and evolution of ancient continental crust. In this study, we implement a transdimensional Bayesian joint inversion of surface wave dispersion and receiver functions to constrain the crustal structure of the Northern Great Plains (NGP). Significant portion of the region is covered by sedimentary basin, generating large, reverberational signals in the receiver functions that could interfere the direct Moho signal. We account for this effect by applying a resonance removal filter on the associated receiver functions. This allows us to retrieve a high-resolution shear velocity model for the NGP with accurate uncertainty estimates. High-velocity lower crust layers are widely observed among various tectonic settings of the NGP with different thicknesses. We further incorporate our seismic constraints with lab-measured wavespeeds of crustal rocks to infer the composition of middle and lower crust. Due to the elevated shear velocities, we infer a mafic lower crust for the Superior Craton, the Trans-Hudson Orogen and the Central Montana area. Using results from active source seismology, as well as geochemical and petrological constraints, we advocate for the presence of a partially eclogitized root for the Trans-Hudson Orogen (THO) lower crust, and argue that the lower crust of the westernmost Superior Craton bounded by the THO has been lost by delamination.

5.1 Introduction

The continental crust is one of Earth's unique features, as it is chemically evolved compared to that of other planets. Unlike oceanic crust, which is predominantly thin (<8 km), basaltic, and young (<200 Myr), the continental crust is thick (34.4 ± 4.1 km on average (Huang et al., 2013)), billions of years old, and is composed of diverse lithologies. While seismologists have constructed several global crustal seismic velocity models (Mooney et al., 1998; Laske et al., 2013), these suffer from assumptions made to reduce the non-uniqueness of seismic inversions and from extrapolations in poorly sampled areas. Furthermore, choices made during inversions – such as fixed parameterization and regularization – can complicate the full quantification of the model uncertainty (e.g. Gao and Lekic, 2018).

Characterizing crustal structure using seismic data is challenging because accurate modeling requires both the absolute seismic velocities and the locations of sharp changes to be constrained well. However, no single data type provides strong constraints on absolute velocities and seismic discontinuities, suggesting that joint inversions of datasets with complementary sensitivity to crustal structure (e.g. receiver functions and surface waves) are desired. Recent development in seismic imaging techniques (Shapiro et al., 2005; Liang and Langston, 2008) and the installation of EarthScope Transportable Array provide an opportunity for better constraining the crustal structure across the US using both data types. In this study, we perform joint inversions of surface wave dispersion and receiver functions to leverage their complementary constraints. We implement a transdimensional Bayesian inversion

(Malinverno, 2002; Bodin & Sambridge 2009; Agostinetti & Malinverno 2010; Bodin et al. 2012) that can handle flexible parameterizations and disparate noise characteristics of the datasets.

We focus the study on the Northern Great Plains of the United States, as it comprises areas with different tectonic histories, including Archean cratons and Paleoproterozoic orogens. The Northern Great Plains is an ideal location to study the formation and evolution of the continental crust. Archean-age Superior craton has some of the oldest crust on Earth, while the Trans-Hudson Orogen formed during the collision between the western Churchill and Superior plates 1.83--1.80 Ga ago (Corrigan et al., 2009). The THO is the best-preserved Paleoproterozoic orogenic belt on Earth as it did not experience any major tectonic activity since its formation, making it an ideal region to study Precambrian plate tectonics. The Archean Medicine Hat Block and Wyoming Craton are bounded to the east by the THO; the two Archean blocks are sutured together by the Great Falls tectonic zone (Gietti, 1966; O'Neill and Lopez, 1985; Mueller et al., 2002). Recent geological studies on igneous rocks of the Great Falls tectonic zone reflect a collisional boundary (Mueller et al., 2002; Harms et al., 2004). Additionally, the region is constrained by high quality data with uniform density from the Transportable Array. Finally, the region is characterized by uniformly low heat flow throughout, making it more reasonable to attribute lateral variations in seismic wavespeeds to composition rather than temperature (Figure 5.1 (b)).

We carry out joint inversions of Ps receiver functions and surface wave dispersion (5-40s) for shear velocity (V_S) profiles down to ~100 km depth beneath individual seismic stations across the region. We find that in order to retrieve accurate

and reliable crustal structure, it is necessary to correct for the effect of low-velocity sedimentary basins on the receiver function data. Upon applying sedimentary reverberation corrections at a subset of stations, we develop a three-dimensional V_S model, TBI-NGP, for the top 100 km of the Northern Great Plains. We systematically analyze similarities and differences between velocity structures across different tectonic settings, and contextualize our findings with previous geochemical and petrological studies to further infer the composition and formation of the Northern Great Plains crust.

5.2 Method and Data

To fully incorporate the data uncertainties and to embrace the nonlinearity of joint inversion of SWD and RFs, we implement a transdimensional Bayesian (TB) approach to invert for 1D seismic velocity profiles. The model is represented by a vector quantity that includes all the model parameters of interest. Similarly, all observed data comprise a vector *d*. The aim of Bayesian inference is to quantify the posterior probability density $p(m|d)$, which is the probability density of the model parameter given the observed data (Smith 1991).

According to Bayes' theorem (Bayes & Price 1763), the posterior probability density of the seismic inversion $p(m|d)$ can be written as:

$$p(m|d) = \frac{p(d|m) \times p(m)}{p(d)} \quad (5.1)$$

Note that the term $p(d)$, evidence, is not a function of of $\boldsymbol{m}$, and should remain constant as we vary the model parameter under the same setting. Then we have:

$$p(m|d) \propto p(d|m) \times p(m) \quad (5.2)$$

where $p(m)$ is the prior, it is used to describe our knowledge about the model parameters before looking at the data (Sivia and Skilling, 2006). We assign uniform prior distributions between 0.5-5.5 km/s for Vs and 1.6-1.9 ratio for Vp/Vs, and assume no scaling relation between them. $p(d|m)$ is the likelihood that quantifies how likely we would be to observe the data if the actual model were described by the set of parameters in vector *m*. We use Mahalanobis distance to describe the consistency between the predicted and observed data:

$$\Phi(m) = (g(m) - d)^T C_D^{-1} (g(m) - d) \quad (5.3)$$

Here $g(m)$ is the predicted data and C_D is the data covariance matrix. The Mahalanobis distance determines the likelihood of the observed data given the model:

$$p(d|m) = \frac{1}{\sqrt{(2\pi)^n |C_D|}} e^{\frac{-\Phi(m)}{2}} \quad (5.4)$$

where n is the number of points in the data vector.

TB inversion is a model-space search approach based on the reversible jump Markov chain Monte Carlo, which allows the parameterization to adapt as needed to fit the data (Malinverno 2002; Bodin & Sambridge 2009). This property of TB inversion is advantageous when jointly inverting SWD and RFs, given their complementary data sensitivity. Our implementation – including a progressive inclusion scheme – is identical to the method described in chapter four. The inversion result yields an ensemble of velocity structures primarily constrained by the observations and can be used to quantitatively infer the subsurface structure. Although in this study we only focus on Vs structures, we invert for the top 100km of both Vs and Vp/Vs to avoid the potential bias from assuming a scaling relation between Vp/Vs.

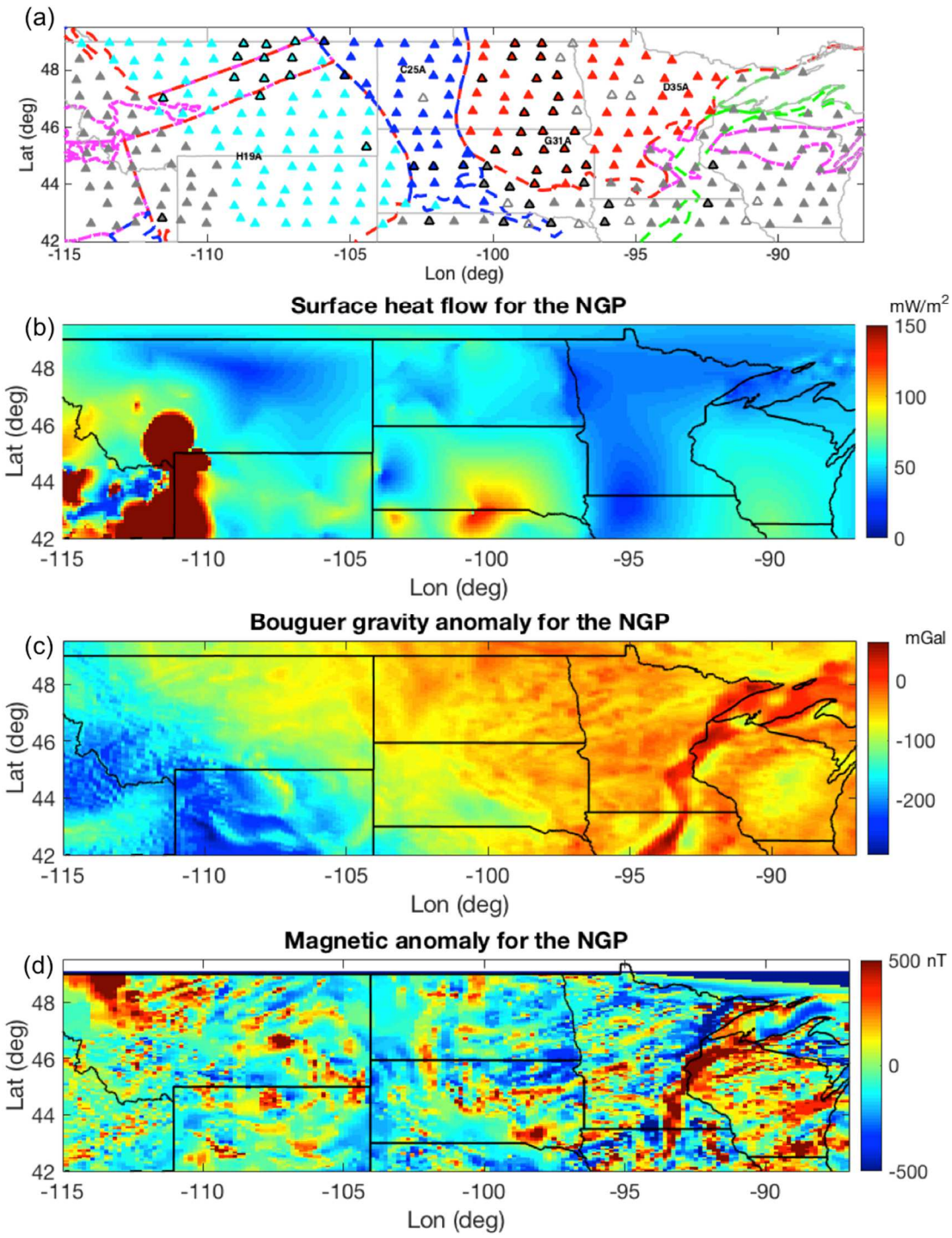


Figure 5.1 (a) Tectonic setting and seismic stations of the study region. Archean Craton (>2.5 Ga), Trans-Hudson Orogen (1.9-1.8 Ga), Mid-Continental Rift (1.2-1.1 Ga) and juvenile volcanic arc (2.0-1.8 Ga) are shown in red, blue, green and magenta dashed lines, respectively (Whitmeyer and Kalstrom, 2007). Transportable Array stations used for transdimensional Bayesian (TB) inversion in this study are shown in triangles. Triangles with black contours are stations where spurious crustal low-

velocity zones are identified and effectively removed when sediment corrections are performed. Triangles with grey edges and no filled color are the stations eliminated in this study due to data qualify. Triangles are colored according to their tectonic settings (Superior Craton -- red; Trans-Hudson Orogen -- blue; Central Montana -- cyan; rest of the stations -- gray). (b) Surface heat flow map of the NGP based on data collection from National Geothermal Data System. (c) Bouguer gravity anomaly map of the NGP (Kucks, 1999). (d) Magnetic anomaly map of the NGP (Bankey et al., 2002)

Figure 5.1 (a) shows the tectonic setting of the Northern Great Plains and the locations of 291 Transportable Array seismic stations used in this study. We use phase velocity maps from Olugboji et al., (2017) and extract both Rayleigh and Love wave dispersion curves between 5 and 40s at the station coordinates. Using ambient noise phase dispersion measurements of Ekstrom (2014), Olugboji et al. (2017) used transdimensional hierarchical Bayesian method to construct phase velocity maps, and associated ensemble solutions, which allow associated uncertainties and tradeoffs to be fully quantified. Figure 5.2 (top panels) shows examples of the surface wave dispersion curves with the associated uncertainties at four stations across the Northern Great Plains.

Earthquakes with Mw > 5.5 in the 30-90° epicentral distance range are selected to compute P-to-s receiver functions. We use an iterative time-domain deconvolution method (Ligorria and Ammon, 1999) based on the data-processing algorithm of Abt et al., (2010). This yields between 66 and 592 RFs at each station. To accurately account for the ray-parameter dependence of RF waveforms, we divide the RFs into 2-6 bins depending on the available number of RFs. Due to the limited number of RFs available in each bin, it is not possible to reliably quantify the full covariance matrix (C_D) describing noise in the RFs. Instead, we approximate C_D at each station and for

each ray-parameter bin with a parameterized form advocated by Kolb and Lekic (2014). Values of the three quantities describing the C_D in this approximation are obtained by minimizing the L1 norm to the average covariance vs. lag time estimated across RFs. We note that this representation of noise is more realistic than commonly assumed white noise (see Figure 4 in Kolb and Lekic, 2014).

Figure 5.2 (lower panels) shows the average RFs in each bin at four stations representative of the study area. At each station, small variations in the RF amplitude are seen between bins due to ray-parameter dependence of transmission and reflection coefficients. Substantial differences in RF waveforms are observed in stations such as C25A-TA, where the complex sediment structure (Gao et al., in review) introduces contrasting direct and reverberational signals at different ray-parameters.

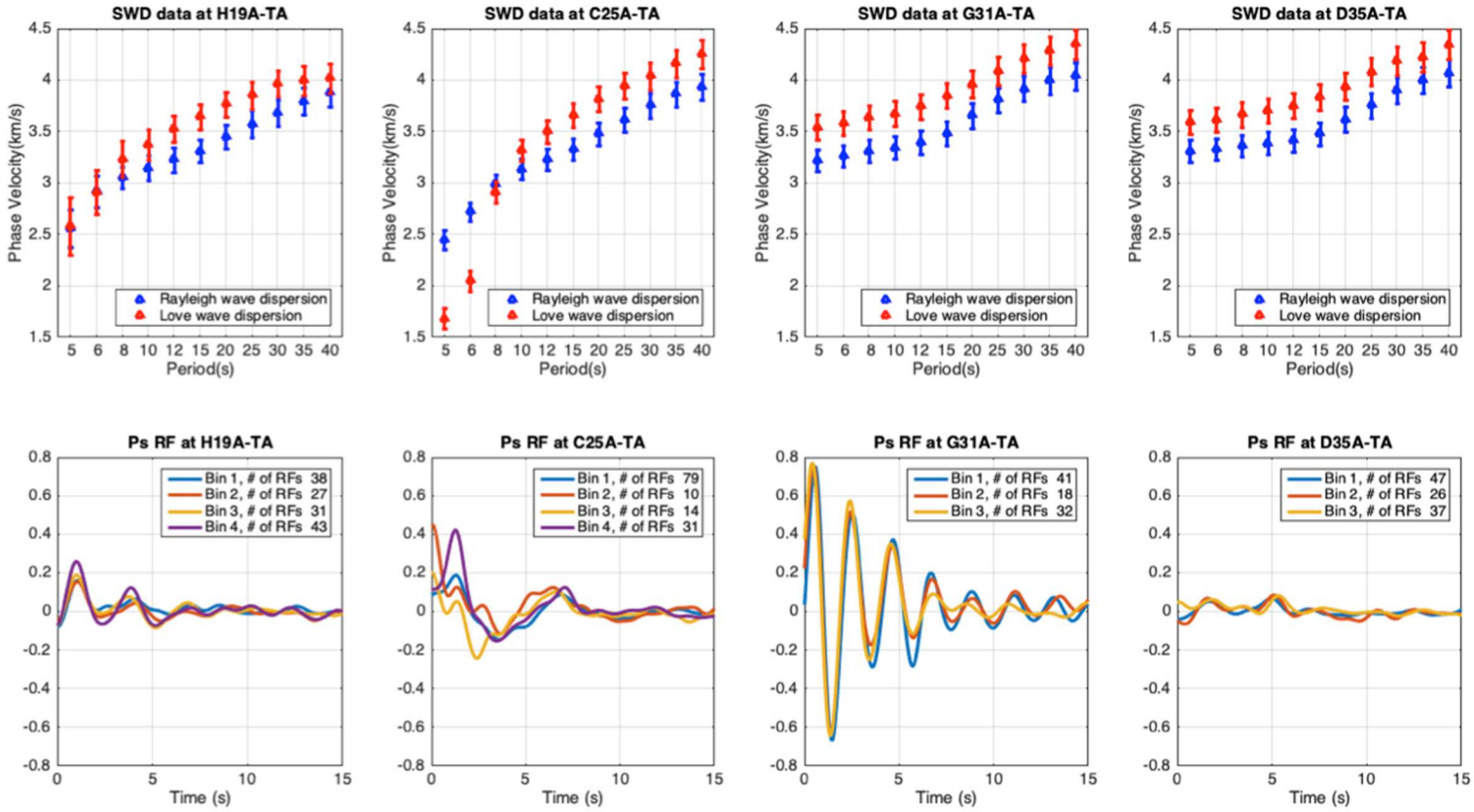


Figure 5.2 Rayleigh and Love dispersion curves (upper panels) from the model of Olugboji et al., 2017 based on ambient noise measurements of Ekstrom, 2016 at Transportable Array (TA) stations H19A, C25A, G31A and D35A. (lower panels) Ps receiver functions at the same four TA stations. Receiver functions are divided into two to six bins during the inversion based on their ray-parameter.

As can be seen in Figure 5.2, RF waveforms also vary drastically across stations. While the impedance-contrast variations from intra-crustal or Moho interfaces could lead to different amplitudes for the associated phases, a low-velocity sedimentary layer on top of the crystalline crust can introduce signals with much higher amplitudes and longer durations. An example can be found in Figure 5.2 by comparing the RFs from station G31A and D35A. Crust to sediment conversion phase and its reverberation dominate the RF waveform for station G31A, and obscured the Moho phase. For station D35A where sediment is absent, the Moho phase can be clearly identified at a lag time ~5 s. It has been shown that the sediment reverberations in the RFs could bias the estimates of crustal thickness (Yeck et al., 2013; Yu et al., 2015). Recently, Gao et al., (in review) demonstrated that reverberations from shallow layers in Ps RFs could produce spurious low-velocity zones when inverted for crustal structure together with surface wave dispersion data.

To correctly account for the effect of sediment reverberations on the RFs, we adopt the approach proposed by Gao et al., (in review), where a resonance-removal filter (Yu et al., 2015; Cunningham & Lekic, in review) is applied to effectively remove near-surface reverberations. For the purpose of comparison, we perform joint inversions of SWD and original RFs for all 291 stations, and then perform joint inversion of SWD and sediment corrected RFs for stations with thin sediment layers (< 2km). Figure 5.1 shows the seismic stations where crustal LVZ appeared when SWD and original Ps RFs are jointly inverted. Among the 291 stations, the posterior of 51 stations resolved LVZs which turned out to be artifacts, demonstrating the

importance of performing sediment corrections for the inversion. We further demonstrate this issue in the cross-sectional view in section 5.4.

5.3 Results

5.3.1 1D Vs Structure and Sediment Map

Figure 5.3 shows the retrieved Vs structures beneath the same four stations as Figure 5.2 using the Transdimensional Bayesian inversion (TBI). The Vs structures beneath the four spatially distant stations demonstrate diverse features at all depths shown in Figure 3. Shallow low-velocity layers are resolved on top of the crystalline crust at H19A, C25A, G31A, but not at D35A. The 10km-thick layer with Vs values between 2.5km/s and 3.2km/s is related to the Yellowstone Hot Spot (Schutt et al., 2008). We consider C25A and G31A as examples of thick and thin sediment structures inside the Williston Basin (See Figure 4, Marshak et al., 2017).

The crustal structures of the four selected stations also differ from each other, preliminary estimate of crustal thickness based on the Vs change along depth ranges from 30km to 50km. Inferring the Moho depth without additional constraint can be challenging due to the potential high-velocity lower crust layer (Schulte-Pelkum et al., 2017). Later in the discussion we further investigate the crustal properties based on velocity profiles grouped by tectonic setting.

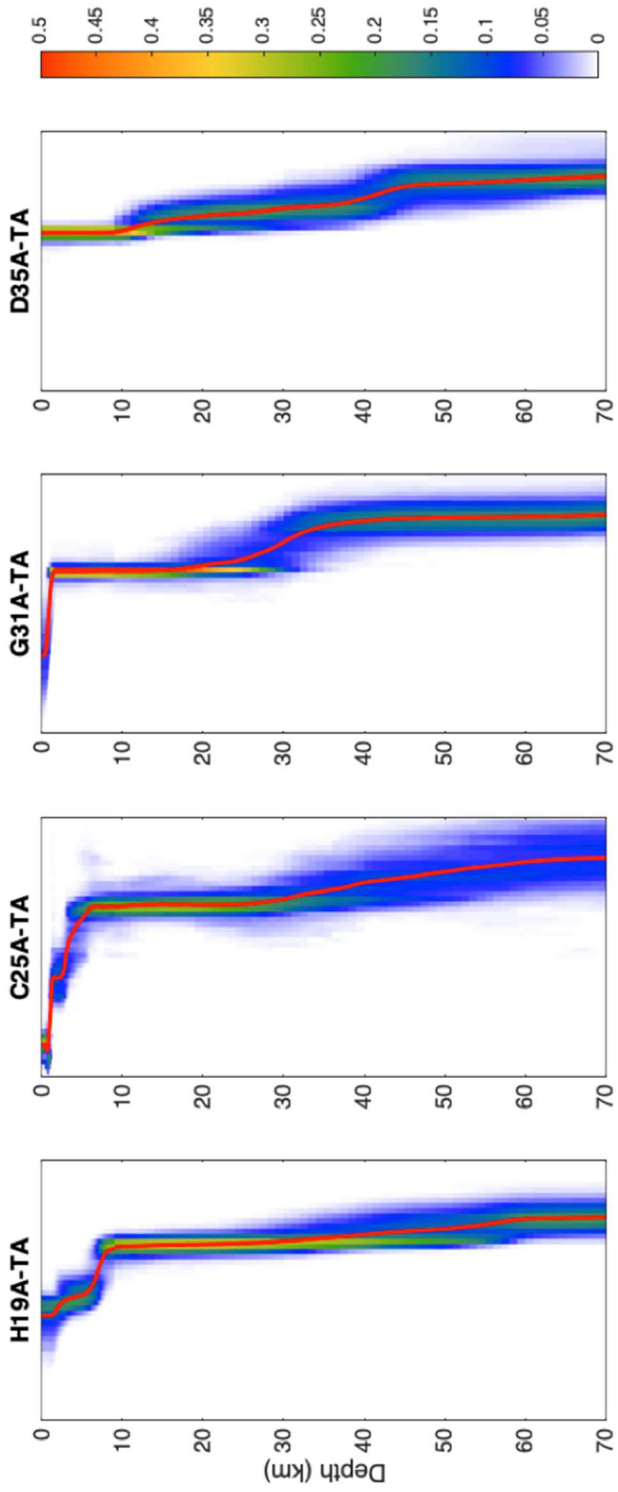


Figure 5.3 V_S depth distribution retrieved using TB joint inversion of SWD and Ps RFs at four TA stations. The ensemble solutions are displayed as probability density functions at each depth, with warmer colors corresponding to higher posterior probabilities, and the solid red line denoting the 5 percent trimmed mean of the posterior.

Benefiting from the adaptive parameterization of TB inversion, we are able to resolve a detailed seismic sediment model of the Northern Great Plains. Moreover, when the resonance-removal filter is constructed, the two-way travel time of S wave is measured and implemented into the inversion as additional prior information, leading to well-constrained sediment structures.

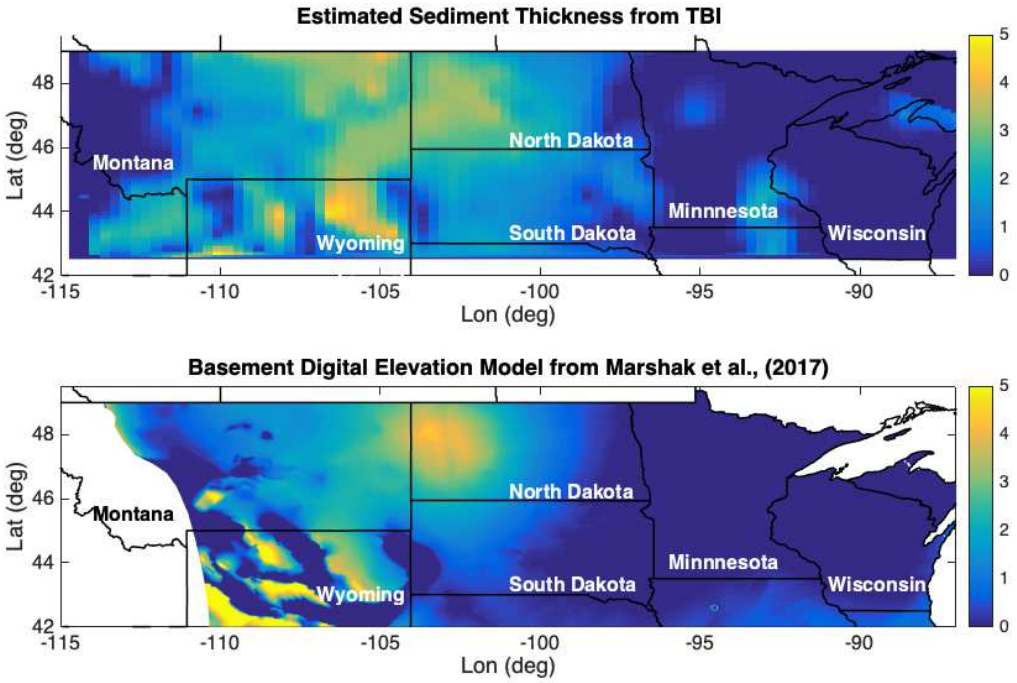


Figure 5.4 Sediment thickness of the Northern Great Plains estimated from TB join inversion using SWD and Ps RFs (upper) compared to basement digital elevation model from Marshak et al., 2017 (lower).

Figure 5.4 shows a comparison between the estimated sediment thickness from this study and the basement digital elevation model from Marshak et al., (2017). We estimate the sediment thickness beneath each seismic station by measuring the depth at which the mean Vs reaches 2.5 km/s (referred as Z2.5, see Campbell and

Bozorgnia, 2008) for the first time. We then interpolate the estimate thickness to the whole area covered by seismic stations using a Natural-neighborhood method. Our sediment-thickness map shows a consistent contour and depth distribution of the Williston Basin compared to Marshak et al., (2017). Our model is also able to capture smaller scale features such as the Big Horn Basin and the Powder River Basin in the northern Wyoming. This shows that passive-source seismology can constrain the large-scale sedimentary basin structure that is comparable to sediment maps that are constructed based on well log data or active-source seismological survey.

5.3.2 Shear Velocity Structure of the Northern Great Plains

Shear velocity maps of the Northern Great Plains around 20, 30 and 60 km depths are shown in Figure 5.5 as representations of the mid-crust, lower-crust and uppermost-mantle, respectively. To demonstrate the bulk property of the crust and uppermost mantle, we calculate the average velocity across a 5 km range for each map.

The most notable feature in the uppermost mantle velocity map is the clear east-west dichotomy. The shear velocity of the western part is significantly slower than the shear velocity of the eastern part, representing the tectonically more active region. Within the western part, smaller scale features such as the low velocity anomaly beneath the Snake River Plain (Braile et al., 1982; Shen et al., 2016) is also resolved.

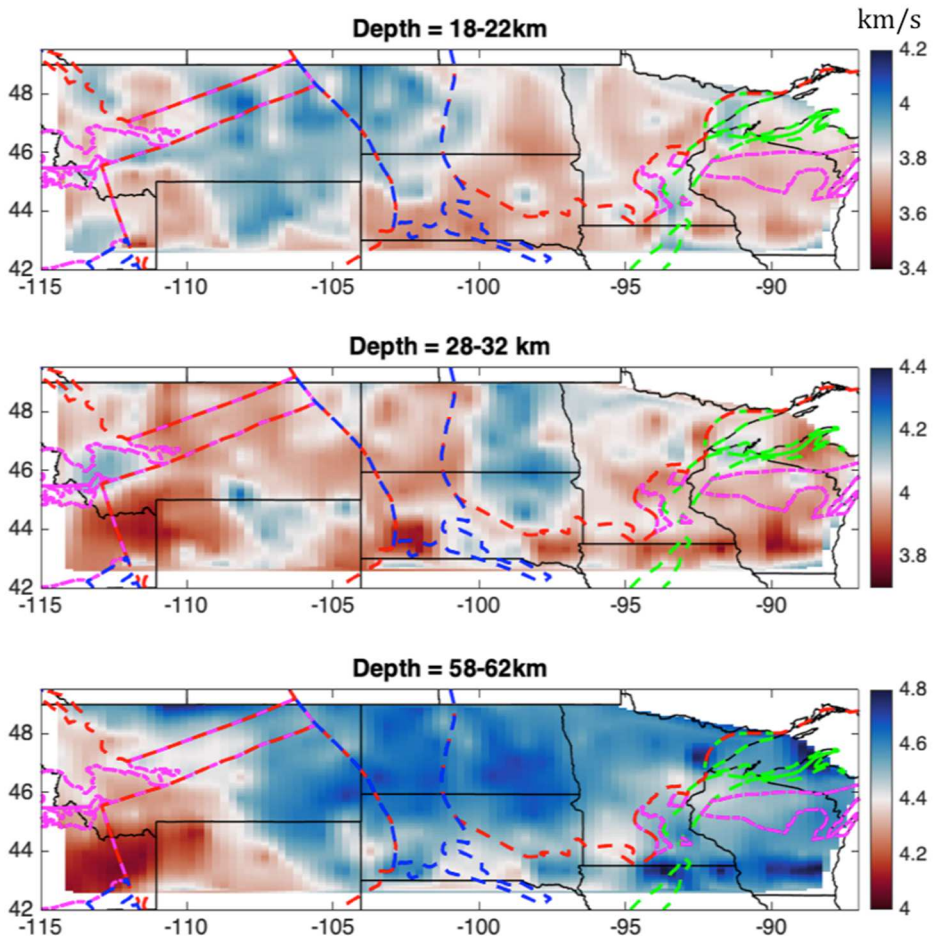


Figure 5.5 Shear velocity map averaged between 18-22 km, 28-32 km, 58-62 km. Archean Craton (>2.5 Ga), Trans-Hudson Orogen (1.9-1.8 Ga), Mid-Continental Rift (1.2-1.1 Ga) and juvenile volcanic arc (2.0-1.8 Ga) are shown in red, blue, green and magenta dashed lines respectively (Whitmeyer and Kalstrom, 2007).

Contrasting crustal structures reflecting different tectonic histories are expected in this region (Whitmeyer & Karlstrom, 2007). Archean terranes (> 2.5 Ga) including the Superior Province, Medicine Hat block and Wyoming Province are the oldest crust in this region. Due to the limitation of the Transportable Array deployment, only the westernmost Superior craton in the US is studied in this work. Superior Craton is bounded to the west by the ca. 1.8 Ga Trans-Hudson Orogen. The

Trans-Hudson Orogen represents the Paleoproterozoic collision between Superior, Hearne, and Wyoming cratons (Hoffman, 1988; Ross and Villeneuve, 2003). The Wyoming Craton and the Medicine Hat block are sutured by the Great Falls tectonic zone (1.86 -- 1.71 Ga, Gietti, 1966; O'Neill and Lopez, 1985; Mueller et al., 2002). In this paper, we refer to this region with two Archean blocks and the volcanic arc between them as Central Montana. Crustal structure of younger regions, such as Mid-continental Rift (1.2 – 1.1 Ga), is also resolved here.

Although the Superior Craton, Medicine Hat Block and Wyoming Craton are all considered to be Archean crust, our V_S map reveals contrasting properties between the westernmost Superior Craton and the other two cratons. At a depth of 20km, the average V_S within Central Montana is about 3.8 km/s (See Figure 5.8(a)). The average V_S within the Superior Craton is about 3.6 km/s (See Figure 5.8(c)), significantly slower than the Vs of Central Montana. At 30 km depth, the Central Montana region shows average V_S around 3.9 km/s (See Figure 5.8(a)). Compared to Central Montana, the eastern part of the Superior Craton in the US shows slightly faster shear velocities between 3.9 and 4 km/s, the western part shows shear velocities larger than 4 km/s. It's also worth noting that at both 20 km and 30 km, the bulk of THO shares a similar velocity with Central Montana.

5.4 Importance of Sediment Correction

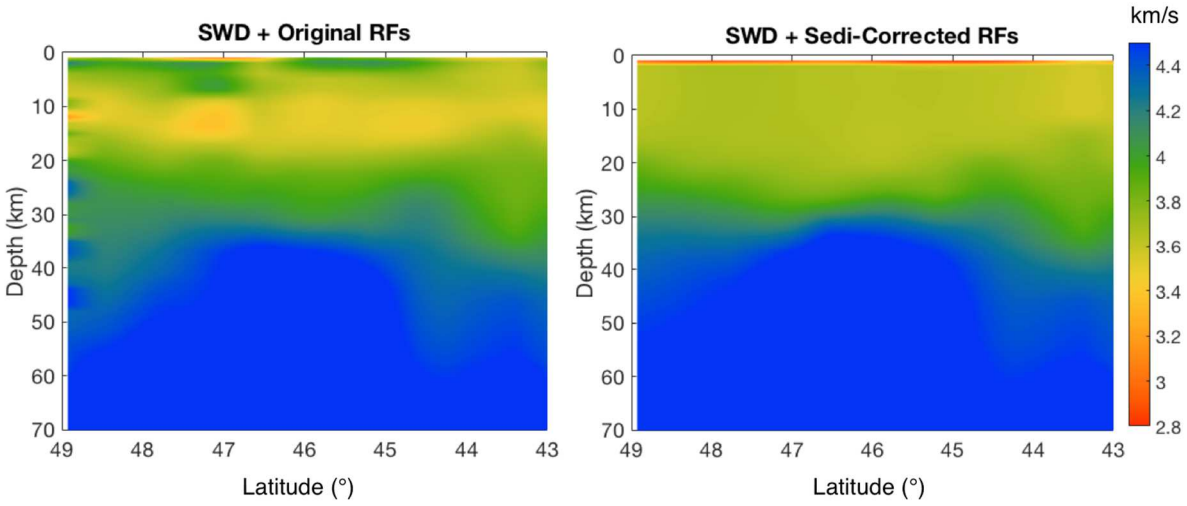
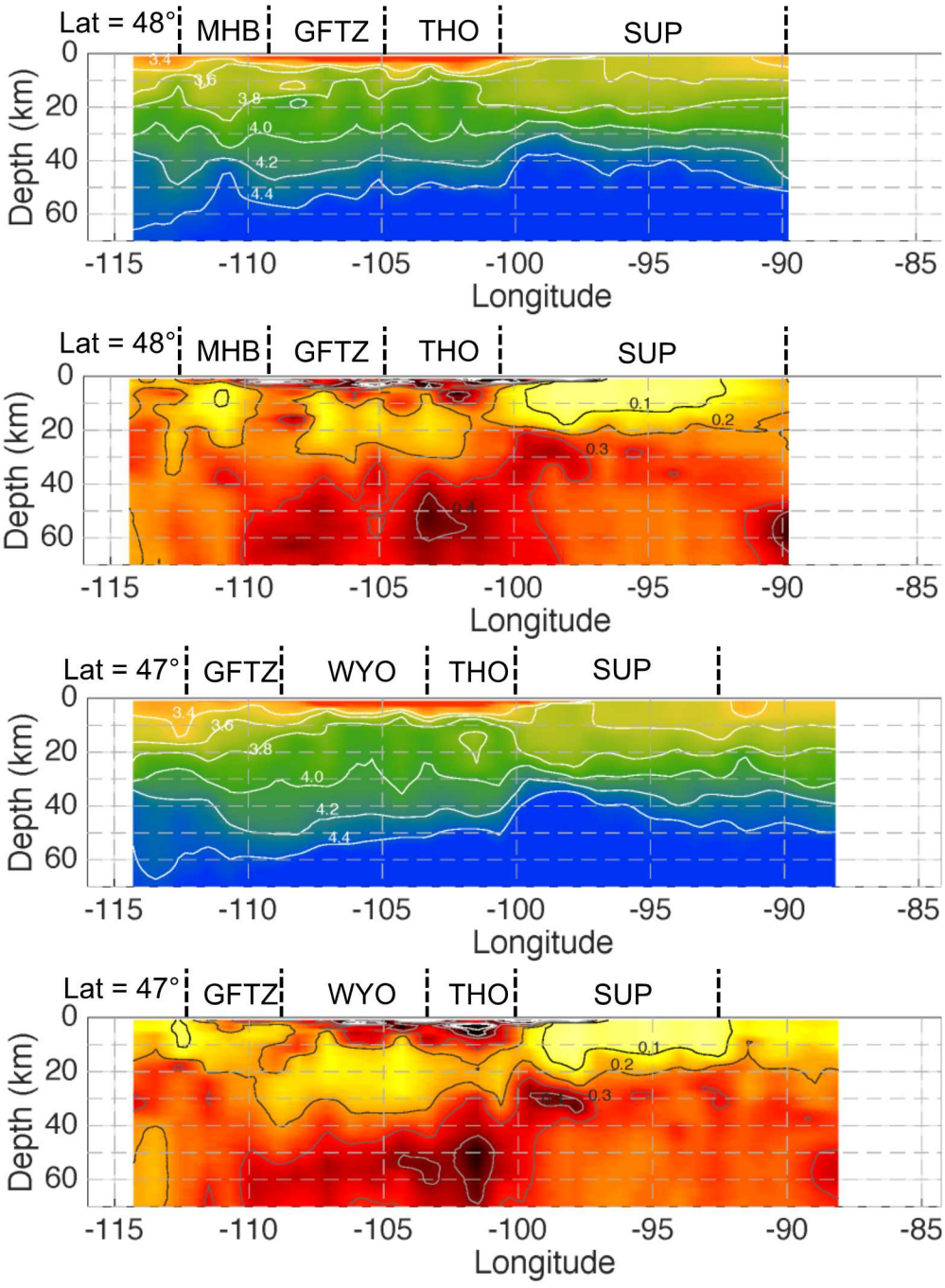


Figure 5.6 Cross-section view along latitude of the shear velocity structure at longitude 98°. TBI of SWD and original Ps receiver functions are shown in left panel; TBI of SWD and sediment corrected Ps receiver functions are shown in right panel.

As is shown in Figure 5.1, 51 out of 291 TA stations from this study initially resolved a crustal LVZ when the sediment reverberation is not explicitly accounted for. All of the LVZs disappear applying a sediment resonance removal filter is applied. To further emphasize the potential significance of this issue, we show two cross-sections of the Vs structures for comparison. In Figure 5.6, the right panel shows the cross-sectional view of Vs structure at longitude 98° from this study, where TBI of SWD and sediment corrected Ps receiver functions are used when thin sediment is present. On the left, we show the same cross-section where TBI of SWD and original Ps RFs are used. The cross section without sediment correction shows a clear LVZ at the depth range between 5 and 20 km. Contrarily, the cross section with sediment correction shows no sign of the LVZ. The sediment corrected cross section also shows different geometry of the Vs transition from 4 km/s to 4.4 km/s. As such transition is often used to infer the Moho, the crustal thickness estimated based on the left panel in

Figure 5.6 could be misleading. The comparison plot in Figure 5.6 shows that when the sediment reverberation in the RFs is not correctly accounted for, seismic inversion could result in artificial features that bias our estimate on crustal structures.

5.5 Discussion



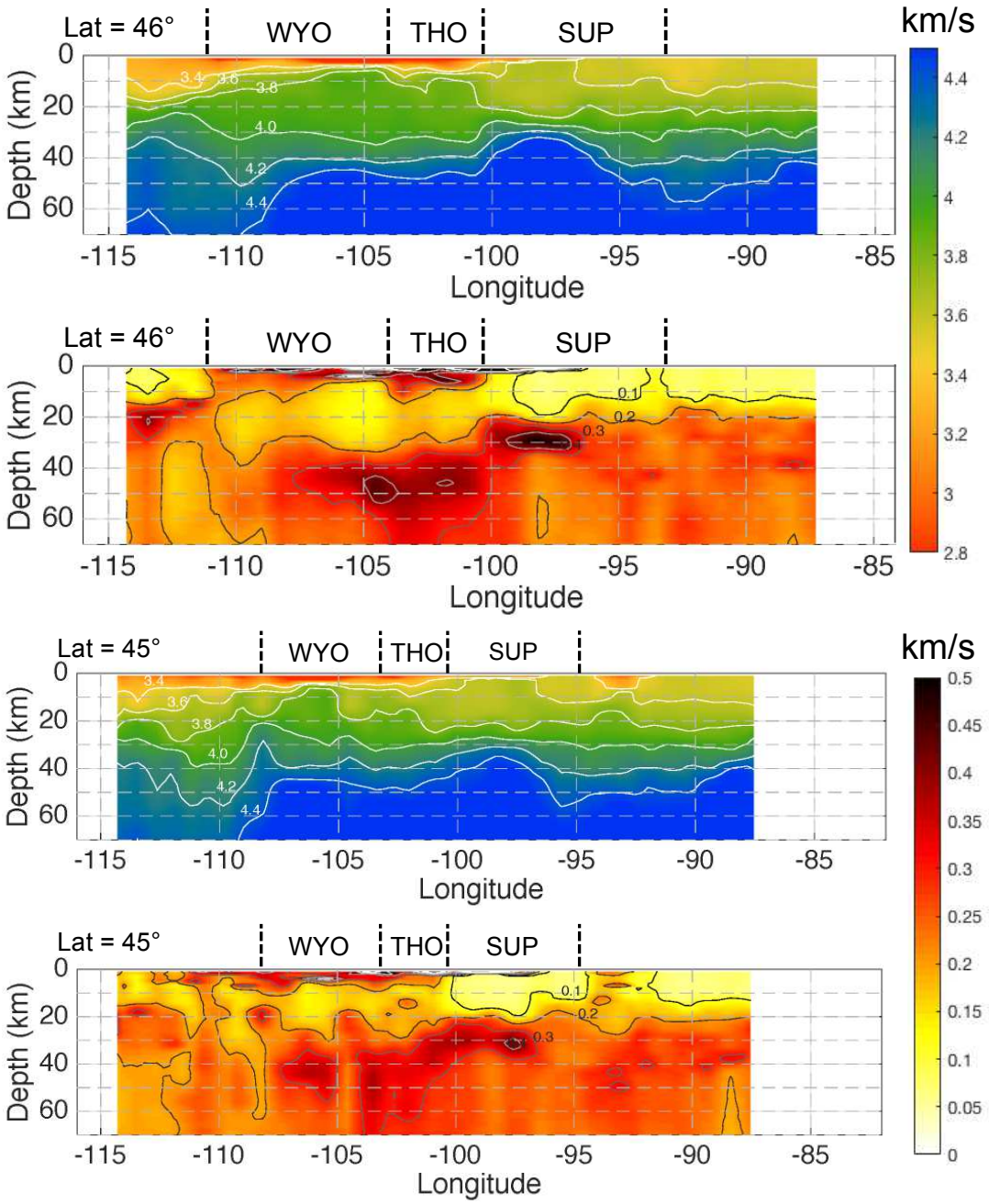


Figure 5.7 Cross-section view of the mean shear velocity structure and the 68% credible interval along latitude of 48°, 47°, 46°, 45°. Tectonic boundaries are indicated on the top of each panel (MHB – Medicine Hat Block; GFTZ – Great Falls Tectonic Zone; WYO – North Wyoming Craton; THO – Trans-Hudson Orogen; SUP – Superior Craton).

Figure 5.7 shows the latitudinal cross-section view across the North Great Plains. We label tectonic boundaries on top of the cross section to help associate velocity structure to different tectonic settings. In addition, we calculate mega-ensembles of Vs for different settings by combining ensemble solutions from three groups of seismic stations (Figure 5.8). The groups of stations, reflecting Central Montana, Trans-Hudson Orogen, and Superior Craton (Figure 5.1), are chosen based on the tectonic provinces from Whitmeyer and Kalstrom (2007). This mega-ensemble allows statistical inferences of Vs structure within different tectonic settings without losing the information of velocity variation from individual station.

5.5.1 Superior Craton

The top 20 km of the Superior Craton shows a relatively gradual increase of Vs from 3.5 km/s to 3.6 km/s. Starting around 20 km, the shear velocities increase to a value of 3.8 km/s. We interpret this change as a transition from upper-middle crust to lower crust due to the relatively abrupt velocity increase (See credible interval plots in Figure 5.7). The Moho depths within the Superior Craton vary from 30 km to 42 km based on TBI-NGP. Huang et al., (2013) estimated the crustal thickness of 37.5 ± 3.1 km for Archean shield based on global crustal models. Generally, Archean crusts are considered to be structurally simple with a clear crust-mantle interface due the lack of tectonic activities. Seismological investigations of the Superior Craton within Canada using receiver functions and surface wave dispersion (Gilligan et al., 2016) constrained the Archean crust to be 38 -- 40 km thick with a sharp Moho. The thin crust with relatively uniform thickness of the Superior Craton has also been attributed to the elevated lower crustal temperature (Flament et al., 2011). As is also observed in

map view (Figure 5.5), the eastern part of the westernmost Superior Craton appears to have a crust thickness of 40 km, while the western part comprises a thinner crust of about 30 km. Combining the observations from Figure 5.7 and Figure 5.8, it is clear that the eastern part of the Superior Craton can be attributed to a typical Archean crust. The small Vs variation above 20 km in Figure 5.7 and Figure 5.8 suggests a uniform upper-mid crust within the Superior Craton. The western part only differ from a typical Archean crust in the lower crust. The thin Superior crust is bounded by the Paleoproterozoic THO (Hoffman et al., 1988) to the south. During the collision between THO and the Superior Craton, the lower crust of Superior Craton is believed to be retained along most of their edges (Hoffman et al., 1989). However, our seismic velocity structure shows an exception where the lower crust of Superior appears to be missing likely through a post-formation delamination process.

5.5.2 Wyoming Craton, Medicine Hat Block, and Great Falls Tectonic Zone

Montana and Wyoming consists of the other two Archean cratons in our study region. Wyoming Craton and Medicine Hat Block are Archean cratons sutured together by the Great Falls tectonic zone. In this study, we analyze these Archean blocks together with the Great Falls tectonic zone as one group (See Figure 1, Figure 8) and compare its Vs structure with the Superior Craton.

The cross-sectional view crustal structure shows that the Archean Crust in Central Montana differs from the southwestern Superior Craton. The Moho depth appears to vary from 40 km to 50 km, significantly deeper than the Moho depth of the southwestern Superior Craton. The active source seismological experiment DeepProbe

(Gorman et al., 2002) estimated the Moho depth between 49 km and 60 km based on compressional wave velocity model, which is deeper than our estimate.

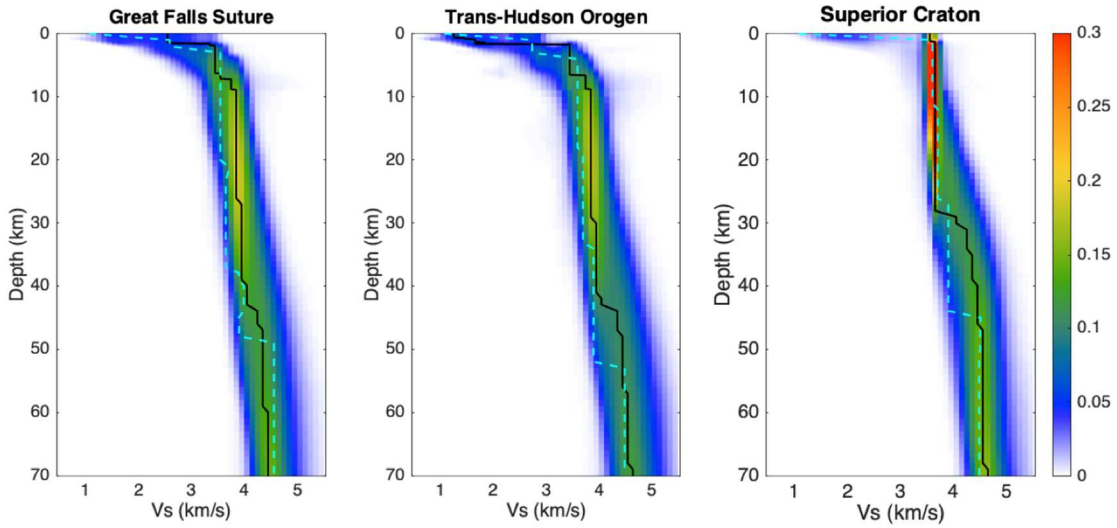


Figure 5.8 Ensemble of V_S posteriors from Great Falls Tectonic Zone (left), Trans-Hudson Orogen (middle) and Superior Craton (right). The ensemble solutions are displayed as probability density functions at each depth, with warmer colors indicating higher probability. The modes of the ensemble are plotted in black lines. Crust 1.0 V_S models of the related regions are plotted in cyan. Note that Crust 1.0 is systematically slower than our model in the lower crust.

Figure 5.8 shows the Vs variation from the mega-ensemble of Central Montana. The mode of Vs increases from 1.0 km/s to 3.5 km/s at the top 10 km, corresponding to the transition from the low-velocity sediment to crystalline crust. Unlike the relatively sharp Vs change around 20 km within the Superior Craton, no sharp Vs change is observed in this region. Vs gradually changes from 3.6 km/s to 3.9 km/s between 10 km and 40 km. Larger Vs variation are also observed at this depth range compared to Figure 5.8(c), suggesting greater lateral heterogeneities at the mid-to-lower crust range for Central Montana compared to the westernmost Superior

Craton. A clear Vs jump occurs around the depth of 45 km, while the Vs change at 60 km does not appear to be abrupt (Figure 5.8).

Previous passive source seismological studies using surface wave dispersion (Bensen et al., 2009), receiver functions (Gilbert, 2012), or joint inversion of the two (Shen et al., 2013; Shen et al., 2016) estimate a Moho depth between 40 km and 50 km for Central Montana, shallower the active source study results. Mahan et al., (2012) discussed the potential reasons of such discrepancy, including limited data constraints from individual inversion and model parameter trade-offs. Schulte-Pelkum et al., (2017) suggested that strong model parameter assumptions made for joint inversion could also lead to misinterpretation of the Moho depth from passive source seismology studies. While joint inversion with refined model parameterization in Schulte-Pelkum et al., (2017) allow for the retrieval of the Moho at deeper depths consistent with active source studies, the uncertainties of the Moho depths are large. This leads Schulte-Pelkum et al., (2017) to interpret the Moho in Central Montana as a transitional interface partially due to metasomatic alteration (Mahan et al., 2012; Downes et al., 2004; Facer et al., 2009). The adaptive parameterization of TB approach allows the inversion to add/remove model parameters primarily based on the data constraints; therefore, no additional prior information is needed to accommodate the complex crustal structure in this region. However, such discrepancies between active source and passive source seismic imaging still exist with our TBI-NGP model.

Crustal xenolith data in Central Montana (Mahan et al., 2012) are more consistent with the Moho depth based on active source studies. Gorman et al., (2002)

estimated a Vp exceeding 7 km/s between 40 and 60 km from active source seismic survey, significantly higher than the average continental crust Vp. For passive source studies using SWD and RFs, the primarily data sensitivity lies in Vs. Figure 5.8(a) shows a Vs between 4.2 and 4.4 km/s at the depth range of 45 km 60 km. The origin and the composition of this anonymously high-velocity lower crust layer have been well studied and discussed by combining the seismic velocity structures and the crustal xenoliths (Mahan et al., 2012; Barnhart et al., 2012; Schulte-Pelkum et al., 2017). The heterogeneous characteristics of the xenoliths as well as several recorded magmatic, rifting, and collisional events suggest the formation of the high-velocity lower crust layer as a result of periodic lower crust addition through magmatic or mechanical processes.

5.5.3 Trans-Hudson Orogen

Previously, active-source seismic surveys (McCamy & Meyer, 1964; Hajnal et al., 1984; Braile, 1989) have been conducted to study the crustal structure of the THO south, including Montana, North Dakota in the US as well as Saskatchewan in Canada. Crustal structures in this region are characterized as thick (45 -- 50 km) and heterogeneous. High-velocity lower crusts with Vp exceeding 7 km/s (also called 7.x layer) are widely seen in this region. The thickness of the 7.x layer ranges from 20 to 25 km (Schulte-Pelkum et al., 2017), significantly thicker than the average of 7.x layer detected from refraction results in the North America. Passive-source seismological studies of the THO structure have recently proliferated due to the deployment of seismic arrays (e.g. HuBLE array, see Bastow et al., 2015). Several receiver function

studies constrained the crustal structure in the northern part of the THO (Thompson et al., 2010; Darbyshire et al., 2013; Gilligan et al., 2016). Crust in north THO has a thickness of 40-45 km; with an elevated Vp/Vs ratio (>1.75, Thompson et al., 2010), suggesting a more mafic crust.

Seismic imaging that includes RFs in the THO can be challenging due to the overlying Williston Basin. We address this issue by applying a resonance-removal filter to the RFs, and then perform the TBI. The Vs mega-ensemble for THO in Figure 5.8 shows a similar structure to the one in Central Montana. In addition, both the map view and the cross-sectional view of these two regions support this observation. The top 10 km of THO Vs structure shows the transition from the low-velocity sediment to the crystalline crust. Using the mode of the ensemble as an estimate for the bulk property, Vs ranges from 3.4 km/s to 3.7 km/s in the topmost crust. No abrupt Vs increase is observed between the depth of 10 km and 30 km, the mode of Vs is 3.8 km/s. Starting from 30 km, Vs increases to reach 4 km/s. Based on the mode estimate, the most abrupt Vs increase below 10 km is observed between 41 km and 44 km, where the Vs increases from 3.9km/s to 4.3 km/s. If we interpret this transition as the Moho, it would lead to a shallower estimate of crustal thickness compared to the active source studies. Consider the widely seen 7.x layer in this region, we interpret such transition around between 30 and 44 km as the change from the mid-crust to the high-velocity lower crust. The thickness of this high-velocity lower crust layer is not well constrained in THO, as the Moho discontinuity appear to be diffused in our Vs mega-ensemble. The crust thickness estimated by active source studies (McCamy and

Meyer, 1964; Braile, 1989) is between 45 -- 50 km, leading to a thickness of 10 – 20 km for the 7.x layer.

The origin of this high-velocity lower crust could have further implication of the formation and evolution of the THO and the Archean crust. Eclogitization has been proposed as a possible explanation in orogenic settings due to its high density and high seismic velocity (Worthington et al., 2013). However, the seismic velocity of pure eclogite (4.5 -- 4.8 km/s, Worthington et al., 2013) is significantly faster than our observation of 4.3 -- 4.5 km/s. Nelson et al., (1992) suggested that eclogizited lower crust could have been delaminated during the process of orogenic collapse. Partially eclogitization might be a more likely scenario for our observation in THO. The presence of partially eclogized layer has been proposed in West Tibet (Zhang et al., 2014) to account for the high seismic velocities and diffused Moho. Gilligan et al., (2016) proposed a partially eclogized root in the lower crust beneath the southern Baffin Island, a region affected by THO, due to the lack of structural characteristics associated with orogenic collapse of THO in the northern Hudson Bay. Baird et al., (1996) suggested that the THO within the US did not experience orogenic collapse. Our seismic velocity structure in THO is compatible with the partially eclogized root hypothesis. Additionally, this partially eclogized root could also help explain the formation and the subsidence history of the Williston Basin. The subsidence of Williston Basin started at 525 MA and lasted for 520 Myr. The subsidence record of Williston Basin shows that the subsidence accelerated sometime after the initiation (Haid, 1991). Hamdani et al., (1993) proposed a model that involves both thermal contraction and lower crust phase change. In this model, the dense eclogite root

resulting from the phase change in the lower crust would allow the acceleration of the subsidence process of the Williston Basin. Given the seismic constraints as well as the subsidence record, it is reasonable to expect a partially eclogized lower crust in the THO.

5.6 Composition of Middle and Lower Crust

The composition of the lower continental crust, especially in the old, stable part of the Earth, could have implication on the formation and evolution of Earth's crust. Unlike the relatively well-studied upper continental crust, the composition and properties of the middle and lower continental crust are more difficult to determine. Inferences of middle and lower continental crust often rely on the combination of geophysical data, xenolith, and exposed terrains recording lower crustal pressure (e.g. Rudnick & Gao, 2003).

There have been substantial efforts aimed at using seismic wavespeed to infer the composition of lower crust (Christensen & Mooney, 1995; Holbrook et al., 1992; Rudnick & Fountain 1995). A typical approach is to compare the seismic wavespeeds inferred by inversion of seismic data to the wavespeeds of crustal rocks measured in the laboratory; the measured crustal sections are then assigned to different tectonic settings and modeled as layered crustal structure. It has been shown in these studies that Vp structure shows clear variations among different tectonic settings. Furthermore, many studies have investigated the potential correlation between rock composition and Vp, Vs, or Vp/Vs. Christensen (1996) reported a correlation between Vp/Vs and SiO_2 content for rocks with SiO_2 content between 55 and 75 wt%.

Recently, Huang et al., (2013) reported the correlation between SiO_2 content and laboratory measured Vp and Vs. While it has been argued that SiO_2 estimated using this approach could bear large uncertainties due to several limitations, it is still a reasonable indicator for mafic lower crust composition with elevated seismic wavespeed (Hacker et al., 2015).

In this paper, we adopt the reported Vs distributions for mafic and felsic endmembers of granulite and amphibolite from Huang et al., (2013), and combine our seismic structure to infer the SiO_2 content of the NGP middle-lower crust. We implement a Bayesian analysis approach to achieve this goal.

To incorporate the variation of the measured wavespeeds, we work directly with the normal distributions for the felsic and mafic endmembers instead of the mean. In Huang et al., (2013), the measured Vs are corrected for temperature and pressure using pressure and temperature derivatives of $2\times 10^{-4}\ \mathrm{km\cdot s^{-1}\cdot MPa^{-1}}$ and $-4\times 10^{-4}\mathrm{km\cdot s^{-1}\cdot {}^{\circ}C^{-1}}$ (Christensen and Mooney, 1995; Rudnick and Fountain, 1995). A typical conductive geotherm equivalent to a surface heat flow of $60\ \mathrm{mW\cdot m^{-2}}$ is assumed for the global average estimate. We follow the same approach, except that we correct the temperature effect by using regional surface heat flow from International Heat Flow Commission Database (IHFC, 2011; see Pollack et al., 1993) instead. We use surface heat flow of $60\ \mathrm{mW\cdot m^{-2}}$, $60\ \mathrm{mW\cdot m^{-2}}$, and $40\ \mathrm{mW\cdot m^{-2}}$ for Central Montana, THO and Superior Craton respectively. We then calculate the associated geotherm using the following equation (Turcotte & Schubert, 2014):

$$T = T_0 + \frac{q_m y}{k} + \frac{(q_0 - q_m)h_r}{k}\left(1 - e^{-y/h_r}\right) \quad (5.5)$$

Here T_0 is the room temperature; $q_m$is the mantle heat flow, we assume that mantle heat flow constribute to 60% of the surface heat flow q_0 (Pollack & Chapman, 1977); *k* is the thermal conductivity of the continental crust, we choose a value of $3.36\ W \cdot m^{-1} \cdot k^{-1}$; *y* is the depth; $h_r$is the reference length scale that equals to 10 km. The calculated regional geotherm are shown in the left panel in Figure 5.9 – 5.11. Given all the information above, we can write the temperature and pressure corrected velocity V_{TP} as:

$$V_{TP} = V_0 + \frac{dV}{dT} \times (T - T_0) + \frac{dV}{dP} \times \Delta P \quad (5.6)$$

We choose to use felsic fraction *f*, to model the SiO_2 content so that when *f* varied from 0 to 1, it represent the SiO_2 endmember from mafic to felsic. According to Bayes' Theorem, if we are interested in the distribution of SiO2 content at certain depth given the seismic wavespeed, we have:

$$P(f|V_s) = \frac{P(V_s|f) \times P(f)}{P(V_s)} \quad (5.7)$$

Here, P(f) is uniform prior between 0 and 1. We sample the whole prior from 0 to 1 with an increment of 0.05. For each value of *f*, we represent the $P(V_s|f)$ as normal distribution with the weighted mean $(1\text{-}f)\times V_{mean\_mafic} + f\times V_{mean\_felsic}$ and variance $(1\text{-}f)^2\times std_{mafic} + f^2\times std_{felsic}$. $P(V_s)$ is a constant as it does not vary with f, therefore, it can be determined by computing the factor needed to make integral of $P(f|V_s)$ unity. We then marginalize along Vs to estimate the PDF of composition at certain depth h:

$$P(f|h) = \int P(f|V_s) \cdot P(V_s|h) dV_s \quad (5.8)$$

Figure 5.9 – 5.11 shows the estimated the felsic fraction for the three regions discussed above using the mode of TBI-NGP assuming amphibolite and granulite

facies. As a comparison, we also compute the felsic fraction constrained by the CRUST1.0 model in the associated location. The Vs model for CRUST1.0 is shown in Figure 8 as well. The crustal thickness from CRUST1.0 is consistent with our estimate except in Central Montana, where the CRUST1.0 places the Moho at 48 km, compared to the TBI-NGP estimate of 60 km.

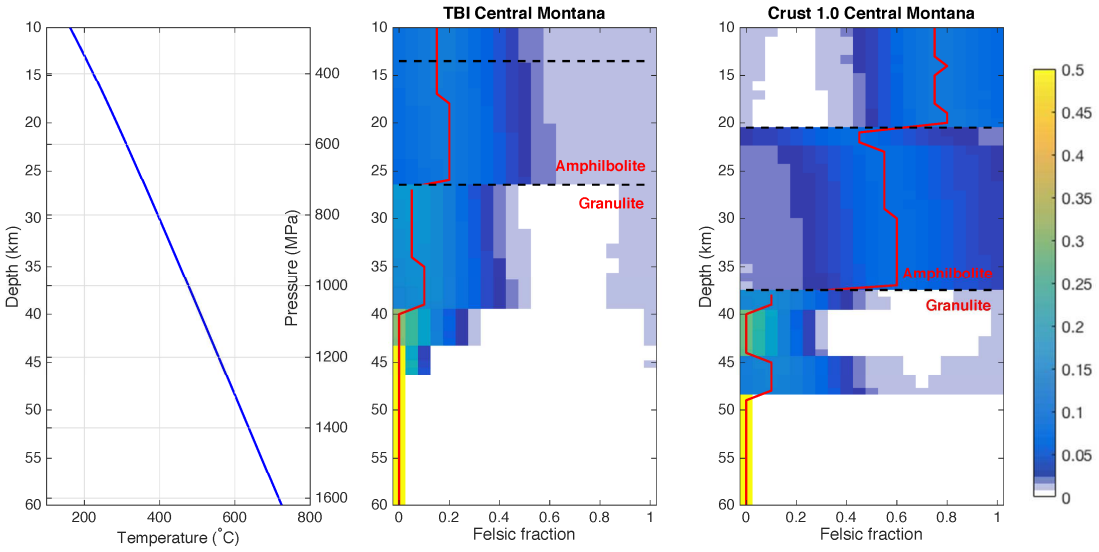


Figure 5.9 Geotherm (left) and felsic fraction estimated based on Huang et al., (2013) for Central Montana using Vs of TBI-NGP (middle) and Crust 1.0 (right). The ensemble solutions are displayed as probability density functions at each depth, with warmer colors indicating higher probability. The modes of the ensemble are plotted in red lines. Black dashed lines indicate the estimated boundaries between upper and middle crust, as well as between middle and lower crust, the latter of which we associate with transition from amphibolite to granulate facies.

CRUST1.0 parameterized the crust as upper, middle and lower crust, allowing us to directly estimate the composition of middle crust at explicit depth range. The depth ranges of middle crust for Central Montana, THO, and Superior Craton are 21 – 37 km, 20 – 33 km, 12 – 25 km, respectively. Since we do not use a fixed parameterization within the crust, no boundary between upper crust and middle crust is imposed as priori information; instead, it can arise in the ensemble solutions as

required by data. Here we used the global average upper crust thickness of 13.5 km based CRUST 1.0 as a reference boundary for the discussion. Combining the observation shown in section 5.3.2, we refer the depth range 13 – 27 km, 13 – 29 km, and 13 – 28 km as middle crust for Central Montana, THO, and Superior Craton (See black dashed lines in Figure 5.9 – 5.11). CRUST 1.0 assigned lower crustal depths between 38 – 49 km, 34 – 52 km, and 26 –44 km for Central Montana, THO, and Superior, respectively. The Moho variation of the TBI-NGP model is discussed in Section 5.5. At mid-crustal depths, temperature-pressure conditions associated with amphibolite metamorphic facies are present; whereas at lower-crustal depth, temperature-pressure conditions associated with granulite metamorphic facies are present (Huang et al., 2013). For all three regions, we assume amphibolite facies above the middle crust to infer the crustal composition given seismic wavespeeds. For the depth range below middle crust to 60 km, we granulite metamorphic facies to infer the crustal composition. We choose a wide depth range to cover the possible variation of crustal depth/thickness.

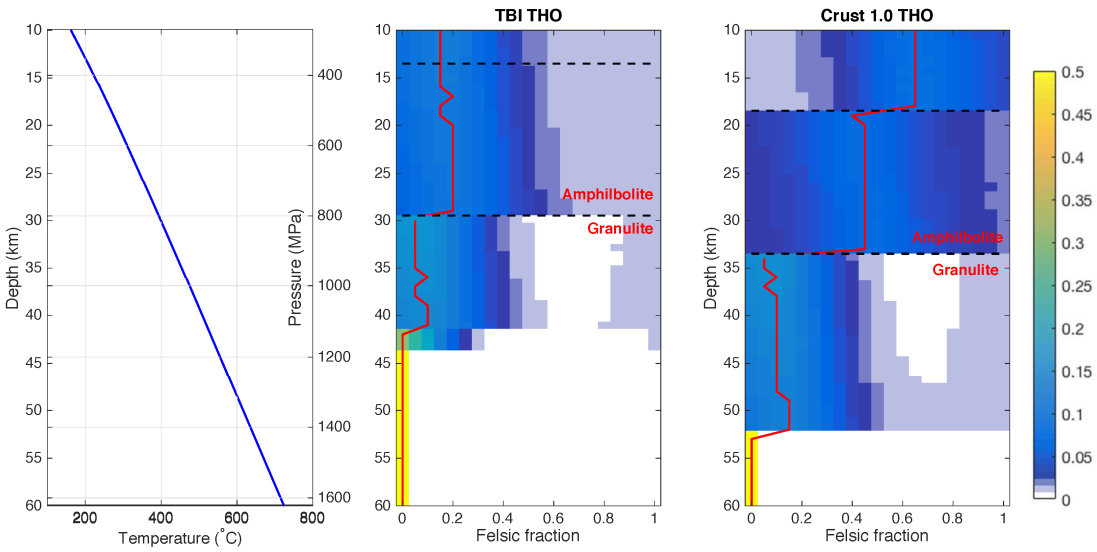


Figure 5.10 Geotherm (left) and felsic fraction estimated based on Huang et al., (2013) for Trans-Hudson Orogen using Vs of TBI-NGP (middle) and Crust 1.0 (right). The ensemble solutions are displayed as probability density functions at each depth, with warmer colors indicating higher probability. The modes of the ensemble are plotted in red lines. Black dash lines indicate the estimated boundaries between upper and middle crust, as well as between middle and lower crust, the latter of which we associate with transition from amphibolite to granulate facies.

For both Central Montana and THO, CRUST 1.0 implies an intermediate composition with large uncertainties, indicating poorly constrained middle crust composition given CRUST 1.0 wavespeed. In contrast, our TBI results prefer a relatively mafic composition for the middle crust of both Central Montana and THO, although large uncertainties in the felsic fraction estimates mean that an intermediate to felsic composition cannot be ruled out. For Superior Craton, CRUST 1.0 velocities imply an intermediate composition with P(f|h) centered on a felsic fraction of 0.5. The TBI prefers a relatively more felsic middle crust, with P(f|h) centered on $f = 0.75$. Given the large uncertainties in both distributions, it is not conclusively determined whether the composition of Superior Craton middle crust is felsic or mafic. This difference in inferred crustal compositions results from the fact that the Vs

reported by CRUST 1.0 in middle crust is significantly slower than the TBI posterior. Given the extrapolations and parameterization assumptions used in constructing a global model like CRUST 1.0, it is reasonable to conclude that compositions implied by the higher resolution TBI results obtained with few prior assumptions should be favored.

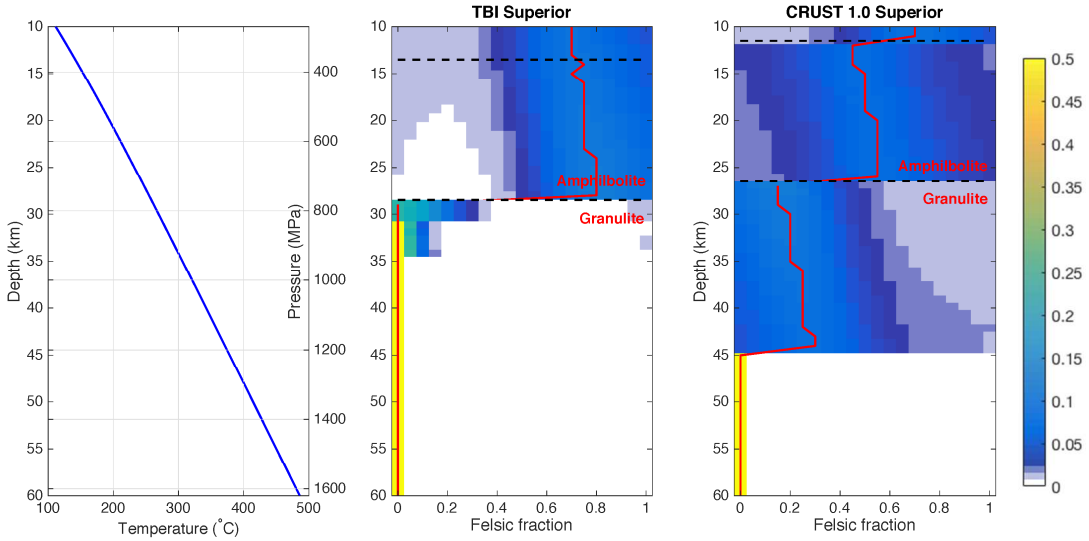


Figure 5.11 Geotherm (left) and felsic fraction estimated based on Huang et al., (2013) for Superior Craton using Vs of TBI-NGP (middle) and Crust 1.0 (right). The ensemble solutions are displayed as probability density functions at each depth, with warmer colors indicating higher probability. The modes of the ensemble are plotted in red lines. Black dash lines indicate the estimated boundaries between upper and middle crust, as well as between middle and lower crust, the latter of which we associate with transition from amphibolite to granulate facies.

In the Superior Craton, both CRUST 1.0 and TBI-NGP results (29 – 42 km) favor a mafic lower crust, with CRUST 1.0 bearing more uncertainties than TBI-NGP due to its slower lower crust velocities. As is pointed out in Hacker et al., (2015), with Vp greater than 7.0 – 7.2 km/s, it is reasonable to infer a mafic composition in the lower crust. Similarly, in this study, when the observed Vs is significantly greater than

the measured mafic granulite Vs, we conclude that the lower crust is mafic (or ultramafic). In THO, both CRUST 1.0 and TBI-NGP prefer a mafic lower crust, with TBI-NGP requiring a mafic composition below 42 km. As is shown in Section 5.5.4, the elevated Vs observed in the lower crust and the subsidence history of Williston Basin favor a partially eclogitized lower crust in THO. A highly mafic composition is consistent with this hypothesis. In Central Montana, the CRUST 1.0 likely underestimated the crustal thickness; therefore, we discuss the lower crust composition at different depth range for different velocity models. Lower crust for CRUST 1.0 (38 – 49 km) prefers a mafic composition with mode felsic fraction between 0 and 0.15. The lower crust retrieved from TBI shows more complex layering within the lower crust. Between 28 and 44 km, the TBI results prefer a mafic composition. Starting from 45 km a mafic composition is required by the observed high Vs. As is discussed in section 5.5.3, a heterogeneous lower crust due to multiple region tectonic events from Archean to Mesoproterozoic time is suggested based on crustal xenoliths study. Barnhart et al., (2012) reported a diverse range modal mineralogy and bulk major element chemical compositions of the lower crust xenoliths, however, the deepest xenoliths have basaltic compositions in general. The pressure-temperature estimates of their xenolith samples also demonstrate a more mafic composition with higher pressure. Therefore, a lower crust that starts with intermediate to mafic composition then transitions into a predominantly mafic composition with increasing depth is a plausible scenario for Central Montana.

5.7 Conclusion

In this chapter, we present a crustal shear velocity model for the Northern Great Plains of the US using surface wave dispersion and receiver function data measured from 291 Transportable Array stations. The probabilistic inversion scheme presents results not with a single best-fit model, but by full probability density functions, allowing statistical inferences on the seismic velocity structures.

A large portion of the North Great Plains is covered by low-velocity sediments that produce strong, reverberational signals in the RFs that can obscure the direct conversion from Moho. In chapter four, we showed that inversion of crustal structure with a thin sediment on top using RFs and insufficiently high frequency SWD could result in spurious crustal LVZ. Here we further emphasize, and illustrate with concrete examples, that sediment reverberations have to be appropriately accounted for in order to retrieve reliable crustal structure. We use parameters describing the sediment reverberation removal filters as additional prior constraints on the joint inversions, which yields more accurate sediment velocity and thickness estimates. Our sediment structure model is consistent and comparable to the industrial active source model as shown in Figure 5.4.

The crustal structures display large variations among different tectonic settings of our studied region. To further incorporate the seismic constraints and the laboratory measured crustal rock wavespeeds, we estimate the SiO_2 content in the middle-lower crust of the Northern Great Plains based on the compiled lab measurement from Huang et al., (2013) assuming amphibolite and granulite facies. We correct for the temperature and pressure effects of the laboratory-measured values for the 10-60 km

depth range, and then infer the felsic fraction given the seismic wavespeeds using a Bayesian inference process. We compare compositional inferences based on our TBI-NGP results with those implied by the CRUST 1.0 model.

In the Archean Superior Craton within US, we resolve a laterally uniform, and vertically highly stratified upper and middle crust. The crustal thickness of the Superior Craton ranges from 30 to 42 km. Our analysis of middle crust composition in the Superior Craton shows that, for both TBI-NGP and CRUST 1.0, a felsic to intermediate composition is preferred. It is worth noting that such inferences are associated with relatively large uncertainties, as shown in Figure 5.9. Assuming granulite facies for the lower crust, seismic wavespeeds of TBI-NGP require a mafic composition for the US portion of the Superior Craton. We observe an anomalous thin lower crust in western part of the Superior Craton compared to the eastern part. This anomalous lower-crustal thinning could result from post-formation delamination processes at the boundary between the Trans-Hudson Orogen and Superior Craton.

The other two Archean blocks in this region, Wyoming Craton and Medicine Hat Block that are sutured by the Great Falls tectonic zone, show velocity structures distinct from those of the Superior Craton. The estimated crustal thickness between 49 and 60 km is significantly great than that in the Superior Craton. This crustal thickening is likely a result of the collision and subduction in the Great Falls tectonic zone. Large Vs variations are also observed in the middle-lower crust of GFS, suggesting greater lateral heterogeneities compared to the Superior Craton. We observe a thick (~30 km), fast (Vs greater than 4 km/s) lower crust in the GFS. The elevated Vs is consistent with the 7.x layer (Vp > 7 km/s) constrained by active-source

studies in this region. Local crustal xenoliths studies (Barnhart et al., 2012; Mahan et al., 2012) propose a model of incrementally assembly of the 7.x layer due to periodic lower crust addition through magmatic and possibly mechanical processes. The SiO_2 content for the lower crust implied by TBI-NGP wavespeeds are consistent with this model, since ultra mafic composition is preferred in the GFS lower crust.

The Paleoproterozoic Trans-Hudson Orogen in the Northern Great Plains shows a thickened (45 – 50 km) crust, likely due to the collision between the THO and the Archean terranes. The 7.x layer at the base of the crust is also widely observed in this region, ranging in thickness between 10 and 20 km. Partially eclogitized root in the lower crust has been proposed beneath the northern part of THO in southern Baffin Island (Gilligan et al., 2016), due to the lack of structural characteristics associated with orogenic collapse of THO in the northern Hudson Bay. We consider this is to be likely for THO in the Northern Great Plains as well, considering the elevated seismic wavespeeds together with the subsidence history of the overlying Williston Basin.

Chapter 6: Summary and Outlook

Understanding the composition, formation evolution of the continental lithosphere remains an important and challenging topic of geoscience. This dissertation contributes to the study of continental lithosphere from a seismological perspective, specifically using a transdimensional Bayesian approach. As is demonstrated by this work and many others (e.g. Bodin and Sambridge, 2009; Dettmer et al., 2010; Ray et al., 2014; Rudolph et al., 2015; Burdick and Lekic, 2017; Olugboji et al., 2017), transdimensional Bayesian geophysical inversion is capable of presenting accurate uncertainty estimates of the inversion results while imposing less-restrictive assumptions. The work carried out in this dissertation focuses on the seismic imaging of continental lithosphere, and consists of systematic synthetic investigation, method development, application, and model interpretation.

The work shown in chapter two and chapter three are a combination of systematic synthetic tests and method development. It investigates the application transdimensional Bayesian inversion when the multiple types of elastic parameters (V_P – V_S – density – ξ) are inverted instead of one type. This is a relatively poorly discussed topic, yet there has been an increasing effort in interpreting the multi-parameter seismic structures in the published literature. In these chapters, I showed that due to the data sensitivity of seismic surface wave observables, the constraints on V_P, density, and radial anisotropy are limited compares to the V_S. The inversion results on these weakly constrained parameters can be biased by certain assumptions of the model parameterization. The synthetic tests and the discussion in chapter one could

help inform parameterization choices for further inversion of surface wave measurements alone and together with complementary observables (e.g. receiver functions).

Figure 6.1 (Panning et al., 2017) shows an example of using TB inversion to study the synthetic test for crustal and lithospheric layering of Mars. The inversion is intended to demonstrate potential product from the Insight lander mission to Mars (Banerdt et al., 2013). With little prior knowledge on the structure of Mars, it is challenging to perform a classical inversion with a restrictive parameterization. TB inversion offers the adaptive parameterization that determines the model complexity mainly from the data themselves. Our joint inversion using synthetic receiver function and Rayleigh wave ZH ratio in Figure 6.1 is able to recover of both Vp and Vs in the crust and mantle lithosphere. This example shows the potential of using TB inversion to constrain planetary interior with limited observation.

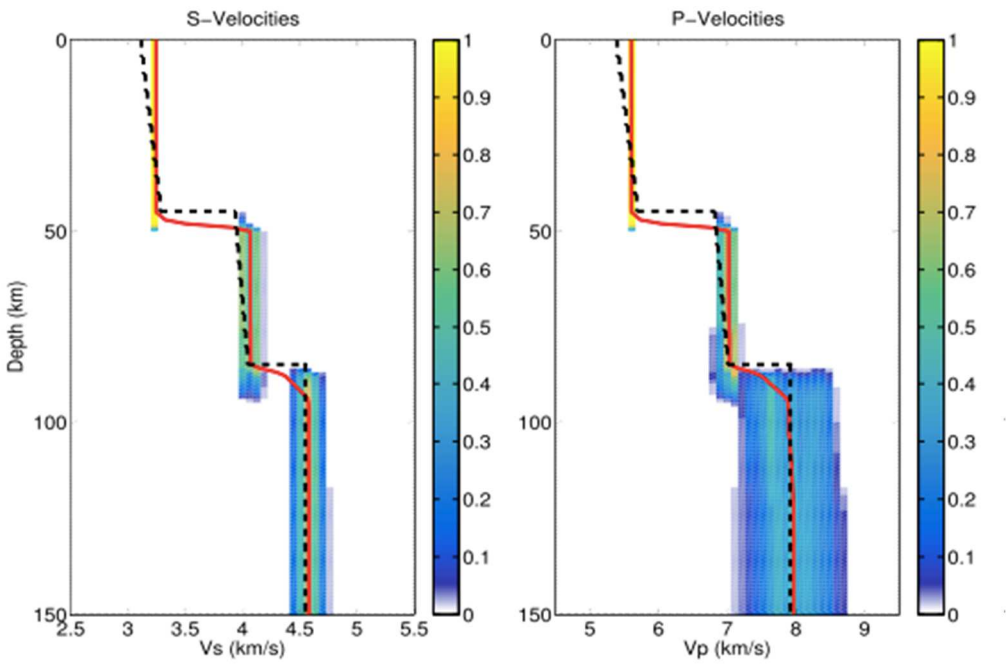


Figure 6.1 Probability density functions (PDFs) for Vs (left) and Vp (right) obtained from a Bayesian inversion of synthetic receiver functions combined with Rayleigh wave ZH ratio measurements. The input model is shown by the black dashed line, and the mean model from each PDF is shown with a red line.

Chapter two also introduced and compared two types of model parameterizations of 1D layered structure for the multi-parameter inversion problems. The independent parameterization imposes less-restrictive assumptions where different elastic parameters do not share identical model geometry. I advocate for this type of parameterization for such multi-parameter inversion as it allows better model estimates with reduced trade-offs. For example, the independent parameterization scheme is well suited for the application of constraining radial anisotropy using surface wave data and targeting disparate V_P and V_S structures such as those associated with α-β quartz transition. Figure 6.2 is another example for the possible disparate V_P and V_S structures of the firn of the southeastern Greenland ice sheet. Active source

surveys have been conducted to study the accumulation of melted snow within the firn pore space (Montgomery et al., 2017). Since V_S is more sensitive to the presence of water compared to V_P, it can be expected to have different structures for V_P and V_S. Using body wave travel time and surface wave dispersion data, the transdimensional Bayesian inversion is able to retrieve the top 80-meter V_P and V_S structure (See figure 6.2). The independent parameterization allows us to infer the V_P/V_S structure and estimate the depth and thickness of the aquifer layer. This ongoing collaborative project can also allow the attenuation to be inverted as another model parameter to take into account the high attenuation features within the aquifer. The transdimensional Bayesian approach is an inversion strategy that could be applied in the broad geophysical inverse problems. The method development part of this dissertation further broadens its application especially for joint inversions targeting retrieval of multi-parameter geophysical models.

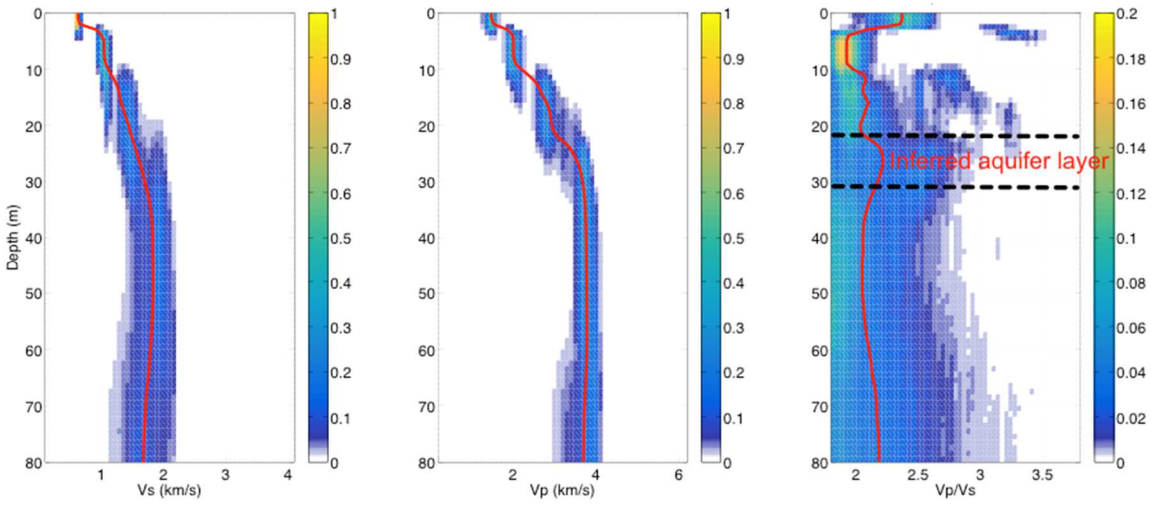


Figure 6.2 Probability density functions (PDFs) for V_S (left) and Vp (middle) obtained from a Bayesian inversion of P-wave travel time and Rayleigh wave dispersion. The Vp/V_S ratio structure (right) is calculated from the V_S and Vp ensemble. The mean model from each PDF is shown with a red line.

Chapter four shows the implementation of transdimensional Bayesian joint inversion with surface wave dispersion and receiver functions. Joint inversions of these two data types (and sometimes including ZH ratio, gravity data) have gain increasing attention in geophysical studies due to the complimentary data sensitivity. In addition to the TB inversion framework and parameterization choices introduced in chapter two, this chapter further demonstrates the data uncertainty quantification, especially for receiver function. A progressive-inclusion inversion scheme is also deployed to promote fast and stable convergence without significantly increasing the computational cost, which is a critical factor for inversions with large seismic array data. As I started conducting inversion with actual data from the Transportable Array, an alarming correlation between the presences of thin sedimentary layer and the apparent retrieval of a crustal low velocity zone is observed. I systematically show that these LVZs are artifacts resulting from large amplitude sediment reverberations combined with insufficient constraints from SWD data. I then demonstrate that the artificial LVZs disappear from joint inversion results when a simple sediment-removal filter is applied to the RFs, allowing reliable estimates of crustal structure. These findings call for previously claimed crustal LVZs obtained from similar data to be re-evaluated (e.g. Horspool et al., 2006; Chen et al., 2015; Li et al., 2017). Therefore, I advocate best practices for constraining crustal structure using joint inversion of SWD and RFs under different shallow-layer scenarios.

In this dissertation, Ps receiver functions are used to constrain the isotropic, flat-layered model. With the increasing effort on modeling the effects of anisotropy

and dipping layers on receiver functions, future work of TB inversion that incorporates receiver function with more complex modeling together with surface wave data could lead to more comprehensive seismic structures. While this dissertation only focuses on the top 100km of the continental lithosphere, straightforward addition of longer-period surface wave dispersion measurements from earthquake waveforms can help improve the resolution of deeper parts of the lithospheric mantle. Joint inversion including Sp receiver functions could also help constrain the structure of the lithospheric mantle, as signals from deeper discontinuities in the lithosphere are more easily interpreted in Sp receiver functions because they do not suffer from contamination by sediment reverberations.

Chapter two, three and four tackled on several technical issues for retrieving reliable seismic imaging of the continental lithosphere using complementary data from large seismic array. As an application, I carried out joint inversion using surface wave dispersion and receiver function data from 291 Transportable Array seismic stations. The inversion results allow me to construct a shear velocity model for the top 100 km of the Northern Great Plains of the United States. The studied region comprises well-preserved tectonic units including three Archean blocks and the Paleoproterozoic Trans-Hudson Orogen. The bulk properties and variations of the seismic velocities within each tectonic unit are analyzed and interpreted along with geochemical and petrological studies in this region. Our TBI-NGP model demonstrates contrasting structures between different tectonic settings. One noticeable feature observed across the majority of the NGP is the high velocity lower crust. The high seismic velocity in the lower crust could indicate a mafic composition, given the correlation between

seismic wavespeeds and the SiO_2 content (e.g. Huang et al., 2013). I applied a Bayesian analysis to incorporate the seismic constraints and the laboratory measured crustal rock wavespeeds from Huang et al., (2013). By assuming amphibolite and granulite facies, I was able to infer the middle and lower crust composition for the different tectonic settings within the NGP. As a comparison, I performed composition estimates for both the TBI model and the reference global crust model CRUST 1.0. Due to the faster lower crust in the TBI model, both Archean blocks and the Trans-Hudson Orogen suggest a mafic composition despite the contrasting crustal thickness. Together with active source seismic survey, geochemical and xenolith studies, the formation of the lower crust in NGP appear to be a results of different tectonic processes including periodic lower crust addition through magmatic and possibly mechanical processes in the Medicine Hat Block and Northern Wyoming Craton; a partially eclogitized root in the lower crust of Trans-Hudson Orogen and a well stratified and possibly partially delaminated Superior Craton.

In chapter five of this dissertation, a crustal shear velocity model for the Northern Great Plains of the United States is presented and also compared to CRUST 1.0. Unlike the global reference model, the TBI-NGP model yields higher resolution due to the dense seismic array and more accurate velocity structure due to the joint inversion. Accurate crustal model could benefit seismic imaging of the mantle, as crust correction is often needed for such study. In figure 6.3, I show a comparison plot to demonstrate that inaccurate crust model could lead to biased mantle structure. In this test, I use a PREM mantle structure down to 250 km, and impose the average TBI-NGP crust model to predict the surface wave dispersion data. The other set of surface

wave dispersion data are predicted using the CRUST 1.0 model imposed on PREM mantle structure. For both sets, I perturb the mantle velocity by ± dln 2%. The figure shows that at certain period range, the surface wave dispersion prediction of Crust 1.0 with a fast mantle could be replace by TBI-NGP with a normal mantle, meaning that the mantle velocities might be biased due to inaccurate crustal structure.

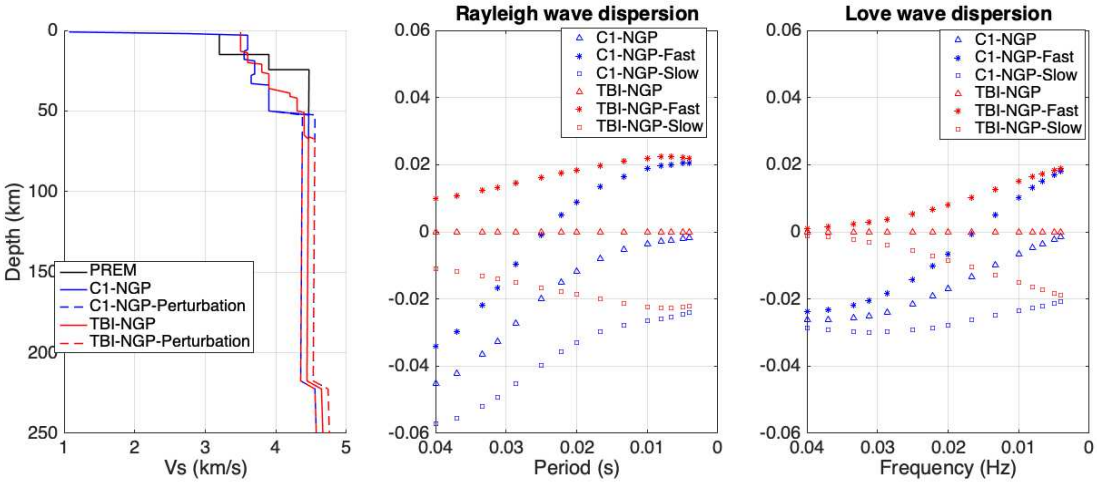


Figure 6.3 (Left) One-dimensional crust and mantle velocity structures of PREM (black), TBI-NGP crust on top of PREM mantle (red), and CRUST 1.0 in NGP on top of PREM mantle (blue). (middle) Rayleigh wave prediction from the crust and mantle structure shown on the left, for both two models, the mantle structures are perturbed by ± dln 2%. (right) Same as the middle panel, but for Love wave dispersion.

Chapter five also demonstrates one approach to infer crustal composition from seismic velocities. The ensemble nature of the TBI-NGP model allows me to incorporate seismic constraints with lab-measured wavespeeds for crustal rocks. An alternative method for estimating composition from seismic velocities is to use thermodynamic modeling to determine the mineral assemblages that crystallize from a particular bulk composition at given pressure and temperature. Seismic velocities (V_P and V_S) can then be estimated from the elastic properties of each mineral based on

mixture theory. This method allows calculation of seismic velocities over a large P-T range with small increment compared to the lab-measurement. Similar Bayesian analysis can be applied for this approach to infer the crustal composition. In the regions where both seismic data and crustal xenoliths are available, the geotherm could also be estimated by incorporating seismic velocities with predicted wavespeeds at different P-T condition. Since the composition of crustal xenoliths can be determined, assuming the xenoliths are representative of the bulk crust composition, one can narrow down the possible P-T condition of the crust based on the agreement between seismic velocities with predicted wavespeeds. A conditional probability like $p(P, T|V_{seismic})$ can be estimated using Bayesian analysis. Given the relatively well-constrained pressure distribution along depths, the geotherm can then be estimated. With the increasing amount of seismic inversion conducted in a probabilistic fashion, it is reasonable to expect that interdisciplinary studies in geoscience would benefit from the incorporation of accurate, complementary knowledge to improve our understand of Earth's composition, formation and evolution.

References

Abt, D. L., Fischer, K. M., French, S. W., Ford, H. A., Yuan, H., & Romanowicz, B. (2010). North American lithospheric discontinuity structure imaged by Ps and Sp receiver functions. *Journal of Geophysical Research: Solid Earth*, *115*(B9).

Agostinetti, N. P., & Malinverno, A. (2010). Receiver function inversion by trans-dimensional Monte Carlo sampling. *Geophysical Journal International*, *181*(2), 858-872.

Aki, K. (1968). Seismological evidences for the existence of soft thin layers in the upper mantle under Japan. *Journal of Geophysical Research*, *73*(2), 585-594.

Aki, K., & Richards, P. G. (2002). *Quantitative seismology* (University Science Books, Sausalito, CA), 2nd ed.

Alterman, Z., Jarosch, H., & Pekeris, C. L. (1961). Propagation of Rayleigh waves in the earth. *Geophysical Journal International*, *4*(Supplement_1), 219-241.

Ammon, C. J., Randall, G. E., & Zandt, G. (1990). On the nonuniqueness of receiver function inversions. *Journal of Geophysical Research: Solid Earth*, *95*(B10), 15303-15318.

Anderson, D. L. (1961). Elastic wave propagation in layered anisotropic media. *Journal of Geophysical Research*, *66*(9), 2953-2963.

Arora, S., Bhattacharya, S. N., & Gogna, M. L. (1996). Rayleigh wave dispersion equation for a layered spherical earth with exponential function solutions in each shell. *pure and applied geophysics*, *147*(3), 515-536.

Aurenhammer, F. (1991). Voronoi diagrams—a survey of a fundamental geometric data structure. *ACM Computing Surveys (CSUR)*, *23*(3), 345-405.

Backus, G. E., & Gilbert, J. F. (1967). Numerical applications of a formalism for geophysical inverse problems. *Geophysical Journal International*, *13*(1-3), 247-276.

Baird, D. J., Nelson, K. D., Knapp, J. H., Walters, J. J., & Brown, L. D. (1996). Crustal structure and evolution of the Trans‐Hudson orogen: Results from seismic reflection profiling. *Tectonics*, *15*(2), 416-426.

Banerdt, W. B., Smrekar, S., Lognonné, P., Spohn, T., Asmar, S. W., Banfield, D., ... & Giardini, D. (2013, March). InSight: a discovery mission to explore the interior of Mars. In *Lunar and planetary science conference* (Vol. 44, p. 1915).

Bankey, V., Cuevas, A., Daniels, D., Finn, C. A., Hernandez, I., Hill, P., ... & Roest, W. (2002). *Digital data grids for the magnetic anomaly map of North America (No. 2002-414)*.

Barnhart, K. R., Mahan, K. H., Blackburn, T. J., Bowring, S. A., & Dudas, F. O. (2012). Deep crustal xenoliths from central Montana, USA: Implications for the timing and mechanisms of high-velocity lower crust formation. *Geosphere*, *8*(6), 1408-1428.

Bastow, I. D., Eaton, D. W., Kendall, J. M., Helffrich, G., Snyder, D. B., Thompson, D. A., ... & Pawlak, A. E. (2015). The Hudson Bay Lithospheric Experiment (HuBLE): insights into Precambrian plate tectonics and the development of mantle keels. *Geological Society, London, Special Publications*, *389*(1), 41-67.

Mr. Bayes, & Price, M. (1763). An essay towards solving a problem in the doctrine of chances. by the late Rev. Mr. Bayes, F.R.S. communicated by Mr. price, in a letter to John Canton, A.M.F.R.S. *Philosophical Transactions (1683-1775)*, 370-418.

Beck, S. L., & Zandt, G. (2002). The nature of orogenic crust in the central Andes. *Journal of Geophysical Research: Solid Earth*, *107*(B10).

Becker, T. W., Kellogg, J. B., Ekström, G., & O'Connell, R. J. (2003). Comparison of azimuthal seismic anisotropy from surface waves and finite strain from global mantle-circulation models. *Geophysical Journal International*, *155*(2), 696-714.

Becker, T. W. (2006). On the effect of temperature and strain-rate dependent viscosity on global mantle flow, net rotation, and plate-driving forces. *Geophysical Journal International*, *167*(2), 943-957.

Becker, T. W. (2008). Azimuthal seismic anisotropy constrains net rotation of the lithosphere. *Geophysical Research Letters*, *35*(5).

Beghein, C., & Trampert, J. (2004). Probability density functions for radial anisotropy: implications for the upper 1200 km of the mantle. *Earth and Planetary Science Letters*, *217*(1), 151-162.

Beghein, C., Yuan, K., Schmerr, N., & Xing, Z. (2014). Changes in seismic anisotropy shed light on the nature of the Gutenberg discontinuity. *Science*, *343*(6176), 1237-1240.

Bensen, G. D., Ritzwoller, M. H., & Shapiro, N. M. (2008). Broadband ambient noise surface wave tomography across the United States. *Journal of Geophysical Research: Solid Earth*, *113*(B5).

Bhattacharya, S. N. (1996). Earth-flattening transformation for P-SV waves. *Bulletin of the Seismological Society of America*, *86*(6), 1979-1982.

Bhattacharya, S. N., & Arora, S. (1997). A flattening transformation for P-SV waves in a transversely isotropic Earth. *Bulletin of the Seismological Society of America*, *87*(5), 1297-1304.

Biswas, N. N., & Knopoff, L. (1970). Exact earth-flattening calculation for Love waves. *Bulletin of the Seismological Society of America*, *60*(4), 1123-1137.

Bodin, T., & Sambridge, M. (2009). Seismic tomography with the reversible jump algorithm. *Geophysical Journal International*, *178*(3), 1411-1436.

Bodin, T., Sambridge, M., Tkalčić, H., Arroucau, P., Gallagher, K., & Rawlinson, N. (2012). Transdimensional inversion of receiver functions and surface wave dispersion. *Journal of Geophysical Research: Solid Earth*,*117*(B2).

Bodin, T., Capdeville, Y., Romanowicz, B., & Montagner, J. P. (2015). Interpreting radial anisotropy in global and regional tomographic models. In The Earth's Heterogeneous Mantle (pp. 105-144). Springer International Publishing.

Bodin, T., Leiva, J., Romanowicz, B., Maupin, V., & Yuan, H. (2016). Imaging anisotropic layering with Bayesian inversion of multiple data types. *Geophysical Journal International*, *206*(1), 605-629.

Boore, D. M., & Toksöz, M. N. (1969). Rayleigh wave particle motion and crustal structure. *Bulletin of the Seismological Society of America*, *59*(1), 331-346.

Boschi, L., & Ekström, G. (2002). New images of the Earth's upper mantle from measurements of surface wave phase velocity anomalies. Journal of Geophysical Research: Solid Earth (1978–2012), 107(B4), ESE-1.

Bozdağ, E., & Trampert, J. (2008). On crustal corrections in surface wave tomography. *Geophysical Journal International*, *172*(3), 1066-1082.

Braile, L. W., Smith, R. B., Ansorge, J., Baker, M. R., Sparlin, M. A., Prodehl, C., ... & Olsen, K. H. (1982). The Yellowstone-Snake River Plain seismic profiling experiment: crustal structure of the eastern Snake River Plain. *Journal of Geophysical Research: Solid Earth*, *87*(B4), 2597-2609.

Braile, L. W. (1989). Crustal structure of the continental interior. *Geophysical framework of the continental United States*, *172*, 285-315.

Brocher, T. M. (2005). Empirical relations between elastic wavespeeds and density in the Earth's crust. *Bulletin of the Seismological Society of America*,*95*(6), 2081-2092.

Burdick, S., & Lekić, V. (2017). Velocity variations and uncertainty from transdimensional P-wave tomography of North America. *Geophysical Journal International*, *209*(2), 1337-1351.

Calò, M., Bodin, T., & Romanowicz, B. (2016). Layered structure in the upper mantle across North America from joint inversion of long and short period seismic data. *Earth and Planetary Science Letters*, *449*, 164-175.

Campbell, K. W., & Bozorgnia, Y. (2008). NGA ground motion model for the geometric mean horizontal component of PGA, PGV, PGD and 5% damped linear elastic response spectra for periods ranging from 0.01 to 10 s. *Earthquake Spectra*, *24*(1), 139-171.

Chai, C., Ammon, C. J., Maceira, M., & Herrmann, R. B. (2015). Inverting interpolated receiver functions with surface wave dispersion and gravity: application to the western US and adjacent Canada and Mexico. *Geophysical Research Letters*, *42*(11), 4359-4366.

Chang, S. J., Ferreira, A. M., Ritsema, J., van Heijst, H. J., & Woodhouse, J. H. (2014). Global radially anisotropic mantle structure from multiple datasets: a review, current challenges, and outlook. *Tectonophysics*, *617*, 1-19.

Chang, S. J., Ferreira, A. M., Ritsema, J., Heijst, H. J., & Woodhouse, J. H. (2015). Joint inversion for global isotropic and radially anisotropic mantle structure including crustal thickness perturbations. *Journal of Geophysical Research: Solid Earth*, *120*(6), 4278-4300.

Chen, M., & Tromp, J. (2007). Theoretical and numerical investigations of global and regional seismic wave propagation in weakly anisotropic earth models. *Geophysical Journal International*, *168*(3), 1130-1152.

Chen, X. (1993). A systematic and efficient method of computing normal modes for multilayered half-space. *Geophysical Journal International*, *115*(2), 391-409.

Chen, Y., Gu, Y. J., Dokht, R. M., & Sacchi, M. D. (2015). Crustal imprints of Precambrian orogenesis in western Laurentia. *Journal of Geophysical Research: Solid Earth*, *120*(10), 6993-7012.

Christensen, N. I., & Mooney, W. D. (1995). Seismic velocity structure and composition of the continental crust: A global view. *Journal of Geophysical Research: Solid Earth*, *100*(B6), 9761-9788.

Christensen, N. I. (1996). Poisson's ratio and crustal seismology. *Journal of Geophysical Research: Solid Earth*, *101*(B2), 3139-3156.

Constable, S. C., Parker, R. L., & Constable, C. G. (1987). Occam's inversion: A practical algorithm for generating smooth models from electromagnetic sounding data. *Geophysics*, *52*(3), 289-300.

Crampin, S., Chesnokov, E. M., & Hipkin, R. G. (1984). Seismic anisotropy—the state of the art: II. *Geophysical Journal International*, *76*(1), 1-16.

Corrigan, D., Pehrsson, S., Wodicka, N., & De Kemp, E. (2009). The Palaeoproterozoic Trans-Hudson Orogen: a prototype of modern accretionary processes. *Geological Society, London, Special Publications*, *327*(1), 457-479.

Darbyshire, F. A., Eaton, D. W., & Bastow, I. D. (2013). Seismic imaging of the lithosphere beneath Hudson Bay: Episodic growth of the Laurentian mantle keel. *Earth and Planetary Science Letters*, *373*, 179-193.

Dettmer, J., Dosso, S. E., & Holland, C. W. (2010). Trans-dimensional geoacoustic inversion. *The Journal of the Acoustical Society of America*,*128*(6), 3393-3405.

Dettmer, J., Benavente, R., Cummins, P. R., & Sambridge, M. (2014). Trans-dimensional finite-fault inversion. *Geophysical Journal International*, *199*(2), 735-751.

Dettmer, J., Dosso, S. E., Bodin, T., Stipčević, J., & Cummins, P. R. (2015). Direct-seismogram inversion for receiver-side structure with uncertain source–time functions. *Geophysical journal international*, *203*(2), 1373-1387.

Diaferia, G., & Cammarano, F. (2017). Seismic Signature of the Continental Crust: What Thermodynamics Says. An Example From the Italian Peninsula. *Tectonics*.

Downes, H., Macdonald, R. A. Y., Upton, B. G., Cox, K. G., Bodinier, J. L., Mason, P. R., ... & Hearn Jr, B. C. (2004). Ultramafic xenoliths from the Bearpaw Mountains, Montana, USA: Evidence for multiple metasomatic events in the lithospheric mantle beneath the Wyoming craton. *Journal of Petrology*, *45*(8), 1631-1662.

Dziewonski, A. M., & Anderson, D. L. (1981). Preliminary reference Earth model. *Physics of the earth and planetary interiors*, *25*(4), 297-356.

Ekström, G., Tromp, J., & Larson, E. W. (1997). Measurements and global models of surface wave propagation. Journal of Geophysical Research: Solid Earth (1978–2012), 102(B4), 8137-8157.

Ekström, G., & Dziewonski, A. M. (1998). The unique anisotropy of the Pacific upper mantle. *Nature*, *394*(6689), 168-172.

Ekström, G. (2011). A global model of Love and Rayleigh surface wave dispersion and anisotropy, 25-250 s. Geophysical Journal International, 187(3), 1668-1686.

Ekström, G. (2014). Love and Rayleigh phase-velocity maps, 5–40 s, of the western and central USA from USArray data. *Earth and Planetary Science Letters*, *402*, 42-49.

Ekström, G. (2017). Short-period surface-wave phase velocities across the conterminous United States. *Physics of the Earth and Planetary Interiors*, *270*, 168-175.

Facer, J., Downes, H., & Beard, A. (2009). In situ serpentinization and hydrous fluid metasomatism in spinel dunite xenoliths from the Bearpaw Mountains, Montana, USA. *Journal of Petrology*, *50*(8), 1443-1475.

Ferreira, A. M., & Woodhouse, J. H. (2007). Observations of long period Rayleigh wave ellipticity. *Geophysical Journal International*, *169*(1), 161-169.

Ferreira, A. M. G., Woodhouse, J. H., Visser, K., & Trampert, J. (2010). On the robustness of global radially anisotropic surface wave tomography. *Journal of Geophysical Research: Solid Earth*, *115*(B4).

Flament, N., Rey, P. F., Coltice, N., Dromart, G., & Olivier, N. (2011). Lower crustal flow kept Archean continental flood basalts at sea level. *Geology*, *39*(12), 1159-1162.

Franklin, J. N. (1970). Well-posed stochastic extensions of ill-posed linear problems. *Journal of Mathematical Analysis and Applications*, *31*(3), 682-716.

Gao, C., & Lekić, V. (2018). Consequences of parametrization choices in surface wave inversion: insights from transdimensional Bayesian methods. *Geophysical Journal International*, *215*(2), 1037-1063.

Gehrmann, R. A., Dettmer, J., Schwalenberg, K., Engels, M., Dosso, S. E., & Özmaral, A. (2015). Trans‐dimensional Bayesian inversion of controlled‐source electromagnetic data in the German North Sea. *Geophysical Prospecting*, *63*(6), 1314-1333.

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

Giletti, B. J. (1966). Isotopic ages from southwestern Montana. *Journal of Geophysical Research*, *71*(16), 4029-4036.

Gilligan, A., Bastow, I. D., & Darbyshire, F. A. (2016). Seismological structure of the 1.8 Ga Trans‐Hudson Orogen of North America. *Geochemistry, Geophysics, Geosystems*, *17*(6), 2421-2433.

Gorman, A. R., Clowes, R. M., Ellis, R. M., Henstock, T. J., Levander, A., Spence, G. D., ... & Kanasewich, E. R. (2002). Deep Probe: imaging the roots of western North America. *Canadian Journal of Earth Sciences*, *39*(3), 375-398.

Gosselin, J. M., Dosso, S. E., Cassidy, J. F., Quijano, J. E., Molnar, S., & Dettmer, J. (2017). A gradient-based model parametrization using Bernstein polynomials in Bayesian inversion of surface wave dispersion. *Geophysical Journal International*, *211*(1), 528-540.

Green, P. J. (2003). Trans-dimensional markov chain monte carlo. *Oxford Statistical Science Series*, 179-198.

Hacker, B. R., Kelemen, P. B., & Behn, M. D. (2011). Differentiation of the continental crust by relamination. *Earth and Planetary Science Letters*, *307*(3-4), 501-516.

Haid, J. H. (1991). *Tectonic subsidence analysis of the Williston Basin* (Doctoral dissertation, University of Saskatchewan).

Hajnal, Z., Fowler, C. M. R., Mereu, R. F., Kanasewich, E. R., Cumming, G. L., Green, A. G., & Mair, A. (1984). An initial analysis of the Earth's crust under the Williston

Basin: 979 Cocrust Experiment. *Journal of Geophysical Research: Solid Earth*, *89*(B11), 9381-9400.

Hamdani, Y., Mareschal, J. C., & Arkani-Hamed, J. (1994). Phase change and thermal subsidence of the Williston basin. *Geophysical Journal International*, *116*(3), 585-597.

Harkrider, D. G., & Anderson, D. L. (1962). Computation of surface wave dispersion for multilayered anisotropic media. *Bulletin of the Seismological Society of America*, *52*(2), 321-332.

Haskell, N. A. (1953). The dispersion of surface waves on multilayered media. *Bulletin of the seismological Society of America*, *43*(1), 17-34.

Herrmann, R. B. (2013). Computer programs in seismology: An evolving tool for instruction and research. *Seismological Research Letters*, *84*(6), 1081-1088.

Hisada, Y. (1994). An efficient method for computing Green's functions for a layered half-space with sources and receivers at close depths. *Bulletin of the Seismological Society of America*, *84*(5), 1456-1472.

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*(3), 1466-1483.

Huang, Z., Su, W., Peng, Y., Zheng, Y., & Li, H. (2003). Rayleigh wave tomography of China and adjacent regions. *Journal of Geophysical Research: Solid Earth*, *108*(B2).

Huang, Y., Chubakov, V., Mantovani, F., Rudnick, R. L., & McDonough, W. F. (2013). A reference Earth model for the heat‐producing elements and associated geoneutrino flux. *Geochemistry, Geophysics, Geosystems*, *14*(6), 2003-2029.

Hoffman, P. F. (1988). United plates of America, the birth of a craton: Early Proterozoic assembly and growth of Laurentia. *Annual Review of Earth and Planetary Sciences*, *16*(1), 543-603.

Hoffman, P. F., Bally, A. W., & Palmer, A. R. (1989). Precambrian geology and tectonic history of North America. *The geology of North America—an overview*, 447-512.

Holbrook, W. S., Mooney, W. D., & Christensen, N. I. (1992). The seismic velocity structure of the deep continental crust. *Continental lower crust*, *23*, 1-43.

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*(1), 99-112.

Kendall, J. M. (2000). Seismic anisotropy in the boundary layers of the mantle. *Earth's Deep Interior: Mineral physics and tomography from the atomic to the global scale*, 133-159.

Kennett, B. L. N., & Kerry, N. J. (1979). Seismic waves in a stratified half space. *Geophysical Journal International*, *57*(3), 557-583.

Kennett, B.L.N. (1983). Seismic wave propagation in stratified media. Cambridge: Cambridge Univ. Press.

Kennett, B. L. N. (1991). The removal of free surface interactions from three-component seismograms. *Geophysical Journal International*, *104*(1), 153-163.

Kim, S., Dettmer, J., Rhie, J., & Tkalčić, H. (2016). Highly efficient Bayesian joint inversion for receiver-based data and its application to lithospheric structure beneath the southern Korean Peninsula. *Geophysical Journal International*, *206*(1), 328-344.

Kind, R., Ni, J., Zhao, W., Wu, J., Yuan, X., Zhao, L., ... & Hearn, T. (1996). Evidence from earthquake data for a partially molten crustal layer in southern Tibet. *Science*, *274*(5293), 1692-1694.

Kolb, J. M., & Lekić, V. (2014). Receiver function deconvolution using transdimensional hierarchical Bayesian inference. *Geophysical Journal International*, *197*(3), 1719-1735.

Kosarev, G. L., Petersen, N. V., Vinnik, L. P., & Roecker, S. W. (1993). Receiver functions for the Tien Shan analog broadband network: Contrasts in the evolution of structures across the Talasso‐Fergana fault. *Journal of Geophysical Research: Solid Earth*, *98*(B3), 4437-4448.

Kucks, R. P. (1999). Bouguer gravity anomaly data grid for the conterminous US: US Geological Survey Digital Data Series DDS-9.

Kullback, S., & Leibler, R. A. (1951). On information and sufficiency. *The annals of mathematical statistics*, *22*(1), 79-86.

Kumar, M. R., & Bostock, M. G. (2008). Extraction of absolute P velocity from receiver functions. *Geophysical Journal International*, *175*(2), 515-519.

Kuo‐Chen, H., Wu, F. T., Jenkins, D. M., Mechie, J., Roecker, S. W., Wang, C. Y., & Huang, B. S. (2012). Seismic evidence for the α‐β quartz transition beneath Taiwan from Vp/Vs tomography. *Geophysical Research Letters*, *39*(22).

Kustowski, B., Ekström, G., & Dziewoński, A. M. (2008). Anisotropic shear-wave velocity structure of the Earth's mantle: A global model. *Journal of Geophysical Research: Solid Earth*, *113*(B6).

Lai, C. G., & Rix, G. J. (1998). *Simultaneous inversion of Rayleigh phase velocity and attenuation for near-surface site characterization* (p. 258). Georgia: School of Civil and Environmental Engineering, Georgia Institute of Technology.

Langston, C. A., & Liang, C. (2008). Gradiometry for polarized seismic waves. *Journal of Geophysical Research: Solid Earth*, *113*(B8).

Laske, G., & Masters, G. (1996). Constraints on global phase velocity maps from long-period polarization data. *Journal of Geophysical Research: Solid Earth*, *101*(B7), 16059-16075.

Laske, G., Masters, G., Ma, Z., & Pasyanos, M. (2013, April). Update on CRUST1. 0—A 1-degree global model of Earth's crust. In *Geophys. Res. Abstr* (Vol. 15, p. 2658). Vienna, Austria: EGU General Assembly.

Ledoit, O., & Wolf, M. (2004). A well-conditioned estimator for large-dimensional covariance matrices. *Journal of multivariate analysis*, *88*(2), 365-411.

Lekić, V., Panning, M., & Romanowicz, B. (2010). A simple method for improving crustal corrections in waveform tomography. *Geophysical Journal International*, *182*(1), 265-278.

Lekić, V., & Romanowicz, B. (2011). Inferring upper-mantle structure by full waveform tomography with the spectral element method. *Geophysical Journal International*, *185*(2), 799-831.

Li, X., G. Bock, A. Vafidis, R. Kind, H.-P. Harjes, W. Hanka, K. Wylegalla, M. Van Der Meijde, and X. Yuan (2003), Receiver function study of the Hellenic subduction zone: Imaging crustal thickness variations and the oceanic Moho of the descending African lithosphere, Geophys. J. Int., 155(2), 733–748.

Li, J., Song, X., Zhu, L., & Deng, Y. (2017). Joint inversion of surface wave dispersions and receiver functions with P velocity constraints: Application to Southeastern Tibet. *Journal Geophysical Research: Solid Earth*, 122, 7291–7310.

Li, M., Zhang, S., Bodin, T., Lin, X., & Wu, T. (2018). Transdimensional inversion of scattered body waves for 1D S-wave velocity structure–Application to the Tengchong volcanic area, Southwestern China. *Journal of Asian Earth Sciences*, *159*, 60-68.

Liang, C., & Langston, C. A. (2008). Ambient seismic noise tomography and structure of eastern North America. *Journal of Geophysical Research: Solid Earth*, *113*(B3).

Ligorria, J. P., & Ammon, C. J. (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*(1), 281-298.

Lin, F. C., Schmandt, B., & Tsai, V. C. (2012). Joint inversion of Rayleigh wave phase velocity and ellipticity using USArray: Constraining velocity and density structure in the upper crust. *Geophysical Research Letters*, *39*(12).

Lin, F.C., Tsai, V.C. and Schmandt, B., (2014). 3-D crustal structure of the western United States: application of Rayleigh-wave ellipticity extracted from noise cross-correlations. *Geophysical Journal International*, p.ggu160.

Love, A.E.H. (1911). Some problems of Geodynamics. Cambridge University Press, London

Luco, J. E., & Apsel, R. J. (1983). On the Green's functions for a layered half-space. Part I. *Bulletin of the Seismological Society of America*, *73*(4), 909-929.

Ma, Z., Masters, G., Laske, G., & Pasyanos, M. (2014). A comprehensive dispersion model of surface wave phase and group velocity for the globe. *Geophysical Journal International*, *199*(1), 113-135.

Mahan, K. H., Schulte‐Pelkum, V., Blackburn, T. J., Bowring, S. A., & Dudas, F. O. (2012). Seismic structure and lithospheric rheology from deep crustal xenoliths, central Montana, USA. *Geochemistry, Geophysics, Geosystems*, *13*(10).

Malinverno, A. (2000). A Bayesian criterion for simplicity in inverse problem parametrization. *Geophysical Journal International*, *140*(2), 267-285.

Malinverno, A. (2002). Parsimonious Bayesian Markov chain Monte Carlo inversion in a nonlinear geophysical problem. *Geophysical Journal International*, *151*(3), 675-688.

Masters, G.; Woodhouse, J.H.; Freeman, G. (2011), Mineos v1.0.2 [software], Computational Infrastructure for Geodynamics, doi: tempMi3, url: **https://geodynamics.org/cig/software/mineos/**

Marshak, S., Domrois, S., Abert, C., Larson, T., Pavlis, G., Hamburger, M., ... & Chen, C. (2017). The basement revealed: Tectonic insight from a digital elevation model of the Great Unconformity, USA cratonic platform. *Geology*, *45*(5), 391-394.

McCamy, K., & Meyer, R. P. (1964). A correlation method of apparent velocity measurement. *Journal of Geophysical Research*, *69*(4), 691-699.

Menke, W., 2012. *Geophysical data analysis: Discrete inverse theory* (Vol. 45). Academic press.

Montagner, J. P., & Anderson, D. L. (1989). Constrained reference mantle model. *Physics of the earth and planetary interiors*, *58*(2), 205-227.

Montagner, J. P., & Tanimoto, T. (1991). Global upper mantle tomography of seismic velocities and anisotropies. *Journal of Geophysical Research: Solid Earth*, *96*(B12), 20337-20351.

Montagner, J. P. (2002). Upper mantle low anisotropy channels below the Pacific Plate. *Earth and Planetary Science Letters*, *202*(2), 263-274.

Montgomery, L.N., Schmerr, N., Burdick, S., Forster, R.R., Koenig, L., Legchenko, A., Ligtenberg, S., Miège, C., Miller, O.L. & Solomon, D.K., (2017). Investigation of firn aquifer structure in southeastern Greenland using active source seismology. *Frontiers in Earth Science*, *5*, p.10.

Mooney, W. D., Laske, G., & Masters, T. G. (1998). CRUST 5.1: A global crustal model at 5×5. *Journal of Geophysical Research: Solid Earth*, *103*(B1), 727-747.

Mooney, W. D. (2015), Crust and lithospheric structure—Global crustal structure, in *Treatise of Geophysics*, 2nd ed., vol. 1, pp. 339–390, Elsevier, Amsterdam.

Mosegaard, K. (1998). Resolution analysis of general inverse problems through inverse Monte Carlo sampling. *Inverse problems*, *14*(3), 405.

Moulik, P., & Ekström, G. (2014). An anisotropic shear velocity model of the Earth's mantle using normal modes, body waves, surface waves and long-period waveforms. *Geophysical Journal International*, *199*(3), 1713-1738.

Mueller, P. A., Heatherington, A. L., Kelly, D. M., Wooden, J. L., & Mogk, D. W. (2002). Paleoproterozoic crust within the Great Falls tectonic zone: Implications for the assembly of southern Laurentia. *Geology*, *30*(2), 127-130.

Nelson, K. D. (1992). Are crustal thickness variations in old mountain belts like the Appalachians a consequence of lithospheric delamination?. *Geology*, *20*(6), 498-502.

Nelson, K. D., Baird, D. J., Walters, J. J., Hauck, M., Brown, L. D., Oliver, J. E., ... & Sloss, L. L. (1993). Trans-Hudson orogen and Williston basin in Montana and North Dakota: New COCORP deep-profiling results. *Geology*, *21*(5), 447-450.

Olugboji, T. M., V. Lekic, and W. McDonough (2017), A statistical assessment of seismic models of the U.S. continental crust using Bayesian inversion of ambient noise surface wave dispersion data, *Tectonics*, 36, 1232–1253.

Owens, T. J., Zandt, G., & Taylor, S. R. (1984). Seismic evidence for an ancient rift beneath the Cumberland Plateau, Tennessee: A detailed analysis of broadband teleseismic P waveforms. *Journal of Geophysical Research: Solid Earth*, *89*(B9), 7783-7795.

O'Neill, J. M., & Lopez, D. A. (1985). Character and regional significance of Great Falls tectonic zone, east-central Idaho and west-central Montana. *AAPG Bulletin*, *69*(3), 437-447.

Özalaybey, 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*(1), 183-199.

Panning, M., & Romanowicz, B. (2004). Inferences on flow at the base of Earth's mantle based on seismic anisotropy. *Science*, *303*(5656), 351-353.

Panning, M. P., Lognonné, P., Banerdt, W. B., Garcia, R., Golombek, M., Kedar, S., ... & Weber, R. (2017). Planned products of the Mars structure service for the InSight mission to Mars. *Space Science Reviews*, *211*(1-4), 611-650.

Pasyanos, M. E., Masters, T. G., Laske, G., & Ma, Z. (2014). LITHO1. 0: An updated crust and lithospheric model of the Earth. *Journal of Geophysical Research: Solid Earth*, *119*(3), 2153-2173.

Petrescu, L., Bastow, I. D., Darbyshire, F. A., Gilligan, A., Bodin, T., Menke, W., & Levin, V. (2016). Three billion years of crustal evolution in eastern Canada: Constraints from receiver functions. *Journal of Geophysical Research: Solid Earth*.

Pollack, H. N., & Chapman, D. S. (1977). On the regional variation of heat flow, geotherms, and lithospheric thickness. *Tectonophysics*, *38*(3-4), 279-296.

Pollack, H. N., Hurter, S. J., & Johnson, J. R. (1993). Heat flow from the Earth's interior: analysis of the global data set. *Reviews of Geophysics*, *31*(3), 267-280.

Ray, A., Key, K., Bodin, T., Myer, D., & Constable, S. (2014). Bayesian inversion of marine CSEM data from the Scarborough gas field using a transdimensional 2-D parametrization. *Geophysical Journal International*,*199*(3), 1847-1860.

Romanowicz, B. (2002). Inversion of surface waves: a review. *International Geophysics Series*, 81(A), 149-174.

Ross, G. M., & Villeneuve, M. (2003). Provenance of the Mesoproterozoic (1.45 Ga) Belt basin (western North America): Another piece in the pre-Rodinia paleogeographic puzzle. *GSA Bulletin*, *115*(10), 1191-1217.

Roy, C., & Romanowicz, B. A. (2017). On the Implications of A Priori Constraints in Transdimensional Bayesian Inversion for Continental Lithospheric Layering. *Journal of Geophysical Research: Solid Earth*.

Rudnick, R. L., & Fountain, D. M. (1995). Nature and composition of the continental crust: a lower crustal perspective. *Reviews of geophysics*, *33*(3), 267-309.

Rudnick, R. L., & Gao, S. (2003). Composition of the continental crust. *Treatise on geochemistry*, *3*, 659.

Rudolph, M. L., Lekić, V., & Lithgow-Bertelloni, C. (2015). Viscosity jump in Earth's mid-mantle. *Science*, *350*(6266), 1349-1352.

Rychert, C. A., & Harmon, N. (2016). Stacked P-to-S and S-to-P receiver functions determination of crustal thickness, Vp, and Vs: The H-V stacking method. *Geophysical Research Letters*, *43*(4), 1487-1494.

Sambridge, M., Bodin, T., Gallagher, K., & Tkalčić, H. (2013). Transdimensional inference in the geosciences. *Phil. Trans. R. Soc. A*,*371*(1984), 20110547.

Schutt, D. L., Dueker, K., & Yuan, H. (2008). Crust and upper mantle velocity structure of the Yellowstone hot spot and surroundings. *Journal of Geophysical Research: Solid Earth*, *113*(B3).

Schulte-Pelkum, V., Mahan, K. H., Shen, W., & Stachnik, J. C. (2017). The distribution and composition of high-velocity lower crust across the continental US: Comparison of seismic and xenolith data and implications for lithospheric dynamics and history. *Tectonics*, *36*(8), 1455-1496.

Shapiro, N. M., & Ritzwoller, M. H. (2002). Monte-Carlo inversion for a global shear-velocity model of the crust and upper mantle. *Geophysical Journal International*, *151*(1), 88-105.

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

Shen, W., Ritzwoller, M. H., & Schulte-Pelkum, V. (2013). 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. *Journal of Geophysical Research: Solid Earth*, *118*(1), 262-276.

Shen, W., & Ritzwoller, M. H. (2016). Crustal and uppermost mantle structure beneath the United States. *Journal of Geophysical Research: Solid Earth*, *121*(6), 4306-4342.

Silver, P. G., & Holt, W. E. (2002). The mantle flow field beneath western North America. *Science*, *295*(5557), 1054-1057.

Simmons, N. A., A. M. Forte, and S. P. Grand (2009), Joint seismic, geodynamic and mineral physical constraints on three‐dimensional mantle heterogeneity: Implications for the relative importance of thermal versus compositional heterogeneity, *Geophys. J. Int.,* 177(5), 1284–1304.

Simons, F. J., Van Der Hilst, R. D., Montagner, J. P., & Zielhuis, A. (2002). Multimode Rayleigh wave inversion for heterogeneity and azimuthal anisotropy of the Australian upper mantle. *Geophysical Journal International*,*151*(3), 738-754.

Sivia, D., & Skilling, J. (2006). *Data analysis: a Bayesian tutorial*. OUP Oxford.

Smith, A. F. (1991). Bayesian computational methods. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, *337*(1647), 369-386.

Tanimoto, T., & Rivera, L. (2005). Prograde Rayleigh wave particle motion. *Geophysical Journal International*, *162*(2), 399-405.

Tanimoto, T., & Rivera, L. (2008). The ZH ratio method for long-period seismic data: sensitivity kernels and observational techniques. *Geophysical Journal International*, *172*(1), 187-198.

Tarantola, A., & Valette, B. (1982). Generalized nonlinear inverse problems solved using the least squares criterion. *Reviews of Geophysics*, *20*(2), 219-232.

Tarantola, A. (2005). *Inverse problem theory and methods for model parameter estimation*. Society for Industrial and Applied Mathematics.

Thomson, W. T. (1950). Transmission of elastic waves through a stratified solid medium. *Journal of applied Physics*, *21*(2), 89-93.

Thompson, D. A., Bastow, I. D., Helffrich, G., Kendall, J. M., Wookey, J., Snyder, D. B., & Eaton, D. W. (2010). Precambrian crustal evolution: seismic constraints from the Canadian Shield. *Earth and Planetary Science Letters*, *297*(3-4), 655-666.

Tkalčić, H., Pasyanos, M. E., Rodgers, A. J., Gök, R., Walter, W. R., & Al‐Amri, A. (2006). A multistep approach for joint modeling of surface wave dispersion and teleseismic receiver functions: Implications for lithospheric structure of the Arabian Peninsula. *Journal of Geophysical Research: Solid Earth*, *111*(B11).

Trampert, J., & Snieder, R. (1996). Model estimations biased by truncated expansions: possible artifacts in seismic tomography. Science, 271(5253), 1257.

Trampert, J., & Woodhouse, J. H. (1996). High resolution global phase velocity distributions. *Geophysical Research Letters*, *23*(1), 21-24.

Trampert, J., & Woodhouse, J. H. (2003). Global anisotropic phase velocity maps for fundamental mode surface waves between 40 and 150 s. *Geophysical Journal International*, *154*(1), 154-165.

Turcotte, D., & Schubert, G. (2014). *Geodynamics*. Cambridge university press.

van Heijst, H. J., & Woodhouse, J. (1999). Global high-resolution phase velocity distributions of overtone and fundamental-mode surface waves determined by mode branch stripping. *Geophysical Journal International*,*137*(3), 601-620.

Wagner, L. S., Fouch, M. J., James, D. E., & Hanson‐Hedgecock, S. (2012). Crust and upper mantle structure beneath the Pacific Northwest from joint inversions of ambient noise and earthquake data. *Geochemistry, Geophysics, Geosystems*, *13*(12).

Ward, K. M., Zandt, G., Beck, S. L., Christensen, D. H., & McFarlin, H. (2014). Seismic imaging of the magmatic underpinnings beneath the Altiplano-Puna volcanic complex from the joint inversion of surface wave dispersion and receiver functions. *Earth and Planetary Science Letters*, *404*, 43-53.

Whitmeyer, S. J., & Karlstrom, K. E. (2007). Tectonic model for the Proterozoic growth of North America. *Geosphere*, *3*(4), 220-259.

Worthington, J. R., Hacker, B. R., & Zandt, G. (2013). Distinguishing eclogite from peridotite: EBSD-based calculations of seismic velocities. *Geophysical Journal International*, *193*(1), 489-505.

Yano, T., Tanimoto, T., & Rivera, L. (2009). The ZH ratio method for long-period seismic data: inversion for S-wave velocity structure. *Geophysical Journal International*, *179*(1), 413-424.

Yao, H., van Der Hilst, R. D., & Maarten, 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*(2), 732-744.

Yao, H., Beghein, C., & Van Der Hilst, R. D. (2008). Surface wave array tomography in SE Tibet from ambient seismic noise and two-station analysis-II. Crustal and upper-mantle structure. *Geophysical Journal International*,*173*(1), 205-219.

Yeck, W. L., Sheehan, A. F., & Schulte‐Pelkum, V. (2013). Sequential H‐κ stacking to obtain accurate crustal thicknesses beneath sedimentary basins. *Bulletin of the Seismological Society of America*, *103*(3), 2142-2150.

Yilmaz, Ö. (2001). Seismic data analysis: Processing, inversion, and interpretation of seismic data. *Society of exploration geophysicists*.

Young, M. K., Tkalcić, H., Bodin, T., & Sambridge, M. (2013). Global P wave tomography of Earth's lowermost mantle from partition modeling. *Journal of Geophysical Research: Solid Earth*, *118*(10), 5467-5486.

Yu, Y., Song, J., Liu, K. H., & Gao, S. S. (2015). Determining crustal structure beneath seismic stations overlying a low‐velocity sedimentary layer using receiver functions. *Journal of Geophysical Research: Solid Earth*, *120*(5), 3208-3218.

Zelt, B. C., & Ellis, R. M. (1999). Receiver-function studies in the Trans-Hudson orogen, Saskatchewan. *Canadian Journal of Earth Sciences*, *36*(4), 585-603.

Zhang, Z., Wang, Y., Houseman, G. A., Xu, T., Wu, Z., Yuan, X., ... & Teng, J. (2014). The Moho beneath western Tibet: Shear zones and eclogitization in the lower crust. *Earth and Planetary Science Letters*, *408*, 370-377.

Zhang, H., Lee, S., Wolin, E., Bollmann, T. A., Revenaugh, J., Wiens, D. A., ... & Stein, S. (2016). Distinct crustal structure of the North American Midcontinent Rift from P wave receiver functions. *Journal of Geophysical Research: Solid Earth*, *121*(11), 8136-8153.

Zhu, L., & Kanamori, H. (2000). Moho depth variation in southern California from teleseismic receiver functions. *Journal of Geophysical Research: Solid Earth*, *105*(B2), 2969-2980.

Zorin, Y. A., Mordvinova, V. V., Turutanov, E. K., Belichenko, B. G., Artemyev, A. A., Kosarev, G. L., & Gao, S. S. (2002). Low seismic velocity layers in the Earth's crust beneath Eastern Siberia (Russia) and Central Mongolia: receiver function data and their possible geological implication. *Tectonophysics*, 359(3-4), 307-327.