

2015-01-01

Development Of 3-D Shear Wave Models Using A Multi-Objective Optimization Scheme

Lennox Emmanuel Thompson

University of Texas at El Paso, lethompson@miners.utep.edu

Follow this and additional works at: https://digitalcommons.utep.edu/open_etd



Part of the [Geophysics and Seismology Commons](#)

Recommended Citation

Thompson, Lennox Emmanuel, "Development Of 3-D Shear Wave Models Using A Multi-Objective Optimization Scheme" (2015). *Open Access Theses & Dissertations*. 1362.

https://digitalcommons.utep.edu/open_etd/1362

This is brought to you for free and open access by DigitalCommons@UTEP. It has been accepted for inclusion in Open Access Theses & Dissertations by an authorized administrator of DigitalCommons@UTEP. For more information, please contact lweber@utep.edu.

DEVELOPMENT OF 3-D SHEAR WAVE MODELS USING A MULTI-OBJECTIVE
OPTIMIZATION SCHEME

LENNOX THOMPSON

Department of Geological Sciences

APPROVED:

Aaron Velasco, Ph.D., Chair

Laura Serpa, Ph.D.

Tom Gill, Ph.D.

Deana Pennington, Ph.D.

Vladik Kreinovich, Ph.D.

Charles Ambler, Ph.D.
Dean of the Graduate School

to my

WIFE and DAUGHTER

with love

DEVELOPMENT OF 3-D SHEAR WAVE MODELS USING A MULTI-OBJECTIVE
OPTIMIZATION SCHEME

by

LENNOX THOMPSON, M.S.

DISSERTATION

Presented to the Faculty of the Graduate School of

The University of Texas at El Paso

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

Department of Geological Sciences

THE UNIVERSITY OF TEXAS AT EL PASO

May 2015

Acknowledgements

I would like to express my deep-felt gratitude to my advisor, Dr. Aaron Velasco of the Geological Science Department at The University of Texas at El Paso, for his advice, encouragement, enduring patience and constant support. He was never ceasing in his belief in me (though I was often doubting in my own abilities), always providing clear explanations when I was (hopelessly) lost, constantly driving me with energy (*Where does he get it?!*) when I was tired, and always, *always* giving me his time, in spite of anything else that was going on. His response to my verbal thanks one day was a very modest, “It’s my job.” I wish all students the honor and opportunity to experience his ability to perform at that job.

I also wish to thank Dr. Vladik Kreinovich, Dr. Rodrigo Romero of the Computer Science Department, both at The University of Texas at El Paso who were extremely helpful in providing the additional guidance and expertise I needed in order to complete this work, especially with regard to the mathematical formulation and 3-D visualizations.

Additionally, I want to thank The University of Texas at El Paso Cybershare center and staff for all their hard work and dedication, providing me the means to complete my degree and prepare for a career as a geophysicist. This includes (but certainly is not limited to) the following individuals:

Ezer Patlan

He made it possible for me to have many wonderful experiences as a student, including the opportunity to work along side my friend for the duration of my Ph.D. studies and assistance with my research.

Dr. Anibal Sosa

His influence, though unbeknownst to him, was one of the main reasons for to want to use an optimization scheme to generate 3-D visualizations of rift systems.

And finally, I must thank my dear wife for putting up with me during the development of this work with continuing, loving support and no complaint and also to my daughter who always bring a smile to my face when I endure the hardships of conducting research for my dissertation.

NOTE: This dissertation was submitted to my Supervising Committee on the March 4, 2015.

Abstract

Evidence of geologic activity still occurring in Rio Grande Rift (RGR) includes quaternary faulting, seismicity, and widening at a small rate. We map the crustal thickness and seismic velocity ratio to create crustal model cross sections that highlights the regional extension of the Southern Rio Grande Rift (SRGR). Specifically, we compute receiver functions and receiver function stacks for 147 USArray and previously collected data, and interpolate the crustal and velocity results using a kriging interpolation scheme. By incorporating gravity, magnetics, receiver functions, velocity models, and the interpretation of seismic reflection/refraction data, we produce a constrained crustal model that characterizes the evolution of the SRGR. In consequence, important questions related to tectonic and lithospheric activity of the Rio Grande Rift remain unresolved. To address some of these geological questions, we developed a 3D crust and upper mantle velocity model using a constrained optimization approach for joint inversion of surface wave and receiver functions. Our 3D models show a thin lower velocity crust anomaly along the southern east portion of the Rio Grande Rift, a persistent low velocity anomaly underneath the Colorado Plateau and Basin and Range province, and another one at depth beneath the Jemez lineament, at the southern RGR.

Furthermore, to make the joint inversion process more robust, we use multi-objective optimization technique: namely, we repeatedly solve the joint inversion problem with different possible combinations of variances. From the mathematical viewpoint, such solutions form a Pareto front of the corresponding multi-objective optimization problem. If a certain geological feature is visible in *all* these solutions, then we can be confident that this feature is also present in the actual solution corresponding to the actual (unknown) values of the variances – i.e., that it is the feature of the actual Earth structure. We applied this more robust optimization technique to the Texas region and we also compared the structure of the upper mantle with other ancient and/or active rift systems within the Texas region.

Our resulting 3-D velocity models will help us to better understand the tectonic history and physical properties of the Earth structure and also determine if an ancient rift system within Texas is reactivated or not.

Table of Contents

	Page
Acknowledgements	iv
Abstract	vi
Table of Contents	viii
Chapter	
1 Geophysical constraints on the crustal structure of the Southern Rio Grande Rift	1
1.1 Introduction	2
1.2 Tectonic Setting	2
1.3 Approach	3
1.3.1 Data	3
1.3.2 Receiver Functions	4
1.4 Kriging Interpolation	8
1.5 Gravity and Magnetic Data	8
1.6 Crustal Models	12
1.7 Discussion and Conclusion	16
1.8 Data Resources	16
2 3D structure of the Rio Grande Rift from 1-D constrained joint inversion of receiver functions and surface wave dispersion	18
2.1 Introduction	19
2.2 Tectonic Setting	20
2.3 Receiver Functions	23
2.4 Surface Wave Dispersion	25
2.5 1-D Constrained Joint Inversion	27
2.6 Joint Inversion Results	31
2.7 Kriging interpolation from 1D velocity profiles	31

2.8	3-D RGR S-Wave Velocity Model	33
2.9	Discussion	39
2.9.1	PDIP methodology	39
2.9.2	Interpretation: Basin and Range, Colorado Plateau, RGR, and Great Plains	41
2.9.3	The Jemez Upwelling and the Southern RGR Anomaly	42
2.10	Conclusions	43
3	Construction of shear wave models by applying multi-objective optimization to multiple geophysical datasets	45
3.1	Introduction	46
3.2	Receiver Functions	48
3.3	Receiver function stacking	49
3.4	Surface Wave Dispersion	51
3.5	Delay Travel Times	54
3.6	Gravity Anomalies	55
3.7	Forward Problem	58
3.8	Inverse Problem	58
3.9	Need for multi-objective optimization	59
3.10	Numerical Algorithm	62
3.11	Primal Dual Interior-Point Method	63
3.12	Conclusion	64
4	A Constrained Multi-Objective Optimization Framework for Joint Inversion . .	67
4.1	Introduction	68
4.2	Need for Multi-Objective Optimization	70
4.2.1	Inverse Problems: Brief Reminder	70
4.2.2	Need to Take Constraints into Account	71
4.2.3	Joint Inversion: Idealized Case	72
4.2.4	Reformulation of the Problem	73

4.2.5	General Case: A Description	74
4.2.6	General Case: Analysis of the Problem	75
4.2.7	General Case: Resulting Optimization Problem	77
4.2.8	Relation to Multi-Objective Optimization	78
4.2.9	How to Generate a “Typical” Solution	80
4.3	Solving the Corresponding Constraint Optimization Problems	80
4.3.1	Linearization	82
4.3.2	Primal Dual Interior-Point Method	83
4.4	Geophysical Datasets	85
4.4.1	Receiver Functions	85
4.4.2	Surface Wave Dispersion	87
4.4.3	Delay Travel Times	89
4.5	Results and Discussion	91
4.6	Conclusion	100
5	Generation of 3-D shear wave based models using 1-D constrained multiobjective optimization	102
5.1	Introduction	103
5.2	Tectonic Setting	104
5.3	Receiver Functions	107
5.4	Surface Wave Dispersion	112
5.5	Joint Inversion	114
5.6	Joint Inversion Results	116
5.7	Kriging interpolation from 1D velocity profiles	116
5.8	3-D Texas S-Wave Velocity Model	118
5.8.1	MOP methodology	132
5.8.2	Interpretation	135
5.9	Discussion	136
5.10	Conclusions	137

References 138

Curriculum Vitae 156

Chapter 1

Geophysical constraints on the crustal structure of the Southern Rio Grande Rift

Evidence of geologic activity still occurring in Rio Grande Rift (RGR) includes quaternary faulting, seismicity, and widening at a small rate. We map the crustal thickness and seismic velocity ratio to create crustal model cross sections that highlights the regional extension of the Southern Rio Grande Rift (SRGR). Specifically, we compute receiver functions and receiver function stacks for 147 USArray and previously collected data, and interpolate the crustal and velocity results using a kriging interpolation scheme. By incorporating gravity, magnetics, receiver functions, velocity models, and the interpretation of seismic reflection/refraction data, we produce a constrained crustal model that characterizes the evolution of the SRGR. Our constrained crustal models show that there is an uplifted lower and upper crust, and shallower Moho in the region of SRGR. This chapter is based on the foundation of the following manuscript:

Thompson, L., A. A. Velasco, and M. Hussein, 2015, Geophysical constraints on the crustal structure of the Southern Rio Grande Rift, *Bulletin of the Seismological Society of America*, (in review).

1.1 Introduction

The Rio Grande Rift (RGR), one of the major continental rift zones, extends ~ 1000 km from central (Leadville) Colorado to south of El Paso, Texas (Figure 1). The southern end of the rift is poorly defined in the geologic record of the rift system (Keller and Baldrige, 1999) and there have been few seismic studies that provide information on the deeper rift structure ([55]; [5]; [34]; [121]; [80]). The southernmost portion of the RGR has recently experienced crustal extension as high as ~ 0.5 mm/yr and appears to be extending more towards the west of the RGR ([144]; [58]; [12]). Geodetic measurements near Socorro, New Mexico failed to detect any significant extension rates across the RGR ([109]; [58]). However, the rift does appear to be widening at a rate of ~ 0.14 mm/yr from 5 Ma to the present ([144]; [58]; [12]). Two key questions remain about the rift system: 1) is it actively deforming along its southern extent ([12]; [53]; [81]); and 2) how does the rift influence the evolution of adjacent areas within the North American Plate ([100])? To address these questions, we define the crustal and upper mantle structure of the southern RGR using data from the EarthScope Transportable Array (USArray), permanent stations in the region, and the LA RISTRA ([143]) experiment to construct an image of the crust-mantle boundary in the southern Rio Grande Rift (SRGR). In particular, we compute and stack teleseismic receiver functions for 147 stations, which allows for crustal thickness and seismic velocities (V_p/V_s) to be determined. We use the receiver function results to model gravity and magnetic data to create crustal models of the SRGR. The models show crustal thinning along the rift axis, suggesting an uplifted lower and upper crust.

1.2 Tectonic Setting

The RGR, a major continental rift, initiated during the late Cenozoic and separates the Proterozoic continental lithospheres of the Great Plains and the Colorado Plateau ([53]; [120]; [82]; [101]; [9]). Rifting and extension began in the late Oligocene or early Miocene

([26]). Two extensional episodes led to the formation of the RGR. An initial stage of extension began at 30-20 Ma, with an association of low-angle faulting and crustal doming. The second phase of extension occurred (3-10 Ma), which involved 10% of extension trending in the E-W direction ([55]; [139]). Wilson (2003) hypothesized that this extension resulted from an upper mantle asthenosphere upwelling and thermal lithosphere erosion. Extension along the western interior portion of North America stimulated the formation of the RGR while leaving the Colorado Plateau from being deformed during this regional extension (e.g., [68]; [139]). The southern RGR has experienced more regional extension than the northern portion of the rift ([53]). Associated volcanism occurred in the western and central parts of the rift, being prominent during the Pliocene and Quaternary ([26]). Work from the Colorado Plateau-Rio Grande Rift-Great Plains Seismic Transect (LA RISTRA) experiment has shown that the central section of the RGR has a low velocity zone ([143]; [136]; [100]; [34]), suggesting that there could be molten material or that the crust is thinning beneath the center of the RGR. The zone of crustal thinning widens southward as does the physiographic expression of the rift. In southern New Mexico, the rift is assumed to have the thinnest crust (less than 30 km) with very high heat flow ([52]). Recent work by [121], shows no deep mantle velocities beyond 250 km, and the crustal thinning is likely a shallower process.

1.3 Approach

1.3.1 Data

We collected three-component seismic data for 147 stations within the area of latitudes between 29° to 36°N and longitudes between -111° to -102°E (Figure 1) from the LA RISTRA temporary seismic experiment, EarthScope transportable array (USArray), United States Geological Survey ANSS backbone stations, and the International Monitoring System (IMS) network stations in the region. We obtain seismograms from the Incorporated

Research Institutes for Seismology (IRIS) Data Management Center (DMC) for teleseismic events from Jan, 2000 to Dec, 2009 that have epicentral distances ranging from 30° - 95° . This includes 1,464 seismic events with a minimum magnitude of 5.5 and 86,388 continuous waveforms. USArray Transportable Array (TA), an element of EarthScope, places a rolling, dense network of over 400 permanent and portable seismographs on a nearly uniform 70 km grid across the U.S. Each station of the TA network collects data for two years before it is re-deployed to a new location ([80]). The USArray was in the region from approximately 2007 to 2010. The nominal spacing (~ 70 km) between USArray stations allows for both lower crustal and upper mantle seismic studies ([112]). The LA RISTRA temporary seismic experiment recorded data for 18 months between August 1999 and May 2001 ([9]; [143]).

1.3.2 Receiver Functions

A receiver function maps the seismic response of the Earth beneath a seismic station to an incoming P-wave by deconvolving the vertical component of a teleseismic earthquake seismogram from the radial component ([60]). The result allows for the identification of converted phases derived from strong impedance contrasts (e.g., the crustal-mantle boundary). The receiver function technique takes advantage of converted seismic P waves into S waves at discontinuities along the ray path ([29]; [4]; [138]), and has been utilized in many studies (e.g., [143]; [143]; [36]; [42]; [59]; [6]; [64]; [112]). For data collection and processing, we use the Standing Order for Data (SOD) to request three component seismograms for P-wave arrivals and for events with a minimum magnitude 5.5, depth in the range of 1600 km and an epicentral distance ranging from 30° - 95° (e.g., [87]; [6]).

We employ the receiver function technique using the iterative deconvolution method of [65] and the stacking approach described in [151]. In receiver function estimation, the foundation of the iterative deconvolution approach is a least squares minimization of the difference between the observed horizontal component seismogram and predicted signal generated by convolution of an iterative updated spike train with the vertical component

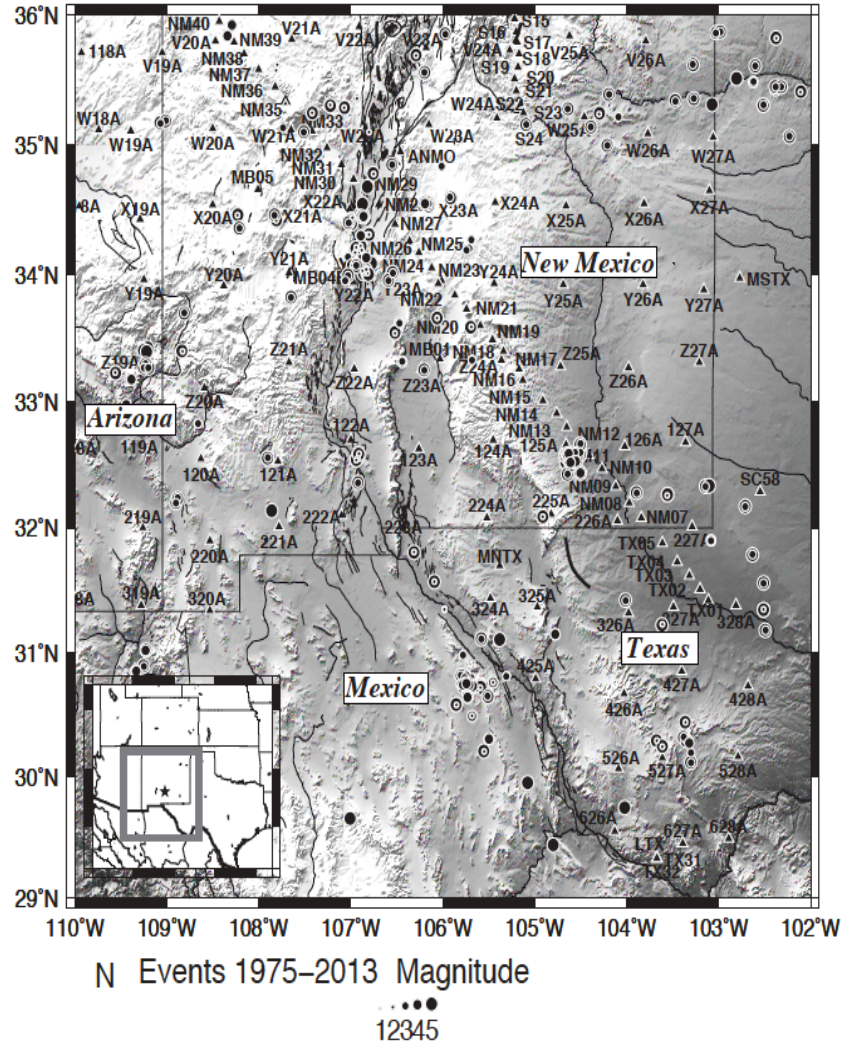


Figure 1.1: Regional topography of the SRGR and sketch of SRGR boundaries. Above the contour plot is a color scale of the elevations in meters. The black triangles signify the stations that we have data for.

seismogram ([65]). We compute receiver functions using the iterative time deconvolution with Gaussian width (Ga) factors of 2.5, 1.75, and 1, which is equivalent to applying low pass filters with cutoff frequencies of 1.2, 0.9, and 0.5 Hz, respectively.

We use the receiver function stacking technique introduced by [151], which estimates the crustal thickness and V_p/V_s ratio based on the radial receiver function. Assuming that no lateral velocity heterogeneities exist, the time separation between the Ps converted wave and the direct P-wave obtained from receiver functions (tPs) can then be used to estimate crustal thickness (H), given the average crustal velocities V_p and a V_p/V_s ratio (κ), and the constant ray parameter p of the incident wave (e.g., [38], [71]). Through the utilization and stacking of multiple events, we consequently increased the signal-to-noise ratio (SNR). This increase in the SNR could be caused by background noise, scattering from crustal heterogeneities, and P-to-S multiple conversions from other velocity discontinuities ([70]). Each of the phases are weighted using the $H - \kappa$ stacking technique and stacked phases are then plotted as a gridded image $s(H, \kappa)$, which reaches a maxima when all three phases (tPs , $tPpPs$, $tPsPs + PpSs$) are stacked coherently along with the correct H and κ ([151]; [103]). Stacking not only enhances the SNR, but also suppresses lateral variations in the vicinity of the recording station ([103]). An example of this technique is shown in Figure 2 for one of the Earthscope USArray stations, 219A.

We manually review each receiver function stack for quality, making sure the maximum solutions do not fall on the bounds of our grid search. This process is repeated until we have a stack that appears robust. If we have less than 10 receiver functions in the stack, we remove the station from our interpolation, described below. We compared our findings with the EarthScope Automated Receiver Survey (EARS) website results, and found inconsistencies that are likely the result of differing quality control parameters and loss of high frequency information ([112]; [143]).

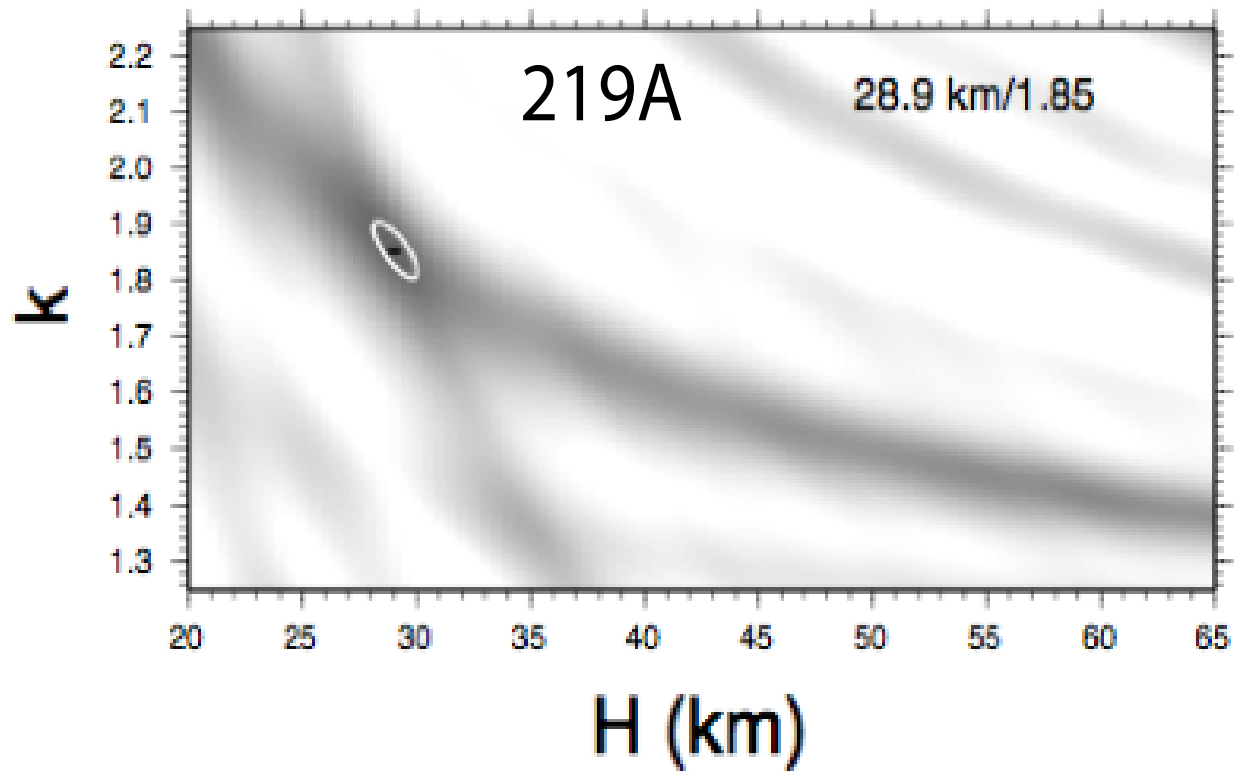


Figure 1.2: Receiver function stack of station 219A, V_p/V_s vs H (km). The black dot with the white circle around it represents the preferred value. Note the multiple shaded regions might result in a poor choice of crustal thickness.

1.4 Kriging Interpolation

Kriging is an optimal interpolation scheme based on regression against observed values of surrounding data points and weighted according to spatial covariance values (e.g., [113]). Kriging interpolation estimates values at a given location as a weighted sum of data values at surrounding locations just like common interpolation algorithms such as inverse distance squared, splines, radial basis functions, triangulation, etc. ([39]). We use a Bayesian kriging interpolation scheme in order to incorporate error associated with modeling to better constraint the mapping of V_p/V_s and crustal thickness ([113]). For this analysis, we krig values of crustal thickness and V_p/V_s using the error function blending functions with 2° (~ 200 km) of influence to guarantee adequate spatial sampling.

Figures 3 and 4 show contour plots of crustal thickness (Moho depth) and V_p/V_s of SRGR from the receiver function stacks using 147 stations from the USArray data set, respectively. Figure 3 shows the crustal thickness in kilometers with black to white colors for thin to thicker crust. In Figure 4, black to white colors represent lower to higher V_p/V_s ratios. As can be seen, the V_p/V_s values or crustal thickness values between the stations do not vary as much as from the EARS results, which suggests that our strict quality restrictions work well for our data set. Figure 4 shows areas with high V_p/V_s and may represent areas of the rift system that are experiencing partial melt, while the crustal thickness of the rift system (Figure 3) appears to be consistent with previous results ([52]).

1.5 Gravity and Magnetic Data

We obtained gravity data from the University of Texas at El Paso (UTEP)-Pan American Center of Earth and Environmental Studies-(PACES) that is currently hosted at the Cyber-ShARE Center of Excellence at UTEP. The gravity data were merged from a variety of surveys and cover the U.S. and the border region (Figure 5). We obtain aeromagnetic data from the U.S. Geological Survey with a grid spacing of 1 km ([8]) (Figure 6). Average error

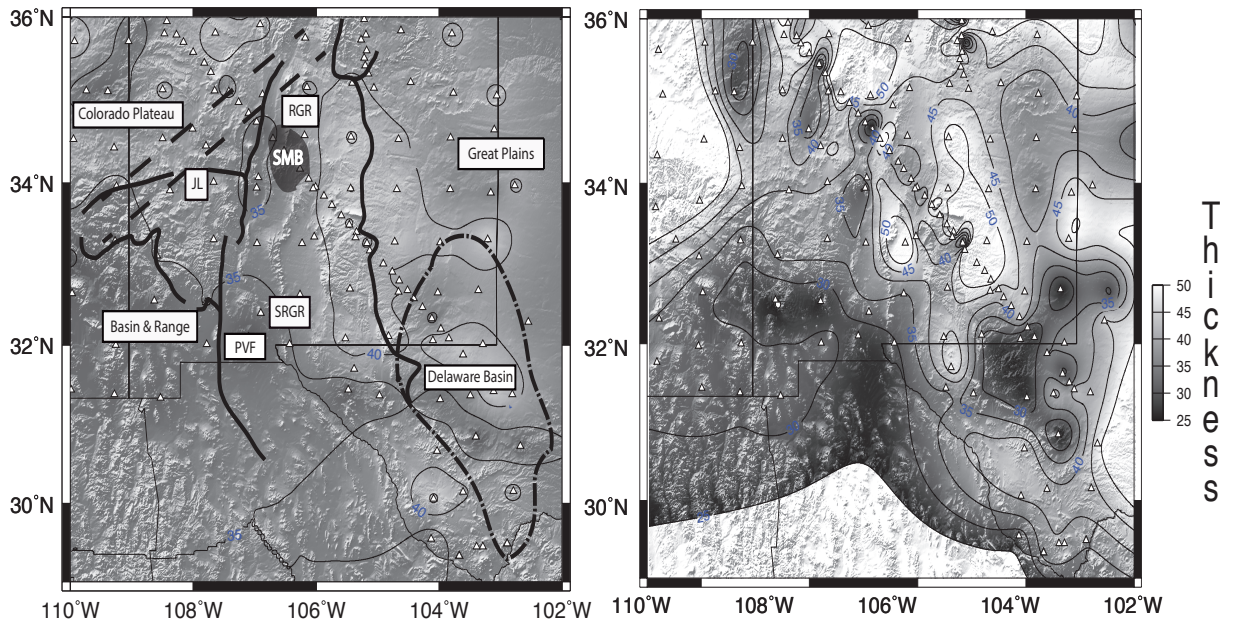


Figure 1.3: Contour plot with the kriging interpolation scheme for crustal thickness of SRGR using USArray and LARISTRA data derived from receiver function stacks. From black to white grayscale represents thin to thicker crust in some areas. The triangles represent the stations used for this analysis. The black dashed lines represents the approximate outline of where we believe the SRGR is deforming based on our results.

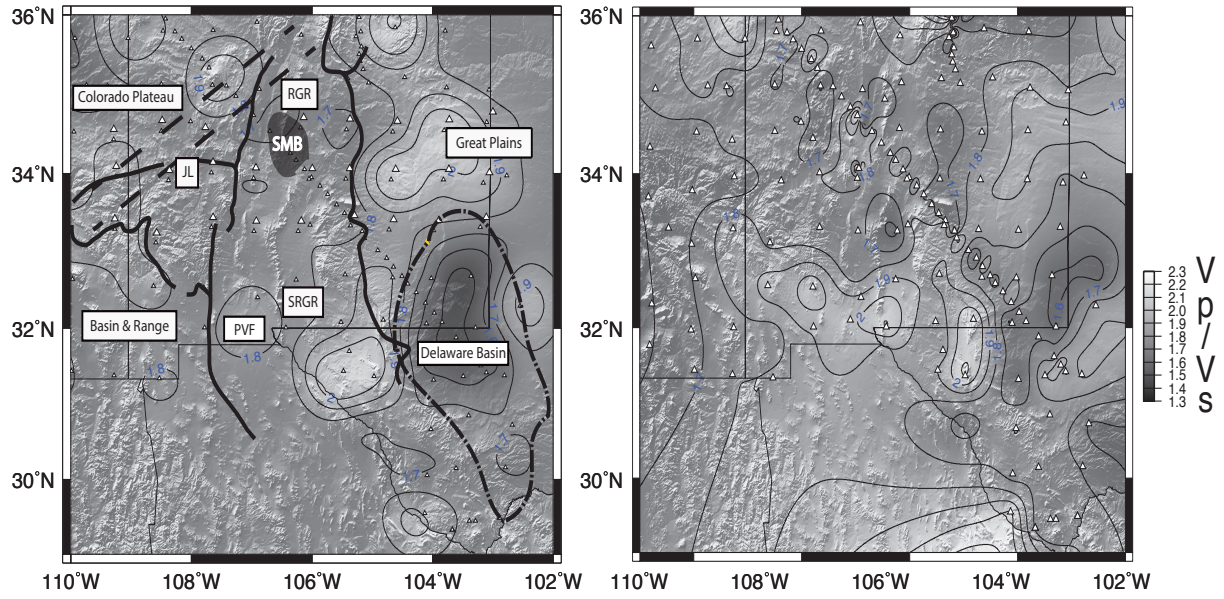


Figure 1.4: Contour plot with the kriging interpolation scheme for V_p/V_s structure of SRGR using USArray and LARISTRA data derived from receiver function stacks. From black to white grayscale represents higher V_p/V_s ratio in some areas. The triangles represent the stations used for this analysis. The black dashed lines represents the approximate outline of where we believe the SRGR is deforming based on our results.

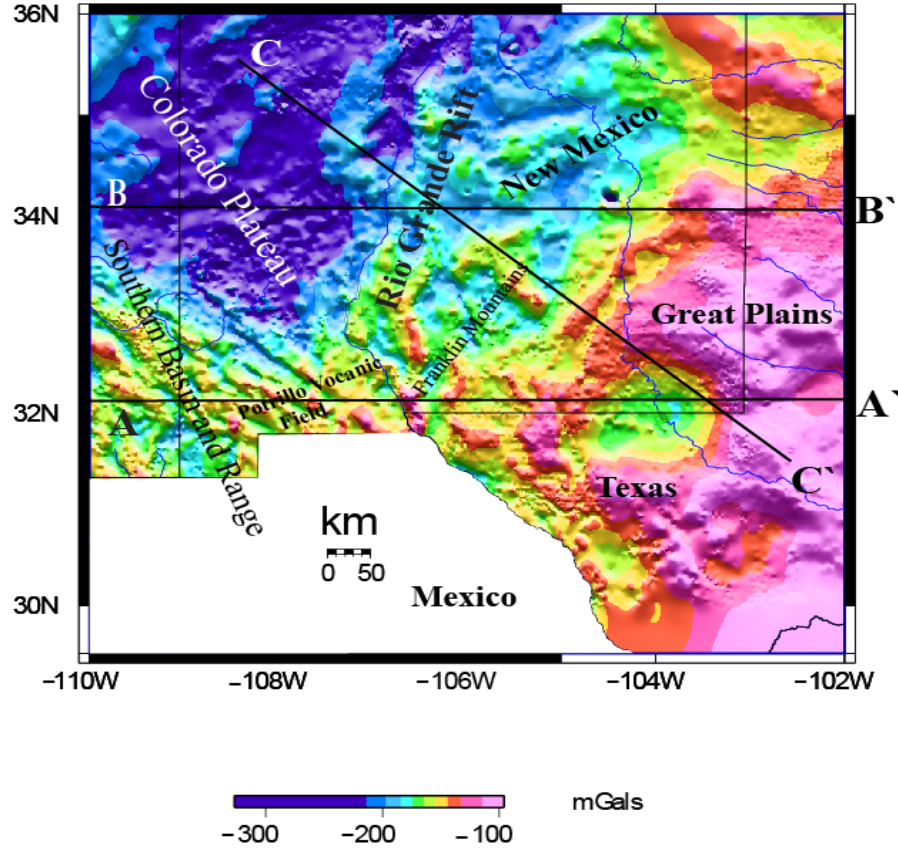


Figure 1.5: Bouguer gravity anomaly map of the Rio Grande and surrounding area. High amplitude gravity anomaly observed in SRGS.

for this data set ranges from 0.05 to 2mGal. Terrain corrections were calculated by [134] using a digital elevation model and a technique based on the approach of [91]. A Bouguer gravity correction was made using 2670 kg/m³ as the reduction density. We used 59,863 Bouguer gravity points to create a Bouguer gravity anomaly map (Figure 5).

The Bouguer gravity values decrease from (−250 mGal) in the Colorado Plateau to (−80 mGal) in the Franklin Mountains, Portillo Volcanic Field (PFV), and the Great plains (Figure 5), gravity values range from −60 to −50 mGal in Delaware basin. Low gravity anomalies are likely caused by unconsolidated sediments in the basins, metasedimentary

and granitic rocks, or by thicker crust or some combination of these features. One feature of interest is the Delaware basin that is located at the western sub basin of the Permian basin, which developed as a cratonic depression in response to Early Pennsylvanian crustal shortening in the Marathon fold and thrust belt ([48];[106]). High gravity associates with thin crust (shallow Moho), near El Paso, and in the central Rio Grande basin.

We filtered the grid by using a low pass, Gaussian, and strike. The purpose of filtering the grid is to shift the anomalies into their actual spatial position and making sure they are not shifted or offset from their location to create a reduce-to-pole magnetic anomaly map (Fig 6) magnetic susceptibility of rocks varies much more than density, magnetic maps appears more complicated than gravity maps ([52]).

1.6 Crustal Models

Using the crustal thickness from the receiver functions, we construct three crustal scale models (Figures 7). We selected these profiles to illustrate the differences along the SRGR and the central RGR, plus compare with previous results along LA RISTRA. We constructed models using 2.5-D modeling of gravity data by a program developed by [126] and further revised by [90] and [19]. Gravity values were extracted from the grid at a 6 km interval. These values were then used as input for the 2.5-D forward modeling program. We determined the depth to the Moho from receiver functions, and the densities for the upper and lower crust and upper mantle were inferred from previous studies ([2],[114], [143], and [5]). We assign a density of 2600 kg/m^3 for the most upper part of the models, although we realize that the density may be as low as 2360 kg/m^3 or as high as 2680 kg/m^3 . It is difficult to place boundaries to separate the different densities of the surface material because of the absence of strong evidence of the exact locations of these boundaries; additionally surface materials are mixed due to transportation or weathering conditions.

Model A-A' (Figure 7) is $\sim 750 \text{ km}$ long and is taken with latitude 32° and covers the southern region of the study area (southern RGR). The depth to the Moho is about 32 km

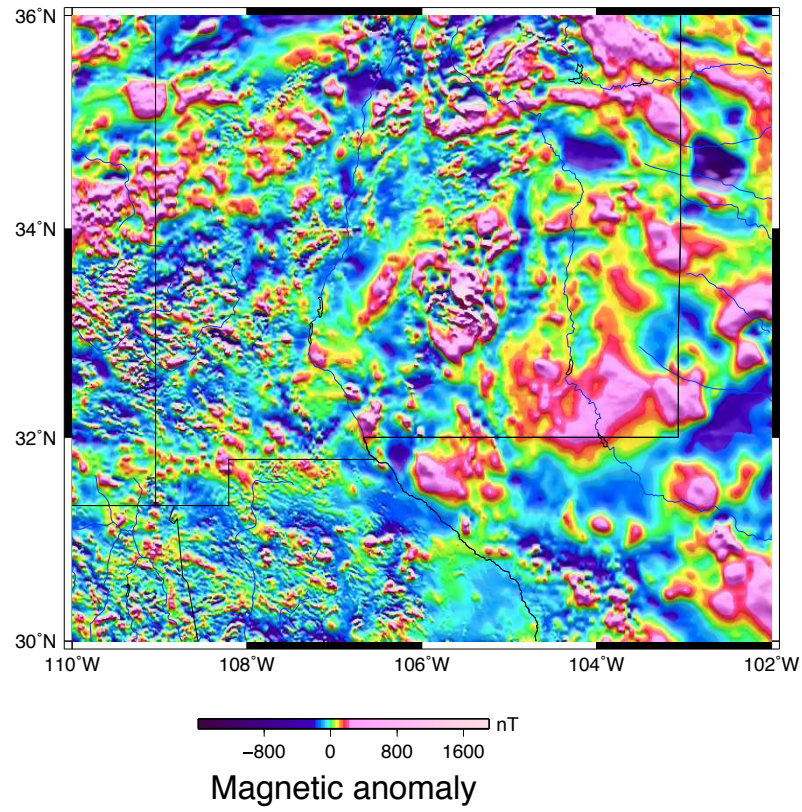


Figure 1.6: Reduced-to-pole magnetic anomaly map. Magnetic anomalies mostly represent shallower structures.

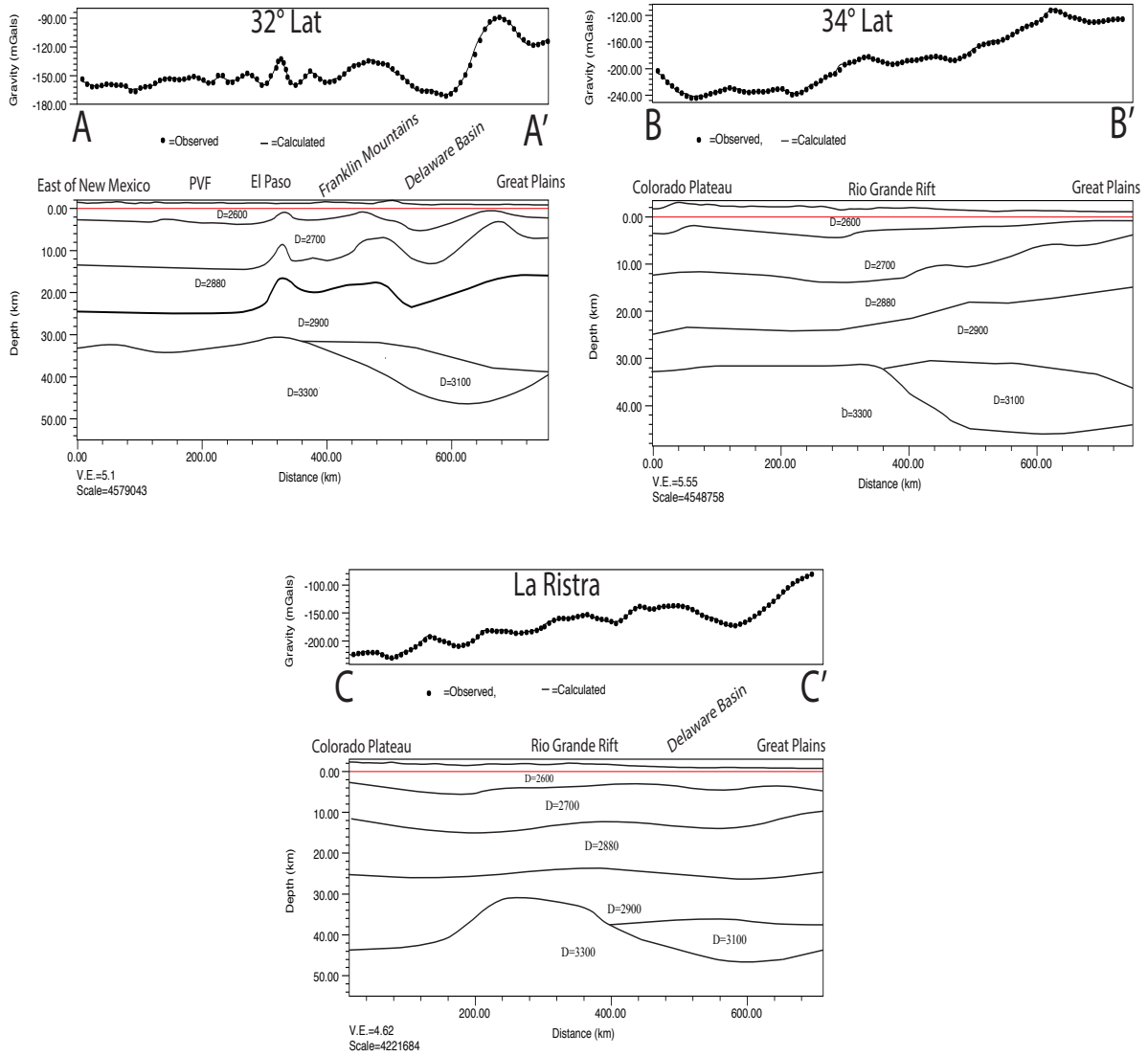


Figure 1.7: Model A-A' is ~ 750 km long coincide with latitude 32° and covers the southern region of the study area. Crustal scale model about 550 km and starts from west New Mexico through SRGR, showing up-lifted crust as an evidence for geothermal activity that causes heating, stretching and rifting. Model B-B' is ~ 740 km long, coincide with latitude 34° , and passes through Colorado Plateau, Socorro magma body and ends at the Great Plains. Model C-C' is ~ 700 km long coincide with Rio Grande Rift Seismic Transect (RISTRA) and cross cut models A-A' and B-B', starts at Colorado Plateau and Delaware basin.

at the starting point (A) of the model, decreases to 30 km in rift basin west of El Paso and deepens to about 45 km under the east flank of the rift and beneath the Delaware basin. Model A-A' shows slightly uplifted Moho, lower and middle crust. The flat-topped shape beneath the area of active upper crust extension suggest that it is primarily the product of magmatic activity in the lower crust and upper mantle ([115]). Beneath El Paso area, upper mantle velocities range from 7.75 to 7.9 km/s, which is consistent with a warm upper mantle and high heat flow values 85 to 125 mWm⁻² associated with the SRGR ([5]).

Model B-B' (Figure 7), approximately 740 km long, is taken at latitude 34°, and passes through Colorado Plateau, Socorro magma body and ends at the Great Plains (central RGR). The depth to the Moho is 34 km at the starting point (B) of the model, decreases to 32 km in the center and deepens significantly to ~48 km under the east flank of the rift. Another explanation for the uplifted crust is Laramide orogeny high velocity bodies and lateral velocity variation at 5 to 10 km depth beneath Potrillo Volcanic Field (PVF) region which show the Laramide basement uplift ([5]). Interferometric synthetic aperture radar (In-SAR) imaging of the central Rio Grande rift (New Mexico, USA) during 1992-1999 reveals a crustal uplift of several centimeters that spatially coincides with the seismologically determined outline of the Socorro magma body [31]. [15] study of the amplitude and frequency content of the COCORP (Consortium for Continental Reflection Profiling) reflections indicates a partially multilayered intrusion. SMB do not reflect any distinguish signature on the Bouguer gravity anomaly map because the inferred location of SMB is dominated by gravity low, and the SMB is deep structure (18-22) km.

Model C-C' (Figure 7), approximately 700 km long, coincides with Rio Grande Rift Seismic Transect (RISTRA) and cross cut models A-A' and B-B', starting at Colorado Plateau and ending at Delaware basin. The depth to the Moho is ~44 km at the starting point (A) of the model, decreases to 30 km in rift basin and deepens to ~47 km under the east flank of the rift and dome in shape. Model C-C' shows a dome shape Moho, the Moho depth is 45 km beneath Colorado Plateau, decreases to 30 km under the Rio Grande rift and deepened to 47 km in the Great Plains.

1.7 Discussion and Conclusion

Our crustal models suggest an uplifted lower and upper crust, and shallower Moho in the region of SRGR; uplifted crust and shallow dome in shape Moho are suggested to be primarily the product of magmatic activity in the lower crust and upper mantle. For model A-A' (Figure 7), our results reveal that the depth to the Moho is about 32 km at the starting point (A) of the model, decreases to 30 km in rift basin west of El Paso, which shows that the southern RGR is still actively deforming. In the case of model B-B' (Figure 7) the depth to the Moho is 34 km at the starting point (B) of the model, decreases to 32 km in the center and deepens significantly to ~ 48 km under the east flank of the rift. Another explanation for the uplifted crust is Laramide orogeny high velocity bodies and lateral velocity variation at 5 to 10 km depth beneath Potrillo Volcanic Field (PVF) region, which show that the Laramide basement uplift and would explain how the RGR influences adjacent provinces within North American plate ([5]). Our receiver function stacks show results that are consistent with rifting in the SRGR. In particular, Figure 3 and 4 show evidence of crustal thinning in the center of the rift, and also a higher V_p/V_s ratio that may signify partial melt and/or higher temperatures in the crust.

Our results have similar characteristics from previous studies in the RGR, but they show an uplifted lower and upper crust, and shallower Moho in the region of SRGR; uplifted crust could primarily be the product of magmatic activity in the lower crust and upper mantle, this is beyond the scope of this research. [121] conclude that the rift is not driven by a deep mantle upwelling, suggesting that RGR rift formation is caused an upper mantle and crustal process.

1.8 Data Resources

Seismic data used in this study can be acquired through the (IRIS) Data Management Center (DMC). We also used seismic data from the EarthScope Transportable Array (US-

Array) (www.earthscope.org) (last accessed June 2013). The EarthScope Automated Receiver Survey (EARS) website results can be found through (<http://www.seis.sc.edu/ears/>) (last accessed January 2012). Gravity data can be accessed through the University of Texas at El Paso (UTEP)-Pan American Center of Earth and Environmental Studies-(PACES) (<http://www.research.utep.edu/paces>) (last accessed July 2014) that is currently hosted at the Cyber-ShARE Center of Excellence at UTEP. Aeromagnetic data used in this study is from the U.S. Geological Survey with a grid spacing of 1 km ([8]).

Chapter 2

3D structure of the Rio Grande Rift from 1-D constrained joint inversion of receiver functions and surface wave dispersion

The Southern terminus of the Rio Grande Rift region has been poorly defined in the geologic record, with few seismic studies that provides information on the deeper Rift structure. In consequence, important questions related to tectonic and lithospheric activity of the Rio Grande Rift remain unresolved. To address some of these geological questions, we collect and analyze seismic data from 147 EarthScope Transportable Array (USArray) and other seismic stations in the region, to develop a 3D crust and upper mantle velocity model. We apply a constrained optimization approach for joint inversion of surface wave and receiver functions using seismic S wave velocities as a model parameter. In particular, we compute receiver functions stacks based on ray parameter, and invert them jointly with collected surface wave group velocity dispersion observations. The inversions estimate 1D seismic S-wave velocity profiles to 300 km depth, which are then interpolated to a 3D velocity model using a Bayesian kriging scheme. Our 3D models show a thin lower velocity crust anomaly along the southern east portion of the Rio Grande Rift, a persistent low velocity anomaly underneath the Colorado Plateau and Basin and Range province, and another one at depth beneath the Jemez lineament, and the southern RGR. Chapter 2 is based on the following publication:

Sosa, A., L. Thompson, A. A. Velasco, R. Romero, and R. Hermann, 2013, 3-D structure of the Southern Rio Grande Rift from 1-D constrained joint inversion of receiver functions and surface wave dispersion, *Earth and Planetary Science Letters*, 402, 127-137.doi:10.1016/j.epsl.2014.06.002

2.1 Introduction

Bounded by the Basin and Range province and the Colorado Plateau to the west and the Great Plains to the East, the Rio Grande Rift (RGR) extends approximately 1000 km from central Colorado to El Paso, Texas (Figure 1). The RGR has recent volcanism, fault scarps, and seismicity (Figure 2) and is widening at a modest rate of about 0.5 mm/yr or less ([12], [58], [144]). Although many studies have focused on the Rift system (e.g., [5], [34], [55], [100], [136], [143]), the southern terminus of the RGR remains poorly defined ([53]) and few seismic studies have provided detailed information on the deeper rift structure ([5], [34], [55]).

Studies of the RGR present several possible Earth models and interpretations, which may be due to the diversity of the methodologies implemented and the specific location of the study ([34], [81], [136]). For example, [52] presented a Moho depth map, based on a compilation of previous studies, and showed significant crustal thinning (~ 28 km) in the southern RGR. [5] conducted a controlled source experiment and developed a detailed profile across the southern RGR, showing a crustal thickness of around 32 km. [34] proposed that small convection cells in the deeper mantle are responsible for recent magmatic and tectonic activity, while [21] propose that rotation of the Colorado Plateau played a role in rift formation.

More recent studies in the region have taken full advantage of data collected by EarthScopes Transportable Array (USArray) to develop models that are derived from a variety of approaches: seismic tomography ([10], [11], [17], [16], [18], [80], [86], [88], [118], [123], [127], [145], [147], [149], [148]), receiver functions ([1], [20], [35], [42], [64], [136], [142], [76]),

joint inversions ([6], [66], [81], [136]). Many models show the main tectonic regions in North America, but do not necessarily focus on rift formation.

Besides the availability of new data and recent studies on the RGR, important questions about the Rift evolution remain unresolved: 1) is it actively deforming along its southern extent ([12], [53], [81])? 2) does it propagate southward?; 3) what is the role of mantle convection in the formation of the Rift? 4) does partial melt and unstable lithosphere composition affect the rifts evolution ([34], [136], [143])? and 5) how does it influence the evolution of adjacent areas within the North American Plate ([100])?

In this chapter, we revisit the driving forces that cause RGR formation using high quality available data and applying a robust inversion/imaging method for an integrated analysis of Earth structure that allows us to create three-dimensional (3D) velocity models. Specifically, we determine the crustal and upper mantle structure of the RGR using the USArray data (www.earthscope.org) along with other data sets. We apply a constrained optimization 1D joint inversion approach ([121]) using receiver functions for 147 USArray and the LA RISTRA (Colorado Plateau/Rio Grande Rift Seismic Transect Experiment ([34], [136]) stations ([143]; [125]), and a high quality surface wave dispersion data set provided ([47]), and then interpolate the results to obtain a 3-D shear velocity model of the region. We find little evidence of deep mantle upwelling driving rifting.

2.2 Tectonic Setting

The RGR, a major continental rift, formed in the late Oligocene or early Miocene ([26]) and separates the Proterozoic continental lithospheres of the western Great Plains and the Colorado Plateau ([9], [53], [82], [101], [120]). An initial stage of extension began at 30-20 Ma, with low-angle faulting and crustal doming. The second phase (3-10 Ma) involved 10% of extension trending in the E-W direction ([54], [139]). [139] hypothesized that this resulted from an upper mantle asthenosphere upwelling and thermal lithosphere erosion. Extension along the western interior portion of North America stimulated the formation

of the RGR, while leaving the Colorado Plateau from being deformed during this period (e.g., [68], [139]). [53] highlight that the southern RGR has experienced more extension than the northern section of the rift, while [58] did not find significant extension across most of the RGR but the southernmost part (as $\sim 0.5 \text{ mm yr}^{-1}$). [12] also show that the extensional deformation is not concentrated in a narrow zone centered on the Rio Grande Rift but rather is distributed broadly from the western edge of the Colorado Plateau well into the western Great Plains.

The Rio Grande rift broadens and changes strike at the US-Mexico border (Figure 2) and has evidence of volcanism and seismicity. Volcanism was prominent in the rift area during the Pliocene and Quaternary ([26]), and also during the Cenozoic especially along the trend called Jemez lineament to the west of the rift ([34]). It is associated to large negative gradients in lithospheric thickness on both sides of the rift ([64]). Although the RGR is not as seismically active as other parts of the North American plate margin, the faults in the region show Quaternary offsets ([24]). The RGR does appear to remain more seismically active than the adjacent Basin and Range Province ([73]) suggesting the two regions may be responding to extensional processes differently. Additionally, the extension directions during rifting appear to have rotated in a clockwise sense since its inception at $\sim 30 \text{ Ma}$ ([54] and references therein), but the causes of this rotation remain unknown.

The LA RISTRA passive experiment results have shown that the center of the RGR has a low velocity zone in the upper mantle ([34], [136], [143]), suggesting that there could be melt material or that the crust is thinning beneath the center of the RGR. The zone of crustal thinning widens southward as does the physiographic expression of the rift. In southern New Mexico, the RGR seems to have experienced more deformation from a geophysical perspective, creating the thinnest crust (less than 30 km) with very high heat flow ([52]). [100] integrated seismic velocities, gravity and xenolith data, to explore temperature and compositional variations together with partial melt content beneath the eastern Colorado Plateau and RGR. They interpreted the results of that study as a product of modified and/or thinned lithosphere. Furthermore, [100] argue that the RGR and southeastern

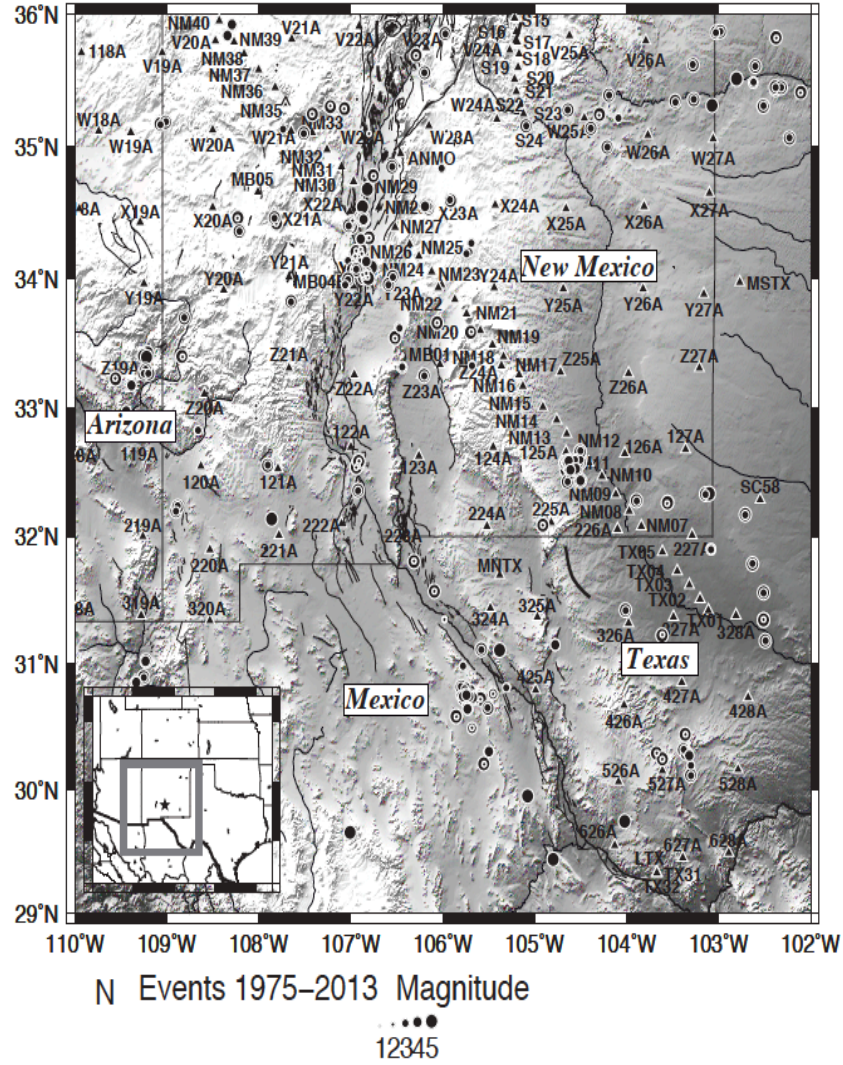


Figure 2.1: Regional topography of the SRGR and sketch of SRGR boundaries. The white triangles signify the stations that we have data for.

Colorado Plateau was underlayed by a low-density upper mantle province, which does not trend along upper crustal tectonic boundaries, but correlates with regions of late Tertiary magmatism. Recently, [6] jointly inverted seismic data constrained by gravity anomalies concluding that the low velocity mantle beneath the RGR indicates that some removal of the lithosphere has occurred.

2.3 Receiver Functions

A receiver function maps the seismic response of the earth beneath a seismic station to an incoming, teleseismic P wave. Deconvolving the vertical component of a teleseismic earthquake seismogram from the radial component (e.g., [60]) results in a receiver function, which then allows for the identification of converted phases corresponding to strong impedance contrasts (e.g., the crustal/mantle boundary). Since we use teleseismic events that arrive to the stations with near-vertical incidence, receiver functions can also be used for imaging deep structure ([36], [42], [59], [64], [112]). Receiver functions can provide valuable information for investigating magma lenses within the crust, determining the Moho depth, other upper-mantle discontinuities ([70]), structure and evolution of the crust ([9]), and rifting extension and magmatism (e.g., [28]).

For this study, we collect three-component seismic data for 147 stations within the area of latitudes between 29° to 36°N and longitudes between -111° to -102°E (Figure 1) from the LA RISTRA portable seismic experiment, EarthScope transportable array (USArray), United States Geological Survey ANSS backbone stations, and the International Monitoring System (IMS) network stations in the region ([125]). The nominal spacing (~ 70 km) between USArray stations allows for both lower crustal and upper mantle seismic studies ([112]). The LA RISTRA experiment recorded data for a year and a half beginning in August 1999. [125] compared their findings with the EarthScope Automated Receiver Survey (EARS) website results (<http://www.seis.sc.edu/ears/>), and found inconsistencies that are likely the result of differing quality control parameters and loss of high frequency

information ([112]; [143]). We utilize the receiver function data set provided by [125], which includes 434 receiver functions stacked in ray parameter bins derived from 1464 teleseismic seismic events with a minimum moment magnitude of 5.5 and occurring from January, 2000 to December, 2009. This data set focuses in the southern Rio Grande Rift, although the area of the imaged region is much larger ([125]).

2.4 Surface Wave Dispersion

In general, surface waves dominate seismograms as the largest amplitude waves from an earthquake and have observed lower frequencies than body waves. Furthermore, surface wave velocities vary depending on the depth sampled by each period, resulting in dispersion. Measuring dispersion of surface waves provides valuable information for studying Earth's crustal and mantle velocity structure ([84], [116], [124]). In particular, Love and Rayleigh wave group dispersion observations generally account for average velocity structure as a function of depth ([51], [72], [116], [124]).

As part of the systematic determination of earthquake moment tensors for North American earthquakes, Saint Louis University measures fundamental mode Love and Rayleigh wave spectral amplitudes and group velocities using a multiple filter analysis from local to regional earthquakes. Tomography is then used to obtain tomographic images of group velocity dispersion for North America with emphasis on the continental United States (Ammon, personal communication; [22]). As of December 2013, there were over 2,020,514 dispersion measurements available for use. Furthermore, the dataset contains dispersion measurements from regional earthquakes, which allows for measurements at shorter periods (less than 15 s) and more sensitivity to shallower Earth structure (upper mantle and crust). We calculate the dispersion curves at each station using the tomographic dispersion model ([47]) for our analysis of the RGR.

To demonstrate the stability of the measurements, we interpolate the surface wave data

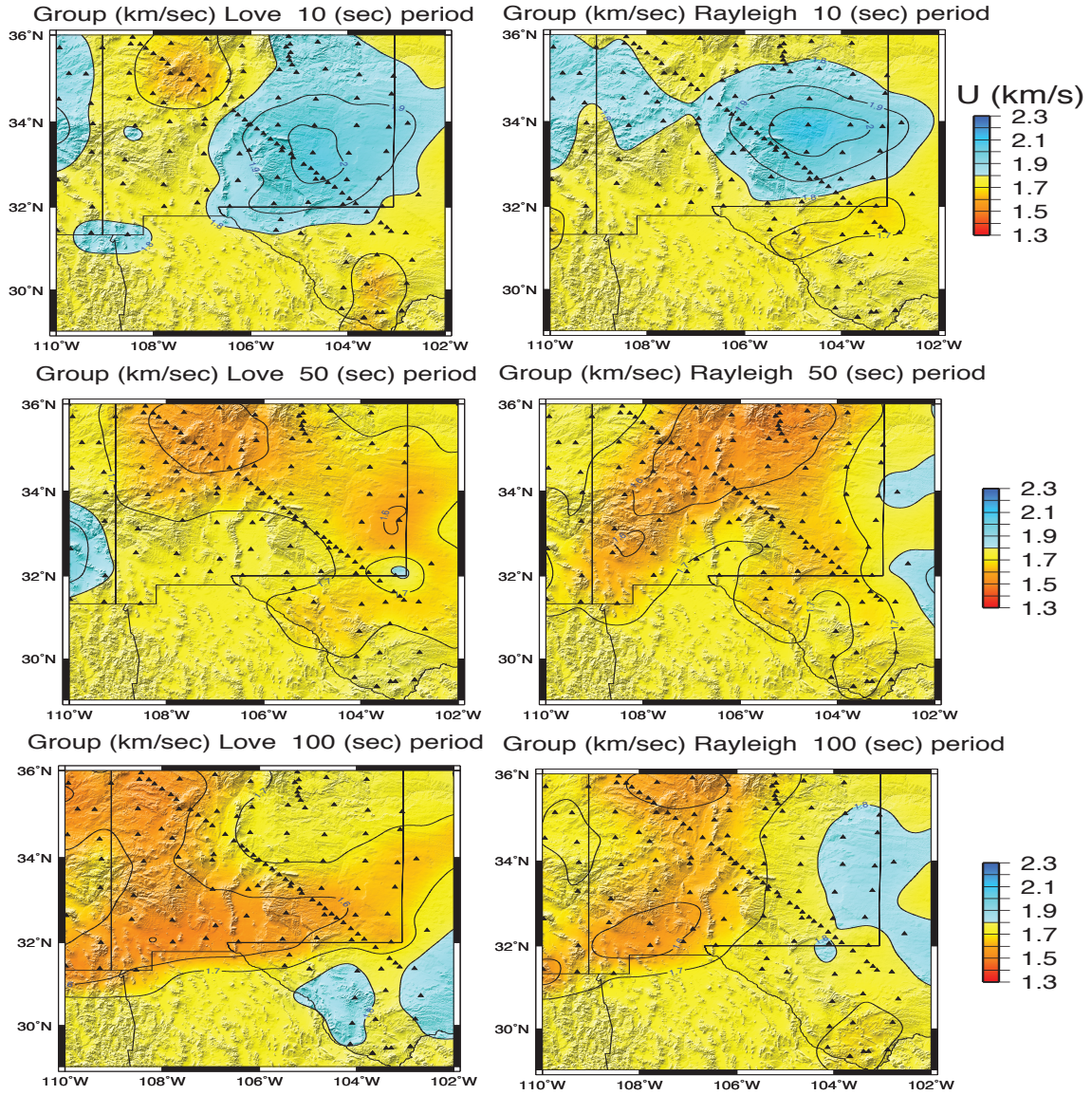


Figure 2.3: We show group velocity maps for the RGR region using 3 different periods (10 s, 50 s, and 100 s). The black triangles represent all the seismic stations that we used to plot the group velocity maps for the rift system. From the group velocity maps, we note that we can resolve the different tectonic provinces at different periods, in particular the slow velocities that persist at longer periods within the Basin and Range province.

for several periods (Figure 3) using Bayesian kriging ([113]). Figure 3 shows group velocity maps for the RGR region using 3 different periods (10 s, 50 s, and 100 s), although we include periods up to 140 s. From the group velocity maps, we distinguish the different tectonic provinces at different periods, in particular the slow velocities that persist at longer periods in the Basin and Range province. To determine how deep the surface wave data can resolve for our joint inversion, we calculate data sensitivity kernels for several stations ([46]). Figure 4 shows the surface waves sensitivity kernels plotted as a function of depth for a suite of periods used in our analysis. From this calculation, we determine that the surface waves resolve to approximately 300 km of depth; thus, we include results from 10 – 300 km in depth for our 3-D models.

2.5 1-D Constrained Joint Inversion

Joint inversion involves the simultaneous optimization of several objective functions, such as the l_2 -norm data misfit. Since the objective function is expected to be less subject to local minima, this approach reduces intrinsic non-uniqueness of the inverse problem ([25]). Some examples of joint inversion studies and data include: cooperative inversion ([67]), weighted schemes for inverting seismic travel times and gravity data ([62]), DC resistivity and seismic data ([33]), receiver functions and surface wave dispersion ([14], [51], [121], [136]), surface wave velocity and gravity observations ([6], [69], [72]), receiver functions, surface wave dispersion, and magnetotelluric data ([77], [78]), and topography, Bouguer anomalies, geoid height, and surface heat flow data ([50]). In most of these studies, the main assumption is that the data sets comprised in the inversion complement each other and sample similar geological boundaries.

In this work, we apply a 1D constrained optimization approach for joint inversion of two complementary data sets, receiver functions and surface waves group dispersion ([51]) using Primal-Dual Interior Point methods as a solver ([121]). Our approach addresses some of the well known numerical difficulties that arise for large-dimensional model spaces by

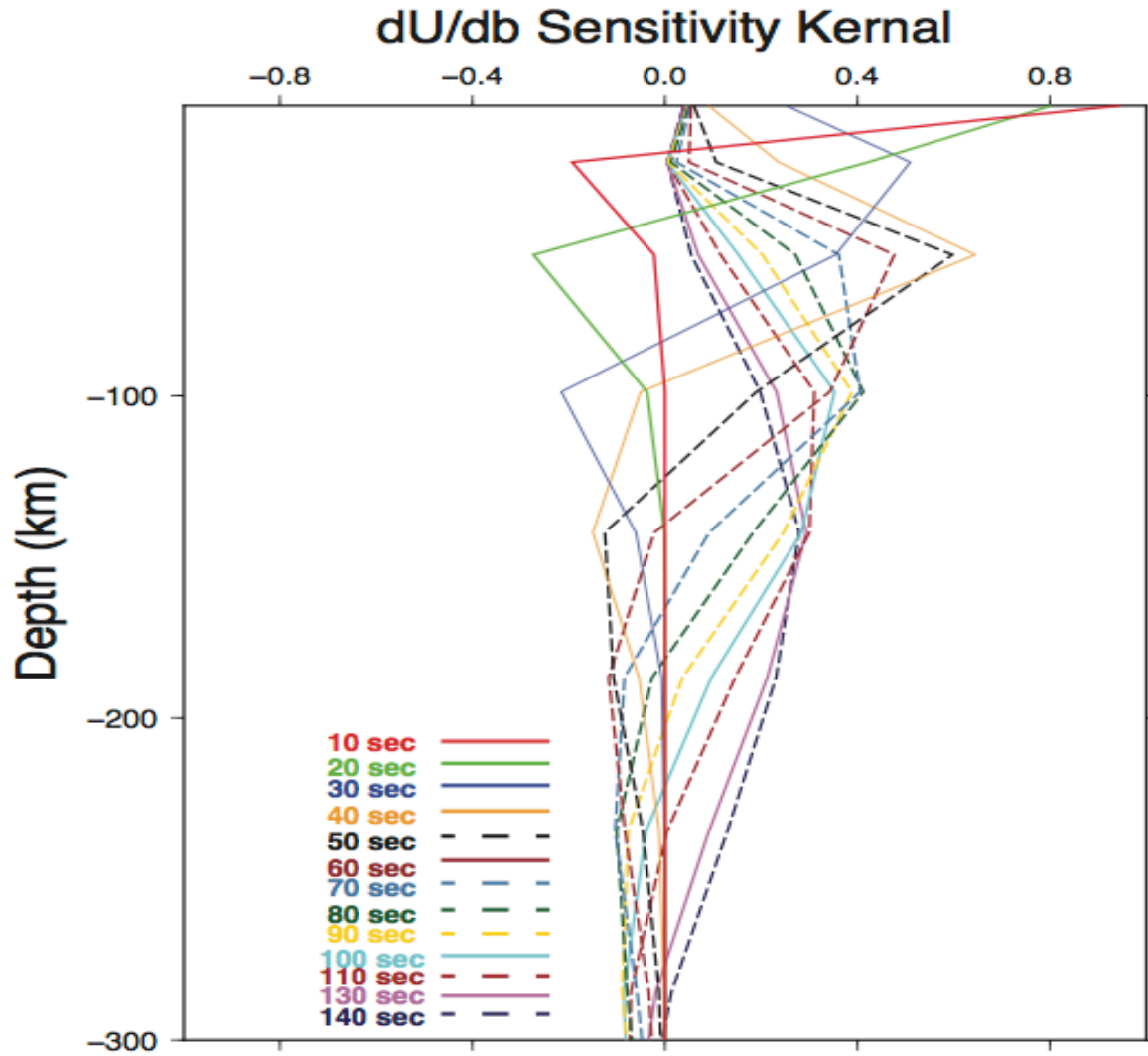


Figure 2.4: Sensitivity kernels for surface wave dispersion for four different stations located at each province included in this study. The surface waves sensitivity kernels are plotted as a function of depth for several periods. Based from the sensitivity kernel plots, we are able to determine what depth we can resolve, which is about 300 km depth.

using inequality constraints to incorporate a priori information and constrain further the geophysical inversion ([121]).

We characterize the Earth's structure by using S-wave velocities as the model parameter. The forward nonlinear operator $F \in R^m$ evaluated at a given velocity $x \in R^n$ provides a prediction of the Earth's response according to the data used as input. For a given observed data vector, $y \in R^m$, we pose the inverse problem as

$$\min_x \frac{1}{2} \left(\|F^{SW}(x) - y^{SW}\|_W^2 + \|F^{RF}(x) - y^{RF}\|_W^2 \right) \quad (2.5.1)$$

where W represents a weighted diagonal matrix used to equalize the contribution of each data set with respect to physical units and number of data points, while accounting for data set influence ([121]). In our case, the forward operator, F , collects both the numerical computation of synthetic waveforms for receiver functions, F^{RF} , ([4]), and the numerical evaluation of surface waves dispersion velocities, F^{SW} , ([72]). We assume a typical uncertainty value σ_i^2 of 0.05 (km/s) for SW , 0.01 (s) for RF observations ([51]), and we accommodate the amount of influence for each data set according to the station data quality. In general, this value is set equal for most of the stations.

Instead of the standard formulation of the inverse problem as in the unconstrained weighted nonlinear least squares ($NLSQ$) setting (2.1), we solve a sequence of linearized constrained LSQ with $F'(x_k)$ as the matrix with the partial derivatives of $F(x) = \begin{bmatrix} F^{SW} \\ F^{RF} \end{bmatrix}$. Therefore we rewrite problem (2.1) as,

$$\begin{aligned} \min_x \frac{1}{2} \|F'(x_k)x + b\|_W^2 \\ \text{s.t. } g(x) \geq 0, \end{aligned} \quad (2.5.2)$$

where $b = F(x_k) - y - F'(x_k)x_k$ is the residual vector. Here, the inequality constraint defined as

$$g(x) = \begin{cases} x - v_{min} \\ v_{max} - x, \gamma \in (0, v_{max} - v_{min}), \\ \eta\gamma - \frac{1}{2} \|Lx\|^2 \end{cases}$$

allows us to add appropriate bounds corresponding to a priori minimum and maximum velocities, i.e. $v_{min} \leq x \leq v_{max}$, while enforcing a roughness model constraint by using a first order discrete derivative operator L ([44]). The idea of a narrow class of models to be used in the inversion may help the originally ill-posed inverse problem to become well-posed ([150]). We apply primal-dual interior point (PDIP) methods to solve problem (2.2), which introduce an intrinsic regularization to the inverse problem making the joint inversion algorithm more robust ([83], [121]). In this method, we define the augmented Lagrangian function associated to (2.2): $\Gamma(x, z) = \frac{1}{2} \|F'(x_k)x + b\|_W^2 - g(x)^T z, z > 0$ where $z \in R^{2n}$ is the Lagrange multiplier corresponding to the inequality constraints.

Interior point methods are based on Newtons method. In our case, the necessary or perturbed Karush-Kuhn-Tucker (KKT) conditions, computed by differentiating with respect to the primal variables x and z , provides the right hand side of a Newtons system. This system can be solved iteratively by using a linesearch strategy ([83]) while enforcing the iterates to stay in a feasible (interior) region as described in [121]. The iterative process proceeds until it is either terminated when the misfit is less than 10^{-6} , or a maximum number of iterations no greater than six is reached, or the difference between iterates fail to differ more than a threshold of 10^{-5} . For all the stations involved in the geophysical inversion, the initial velocity model, x_0 , corresponds to the AK-135 model of [56], starting at 10 km depth and distributed at a 2 km interval up to 70 km depth, then at a 5 km interval up to 250 km and finally at 10 km until 300 km.

Since any inversion algorithms produces non-unique results, with ours being no exception, it seems helpful to being with other information, such as known geological constraints. By incorporating explicit velocity bound constraints with a measure of roughness into the inversion model, and for some of the stations adding a regularization term, our approach can thus produce a better constrained model while having more stable inversions ([121]).

2.6 Joint Inversion Results

We perform 1D joint inversions using our PDIP approach for 147 stations from USArray and LA RISTRA experiment. In general, each independent joint inversion includes at least 3 receiver function bins created according to an average ray parameter, with a width of approximately 0.01 s/km between 0.04 s/km and 0.07 s/km. The average ray parameter was determined by taking the mean value of the maximum and minimum ray parameter for each station before being used for stacking. The number of receiver functions employed to create these stacks depends on the station, but in general is not less than 25 per ray parameter. Each receiver function consists of 820 data points for a time range from -5 to 80 seconds. We also include fundamental mode Love and Rayleigh group velocities with 50 to 65 dispersion measurements, with periods between 5 to 140 seconds. Since the station spacing of the USArray is about 70 km, we anticipate lateral resolution of that order for each individual 1D inversion in the southern RGR region.

As supplementary material, refer to our paper [121], to see the four examples (Figures S1-S4) that are within the distinct tectonic provinces for crustal and upper mantle 1D velocity structure computed by using the constrained joint inversion algorithm: station 118A in the Basin & Range province (Figure S1), station NM26 in the center of the RGR (Figure S2), station V18A in the Colorado Plateau (Figure S3), and station W26A in the Great Plains (Figure S4). The figures show the fit to the RF observations and the Love and Rayleigh wave group dispersion curves, plus the final model approximation provided by the inversion. Velocity values were extracted from layered models as described before.

2.7 Kriging interpolation from 1D velocity profiles

Since our ultimate goal is to create a 3D Earth structure model of the Rio Grande Rift region, we use the 1D S wave velocity profiles of each station as input data for a kriging interpolation algorithm ([113]). In general, interpolation algorithms estimate values by using

a weighted sum of surrounding data. Kriging represents an example of a computationally efficient interpolation technique that allows the incorporation of uncertainty on the predicted values. We implement a Bayesian kriging approach that integrates variable spatial damping, a useful tool to control the kriged solution in extrapolation zones where few or no data is available ([113]). In our case, the station spacing within our region represents a 2D spatial grid, with each station now having a depth varying 1-D velocity structure. We thus can estimate the unknown velocities of the 2-D grid at different depths based on the known velocities, thus creating our 3-D model.

Initially, we remove an appropriate trend prior to applying kriging ([113]), which, in our case, corresponds to the mean of the velocities at a certain depth. A spatially damped kriging estimator then incorporates variable damping and measurement error multiplied by a unit-normalized function, which decreases noise values to zero according to the predictions point relative distance. As a result, we obtain a smoothly damping effect over the predicted velocities that varies according to each velocity node and its surroundings. For our results, we choose the blending functions of 2° to guarantee good spatial sampling.

Interpolating the 1-D profiles by means of kriging can help us to illuminate better the Earth structure beneath each station in the RGR. Schematically, if each station had perfect azimuthal coverage, the region below each station would have cone shaped raypaths, where at a certain depth (that depends on station spacing), the raypaths at adjacent stations begin to overlap, providing us with full subsurface structure coverage. Before this depth, we expect that the surface wave group dispersion information obtained from regional earthquakes can improve the average crustal velocity structure and also the vertical resolution ([112]). In this fashion, we account for velocity structure resolution avoiding additional inversions by grouping the 1D profiles depending on azimuthal range as indicated by [6]. Generally, the upper mantle of a tectonically active region is expected to exhibit 3D heterogeneities with a length scale smaller than both the lateral resolution of surface waves and vertical resolution of receiver functions ([84]). Therefore, the models obtained by using these two data sets should resolve the main features beneath the region of study.

2.8 3-D RGR S-Wave Velocity Model

Figures 5 through 8 show different perspectives of the resulting 3D crustal and upper mantle structure images, including the profiles presented in Figure 2. For Figure 5, we use a color scale (maximum to minimum velocities) to highlight crustal anomalies and a different color scale in Figures 6 through 8 with a reduced color spectrum varying from 4.0 km/s to 5 km/s, similar to that used by [136] to highlight mantle anomalies.

Figure 5 shows depth slices from 10 to 35 km that highlight crustal anomalies in our model. The rift can be seen in the crust with low velocities along its axis from north to south, from 10 to 30 km in depth. The rift appears to continue southeast along the Texas-Mexico border, yet we have no resolution south of the border. At 35 km, we have no signature of the rift, as we likely are in the upper mantle for most of the region.

Figure 6 shows our 3-D velocity model at cross-section A-A' and B-B'. The ~ 740 km long cross section A-A' coincides with latitude 34° , and passes through Colorado Plateau, Socorro Magma Body (SMB) and ends at the Great Plains. Cross-section B-B' coincides with latitude 32° and covers the southernmost part of the study area. Both cross sections show slightly uplifted Moho beneath the Basin & Range province, and a lower and middle crust that might be related to magmatic activity in the upper mantle. [34] relates this activity to convection in the deeper mantle. This area of active upper crust extension is suggested to be primarily the product of magmatic activity in the lower crust and upper mantle ([115]). Beneath the Colorado Plateau, Figure 6 also shows a mantle lid between 100 – 150 km as in [34] and [136], and the presence of cold mantle lithosphere about 200 km below the Great Plains. [131] explained thermally these strong fast anomalies at the western edge of the Great Plains, as cold downwelling lithosphere destabilized by small-scale convection.

To further investigate the upper mantle of our model, Figure 7 shows depth slices from 50 to 300 km depth in 50 km intervals. At 50 km and 100 km depth, a low velocity anomaly

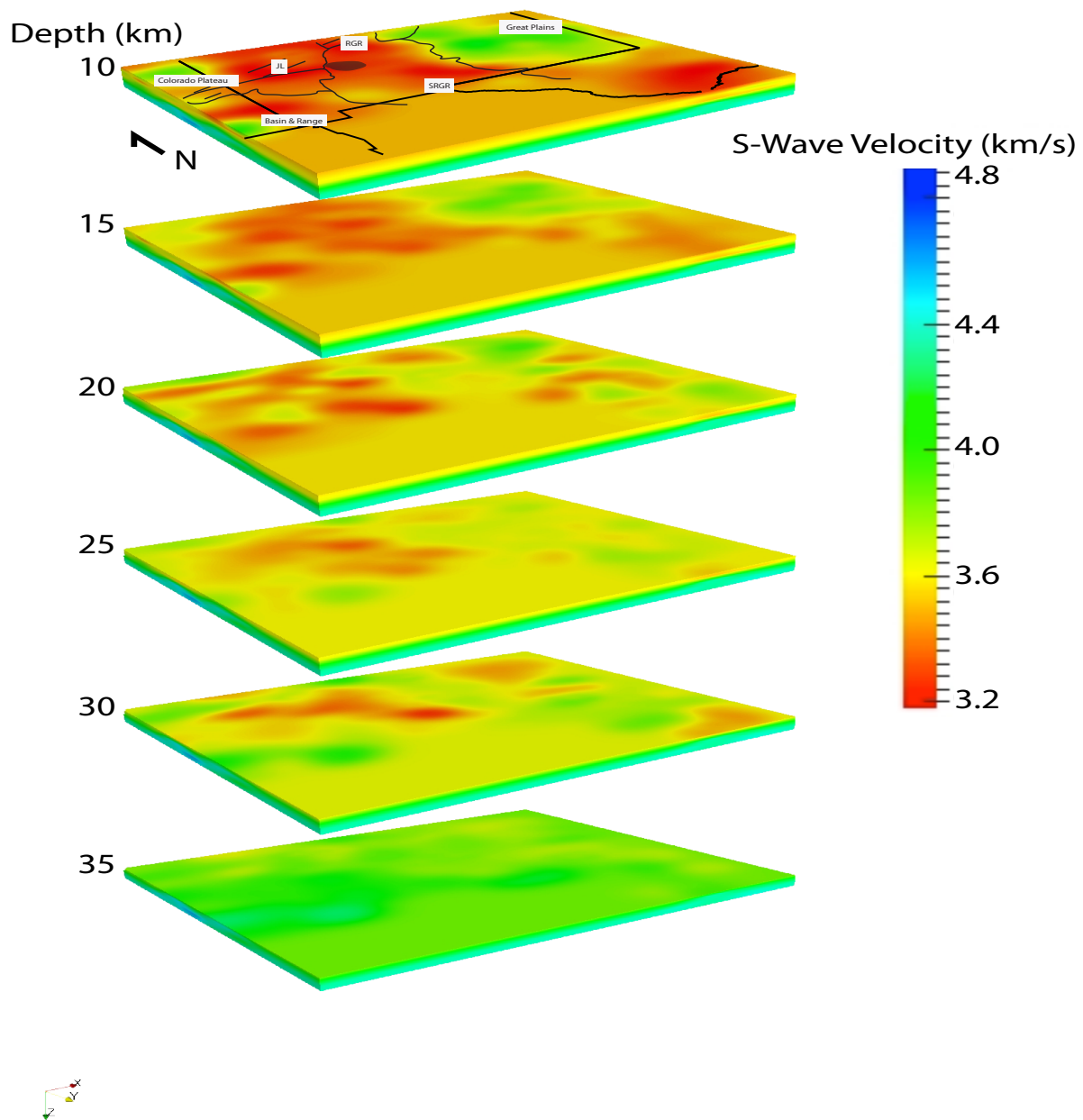


Figure 2.5: Crustal slices at different depths from 10 km up to 35 km illustrating various S wave velocity models. We use a color scale (maximum to minimum velocities) to highlight crustal anomalies within the Basin & Range, Jemez Lineament (JL), Great Plains, and Southern Rio Grande Rift (SRGR) indicated by the labels on top of the first shear wave model at 10 km depth. We can clearly see a distinct pattern that is consistent throughout the different tectonic provinces in the crustal slices from 10 km to 35 km depth.

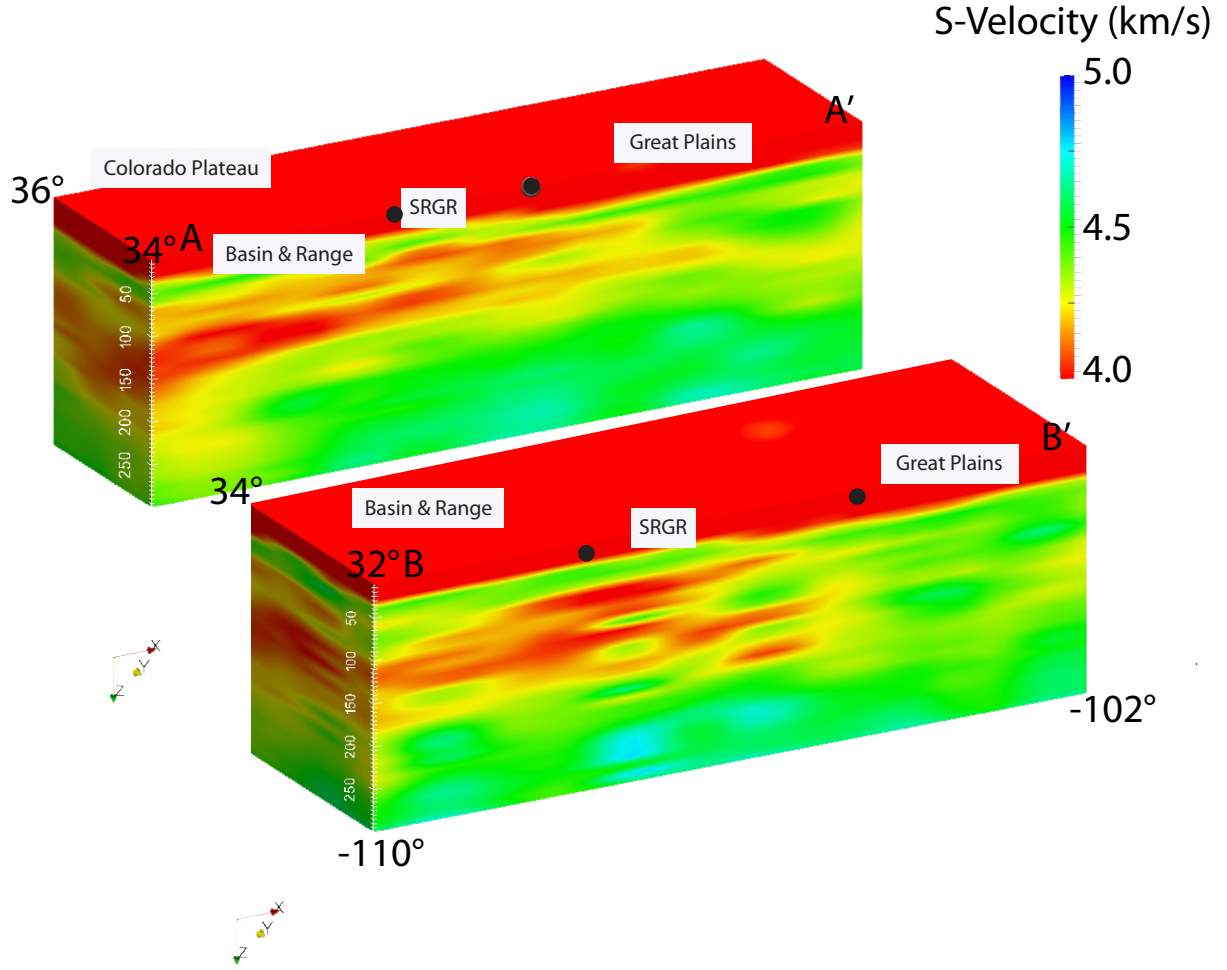


Figure 2.6: (Top) Cross-section A-A' at latitude 34° shows a clear distinction between the Colorado Plateau, Socorro Magma Body at the center of the RGR (middle spheres), and the Great Plains. We find that near the CP there is a mantle lid between 100-150 km as in Gao et al. (2004) and West et al. (2004a), and the presence of cold mantle lithosphere about 200 km below the Great Plains. Anomalous high velocities begin to appear right below the RGR and continue east of the GP between the depths of 200-300 km. (Bottom) Cross-section B-B' at latitude 32° covering the southernmost part of the RGR region. We image a low velocity zone that begins to appear beneath the RGR extending to the west below the B & R and Colorado Plateau. Both cross sections show slightly uplifted Moho beneath the Basin & Range province, and a lower and middle crust that might be related to magmatic activity in the upper mantle.

is present north of 34° and appears affiliated with the Jemez lineament. However, another low velocity anomaly in the southern RGR appears from 50 to 150 km. This anomaly may be connected with the Jemez lineament, but its persistence and narrowness signifies a possible upwelling. These anomalies appear to terminate at 200 km depth, with no strong signatures from 200 – 300 km.

Cross-section C-C' (Figure 8) is ~ 700 km long coincide with the southern part of the Rio Grande Rift (LA RISTRA) experiment and cross cut the profiles A-A', and B-B'. The transect begins at Colorado Plateau and passes through the SMB and ends west of the southern RGR. We identify an upper mantle low velocity feature beneath the Jemez lineament, which may originate outside our study area and is present in Figure 7 north of 34° latitude. [34] discussed the possible presence of an eastward flow of mantle from the Jemez region across the rift with sinking beneath the Great Plains. Furthermore, [136] state that to the west of the Great Plains, an asthenospheric low velocity channel underlies the region and extends to 150 km depth. This low velocity zone continues further to the Colorado Plateau at 100 – 200 km depth. This is similar to what [136] describes, except that we do identify it directly beneath the Rift and it does not form an inverted U-shape. Instead, we see an oblique body that may continue outside our study region, which we term the Jemez Upwelling, that is present in profiles A-A' and B-B'. The difference between our results and [136] is likely the result of different approaches and assumptions, but the 3-D nature of our results further clarify the anomaly, which we discuss below.

Reviewing the 3-D models in Figures 6 and 8, we find no evidence for a deep mantle source under the Rift, implying that it is not currently driven by deep mantle upwelling ([143]; [136]). However, [81] suggests that this interpretation was based on static images that could not assess the level of flow below of the RGR. Based on their flow simulations, [81] proposes that the mantle flow below the RGR is associated with thermal upwelling. As in previous studies, we image not only the low velocity mantle beneath the RGR ([34]; [100]; [136]; [142]), but also sharp changes between the RGR and the two surrounding provinces of Basin & Range and Colorado Plateau ([6]). Along the RGR, we show a persistent low

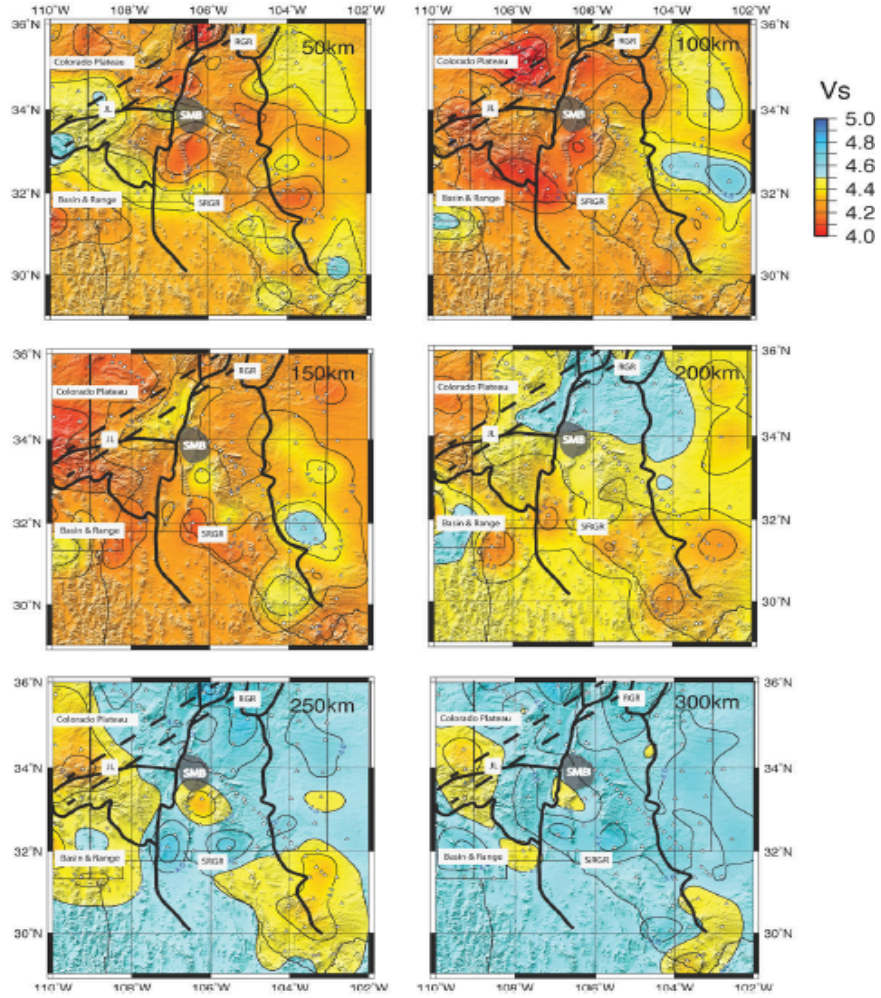


Figure 2.7: Shear wave velocity maps at different depths starting at 50 km up to 300 km. We identify an upper mantle low velocity feature beneath the Jemez lineament, which may originate outside our study area (see Figure 7). At 50 km and 100 km depth, a low velocity anomaly is present north of 34° and seems to be affiliated with the Jemez lineament. However, another low velocity anomaly in the southern RGR appears from 50 to 150 km.

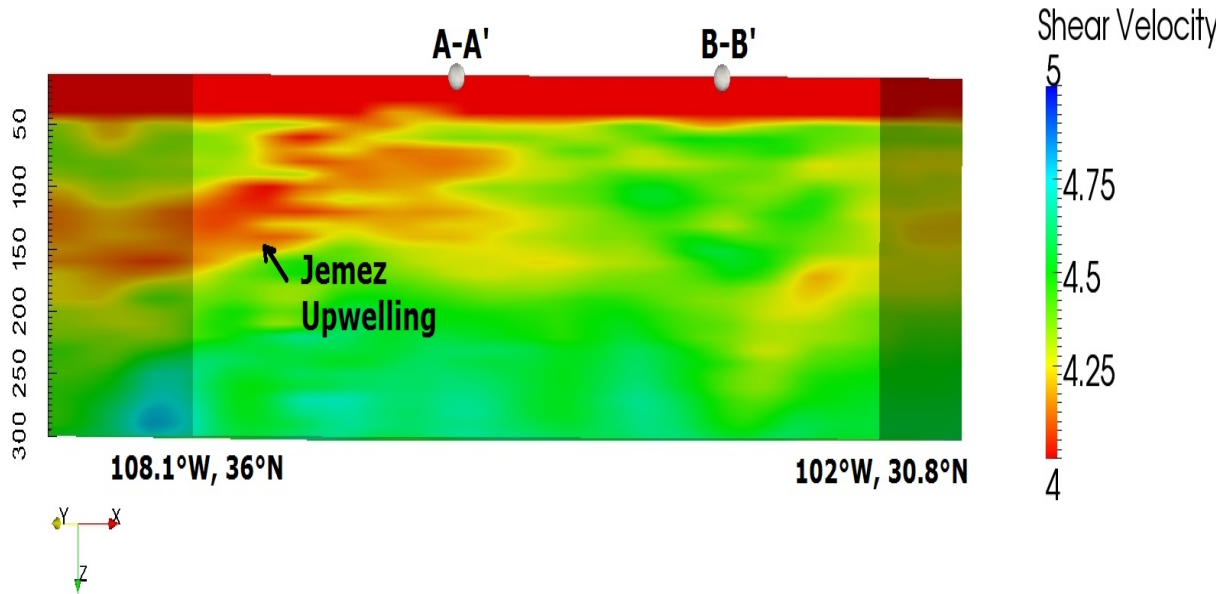


Figure 2.8: Cross-section C-C' coincides with the southern part of the LA RIS-TRA passive experiment. Seismically fast mantle underlies the RGR and relatively slow mantle is seen beneath the Mt. Taylor and Colorado Plateau. We identify an upwelling beneath the Jemez Lineament, which we term the Jemez Upwelling. The anomalously high velocities beneath the RGR appear to decrease at both sides of the rift, particularly beneath the GP portion.

velocity zone that does not originate deeper than 200 km.

2.9 Discussion

Comparing our 3-D model directly to others can be difficult and requires complete visualization of all models in one figure (e.g., [10]; [88]). Since our models were derived from 1-D inversions that included two complementary data sets and then interpolated to 3-D, we expected to find differences in our 3-D model compared to previous studies. We also focus solely in this region, while many of the other models include much larger regions ([35]; [64]; [81]; [119]; [131]). Thus, we show features relevant to RGR evolution, and highlight those that we believe are fully consistent with most recent and past models in the southern Rio Grande Rift obtained by using different data sets and techniques. We discuss below the impact of the PDIP approach, the new model for the southern RGR, and finally highlight the main new feature not highlighted in other work: the Jemez Upwelling and the SRGR low velocity anomaly.

2.9.1 PDIP methodology

We implement a new approach for joint inversion of receiver function and surface wave group dispersion data based on constrained optimization ([121]). We create independent 1D Earth velocity profiles of upper mantle velocity structure along the southern RGR. In general, no smoothing or damping factors were used to stabilize the inversions. This is an advantage when compared to other standard techniques, which often require tuning of several regularization parameters. It is well known that for severely ill-posed inversions that may appear for some stations, determining the optimum regularization parameters remain as a difficult and often speculative task ([121]; [150]). There are several strategies to choose these regularization parameters ([44]; [150]), and some recent advances on trans-dimensional inversion to include them directly as inversion parameters ([14]). However, there is not yet an analytical or automatic way to find the best parameter for all particular

cases ([44]; [78]). The PDIP approach reduces the subjectivity of these selections, since the simple inclusion of a damping parameter stabilizes the inversion ([121]).

Some stations have inversion convergence issues mostly due to the absence of good quality RF data to identify absolute S-wave velocities and sharp discontinuities. Those stations were independently tuned by using several damping factors. We note that to obtain improvements, a damping factor of 0.1 and an influence parameter equal to 0.25 was necessary. Furthermore, we found that in some cases even modifying the influence parameter to be 0.75, the damping factor did not improve the numerical results for the problematic stations. We discuss with detail the impact of regularization (damping/smoothing) on our constrained optimization approach in [121].

[6] use a similar approach combining results obtained by independent joint inversion of surface wave phase velocities and receiver function information, to compute 1D S wave velocity profiles for the Colorado Plateau. They extracted common features of nearby stations that were reconciled with observed gravity anomalies. The anomalies are then established by using empirical relations to density structure. By creating 3D images of Earth structure based on independent inversions that produce 1D velocity profiles, results are likely inferior compared to a full 3D parameter approach ([6]). However, a 3D inversion involves a high volume of information and a great number of parameters, which may be numerically intractable. PDIP methods have been proven to be successful in solving large-scale problems ([83]), and we will explore the migration of our technique to a full 3D inversion in future work.

We can adjust our inversion approach by tuning the kriging interpolation blending function according to the geological province where the stations are located, e.g. Basin and Range, RGR, Great Plains, as an attempt to enhance the performance of the inversion algorithm. However, our approach allows us to begin with a standard initial model (ak135: [56]), and produce coherent independent inversions that could be combined for a consistent 3D structure. This represents another difference to the approach presented by [6], where they weight the contribution of each inversion separately to select a suitable initial model.

2.9.2 Interpretation: Basin and Range, Colorado Plateau, RGR, and Great Plains

[100] argue that the RGR and southeastern Colorado Plateau underlay by a low-density upper mantle province, which does not trend along upper crustal tectonic boundaries, but correlates with regions of late Tertiary magmatism. The Rio Grande Rift system is characterized by anomalously high topography ([100]) and its crustal thickness is thick along its eastern flank and thin in its center ([52]). As in [6], we show sharp changes in crustal thickness that distinguish between the Colorado Plateau and surrounding provinces dominated by extension, e.g. BR, GP and RGR. [100] showed that low wave speed zones are broad in the north and narrow in the south. We find a low velocity zone at depths from 100 – 50 km that coincides with the southernmost RGR. [96] concluded that the LA RISTRA profile experiment showed evidence of mantle convection and partial melting in the crust and upper mantle. Although we have no conclusive evidence of melting, we see higher velocities in some regions, which may relate to a cold mantle upwelling (Figures 6) at latitude 34° and a cold downwelling along the LA RISTRA profile (Figure 8).

We image similar S wave velocities as the LA RISTRA profile, yet we do not have a bow-tie low velocity anomaly ([34],[136],[143]). The layering seen in Figure 8 may be evidence of delamination. Initially, delamination was proposed as one of the main alternative mechanisms of lithospheric recycling in continental collision areas ([13]). If thin lithosphere comes together with high heat flow, this may suggest that the lower crust has been removed by delamination ([63]). The delamination process leads to considerable thickening of the crust in areas with weak lithosphere, which can distinguish the Great Plains and RGR provinces in Figure 7. [9], however, suggested that crustal thinning of the BR was a result of simple stretching of the original crust rather than delamination. We do find thin crust under this province, but we cannot conclude whether the origin of such feature is either delamination or simply stretching of the crust.

2.9.3 The Jemez Upwelling and the Southern RGR Anomaly

The Jemez lineament, an alignment of volcanic centers that extends about 800 km (Figure 2) ([3], [37]), represents the most representative signature of primary volcanism in the region of the Rio Grande Rift ([122]). The Jemez lineament formed from the Miocene to Holocene in which is thought to be a reactivated Precambrian structure ([3], [37]). Using the results of a reconnaissance experiment of lateral variation for P-wave velocities, [122] identified a low velocity zone along the Jemez lineament in the depth range from 50 – 200 km and interpreted the zone to be partial melt. However, the resolution of their model lies beneath the ~ 200 -km-long segment from Mount Taylor through the Jemez volcanic center, and up to 160 km depth only. This feature shows up in other tomography models in the region, where a low velocity anomaly is present until about 200 – 300 km in depth ([34]; [81]; [131]).

We identify the same low velocity zone to a 200 km depth. With a 3D perspective, we link the low velocity zone to an upwelling sheet of hot (low velocity) material, which we term the Jemez upwelling, beneath the Jemez lineament originating beneath the Colorado Plateau, and that likely feeds the volcanic centers in the region (Figure 7). This feature may be linked to the southern RGR low velocity anomaly (Figure 6) at 100 km depth along the axis of the rift. However, the southern RGR anomaly appears to be narrower than the Jemez upwelling and may extend up to 150 km. [81] suggests that strong upwelling impacts the base of the lithosphere at an oblique angle east of the Colorado Plateau and directly below the RGR. The planar body parallels the Jemez lineament and is oriented perpendicular to the LA RISTRA transect (C-C').

Along the C-C' transect, our results are similar to previous images for the LA RISTRA experiment, where a significant transition between high velocities beneath the Colorado Plateau and the Great Plains is revealed as well as a broad low velocity zone beneath the RGR ([34]; [143]). Furthermore, our 3D images allow us to define the dimensions of the low velocity zone in our region and to link this Jemez upwelling to the Jemez lineament. However, due to the northern extent of our model, we do not know if the low velocity zone

is persistent beneath the Colorado Plateau ([34]; [136]).

The Socorro Magma Body (SMB) within central RGR is one of the largest active intrusions in the Earth’s continental crust, and is associated to steady central uplift ([89]). Also, the SMB has been linked to strong magma influence, e.g. diffusion of fluid moving upward from depth, due to an underlying low-velocity molten layer. [100] associated this low velocity zone to a possible combination of partial melt, temperature and compositional variations. [102] suggested that there is not a strong direct magmatic influence in the seismic activity of the SMB, and that this activity is more prone to be associated with characteristics of a continental rift, like preexisting highly fractured crust. Furthermore, the strongest velocity variations located in the upper 200 – 300 km of mantle beneath the magmatically and tectonically active Rio Grande Rift and Basin and Range show a clear relation between tectonic province and mantle velocity beneath the stable Colorado Plateau and Great Plains. This distinction coincides with that presented by [34] and [143]. The RGR can clearly be seen as a strong crustal feature in our models (Figure 4), while the two major low velocity zones in our model (the Jemez Lineament and the southern RGR anomaly) highlight an upper mantle process that may be the result of small-scale convection (Figures 5 – 8). However, because a deep mantle upwelling underneath the RGR does not appear, the mechanism for rift formation remains ambiguous.

2.10 Conclusions

We present a new model of crustal and upper mantle structure beneath the southern RGR. Separate joint inversions were performed for 147 Earthscope USArray and LA RISTRA stations. We create a generic framework that connects a constrained optimization joint inversion algorithm with a Bayesian interpolation scheme for high-resolution imaging of Earth structure. Furthermore, this scheme efficiently provides a robust alternative to extend simultaneous independently created 1D S wave velocity models, to produce 3D images of Earth’s structure compared to full 3D inversions. Our framework generates a continu-

ous and smooth 3D velocity model of the Rift system, revealing the complexities of the southern RGR and helping us to better characterize its crustal and upper mantle velocity structure. We present evidence of crustal thinning in the center of the Rift, and no evidence of a deep mantle upwelling driving the RGR ([34]). We identify the boundaries between the provinces of B&R, CP, GP and RGR, and an upwelling sheet of low velocity material, that we term the Jemez upwelling. We also identify a southern RGR anomaly that may be connected with the Jemez upwelling. The resulting 3D models show a thin, lower velocity lower crust along the southern east portion of the Rio Grande Rift, plus a low velocity lithosphere underneath the Colorado Plateau and Basin and Range province which may be attributed to high crustal temperatures ([11]). We have no evidence of a deep mantle plume that drives the current rifting process, and all velocity anomalies are shallower than 200 km.

Chapter 3

Construction of shear wave models by applying multi-objective optimization to multiple geophysical datasets

For Texas, our main purpose is to obtain a better understanding of the Earth's tectonic processes within the region, which requires us to analyze the Earth structure. We expand on a constrained optimization approach for a joint inversion least-squares (LSQ) algorithm to characterize a one-dimensional Earth's structure of Texas with the use of multiple geophysical data sets. We employed a joint inversion scheme using multiple geophysical datasets for the sole purpose of obtaining a three-dimensional velocity structure of Texas in order to identify an ancient rift system within Texas. In particular, we use data from the USArray, which is part of the EarthScope experiment, a 15-year program to place a dense network of permanent and portable seismographs across the continental United States. Utilizing the USArray data has provided us with the ability to image the crust and upper mantle structure of Texas. We simultaneously inverted multiple datasets from USArray data, to help us to better obtain an estimate of the true Earth structure model. We prove through numerical and experimental testing that our Multi-Objective Optimization (MOP) scheme performs inversion in a more robust, and flexible manner than traditional inversion approaches. The work done in Chapter 3 is based on the following publication:

Thompson, L., A. A. Velasco, and V. Kreinovich.: Construction of shear wave models by applying multi-objective optimization to multiple geophysical data sets. *Handbook of Springer Proceedings in Mathematics & Statistics (PROMS)*, Springer, Current Trends in

3.1 Introduction

For this paper, we propose to combine multiple geophysical datasets for the purpose of assisting us in better determining physical properties of the Earth structure. By simultaneously inverting multiple datasets, we obtain a better estimate of the true Earth structure. In general, there are two reasons why the estimated Earth structure model differs from the true Earth structure. The first reason is the inherent non-uniqueness of the inverse problem that causes several (usually infinitely many) models to satisfy the data. The second reason is that real geophysical data is always affected by noise, which introduces error associated with the estimation of the Earth structure model after inversion. By jointly inverting multiple geophysical data sets, we reduce the inherent non-uniqueness typical for the geophysical datasets (e.g., receiver functions, surface wave dispersion, teleseismic delay travel times, and gravity) individually [133],[25]. For this research, we use receiver functions, surface wave dispersion measurements, and P-wave travel times to characterize the crust and upper mantle structure of the Texas region.

In general, geophysical data sets such as receiver functions are suited to constrain the depth of discontinuities and are sensitive to relative changes in S-wave velocities in different layers. Surface waves measurements on the other hand, constrain the absolute shear velocities between discontinuities whereas receiver functions are unable to do that [117],[51], [72], [116], [124]. Seismic first-arrival travel times and gravity data are complementary to each other because one can recover the causative slowness and density distributions of the Earth structure [62]. The complementary information provided by the following datasets, reduces the inherent ambiguity or non-uniqueness of performing inversion (e.g., [40], [25], [77], [78], [66], [14]). By jointly inverting seismic data along with gravity data, we will be able to overcome the difficulties of non-uniqueness and be able to facilitate the construction of the true Earth model.

When we process a single data set (e.g., Surface Wave Dispersion), we use the least squares method to find the best-fit model. For multiple data sets (e.g., Surface Wave Dispersion and Receiver Functions), if we knew the variance (uncertainty of data) of the different measurements of the multiple data sets, we could still be able to use the least squares approach to finding the model space. In practice, we only have an approximate knowledge of the variances. So, instead of producing a single model, we want to generate several models corresponding to different possible variances. Once several models corresponding to different possible variances are computed, we can then proceed to select the most geophysical meaningful model from the Pareto Front [57]. The reason we will use an optimization technique is to find the best possible solution for nonlinear geophysics inverse problem. For example, in geophysics, most inverse problems require finding some minimization and that is why we will use an optimization technique called MultiObjective Optimization Problem (MOP). The MOP technique generates several possible models. This is what sets it apart from other various joint inversion techniques. We will be able to select the final solution from a population of alternative solutions from the model space. Such methods are described in [107],[108], [57].

There are two types of seismic waves that travel through the Earth: the body waves and the surface waves. Both types of waves give us different sensitivities and information about the Earth structure, since they are sampling the interior and surface of the medium with different velocities and directions. The information collected from the body waves travels deeper into the Earth and translates into teleseismic P-wave receiver functions. In order to obtain information about the Earth surface, surface waves are analyzed, in our case, by means of surface waves dispersion. On one hand we have receiver functions, which resolve discontinuities (impedance contrasts) in seismic velocities, and provide good measurement of crustal thickness, without providing a good average of shear wave velocity. On the other hand, we have surface (Love and Rayleigh) waves whose energy is concentrated near the Earth's surface, and provide good average of absolute shear wave velocity, without a good shear-wave velocity contrasts in layered structures [51], [72], [116], [124], [22], [84]. There-

fore these two data sets can be considered as complimentary and consistent, as long as we sample the same medium. Hence, we expect a mutually consistent estimate of the Earth’s structure. Since both data sets are sensitive to shear wave velocity structure [51], we can assume a forward operator F depending nonlinearly on our model parameter $x \in \mathbf{R}^n$ that represents the different shear velocities of a half space with n horizontal layers (a standard way of modeling Earth’s structure). In the next subsections we explain in more detail the nonlinear relationship with respect to shear wave velocities of this operator and the techniques used to compute each synthetic dataset.

3.2 Receiver Functions

A receiver function is simply a time series representation of the Earth’s response relative to an incoming P-wave propagating near a recording station. Positive or negative spike amplitudes represent positive or negative seismic velocity contrast. A receiver function technique can model the structure of the earth by using seismograms from three component (vertical, north, and east) seismic stations from teleseismic earthquakes. The receiver function technique takes advantage of the fact that part of the energy of seismic P waves is converted into S waves at discontinuities along the ray path [29], [9] and has been utilized in many studies (e.g., [143], [6], [143], [42]). For data collection and processing, we use the Standing Order for Data (SOD) [87], [6] to request three component seismograms for P-wave arrivals and for events with a minimum magnitude 5.5, depth in the range of 1600 km, and an epicentral distance ranging from 30° to 95° (e.g., [6]).

Receiver functions were first applied in the late 1970s at solitary stations to obtain local one-dimensional structural estimates [60]. Since then, there was an increase in the number of stations deployed seismic experiments. It is now possible to generate detailed two or three-dimensional images of structures, such as the moho and upper mantle transition zone discontinuities near 410 km and 670 km depth (e.g., [139]).

Receiver functions are derived using deconvolution, a mathematical method used to

filter a signal and isolate the superimposed harmonic waves. Specially, receiver functions are calculated by deconvolving the vertical component of a seismogram from the radial component, resulting in the identification of converted phases where there is an impedance contrast (crustal-mantle boundary) [116].

3.3 Receiver function stacking

We used the receiver function stacking technique introduced by [151], which estimates the crustal thickness and a V_p/V_s ratio based on the radial receiver function. This technique is the standard approach used by EARS. Assuming that no lateral velocity heterogeneities exist, the time separation between the Ps converted wave and the direct P-wave obtained from receiver functions (t_{Ps}) can then be used to estimate crustal thickness (H), given the average crustal velocities V_p and a V_p/V_s ratio (κ), and the constant ray parameter p of the incident wave (e.g., [38]). The trade-off between the thickness and the crustal velocities presents an ambiguity that can be reduced by using the later multiple phases $tp_p p_s$ and $tp_s p_s + p_p s_s$, which provide additional constraints to both V_p/V_s and the crustal thickness (e.g., [38], [151]). Using and stacking multiple events helps to increase the signal-to-noise ratio (SNR), which may be caused by background noise, scattering from crustal heterogeneities, and P-to-S multiple conversions from other velocity discontinuities [70]. The H - κ domain stacking weights each phase and plots the stacked phases as a gridded image $s(H, \kappa)$, which reaches a maximum when all three phases (tp_s , $tp_p p_s$, $tp_s p_s + p_p s_s$) are stacked coherently with the correct H and κ [151]. The main advantage of this grid-search based technique is that (1) large amounts of receiver functions can be processed without the need of picking P_s arrival times, and (2) the stacking results in an enhancement of the signal/noise ratio and a suppression of lateral variations in the vicinity of the recording station [70]. We will use this technique to derive an average crustal model including H and V_p/V_s (κ). An example of this technique is shown in Figure 2 for one of the Earthscope USArray stations, 219A. The dark dot with the white circle around the dot represents the

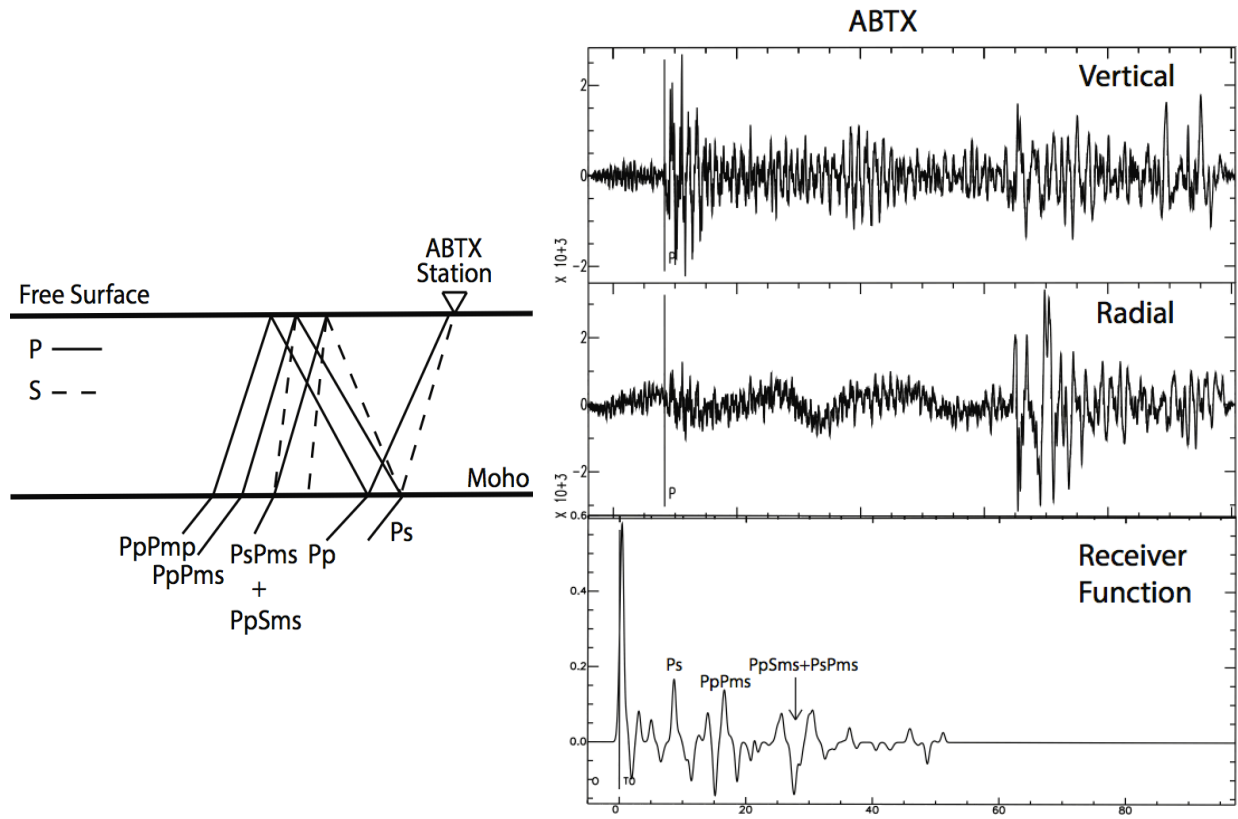


Figure 3.1: (Left) Illustration of a simplified ray diagram, which identifies the Ps , converted phases, which comprise the receiver function for a single layer. (Right) Vertical and radial seismograms and the corresponding receiver function resulting from the deconvolution of the vertical component from the radial component.

possible solution in H and V_p/V_s space (Figure 2).

3.4 Surface Wave Dispersion

Surface waves in general differ from body waves in many respects they travel slower, lower frequencies, largest amplitudes, and their velocities are in fact dependent on frequency [116]. The surface wave velocities vary with respect to depth being sampled by each period of the surface wave. The sampling by each period of the surface wave is known as dispersion [121]. Valuable information can be inferred by measuring surface wave dispersion because it will allow you to be able to better understand the Earth's crustal and mantle velocity structure [84], [121], [61]. In particular, Love and Rayleigh wave group dispersion observations generally account for average velocity structure as a function of depth ([51], [72]). The dispersion curves for surface waves are extracted from station records of three component seismograms for different frequencies and distances, by using reduction algorithms that rely on spectral analysis techniques. The important fact here is that, based on Rayleigh's principle, surface wave velocities are more sensitive to S wave velocity, although they are also theoretically sensitive to P wave velocity and density. The Rayleigh's principle states that the phase velocity perturbation, denoted by $\frac{\delta c}{c}$, can be viewed as a function of $(K_\alpha, K_\beta, K_\rho)$, the sensitivity coefficients for P wave velocity, S wave velocity and density, respectively, i.e.

$$\frac{\delta c(T)}{c(T)} = \int \left(K_\alpha \frac{\delta \alpha(z)}{\alpha(z)} + K_\beta \frac{\delta \beta(z)}{\beta(z)} + K_\rho \frac{\delta \rho(z)}{\rho(z)} \right) \quad (3.4.1)$$

where T is the period and z is the depth. By investigating sensitivity function variation in depth, the relative contribution of each property to dispersion can be shown. This subject is beyond the scope of our work, thus we just mention here that such analysis allows geophysicists to show that the relative contribution of P wave velocity, and density to dispersion is smaller than the one for S wave velocity [51]. This is, surface wave dispersion is much more sensitive with respect to S wave velocity, and therefore we have established the dependence of this data set on shear wave velocity.

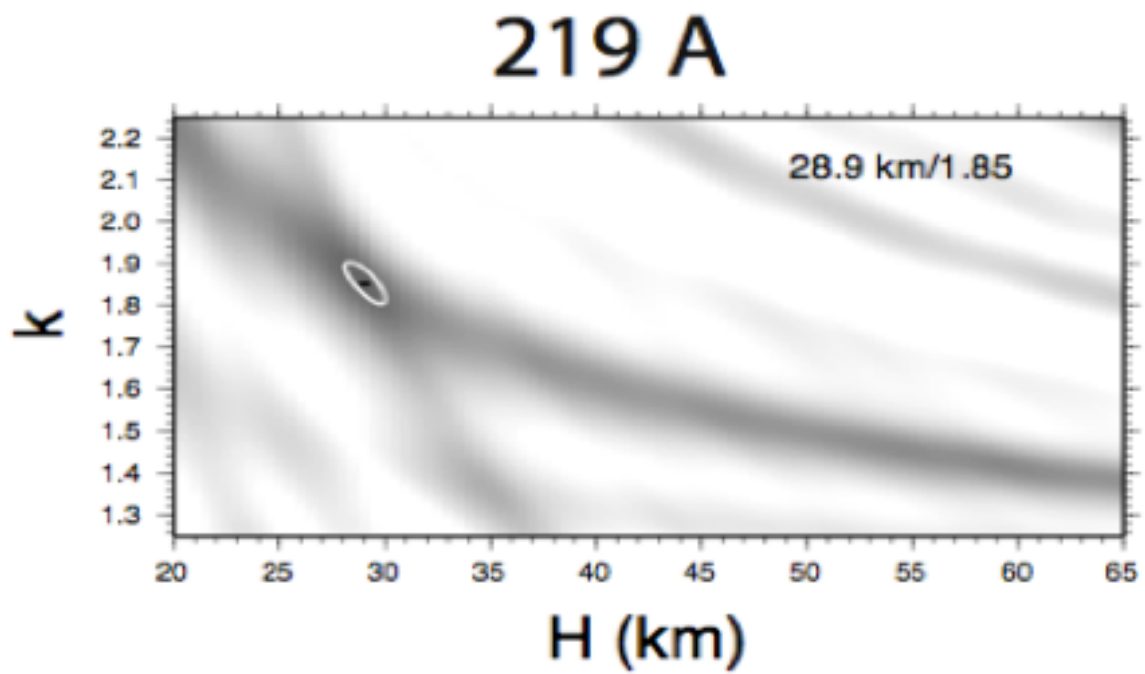


Figure 3.2: This is a receiver function stack of station 219A, V_p/V_s vs H (km). The black dot with the white circle around it represents the preferred value. Note the multiple shaded regions might result in a poor choice of crustal thickness.

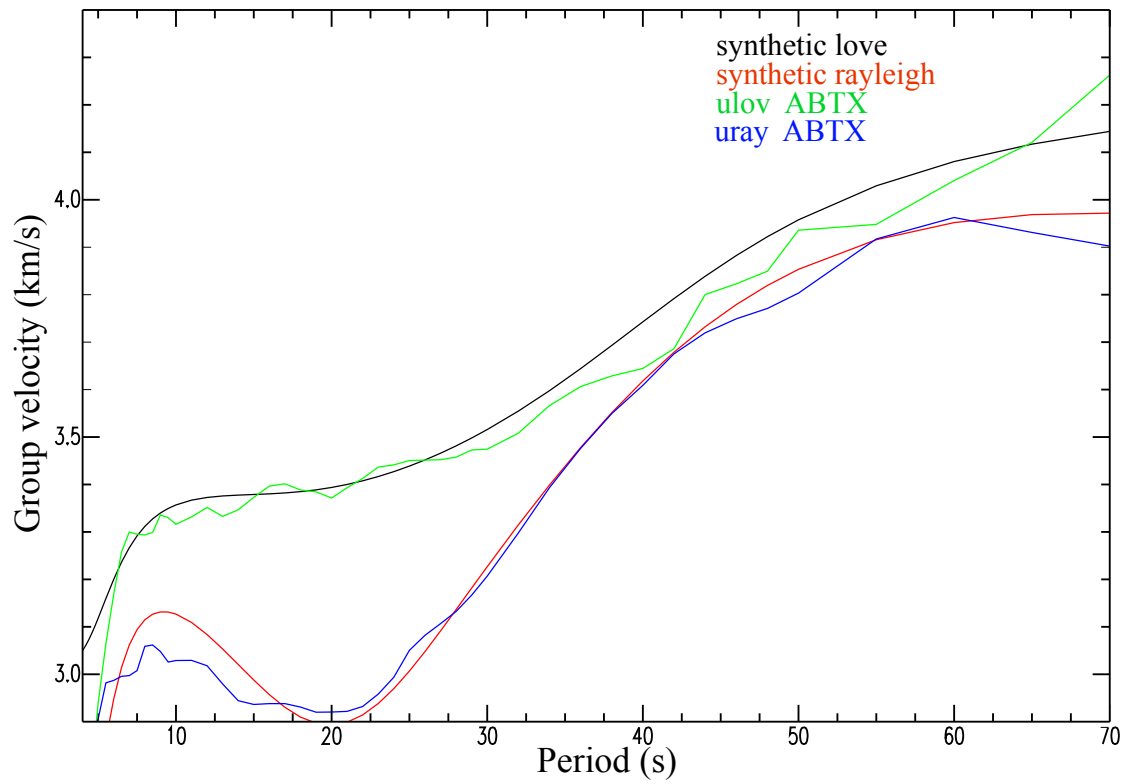


Figure 3.3: Surface wave dispersion curves (Love & Rayleigh) for station ABTX using real data.

3.5 Delay Travel Times

The traveltimes T between a source and receiver along a ray L is given in integral form for a velocity field as

$$T = \int_L \frac{ds}{v(s)} \quad (3.5.1)$$

where s is the position vector in 2D or 3D media. Travel times are considered a nonlinear inverse problem given the relationship between the measured data (travel times) and the unknown model parameters (the velocity field). However, by transforming variables to use slowness, the reciprocal of velocity, instead of velocity as the unknown, a seemingly linear inversion problem is created:

$$\int_L \Delta u(s) ds = \Delta T = T_{obs} - T_{pred} \quad (3.5.2)$$

However, the ray is also dependent on the velocity (or slowness) model, thus making the inverse problem nonlinear regardless of what form of model variable or parameterization is used. If the medium is subdivided into blocks, the path length l_j in the j -th block and can be discretize to

$$\Delta T = \sum_j l_j \Delta u_j \quad (3.5.3)$$

The model can be parameterized any number of ways using velocity or slowness, and cells, nodes, or splines, since the problems nonlinearity must be dealt with regardless of the parameterization. Most often a linearized gradient approach is applied in which a starting model is used and both the model and rays are updated over a series of iterations with the hope that there will be convergence to an acceptable model (the final model). The model is almost always discretized using cells, nodes, or other interpolating functions; in the latter two cases, the discrete model parameters are the coefficients of the interpolating functions. For the formulation of travel times for a tomography problem, the model is parameterized using constant-slowness cells, in which case the equation for the i -th data becomes

$$\Delta T_i = \sum_j l_{ij} \Delta u_j \quad (3.5.4)$$

where l_{ij} is the length of the i -th ray in the j -th model cell and Δu_j is the slowness in the j -th cell. In this case the path length of each ray in a block, l_{ij} is the partial derivative, $\partial T_i / \partial u_j$ of the travel time with respect to the slowness of that block [124].

3.6 Gravity Anomalies

In geophysics gravity anomalies are generally defined as the difference between observed gravity field and the field of a reference model. Depending on the reference gravity model, two different types of anomaly variations are considered: gravity anomalies and gravity disturbances. The geodetic gravity anomaly is defined as the difference between gravity on the geoid and normal gravity on the reference ellipsoid [45]. On the other hand, the gravity disturbance is defined as the difference of the fields at the same point on the reference ellipsoid. It has been demonstrated that the gravity disturbances are more appropriate for geophysical purposes (e.g., [41]). In any case, its necessary to take into account the difference in the interpretation.

The observed gravity anomalies reflect the effect of density variations relative to the homogeneous reference model. Interpretation of the gravity anomalies implies an estimation of the density heterogeneities. The density model should reproduce the observed gravity field, taking into account that the observations may be affected by measurement errors. Density heterogeneity of the Earth, associated with thermal and compositional variations or with deflected boundaries separating layers of different density, is one of the main factors that control dynamic processes and deformations at both shallow and deep levels. Therefore, interpretation of the gravity anomalies or gravity modeling is one of the principal methods, which help to understand the nature and origin of the tectonic processes and the Earths dynamics.

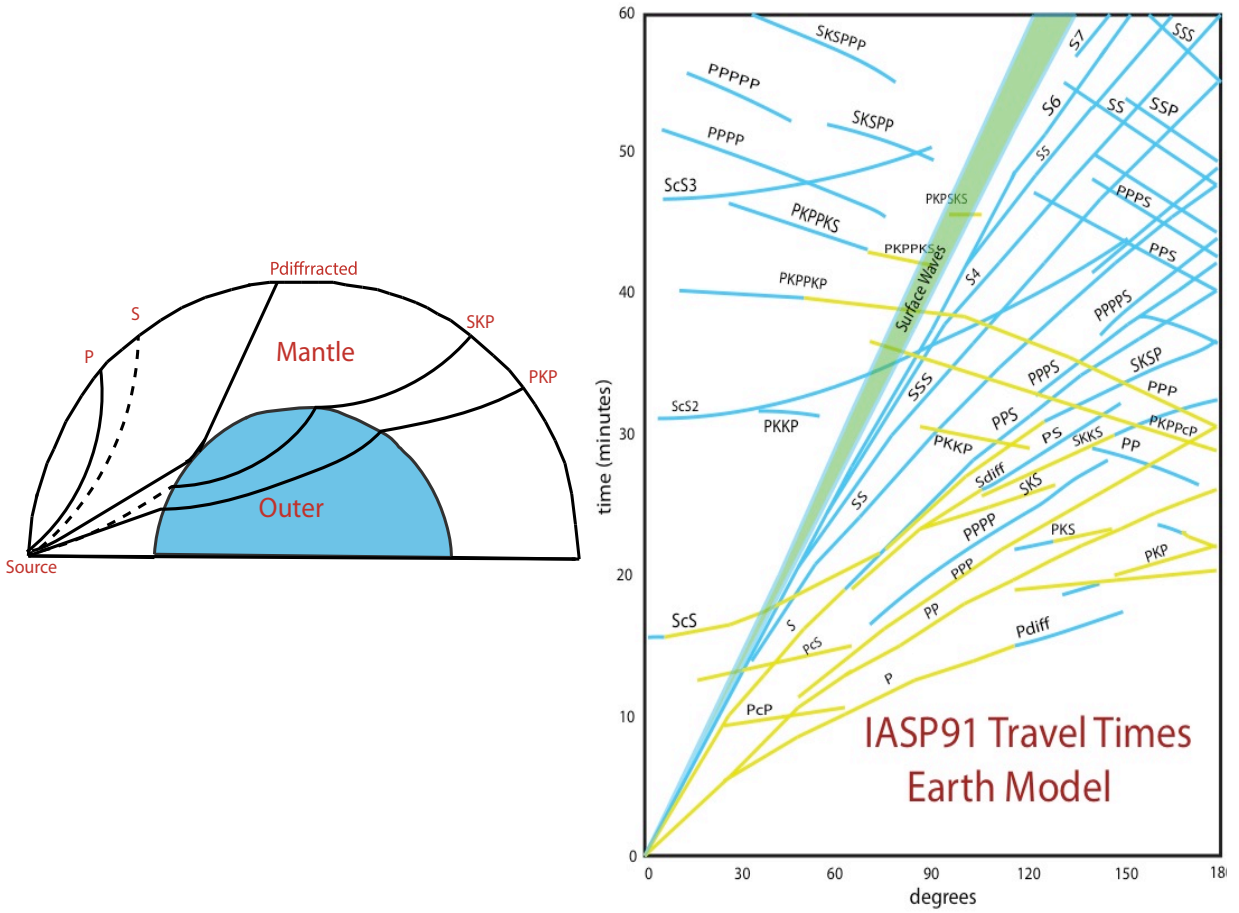


Figure 3.4: (Left) Halfspace showing different P and S waves created within the Earth. (Right) Different seismic phases within the Earth. Travel times are from the Array Network Facility (ANF) seismic catalog.

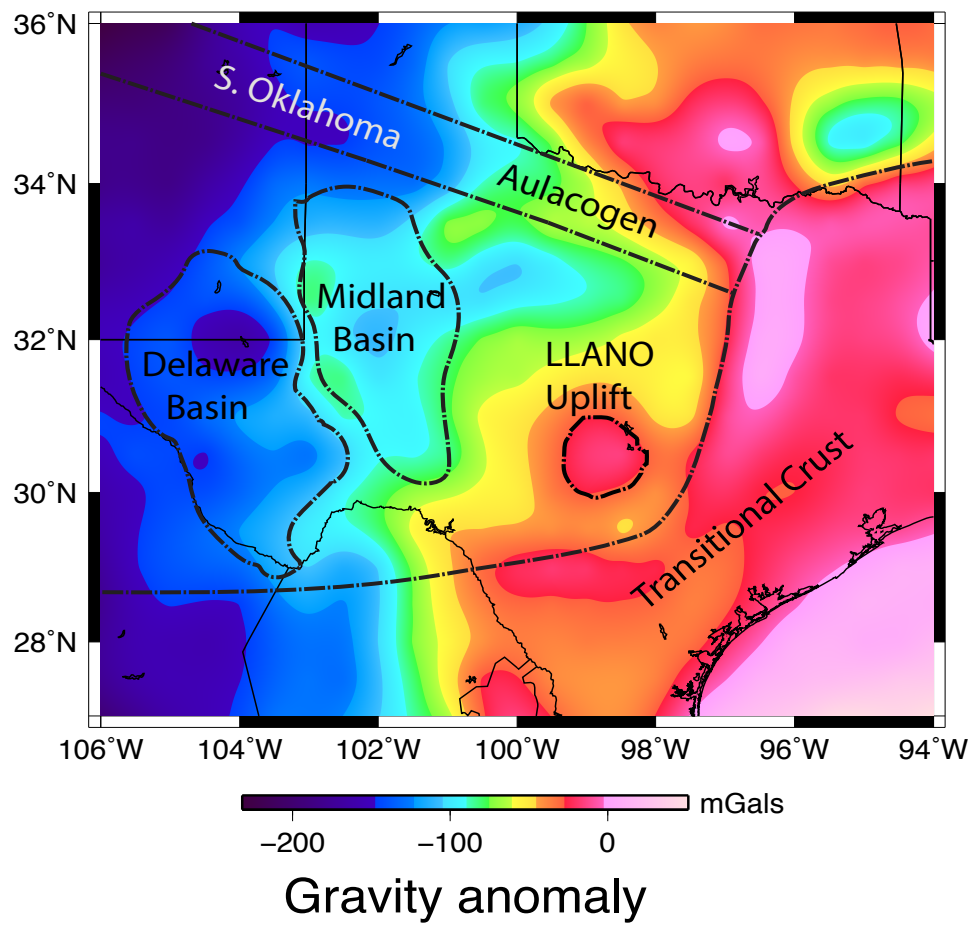


Figure 3.5: Bouguer gravity anomaly map of the Texas region and surrounding area. High amplitude gravity anomaly observed in Texas.

3.7 Forward Problem

If we know the layered shear velocity distribution $x = (x_1, \dots, x_n)$ at n different horizontal layers, then we can evaluate the measured quantities $y = (y_1, \dots, y_m)$ (e.g., the travel times) by applying an appropriate nonlinear operator $F(x)$ that uses the velocities x to predict the Earth's response $y = F(x)$;

$$F(x) = (F_1(x), \dots, F_m(x)) \in \mathbb{R}^m, x = (x_1, \dots, x_n) \in \mathbb{R}^n \ (m \gg n) \quad (3.7.1)$$

The operator F relates the data space and the model space. In other words, if we know the velocity model x , then we can predict the Earth's response based on the velocity model.

3.8 Inverse Problem

Given an observed data vector $y \in \mathbb{R}^m$, we want to find the unknown model x such that $F(x)$ approximates y as much as possible. For each specific type T of observations, this means that we are minimizing

$$\min_x \|F^T(x) - y^T\|^2 = \min_x (F_i^T(x) - y_i^T)^2 \quad (3.8.1)$$

to match measurements of different types, researchers traditionally use weighted non-linear least squares method (LSQ). For example, to simultaneously match the teleseismic receiver functions (RF), surface wave dispersion velocities (SW), travel times (TT), and gravity (GR), we minimize $\min_x J$, where

$$J = w_{RF}^2 \|F^{RF}(x) - y^{RF}\|^2 + w_{SW}^2 \|F^{SW}(x) - y^{SW}\|^2 + w_{TT}^2 \|F^{TT}(x) - y^{TT}\|^2 + w_{GR}^2 \|F^{GR}(x) - y^{GR}\|^2 \quad (3.8.2)$$

This minimization problem can be reformulated as

$$\min_x \|F(x) - y\|^2 \quad (3.8.3)$$

where

$$F(x) = W \begin{pmatrix} F^{SW}(x) \\ F^{RF}(x) \\ F^{TT}(x) \\ F^{GR}(x) \end{pmatrix} \in \mathbb{R}^m,$$

$$y = W \begin{pmatrix} y^{SW} \\ y^{RF} \\ y^{TT} \\ y^{GR} \end{pmatrix} \in \mathbb{R}^n$$

and

$$W = \text{diag}(w_i), w_i = \sqrt{\frac{\eta_1}{\sigma_i^2 p}}, i = 1, \dots, p, w_i = \sqrt{\frac{\eta_2}{\sigma_i^2 q}}, i = p + 1, \dots, p + q,$$

$$w_i = \sqrt{\frac{\eta_3}{\sigma_i^2 r}}, i = p + q + 1, \dots, p + q + r, \quad (3.8.4)$$

$$w_i = \sqrt{\frac{1 - \eta_1 - \eta_2 - \eta_3}{\sigma_i^2 s}}, i = p + q + r + 1, \dots, m = p + q + r + s$$

with W a weighted diagonal matrix used to equalize the contribution of each dataset with respect to physical units and number of data points, $\eta_i \in [0, 1]$ are influence parameters that measures the reliability of each dataset used for the inversion, σ_i^2 is the approximate standard deviation of each point, and p , q , r and s are the number of RF, SW, TT, and GR observations [121].

3.9 Need for mult-objective optimization

In practice, we do not know the exact values of the influence parameters. For different values of the influence parameters, we get, in general, different velocity distributions x ; some of these velocity models are geophysically meaningful, some are not (e.g., some models x predict higher velocities in the crust and lower velocities in the mantle contrary to geophysics).

Traditionally, researchers avoid non-physical non-smooth velocity models by adding a regularization term $\lambda||Lx||^2$ to the minimized function [128]. The problem with this term is that it is not clear how to select λ , and different values of λ lead to different solutions; see, e.g., [44] and [132].

In this work, instead of using regularization, we explicitly formulate constraints that need to be satisfied, for example, the desired smoothness can be described as a bound on $|x_i - x_j| \leq \Delta$ on the difference between velocities x_i and x_j at nearby locations. Then, we find the model x for which $J(x)$ is the smallest under these constraints. Additionally, we include bounds $a \leq x \leq b$ on the velocities at different depths. In geophysical applications, it is crucial to keep the physical parameters within appropriate bounds.

So, instead of selecting a single combination of influence parameters (and thus, of weights), we propose to use multi-objective optimization (MOP); namely, we generate all possible models x corresponding to different combinations of weights, and then we use one of the MOP criteria to select the most promising model [107],[108], [57].

In this case, we want to minimize the four criteria $f_1(x) = ||F^{RF}(x) - y^{RF}||^2$, $f_2(x) = ||F^{SW}(x) - y^{SW}||^2$, $f_3(x) = ||F^{TT}(x) - y^{TT}||^2$, $f_4(x) = ||F^{GR}(x) - y^{GR}||^2$. First, we find the Pareto optimal set P^* , i.e., the set of all feasible solutions x for which there is no other feasible solution x' which is better with respect to all criteria $f_1(x') < f_1(x), \dots, f_k(x') < f_k(x)$.

Definition: (Pareto Optimal Set) *For a given multi-objective problem*

$F(x) = (f_1(x), \dots, f_k(x))$, *the Pareto Optimal Set $P^* \triangleq$, is defined as:*

$$P \triangleq \{x \in \Omega | \neg \exists x' \in \Omega (F(x') \leq F(x))\} \quad (3.9.1)$$

It is known that elements of the Pareto set can be obtained by solving the one-objective (scalar) optimization problem.

$$\min_{x \in X} f(x) = \sum_{i=1}^k w_i f_i(x) \quad (3.9.2)$$

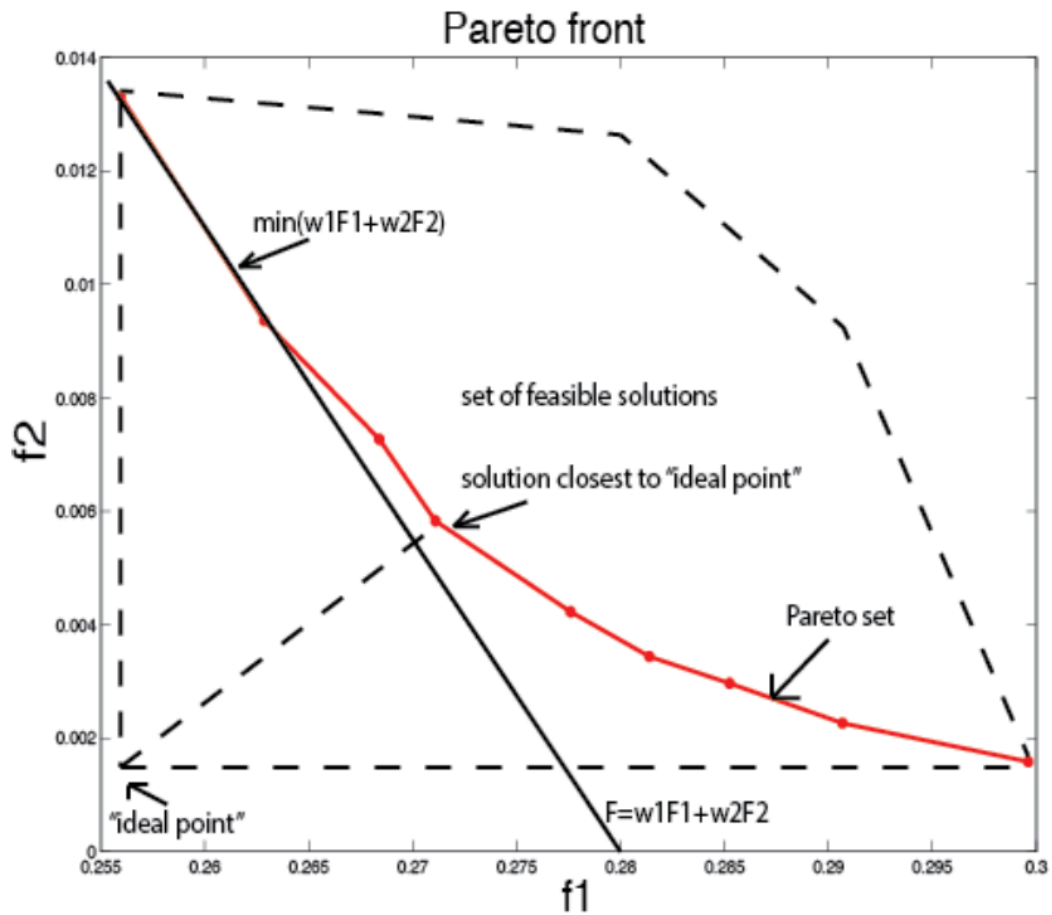


Figure 3.6: Illustration of the solution set or Pareto front, which is, defined as the weights times the perspective objective functions.

where $w = (w_1, \dots, w_k) \geq 0$ is the vector of weighting coefficients assigned prior to the solution of the problem. So, in our computations, we try all possible combinations of weights, and we find all solutions x corresponding to different combinations. For each criterion f_i , we then find the smallest value f_i^{\min} and the largest value f_i^{\max} . The smallest values form an ideal point $f^{\min} = (f_1^{\min}, \dots, f_4^{\min})$. We then select a solution x which is the closest to this ideal point. Specifically, we normalize each differences $f_i(x) - f_i^{\min}(x)$ to the interval (0,1) by dividing it by $f_i^{\max}(x) - f_i^{\min}(x)$, and then we minimize the corresponding normalized distance. In other terms, we select a solution x for which the distance

$$d^2(f^{\min}, f(x)) = \sum_{i=1}^k \left(\frac{f_i(x) - f_i^{\min}(x)}{f_i^{\max}(x) - f_i^{\min}(x)} \right)^2 \quad (3.9.3)$$

3.10 Numerical Algorithm

First, we use a first order Taylor approximation of the operator F around some suitable model \bar{x}_k :

$$F(x) \cong F(\bar{x}_k) + F'(\bar{x}_k)\Delta x = F(\bar{x}_k) + F'(\bar{x}_k)(x - \bar{x}_k), \quad (3.10.1)$$

where $F'(\bar{x}_k)$ is the matrix formed by the partial derivatives of F . Therefore, we rewrite the problem (9) as

$$\begin{aligned} \min_x \quad & \frac{1}{2} \|F'(\bar{x}_k)x + r(\bar{x}_k)\|^2 \\ \text{s.t.} \quad & g(\bar{x}_k) \geq 0 \end{aligned} \quad (3.10.2)$$

$$g(\bar{x}_k) = \begin{pmatrix} \bar{x}_k - a \\ b - \bar{x}_k \end{pmatrix}$$

where $r(\bar{x}_k) = F(\bar{x}_k) - y - F'(\bar{x}_k)\bar{x}_k$, and $g(\bar{x}_k)$ is a vector of constraints, including constraints $x_i - a_i \geq 0$ and $b_i - x_i \geq 0$ that describe the bounds $a_i \leq x_i \leq b_i$ on velocities x_i at different layers.

3.11 Primal Dual Interior-Point Method

To implement the Primal Dual Interior-Point method ([121];[83]), we first rewrite our problem in a standard form as follows:

$$\begin{aligned} \min_x \quad & \frac{1}{2} \|F'(\bar{x}_k)x + r(\bar{x}_k)\|^2 \\ \text{s.t.} \quad & g(\bar{x}_k) - s = 0 \\ & s \geq 0 \end{aligned} \tag{3.11.1}$$

where $s \in \mathbb{R}^{2n}$ is a slack variable. Then we define the Lagrange function associated to problem (16) as:

$$l(\bar{x}_k, z, s, w) = \frac{1}{2} \|F'(\bar{x}_k)x + r(\bar{x}_k)\|^2 - (g(\bar{x}_k) - s)^T z - s^T w \tag{3.11.2}$$

with the Lagrangian multipliers $z, w \in \mathbb{R}^{2n}$, $(z, w) \geq 0$. For a given perturbation parameter $\mu > 0$, the perturbed Karush-Kuhn-Tucker (KKT) or necessary conditions are given by:

$$\hat{F}(\bar{x}_k, z, s, w) = \begin{pmatrix} F'(\bar{x}_k)^T(F'(\bar{x}_k)x + r(\bar{x}_k)) - \nabla g^T(\bar{x}_k)z \\ g(\bar{x}_k) - s \\ z - w \\ SWe - \mu e \end{pmatrix} = 0 \tag{18}$$

where

$$\hat{F} : \mathbb{R}^{n+2n+2n} \longrightarrow \mathbb{R}^{n+2n+2n} \quad S = \text{diag}(s_1, \dots, s_{2n}), \quad W = \text{diag}(w_1, \dots, w_{2n})$$

and $e = (1, \dots, 1) \in \mathbb{R}^{2n}$. It is easy to see that $z = w$, hence the perturbed KKT system (18) is rewritten as

$$\hat{F}(x, z, s, w) = \begin{pmatrix} F'(\bar{x}_k)^T(F'(\bar{x}_k)x + r(\bar{x}_k)) - \nabla g^T(\bar{x}_k)z \\ g(\bar{x}_k) - s \\ SZe - \mu e \end{pmatrix} = 0 \tag{19}$$

thus the Jacobian associated to (19) is then computed as

$$F' \begin{pmatrix} x \\ z \\ s \end{pmatrix} = \begin{bmatrix} F'(\bar{x}_k)^T F'(\bar{x}_k) & -\nabla g^T(\bar{x}_k) & 0_{n \times n} \\ \nabla g(\bar{x}_k) & 0_{n \times m} & -I_{m \times m} \\ 0_{m \times n} & S & Z \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta z \\ \Delta s \end{bmatrix} = - \begin{bmatrix} \nabla_x l(x, z, s) \\ g(\bar{x}_k) - s \\ SZe - \mu e \end{bmatrix} \quad (20)$$

System (20) can be reduced further by eliminating the third block of equations as follows. From the last block of equation in (20) we have

$$S\Delta z + Z\Delta s = -SZe + \mu e,$$

therefore

$$Z\Delta s = -SZe + \mu e - S\Delta z$$

$$\Delta s = -s + \mu Z^{-1}e - Z^{-1}S\Delta z,$$

and then

$$\nabla g^T(\bar{x}_k)\Delta x - \Delta s = \nabla g^T(\bar{x}_k)\Delta x + s - \mu Z^{-1}e + Z^{-1}e + Z^{-1}S\Delta z = -\nabla g^T(\bar{x}_k)x + s$$

$$\nabla g^T(\bar{x}_k)\Delta x + Z^{-1}S\Delta z = \mu Z^{-1}e - g(\bar{x}_k)$$

which allow us to write the reduced linear system

$$\begin{bmatrix} -F'(\bar{x}_k)^T F'(\bar{x}_k) & \nabla g^T(\bar{x}_k) \\ \nabla g(\bar{x}_k) & Z^{-1}S \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta z \end{bmatrix} = \begin{bmatrix} \nabla_x l(x, z, s) \\ Z^{-1}\mu e - g(\bar{x}_k) \end{bmatrix} \quad (21)$$

3.12 Conclusion

In summary, for this study we propose to utilize the MOP technique to perform joint inversion of multiple data sets (Receiver functions and Surface Wave Dispersion). We will incorporate different weights in the MOP inversion scheme in order to map the Pareto Set (Solution Space) of receiver functions and surface wave dispersion measurements. We used the MOP technique to help characterize the crust and upper mantle of an ancient rift

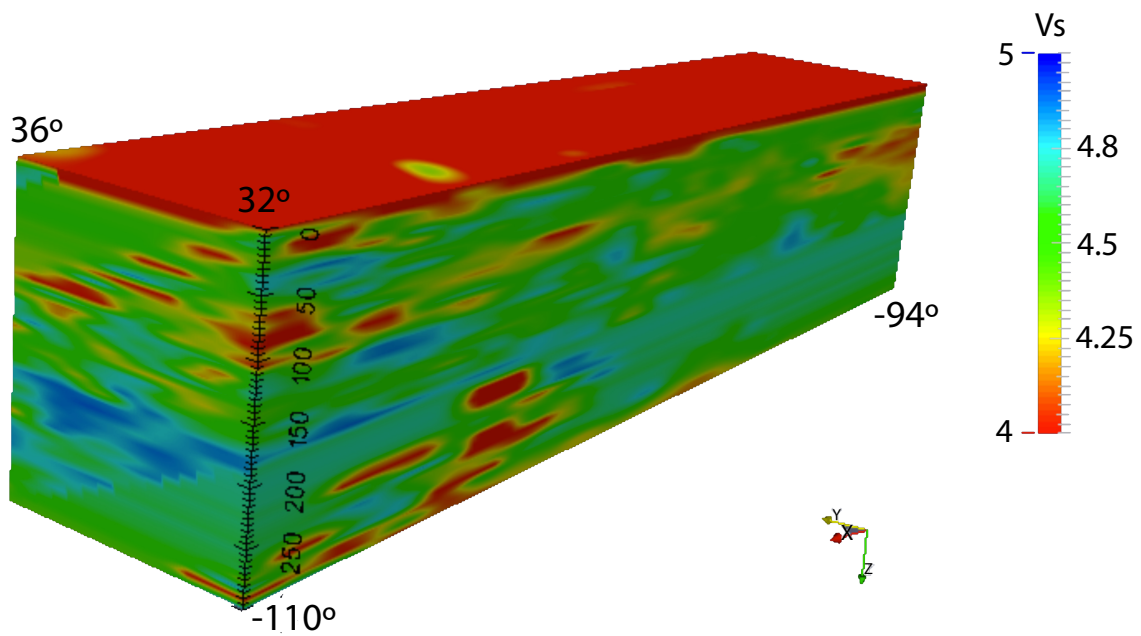


Figure 3.7: 3D shear wave model utilizing three geophysical datasets using MOP technique. Blue represents high velocities & red represents low velocities.

system in Texas using seismic data from USArray and Earthscope Network. We will extend the Primal Dual Interior Point Method (PDIP) algorithm with the MOP scheme in order to obtain high-resolution 3D imagery of Texas using teleseismic receiver functions, surface wave dispersion measurements, delay travel times, and gravity. We chose this optimization approach because we want to find the best possible solution for our nonlinear geophysics inverse problem. In geophysics, most inversion problems require finding some minimization. The optimization technique that we chose to solve our non-linear inverse problem requires the search of the global minimum and this technique will be able to define the entire solution based from using different weights to map the Pareto Set. From the Pareto Set, the MOP technique performs a direct search method that basically selects the final solution from a set of alternative solutions from the model space [107],[108], [57]. For future work, we plan to incorporate gravity into our 3D model to be able to obtain a more constrained earth structure model of Texas, which will allow us to help answer questions such as if the rift system is still actively deforming and how does the rift influence the evolution of adjacent areas within the North American Plate.

Chapter 4

A Constrained Multi-Objective Optimization Framework for Joint Inversion

Different geophysical data sets such as receiver functions, surface wave dispersion measurements, and first arrival travel times, provide complementary information about the Earth structure. To utilize all this information, it is desirable to perform a *joint inversion*, i.e., to use all these datasets when determining the Earth structure. In the ideal case, when we know the variance of each measurement, we can use the usual Least Squares approach to solve the joint inversion problem. In practice, we only have an approximate knowledge of these variances. As a result, if a geophysical feature appears in a solution corresponding to these approximate values of variances, there is no guarantee that this feature will still be visible if we use the actual (somewhat different) variances.

To make the joint inversion process more robust, it is therefore desirable to repeatedly solve the joint inversion problem with different possible combinations of variances. From the mathematical viewpoint, such solutions form a Pareto front of the corresponding multi-objective optimization problem. If a certain geological feature is visible in *all* these solutions, then we can be confident that this feature is also present in the actual solution corresponding to the actual (unknown) values of the variances – i.e., that it is the feature of the actual Earth structure.

In Chapter 4, we describe the corresponding constrained multi-objective optimization framework for multiple geophysical data sets. Chapter 4 is based on the work done in the

following publication:

Thompson, L., A. A. Velasco, and V. Kreinovich.: A constrained multi-objective optimization framework for multiple geophysical datasets, *Journal of Mathematical Geosciences* (2015), (in review).

4.1 Introduction

In this chapter, we describe how to combine multiple geophysical datasets for the purpose of assisting in better determining physical properties of the Earth structure. The need for combining different datasets comes from the fact that different datasets provide complementary information about the Earth structure. By jointly inverting multiple geophysical datasets, we combine these complementary pieces of information and thus, we get a more accurate description of the Earth structure; see, e.g., ([133], [51], [117], [25], [72], [116], [124]).

Specifically, we combine receiver functions, surface wave dispersion measurements, and P -wave travel times. The need to use different datasets comes from the fact that there are two types of seismic waves that travel through the Earth: the body waves and the surface waves. Both types of waves provide different sensitivities and information about the Earth structure, since they are sampling, correspondingly, the interior and surface of the Earth. The information collected from the body waves travels deeper into the Earth and translates into teleseismic P -wave receiver functions. In order to obtain information about the Earth surface, surface waves are analyzed, in our case, by means of surface waves dispersion. Receiver functions resolve discontinuities (impedance contrasts) in seismic velocities, and provide good measurement of crustal thickness, without providing a good average of shear wave velocity. On the other hand, we have surface (Love and Rayleigh) waves whose energy is concentrated near the Earth's surface, and provide good average of absolute shear wave velocity, without a good shear-wave velocity contrasts in layered structures ([51], [72], [116], [124], [22], [84]). Therefore these two data sets provide complementary information

about the Earth structure. Seismic first-arrival travel times are complementary to the other datasets because the travel time enable us to recover the causative slowness of the Earth structure ([62]).

For each dataset, we usually know the relative variance (uncertainty of data) of different measurement results from this dataset, and thus, we can use the Least Squares method to find the corresponding Earth model. For multiple datasets, we can sometimes still use the Least Squares Method to process all these datasets – provided that we know the variances of different measurements from different datasets. In practice, however, we usually only have an approximate knowledge of these variances. So, instead of producing a single model, several models corresponding to different possible variances are generated. If all these models – corresponding to different possible values of variances – contain a certain geophysical feature, then we can be certain that this feature is also present in the actual Earth model (which corresponds to the actual (unknown) values of the variances).

From the mathematical viewpoint, the task of computing all these models is equivalent to computing the Pareto front of the corresponding Multi-Objective Optimization problem [57], where different objective functions correspond to different datasets.

In addition to producing all these models, it is also desirable to produce a “typical” model, so that we only look for features which are present on this typical model. In this paper, we use methods for selecting such a typical model as described in [107], [108], [57].

This paper has the following structure. In Section 2, we describe, in detail, the need for multi-objective optimization. In Section 3, we show how to solve the corresponding optimization problems. In Section 4, we briefly describe the corresponding geophysical datasets. The results of applying multi-objective optimization technique to these datasets are shown and discussed in Section 5. Finally, Section 6 contains conclusions.

4.2 Need for Multi-Objective Optimization

4.2.1 Inverse Problems: Brief Reminder

In many real-life situations, we are interested in the values of the quantities x_1, x_2, \dots , which are difficult (or even impossible) to measure directly. For example, in geophysics, one of our main objectives is to find the shear velocities x_i at different 3-D points i (i.e., at different depths and at different geographic locations). To find these values, we measure easier-to-measure quantities y_1, y_2 , etc., which are related to $x = (x_1, \dots, x_n)$ by a known relation $y = F(x)$, and then use the measured quantities $y = (y_1, \dots, y_m)$ to find the desired values x .

In particular, in geophysics, to find shear velocities x_i , we can calculate the teleseismic receiver functions y^{RF} , surface wave dispersion velocities y^{SW} , travel times y^{TT} , etc. For each of these types of data, if we know the velocity model x , then we can predict the Earth's response by using the corresponding (known) operator F : $y^{RF} = F^{RF}(x)$, $y^{SW} = F^{SW}(x)$, $y^{TT} = F^{TT}(x)$, etc.

Measurements are never absolutely accurate, there is always some measurement inaccuracy, there is always some level of noise preventing us from measuring the corresponding quantities exactly. A usual way to estimate parameters in the presence of noise is to use the Least Squares method, i.e., to find the values x that minimize the expression

$$\sum_{i=1}^m \frac{(F_i(x) - y_i)^2}{\sigma_i^2}, \quad (4.2.1)$$

where σ_i is the standard deviation of the noise (measurement inaccuracy) of the i -th measurement.

In some cases, all the available data points come from measurements of the same type, obtained by using the same methodology and similar instrumentation. For example, we may only have travel times, or only surface wave dispersion velocities. In such cases, it is reasonable to assume that all these measured values have the same accuracy, i.e., that $\sigma_i = \text{const.}$ Under this assumption, minimizing the expression (4.2.1) is equivalent to

minimizing the sum of the squares

$$\|F(x) - y\|^2 \stackrel{\text{def}}{=} \sum_{i=1}^m (F_i(x) - y_i)^2. \quad (4.2.2)$$

4.2.2 Need to Take Constraints into Account

Sometimes, the models x obtained by an appropriate minimization are not physically meaningful. For example, in geophysics, some models x predict higher velocities in the crust and lower velocities in the mantle, contrary to known geophysical models. In other cases, when we expect a smooth dependence on a signal on time, the reconstructed signal x can exhibit abrupt non-physical changes.

To avoid such non-physical solutions, it is desirable to explicitly take into account the corresponding constraints. For example, in most practical problems, there are known physical bounds on the values of the quantities x_i . In particular, in geophysics, for each depth, we can approximate the lower bound and the upper bound on possible values of shear velocity x_i at this depth.

In precise terms, for every i , we know the bounds a_i and b_i such that $a_i \leq x_i \leq b_i$.

Restrictions on smoothness can be described as known bounds Δ_{ij} on the difference between the values x_i and x_j for nearby points of time or at nearby spatial locations: $-\Delta \leq x_i - x_j \leq \Delta_{ij}$.

Under the corresponding constraints, the optimization problem (4.2.2) takes the form

$$\begin{aligned} \min_x \|F(x) - y\|^2 \\ \text{s.t. } g(x) \geq 0. \end{aligned} \quad (4.2.3)$$

where $g(x)$ is a vector consisting of the corresponding constraints. For example, to describe bounds a_i and b_i on the values x_i , we use constraints $x_i - a_i \geq 0$ and $b_i - x_i \geq 0$.

Constraints corresponding to smoothness can also be expressed in the form $g(x) \geq 0$, with the corresponding components of the vector $g(x)$ having the form $\Delta_{ij} - (x_i - x_j)$ and $(x_i - x_j) - (-\Delta_{ij})$.

Comment. Traditionally, researchers avoid non-physical non-smooth velocity models by adding a regularization term $\lambda \|Lx\|^2$ to the minimized function; see, (e.g., [128]). The problem with this term is that it is not clear how to select λ , and different values of λ lead to different solutions; see, (e.g., [44],[132]). Because of this problem, in this paper, instead of using regularization, we explicitly formulate constraints that need to be satisfied. For example, the desired smoothness is described as a bound on the differences $x_i - x_j$.

4.2.3 Joint Inversion: Idealized Case

As noted earlier, measurements of different type usually provide complementary information and it is, therefore, beneficial to use measurement results of all the types.

When we use measurements of different types t , t' , etc., then while it is reasonable to assume that all the measurements i of the same type t have the same standard deviation $\sigma_i = \sigma^t$, standard deviations of measurements of different types are, in general, different: $\sigma^t \neq \sigma^{t'}$. Let us first consider the idealized case, when we assume that we know the accuracy σ^t of measurements of type t .

In this case, we can still use the Least Squares expression (4.2.1) to find the desired model x . By grouping together measurements of different type, we can rewrite the expression (4.2.1) in the following form:

$$\sum_{i=1}^m \frac{(F_i(x) - y_i)^2}{\sigma_i^2} = \sum_{t=1}^T \sum_{i \in t} \frac{(F_i(x) - y_i)^2}{(\sigma^t)^2} = \sum_{t=1}^T \frac{1}{(\sigma^t)^2} \cdot \left(\sum_{i \in t} (F_i(x) - y_i)^2 \right) \quad (4.2.4)$$

where T is the total number of different types of measurements, and the notation $i \in t$ indicates that the i -th measurement is of type t .

We can rewrite this expression as

$$\sum_{t=1}^T c_t^2 \cdot \|F^t(x) - y^t\|^2, \quad (4.2.5)$$

where $c_t \stackrel{\text{def}}{=} \frac{1}{\sigma^t}$, y^t is the list (tuple) consisting of all measurements of type t , and $F^t(x)$ is the list consisting of all the corresponding values $F_i(x)$.

To find the desired solution x , we must minimize this expression under the constraint $g(x) \geq 0$.

4.2.4 Reformulation of the Problem

The more measurements of a given type t we have, the larger the contribution of these measurements to the solution. In general, all the terms $\frac{(F_i(x) - y_i)^2}{\sigma_i^2}$ in the sum (4.2.4) are approximately of the same type, so we can conclude that the joint contribution of all the measurements of type t is proportional to the number m_t of all measurement results of this type.

To compare the importance of measurements of different types, it is useful to define *relative importance* as a ratio of the importance m_t of this type to the overall value $\sum_{t'} m_{t'}$, i.e., as $\eta_t \stackrel{\text{def}}{=} \frac{m_t}{\sum_{t'} m_{t'}}$. These *influence parameters* η_t are non-negative numbers that add up to 1.

To better understand the minimized expression (4.2.5), it makes sense to show the explicit dependence on the influence parameters. This can be done, e.g., by making sure that each term in the new formula is proportional to η_t ; this way we will see that the larger the influence parameter, the larger the influence of this term. To do that, we replace each term c_t^2 with $\eta_t \cdot k_t^2$. From the condition that $c_t^2 = \eta_t \cdot k_t^2$, we conclude that $k_t^2 = \frac{c_t^2}{\eta_t}$. Since $c_t = \frac{1}{\sigma^t}$ and $\eta_t = \frac{m_t}{\sum_{t'} m_{t'}}$, we have

$$k_t = \sqrt{\frac{c_t^2}{\eta_t}} = \sqrt{\frac{1}{(\sigma^t)^2 \cdot m_t}} \cdot \sqrt{\sum_{t'} m_{t'}}. \quad (4.2.6)$$

Thus, each term $c_t^2 = \eta_t \cdot k_t^2$ has the form $c_t^2 = w_t^2 \cdot C$, where $w_t \stackrel{\text{def}}{=} \sqrt{\frac{\eta_t}{(\sigma^t)^2 \cdot m_t}}$ and $C \stackrel{\text{def}}{=} \sum_{t'} m_{t'}$. So, the minimized expression (4.2.5) takes the form

$$C \cdot \sum_{t=1}^T w_t^2 \cdot \|F^t(x) - y^t\|^2. \quad (4.2.7)$$

The location of the minimum does not change if we divide all the values of a function by the same positive coefficient C . Therefore, minimizing the expression (4.2.7) is equivalent to minimizing a simpler expression

$$\sum_{t=1}^T w_t^2 \cdot \|F^t(x) - y^t\|^2. \quad (4.2.8)$$

4.2.5 General Case: A Description

In the previous sections, we considered the ideal case, when we know the exact variance σ_i^2 of each measurement i . In this case, we can use the usual Least Squares approach to solve the joint inversion problem.

In practice, we only have an *approximate* knowledge of these variances. For example, for each measurement type t , we only know an approximate value $\tilde{\sigma}^t$ of the corresponding standard deviation σ^t .

A traditional approach to such situations is to use these approximate values $\tilde{\sigma}^t$ and solve the corresponding optimization problem. The problem with this approach is that, if a geophysical features appears in the solution corresponding to these approximate values of variances, there is no guarantee that this feature will still be visible if we use the actual (somewhat different) variances.

It is desirable to separate artifacts that are due to the specific choice of variances from the phenomena that occur no matter what variances we use. For this purpose, we wish to repeatedly solve the joint inversion problem with different possible combinations of variances. If a certain geological feature is visible in *all* these solutions, then we can be confident that this feature is also present in the actual solution corresponding to the actual (unknown) values of the variances – i.e., that it is the feature of the actual Earth structure.

4.2.6 General Case: Analysis of the Problem

In the previous sections, we described the need to minimize the expression

$$\sum_{t=1}^T \frac{1}{(\sigma^t)^2} \cdot \|F^t(x) - y^t\|^2 \quad (4.2.9)$$

under the condition $g(x) \geq 0$, where σ^t are the known standard deviations. We showed that this equivalent to minimizing the expression (4.2.8), where $w_t = \sqrt{\frac{\eta_t}{(\sigma^t)^2 \cdot m_t}}$ and $\eta_t = \frac{m_t}{\sum_{t'} m_{t'}}$.

In situations when we only know an approximate value $\tilde{\sigma}^t$, the traditional approach would be to use this approximate value, i.e., to minimize the expression

$$\sum_{t=1}^T \frac{1}{(\tilde{\sigma}^t)^2} \cdot \|F^t(x) - y^t\|^2, \quad (4.2.10)$$

or, equivalently, to minimize the expression (4.2.8), in which

$$w_t = \sqrt{\frac{\eta_t}{(\tilde{\sigma}^t)^2 \cdot m_t}}, \quad (4.2.11)$$

with the same values $\eta_t = \frac{m_t}{\sum_{t'} m_{t'}}$ of the influence parameters.

As we have mentioned, a more appropriate approach is to minimize expressions (4.2.9) corresponding to *all possible* combinations of standard deviations σ^t . Let us show that:

- each such constraint minimization problem can be equivalently reformulated into the form (4.2.8) with the weights (4.2.11) if we select *different* values of the influence parameters $\eta_t > 0$, and that
- for each combination of influence parameters $\eta_t > 0$ with $\sum_t \eta_t = 1$, there exist values $\sigma^t > 0$ for which the corresponding optimization problem (4.2.8) is equivalent to the original optimization problem (4.2.9).

Indeed, for each combination of values σ^t , let us take

$$\eta_t \stackrel{\text{def}}{=} \frac{1}{c} \cdot \frac{m_t \cdot (\tilde{\sigma}^t)^2}{(\sigma^t)^2},$$

where the normalization coefficient c is chosen as

$$c \stackrel{\text{def}}{=} \sum_t \frac{m_t \cdot (\tilde{\sigma}^t)^2}{(\sigma^t)^2}.$$

In this case, $\eta_t > 0$ and $\sum_t \eta_t = 1$. For the corresponding weights (4.2.11), we get

$$w_t^2 = \frac{\eta_t}{(\tilde{\sigma}^t)^2 \cdot m_t} = \frac{c}{(\sigma^t)^2}. \quad (4.2.12)$$

Thus, minimizing the expression (4.2.8) is equivalent to minimizing the expression

$$c \cdot \sum_{t=1}^T \frac{1}{(\sigma^t)^2} \cdot \|F^t(x) - y^t\|^2,$$

which, in its turn, is equivalent to minimizing the expression (4.2.9).

Vice versa, for each combination of weights $\eta_t > 0$ for which $\sum_t \eta_t = 1$, we can take

$$\sigma^t = \sqrt{\frac{m_t}{\eta_t \cdot \sum_{t'} m_{t'}}} \cdot \tilde{\sigma}^t. \quad (4.2.13)$$

For these standard deviations σ^t , the expression (4.2.9) takes the form

$$\sum_{t=1}^T (w'_t)^2 \cdot \|F^t(x) - y^t\|, \quad (4.2.14)$$

where we denoted

$$(w'_t)^2 \stackrel{\text{def}}{=} \sum_{t'} m_{t'} \cdot \frac{\eta_t}{(\tilde{\sigma}^t)^2 \cdot m_t}. \quad (4.2.15)$$

From the formula (4.2.11), we conclude that $w_t^2 = \frac{\eta_t}{(\tilde{\sigma}^t)^2 \cdot m_t}$ and thus, that the formula (4.2.15) takes the form $(w'_t)^2 = C \cdot w_t^2$, where $C \stackrel{\text{def}}{=} \sum_{t'} m_{t'}$. Thus, the expression (4.2.14) takes an equivalent form

$$C \cdot \sum_{t=1}^T w_t^2 \cdot \|F^t(x) - y^t\|, \quad (4.2.16)$$

and the minimization of this expression is indeed equivalent to minimizing the expression (4.2.8).

The equivalence is proven.

Comment. The above analysis holds when we know the approximate values $\tilde{\sigma}^t$ of the corresponding accuracies σ^t , but we know no guaranteed bounds on these accuracies. In some practical situations, in addition to the approximate values $\tilde{\sigma}^t$, we also know the bounds $\underline{\sigma}^t$ and $\bar{\sigma}^t$, for which $\underline{\sigma}^t \leq \sigma^t \leq \bar{\sigma}^t$. In this case, instead of all possible values η_t of the corresponding influence parameter, we only need to consider possible values – and we can use the above formulas relating η_t with σ^t to transform bounds on σ^t into the corresponding bounds on η_t .

4.2.7 General Case: Resulting Optimization Problem

So, we arrive at the following equivalent reformulation of our problem: for all possible combination of influence factors $\eta_t > 0$ for which $\sum_t \eta_t = 1$, we compute the weights

$$w_t = \sqrt{\frac{\eta_t}{(\tilde{\sigma}^t)^2 \cdot m_t}},$$

and then minimize the expression

$$\sum_{t=1}^T w_t^2 \cdot \|F^t(x) - y^t\|^2$$

under the corresponding constraints $g(x) \geq 0$.

This minimization problem can be reformulated as

$$\min_x \|F(x) - y\|^2, \tag{4.2.17}$$

where

$$F(x) = W \begin{pmatrix} F^1(x) \\ F^2(x) \\ \cdots \\ F^t(x) \\ \cdots \\ F^T(x) \end{pmatrix} \in \mathbb{R}^m, \quad y = W \begin{pmatrix} y^1 \\ y^2 \\ \cdots \\ y^t \\ \cdots \\ y^T \end{pmatrix} \in \mathbb{R}^n$$

and $W = \text{diag}(w_t)$ is a diagonal matrix whose elements will be called *weights*.

In particular, when we reconstruct the values of shear velocities from Receiver Functions (RF), Surface Waves (SW), and Travel Times (TT), the corresponding minimize functional $J(x)$ takes the form

$$J = w_{RF}^2 \|F^{RF}(x) - y^{RF}\|^2 + w_{SW}^2 \|F^{SW}(x) - y^{SW}\|^2 + w_{TT}^2 \|F^{TT}(x) - y^{TT}\|^2,$$

i.e., equivalently, form (4.2.17), with

$$F(x) = W \begin{pmatrix} F^{SW}(x) \\ F^{RF}(x) \\ F^{TT}(x) \end{pmatrix} \in \mathbb{R}^m, \quad y = W \begin{pmatrix} y^{SW} \\ y^{RF} \\ y^{TT} \end{pmatrix} \in \mathbb{R}^n$$

and

$$\begin{aligned} W = \text{diag}(w_i), w_i &= \sqrt{\frac{\eta_1}{(\tilde{\sigma}_i)^2 p}}, i = 1, \dots, p, w_i = \sqrt{\frac{\eta_2}{(\tilde{\sigma}_i)^2 q}}, i = p+1, \dots, p+q, \\ w_i &= \sqrt{\frac{1 - (\eta_1 + \eta_2)}{(\tilde{\sigma}_i)^2 r}}, i = p+q+1, \dots, p+q+r, \end{aligned} \quad (4.2.18)$$

$\tilde{\sigma}_i$ is the approximate standard deviation of each point, and p , q , and r are the number of RF, SW, and TT observations [121].

4.2.8 Relation to Multi-Objective Optimization

When we have observations of only one type t , then, to find the corresponding model, we minimize the function $\|F^t(x) - y^t\|^2$. Minimizing this function is equivalent to minimizing the expression

$$f_t(x) \stackrel{\text{def}}{=} \frac{1}{(\tilde{\sigma}^t)^2 \cdot m_t} \cdot \|F^t(x) - y_t\|^2.$$

In situations when we have observations of different type, and when we only know the approximate values $\tilde{\sigma}^t$ of the corresponding accuracies, we need to minimize expressions

(4.2.8), i.e., equivalently, expressions of the type

$$\sum_t \eta_t \cdot f_t(x). \quad (4.2.19)$$

corresponding to all possible combinations $\eta_t > 0$ for which $\sum_t \eta_t = 1$.

It is known that, under reasonable conditions, the resulting set of solutions can be described in terms of the corresponding *multi-objective optimization problem* (MOP), namely, the problem of optimizing

$$f(x) \stackrel{\text{def}}{=} (f_1(x), f_2(x), \dots, f_t(x), \dots, f_T(x)).$$

For solving multi-objective problems, a natural idea is to generate the *Pareto optimal set* (also known as *Pareto front*), i.e., the set of all the values x for which it is not possible to improve all the criteria $f_i(x)$. In precise terms, the Pareto optimal set $P(x)$ is defined as

$$P(x) = \{x \in \Omega : \nexists x' \in \Omega (f(x') < f(x)), \quad (4.2.20)$$

where Ω is the set of all possible solutions that satisfy the corresponding constraints, and $f(x') < f(x)$ means that

$$\forall t (f_t(x') \leq f_t(x)) \& \exists t (f_t(x') < f_t(x)). \quad (4.2.21)$$

Under reasonable conditions, elements of the Pareto set can be obtained by finding the minima of all the functions (4.2.19) corresponding to all possible weights η_t adding to 1, and, vice versa, each such minimum is an element of the Pareto set $P(x)$.

In these terms, we can say that what we want in the general case, when we only know the approximate values of the corresponding accuracies, is to find the Pareto set of the multi-objective problem in which we minimize the criteria

$$f_t(x) = \text{const} \cdot \|F^t(x) - y_t\|^2$$

corresponding to measurements of different types t .

In particular, in our geophysical example, we want to minimize the three criteria $f_{FR}(x) = \text{const} \cdot \|F^{RF}(x) - y^{RF}\|^2$, $f_{SW}(x) = \text{const} \cdot \|F^{SW}(x) - y^{SW}\|^2$, and $f_{TT}(x) = \text{const} \cdot \|F^{TT}(x) - y^{TT}\|^2$.

4.2.9 How to Generate a “Typical” Solution

When we have data of several types, and we only know approximate values of the corresponding accuracies, our recommendation is to generate solutions corresponding to all possible combinations of actual accuracies. The number of such solutions is huge, and meaningfully analyzing all these solutions is difficult. It is therefore desirable to select, among these solutions, a solution which is, in some sense, typical.

The expectation is that, in general, this “typical” solution will only have features that all other solutions have. Thus, when we look for features common for all possible solutions, a good idea is to first analyze this typical solution, and then to check whether the features that we found on this solution are indeed present in all other solutions as well.

In multi-objective optimization, there are several possible ways of generating such a “typical” solution; see, e.g., ([107], [108], [57]). For example, once we find all solutions x corresponding to different combinations, then, for each criterion f_t , we can find the smallest value f_t^{\min} and the largest value f_t^{\max} . The smallest values form an *ideal point* $f^{\min} = (f_1^{\min}, \dots, f_t^{\min}, \dots, f_T^{\min})$. We then select a solution x which is the closest to this ideal point. Specifically, we normalize each differences $f_t(x) - f_t^{\min}(x)$ to the interval (0,1) by dividing it by $f_t^{\max}(x) - f_t^{\min}(x)$, and then we minimize the corresponding normalized distance. In other terms, we select a solution x for which the distance

$$d^2(f^{\min}, f(x)) = \sum_{t=1}^T \left(\frac{f_t(x) - f_t^{\min}(x)}{f_t^{\max}(x) - f_t^{\min}(x)} \right)^2 \quad (4.2.22)$$

is the smallest possible.

4.3 Solving the Corresponding Constraint Optimization Problems

In the proposed approach, we need to solve several minimization problems, corresponding to different combinations of the influence parameters η_t . For each such combination, we

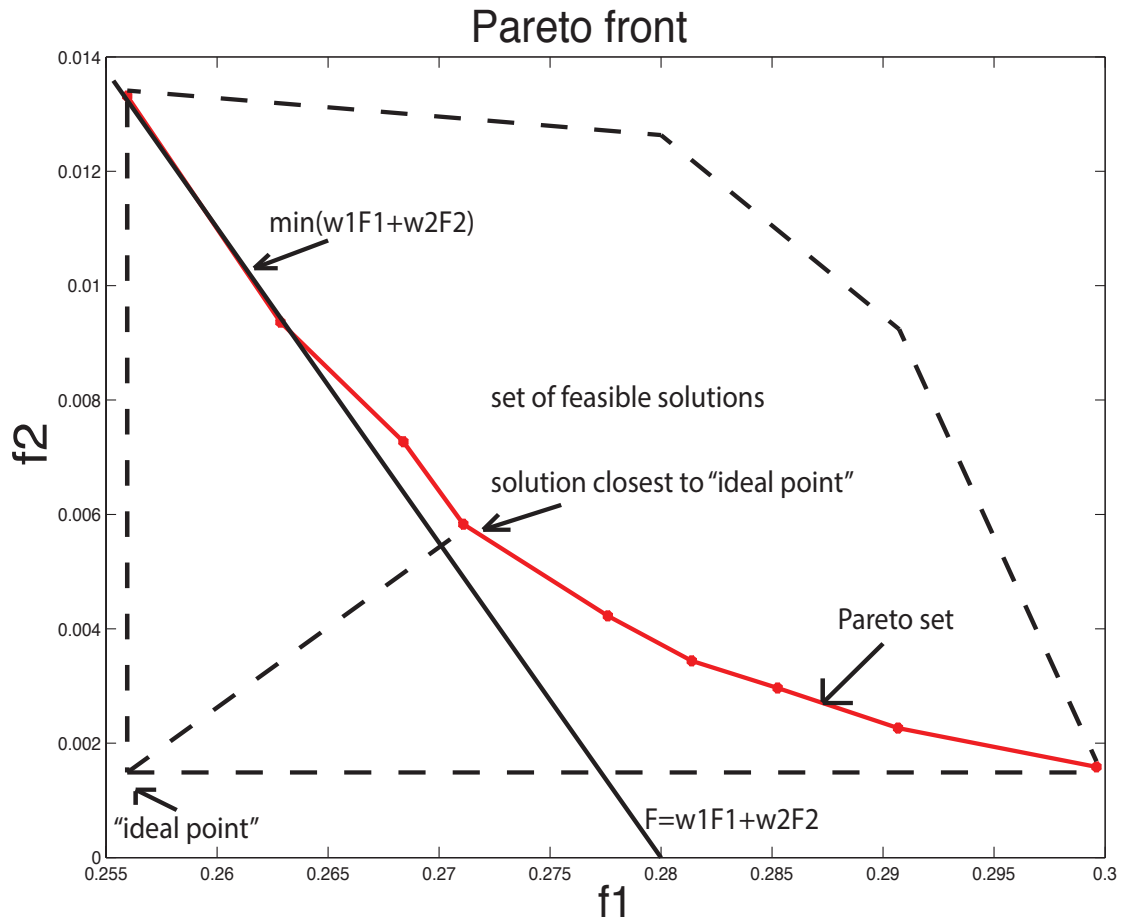


Figure 4.1: Illustration of the solution set or Pareto front, which is, defined as the weights times the perspective objective functions.

need to find the value $\|F(x) - y\|^2$ under the constraint $g(x) \geq 0$. Let us show how to solve the corresponding constraint optimization problems.

4.3.1 Linearization

In most practical situations, we know the approximate values \bar{x}_1 of the corresponding quantities x , i.e., values for which $x \approx \bar{x}_0$. Since these values are close to each other, the difference $x - \bar{x}_0$ is small and thus, we can expand the functions $F(x)$ and $g(x)$ into Taylor series in this difference and safely ignore terms which are quadratic (or higher order), and only keep the first order Taylor approximation. Once we solve this first-order approximation problem, we get a better approximation \bar{x}_1 . We can use this approximation as the basis of a new linearization and get an even more accurate approximation, etc.

On each of these iterations, we start with an approximate model \bar{x}_k , and then we use the first order Taylor approximation of the operators F and g around \bar{x}_k :

$$F(x) \approx F(\bar{x}_k) + F'(\bar{x}_k)\Delta x = F(\bar{x}_k) + F'(\bar{x}_k)(x - \bar{x}_k), \quad (4.3.1)$$

$$g(x) \approx g(\bar{x}_k) + \nabla g^T(\bar{x}_k)(x - \bar{x}_k), \quad (4.3.2)$$

where $F'(\bar{x}_k)$ is the matrix formed by the partial derivatives of F and ∇g is the matrix formed by the partial derivatives of different components of $g(x)$.

When we substitute these expressions into our constraint optimization problem, the minimized function becomes quadratic. By the chain rule, the derivative of a quadratic function $(f(x))^2$ is equal to $2 \cdot f(x) \cdot f'(x)$; to simplify the resulting formulas, we can thus replace the original expression $(f(x))^2$ with $\frac{1}{2} \cdot (f(x))^2$, then the derivative is simpler: $2 \cdot f(x) \cdot f'(x)$. So, to simplify our formulas, instead of the problem of minimizing the expression $\|F(x) - y\|^2$, we consider the equivalent problem of minimizing the expression $\frac{1}{2} \cdot \|F(x) - y\|^2$ under the constraint $g(x) \geq 0$.

Substituting the expressions (4.3.1) and (4.3.2) in the corresponding constraint optimization problem, we get the following linearized constraint optimization problem:

$$\min_x \frac{1}{2} \|F'(\bar{x}_k)x + r(\bar{x}_k)\|^2$$

$$\text{s.t. } g(\bar{x}_k) + \nabla g^T(\bar{x}_k)(x - \bar{x}_k) \geq 0, \quad (4.3.3)$$

where $r(\bar{x}_k) \stackrel{\text{def}}{=} F(\bar{x}_k) - y - F'(\bar{x}_k)\bar{x}_k$.

4.3.2 Primal Dual Interior-Point Method

To solve the linearized problem (4.3.3), we will use the Primal Dual Interior-Point method; see, e.g., ([121], [83]). To use this method, we first reformulate our problem in a standard form as follows:

$$\begin{aligned} \min_x \quad & \frac{1}{2} \|F'(\bar{x}_k)x + r(\bar{x}_k)\|^2 \\ \text{s.t. } \quad & g(\bar{x}_k) + \nabla g^T(\bar{x}_k)(x - \bar{x}_k) - s = 0 \\ & s \geq 0 \end{aligned} \quad (4.3.4)$$

where components of $s \in \mathbb{R}^{2n}$ are called *slack variables*.

Then we define the Lagrange function associated to problem (4.3.4) as:

$$l(\bar{x}_k, z, s, w) = \frac{1}{2} \|F'(\bar{x}_k)x + r(\bar{x}_k)\|^2 - (g(\bar{x}_k) + \nabla g^T(\bar{x}_k)(x - \bar{x}_k) - s)^T z - s^T w \quad (4.3.5)$$

with the Lagrangian multipliers $z, w \in \mathbb{R}^{2n}$, $(z, w) \geq 0$. For a given perturbation parameter $\mu > 0$, the perturbed Karush-Kuhn-Tucker (KKT) or necessary conditions are given by:

$$\hat{F}(\bar{x}_k, z, s, w) = \begin{pmatrix} F'(\bar{x}_k)^T(F'(\bar{x}_k)x + r(\bar{x}_k)) - \nabla g^T(\bar{x}_k)z \\ g(\bar{x}_k) + \nabla g^T(\bar{x}_k)(x - \bar{x}_k) - s \\ z - w \\ SWe - \mu e \end{pmatrix} = 0, \quad (4.3.6)$$

where

$$\hat{F} : \mathbb{R}^{n+2n+2n} \longrightarrow \mathbb{R}^{n+2n+2n} \quad S = \text{diag}(s_1, \dots, s_{2n}), \quad W = \text{diag}(w_1, \dots, w_{2n}) \quad (4.3.7)$$

and $e = (1, \dots, 1) \in \mathbb{R}^{2n}$.

The formula (4.3.6) implies, in particular, that $z - w = 0$, and thus $z = w$. Hence the perturbed KKT system (4.3.6) is rewritten as

$$\hat{F}(x, z, s, w) = \begin{pmatrix} F'(\bar{x}_k)^T(F'(\bar{x}_k)x + r(\bar{x}_k)) - \nabla g^T(\bar{x}_k)z \\ g(\bar{x}_k) - s \\ SZe - \mu e \end{pmatrix} = 0; \quad (4.3.8)$$

thus the Jacobian associated to (4.3.8) is then computed as

$$F' \begin{pmatrix} x \\ z \\ s \end{pmatrix} = \begin{bmatrix} F'(\bar{x}_k)^T F'(\bar{x}_k) & -\nabla g^T(\bar{x}_k) & 0_{n \times n} \\ \nabla g(\bar{x}_k) & 0_{n \times m} & -I_{m \times m} \\ 0_{m \times n} & S & Z \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta z \\ \Delta s \end{bmatrix} = - \begin{bmatrix} \nabla_x l(x, z, s) \\ g(\bar{x}_k) - s \\ SZe - \mu e \end{bmatrix}$$

The system (4.3.9) can be simplified further by eliminating the third block of equations as follows. From the last block of equation in (4.3.9), we have

$$S\Delta z + Z\Delta s = -SZe + \mu e,$$

therefore

$$Z\Delta s = -SZe + \mu e - S\Delta z$$

$$\Delta s = -s + \mu Z^{-1}e - Z^{-1}S\Delta z,$$

and then

$$\nabla g^T(\bar{x}_k)\Delta x - \Delta s = \nabla g^T(\bar{x}_k)\Delta x + s - \mu Z^{-1}e + Z^{-1}e + Z^{-1}S\Delta z = -\nabla g^T(\bar{x}_k)x + s$$

$$\nabla g^T(\bar{x}_k)\Delta x + Z^{-1}S\Delta z = \mu Z^{-1}e - g(\bar{x}_k)$$

which allow us to write the reduced linear system

$$\begin{bmatrix} -F'(\bar{x}_k)^T F'(\bar{x}_k) & \nabla g^T(\bar{x}_k) \\ \nabla g(\bar{x}_k) & Z^{-1}S \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta z \end{bmatrix} = \begin{bmatrix} \nabla_x l(x, z, s) \\ Z^{-1}\mu e - g(\bar{x}_k) \end{bmatrix}$$

4.4 Geophysical Datasets

4.4.1 Receiver Functions

A receiver function is simply a time series representation of the Earth’s response relative to an incoming *P*-wave propagating near a recording station. A representation of a receiver function is indicated in figure 2, with different incoming *P*-to-*S* converted waves and a time series representation of the Earth response underneath a seismic station. Positive or negative spike amplitudes represent positive or negative seismic velocity contrast. A receiver function technique can model the structure of the earth by using seismograms from three component (vertical, north, and east) seismic stations from teleseismic earthquakes. The receiver function technique takes advantage of the fact that part of the energy of seismic *P* waves is converted into *S* waves at discontinuities along the ray path ([9], [29]), and has been utilized in many studies; see, e.g., ([143], [143], [6], [42]). For data collection and processing, we use the Standing Order for Data (SOD) [87], [6] to request three component seismograms for *P*-wave arrivals and for events with a minimum magnitude 5.5, depth in the range of 1–600 km, and an epicentral distance ranging from 30° to 95°; see, e.g., [6].

Receiver functions were first applied in the late 1970s at solitary stations to obtain local one-dimensional structural estimates [60]. Since then, there was an increase in the number of stations that were deployed for seismic experiments. It is now possible to generate detailed two or three-dimensional images of structures, such as the Moho and upper mantle transition zone discontinuities near 410 km and 670 km depth; see, e.g., [139].

Receiver functions are derived using deconvolution [65], a mathematical method used to filter a signal and isolate the superimposed harmonic waves. Specially, receiver functions are calculated by deconvolving the vertical component of a seismogram from the radial component, resulting in the identification of converted phases where there is an impedance contrast (crustal-mantle boundary) [116].

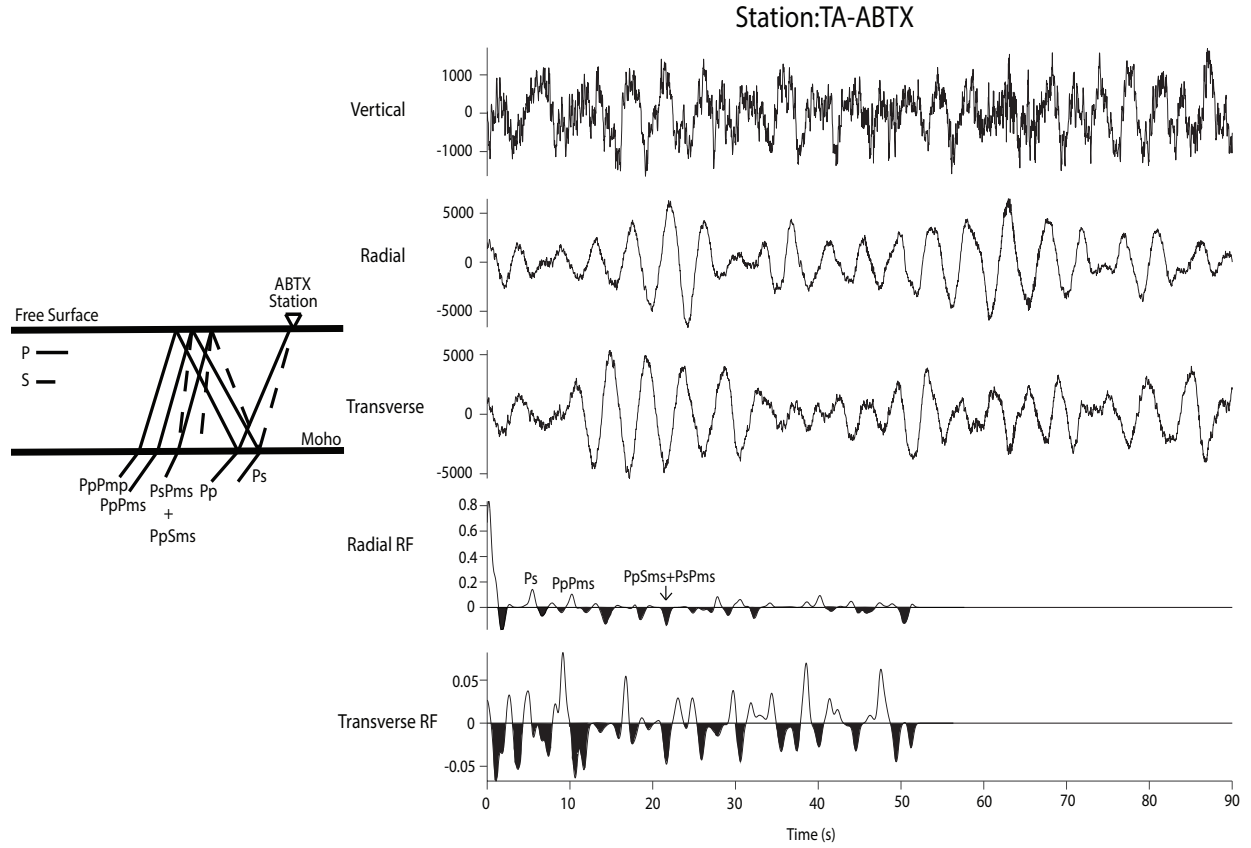


Figure 4.2: (Left) Illustration of a simplified ray diagram, which identifies the Ps , converted phases, which comprise the receiver function for a single layer. (Right) Vertical and radial seismograms and the corresponding receiver function resulting from the deconvolution of the vertical component from the radial component.

4.4.2 Surface Wave Dispersion

Surface waves in general differ from body waves in many respects: they travel slower, lower frequencies, largest amplitudes, and their velocities are in fact dependent on frequency [116]. The surface wave velocities vary with respect to depth being sampled by each period of the surface wave. The sampling by each period of the surface wave is known as dispersion [116]. Valuable information can be inferred by measuring surface wave dispersion because it will allow better understanding of the Earth's crustal and mantle velocity structure [61], [84], [121]. In particular, Love and Rayleigh wave group dispersion observations generally account for average velocity structure as a function of depth [51], [72]. The dispersion curves for surface waves are extracted from station records of three component seismograms for different frequencies and distances, by using reduction algorithms that rely on spectral analysis techniques. The important fact here is that, based on Rayleigh's principle, surface wave velocities are more sensitive to S wave velocity, although they are also theoretically sensitive to P wave velocity and density. Figure 4 provides an example of teleseismic Rayleigh waveforms for the 12/01/2010 event.

The Rayleigh's principle states that the phase velocity perturbation, denoted by $\frac{\delta c}{c}$, can be viewed as a function of $(K_\alpha, K_\beta, K_\rho)$, the sensitivity coefficients for P wave velocity, S wave velocity and density, respectively, i.e.,

$$\frac{\delta c(T)}{c(T)} = \int \left(K_\alpha \frac{\delta \alpha(z)}{\alpha(z)} + K_\beta \frac{\delta \beta(z)}{\beta(z)} + K_\rho \frac{\delta \rho(z)}{\rho(z)} \right) \quad (4.4.1)$$

where T is the period and z is the depth. By investigating sensitivity function variation in depth, the relative contribution of each property to dispersion can be shown. This subject is beyond the scope of our work, thus we just mention here that such analysis allows geophysicists to show that the relative contribution of P wave velocity, and density to dispersion is smaller than the one for S wave velocity [51]. That is, surface wave dispersion is much more sensitive with respect to S wave velocity, and therefore we have established the dependence of this data set on shear wave velocity.

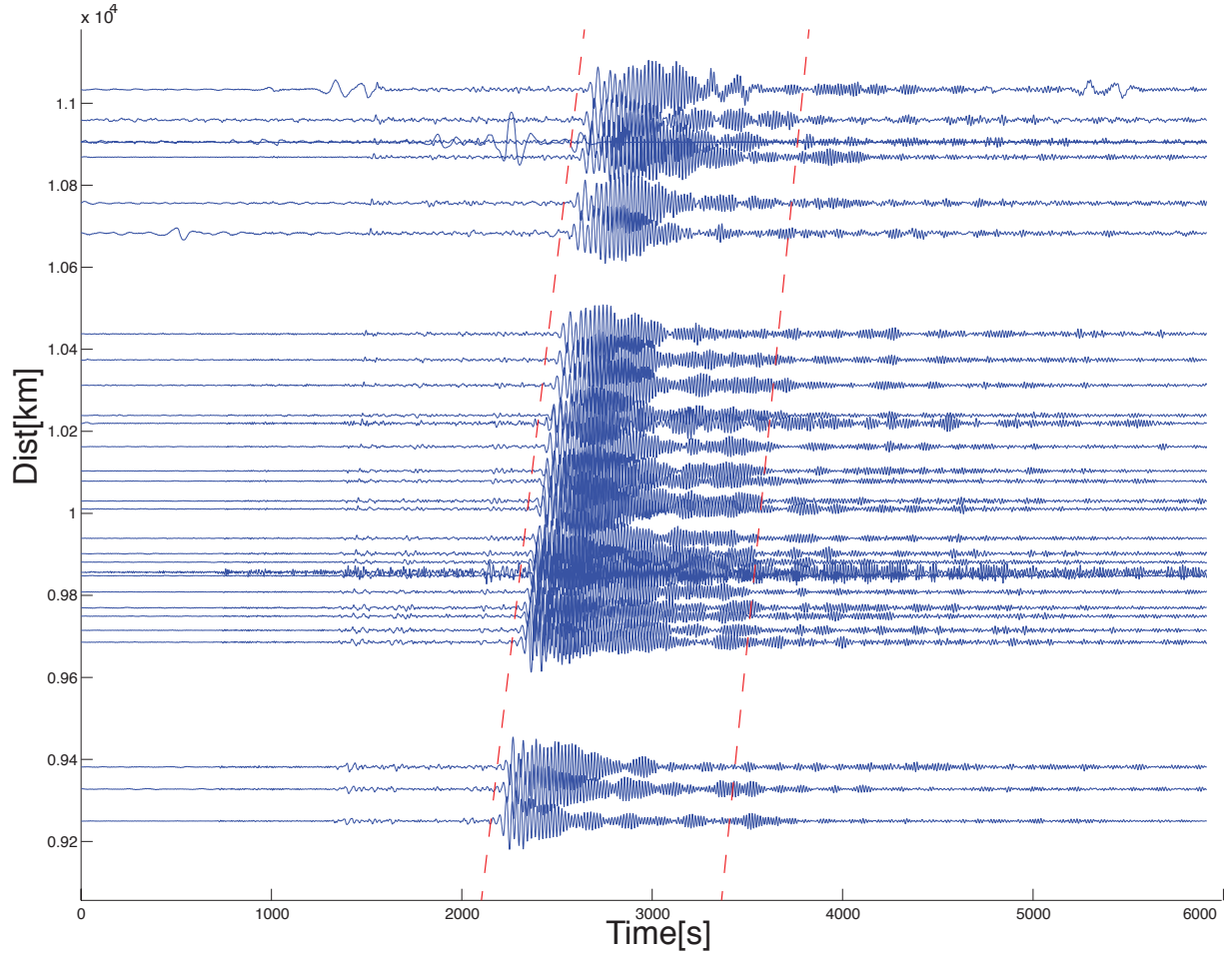


Figure 4.3: Example of surface wave Rayleigh waveforms for all stations used for this research. Dashed red line illustrates the window of where the phase curve passes through the stations.

For this research, a Matlab-based software package – that automatically downloads, analyzes, and measures phase as well as amplitude of surface waves to generate surface-wave tomography maps – was used to construct figure 5 describing tomographic images of the Texas region. The Automated Surface-Wave Phase-Velocity Measuring System (ASWMS) was the matlab package developed by [49], which essentially is an automated cross-correlation based method to generate surface-wave tomography of the entire U.S. The ASWMS tool was used to see what geological signatures that we can resolve using surface wave phase data.

4.4.3 Delay Travel Times

For this research, we used TauP toolkit ($\tau(p)$) and Antelope (BRTTO) database program to acquire the delayed travel times from the Array Network Facility (ANF) seismic catalog. Figure 6, shows an example of some of the delayed travel time data used for this study that were acquired from the ANF seismic catalog. The TauP toolkit program that we used to acquire the delayed travel times, uses the spherical Earth geometry into the computation. The TauP toolkit uses the relation of Snell's law, according to which, for each ray k , the ratio $\frac{\sin i_{kj}}{x_{kj}} = \frac{\Delta T_k}{l_{kj}}$ remains constant along the k -th ray, i.e., does not depend on j . This ratio is known as the constant ray parameter, and is usually denoted by p_k . We use this law to compute delayed travel times ΔT_k (i.e., components of the vector $F^{TT}(x)$):

$$\Delta T_k = \sum_{j=0}^n \frac{h_j}{x_{kj} \cos(i_{kj})}, \quad (4.4.2)$$

where h_j is the thickness of the j -th layer.

By using the Snell's law as mentioned earlier, the incidence angle i_{kj} can be rewritten as:

$$i_{kj} = \sin^{-1}(p_k x_j) \quad (4.4.3)$$

Equation (4.4.2) can be rewritten as

$$\Delta T_k = \sum_{j=0}^n \frac{h_j}{x_{kj} \cos(i_{kj})} = \sum_{j=0}^n h_j (x_{kj} \cos(\sin^{-1}(p_k x_{kj})))^{-1} \quad (4.4.4)$$

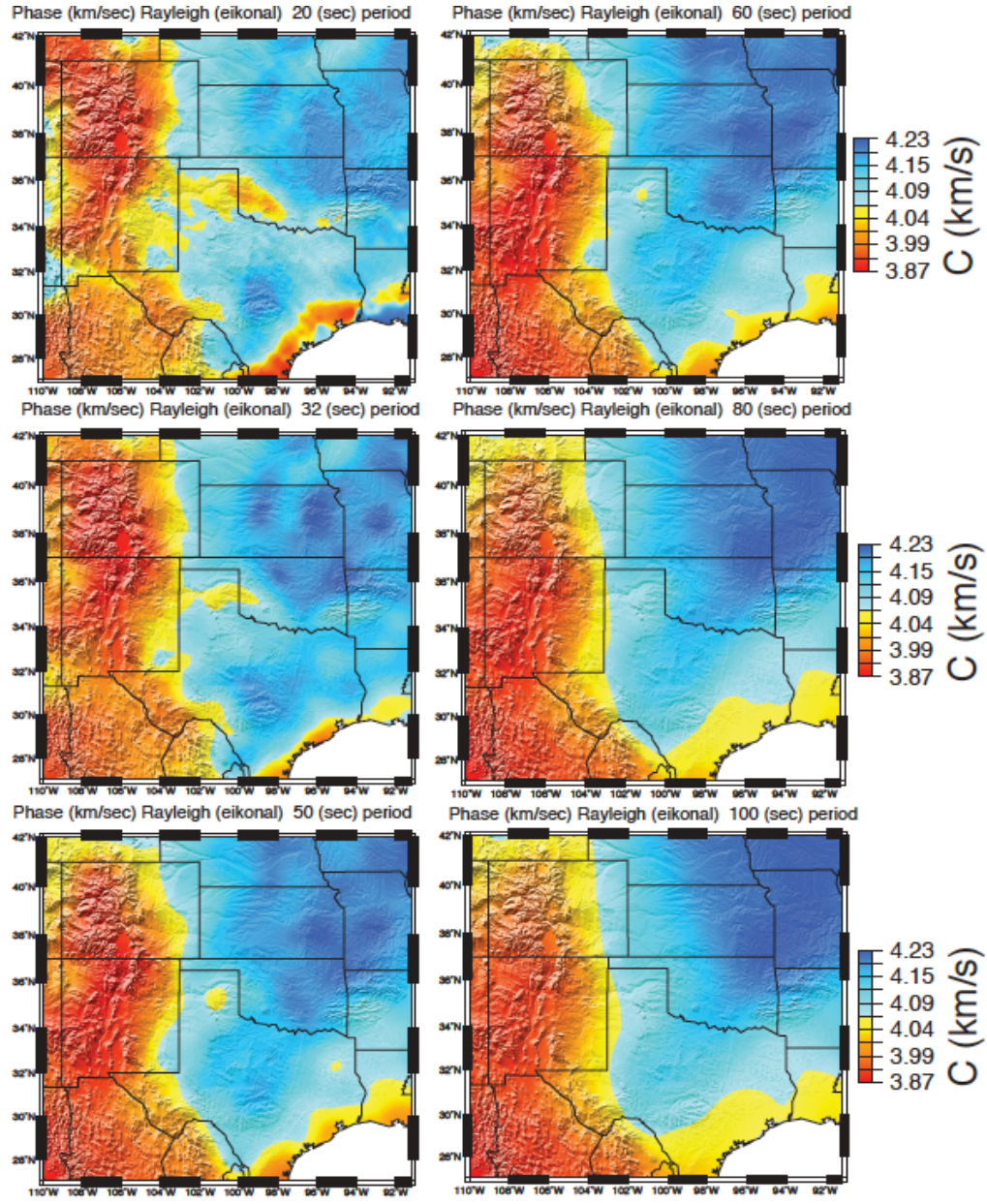


Figure 4.4: Surface wave phase map (Rayleigh) from USArray data used for this research. Red colors indicate low phase velocities and blue colors as high phase velocities for the Texas region.

Using Snell's law and rewriting equation (4.4.2), we obtained the partial derivatives which are needed to use the primal dual interior point method mentioned earlier:

$$\frac{\partial \Delta T_k}{\partial x_{kj}} = \frac{-h_j}{x_{kj}^2 \cos(\sin^{-1}(p_k x_{kj}))^2} \left(\cos(\sin^{-1}(p_k x_{kj})) - \frac{(p_k x_{kj})^2}{\sqrt{1 - (p_k x_{kj})^2}} \right) \quad (4.4.5)$$

When ray paths between the source and the receiver are short enough, Earth's curvature is known to be negligible, which provides us with the importance of utilizing equation (4.4.5) for our computation of partial derivatives of T .

4.5 Results and Discussion

Based on the joint inversion results using multiple geophysical datasets, the compatiability of the datasets provides better estimates of the target model based on numerical experiments with the datasets and the synthetic rift model. In order to illustrate how receiver functions and surface wave dispersion velocities complement each other, we present in figures 6-8 the inversion results for the data sets created from the rift velocity model.

In figure 6, the joint inversion of both RFs and TT data sets gives a better approximation to the target model as expected based on the compatiability between the two datasets. The single inversion of receiver functions (left top) identifies the velocity interfaces (fig 6), while single inversion of surface waves (fig 7) gives information on the average velocities at different depths. For figure 7, the joint inversion of SWs and TT also provides a better approximation of the target model indicated in red.

The joint inversion of RFs, SWs, and TT provides more accuracy and compatiability when performing joint inversion of the three geophysical datasets as shown in figure 8.

A rift velocity model was used as our initial test for the MOP joint inversion scheme. For the MOP joint inversion scheme, a comparison was done with the rift velocity model and

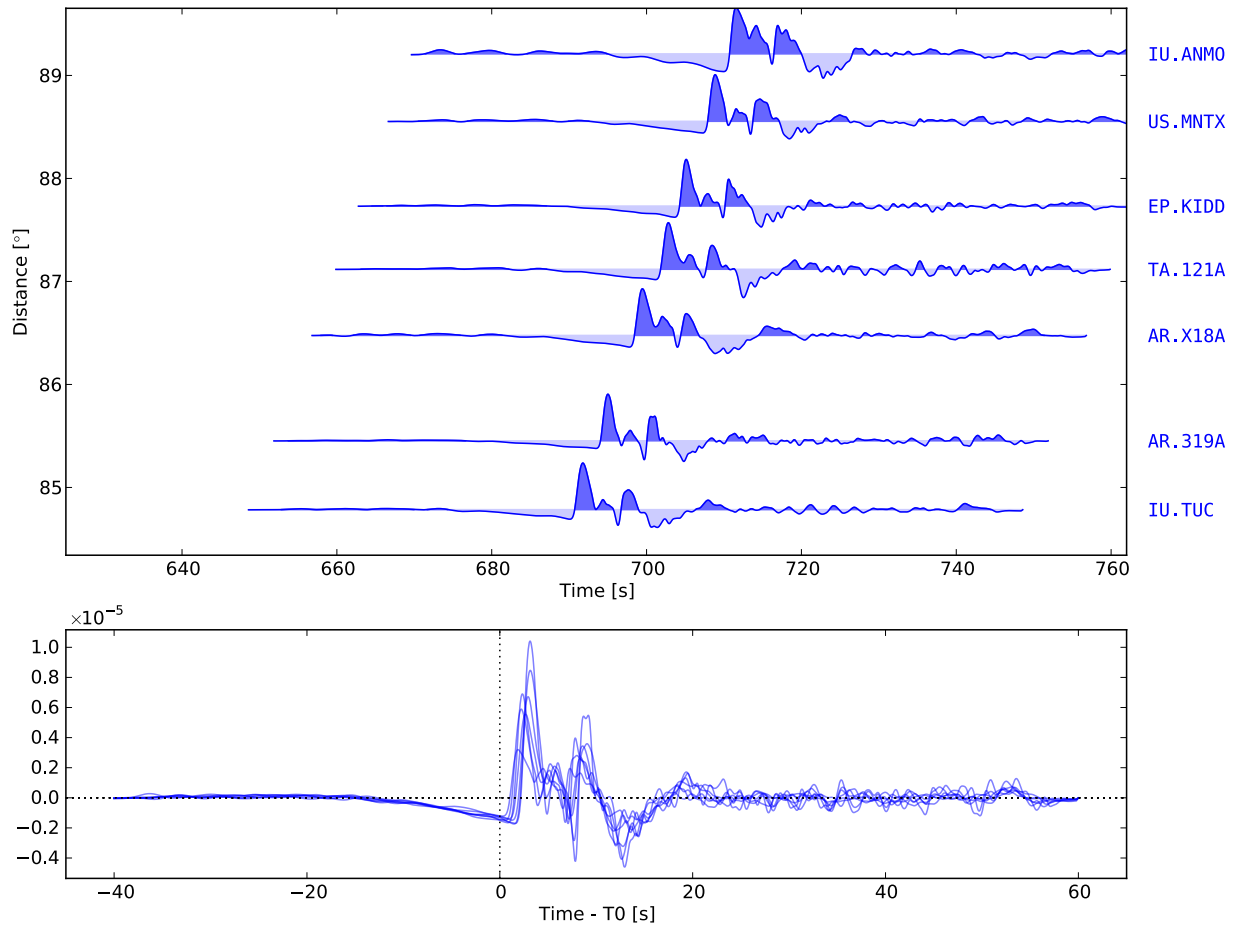


Figure 4.5: Example plot shows P -waves when delay times are measured travel times (Dist (deg) vs. Time(sec)) for different USArray Stations that were used for this research from 09/15/2011 event. (Second) plot shows the difference between the predicted and measured travel times and illustrates the P -wave delay.

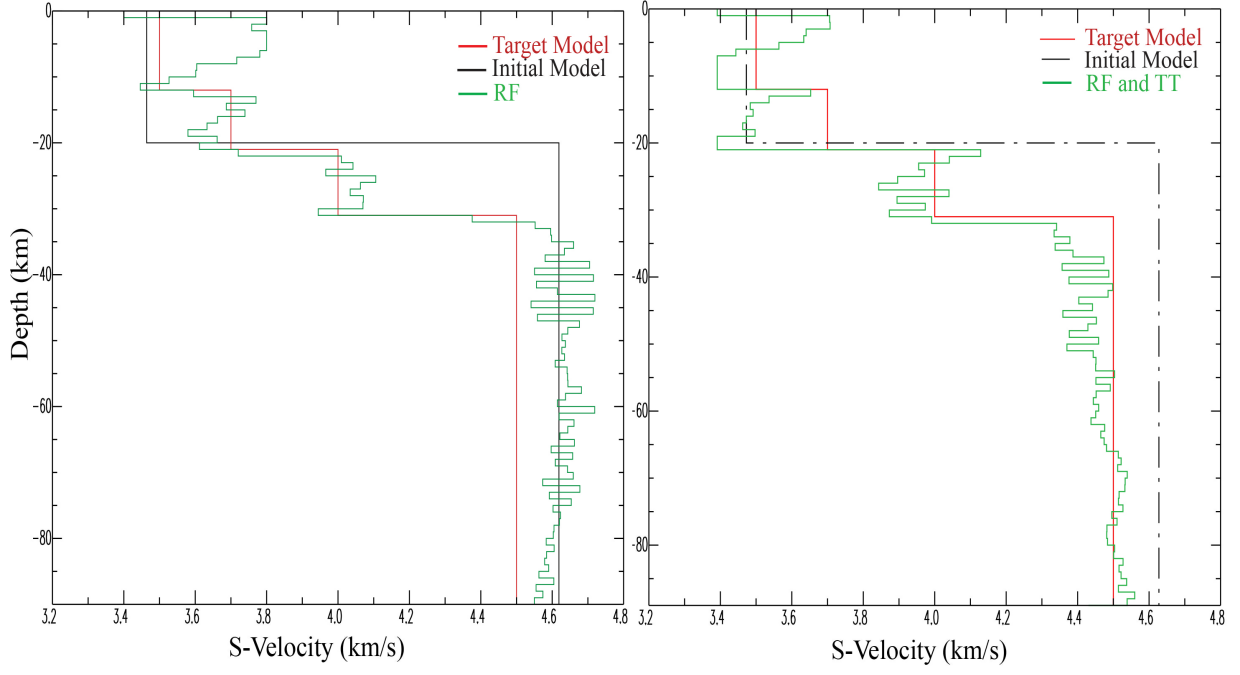


Figure 4.6: (Left) Single data inversion of receiver functions and synthetic rift model. (Right) joint inversion using receiver functions and travel times to obtain a better estimate of the rift model (target model).

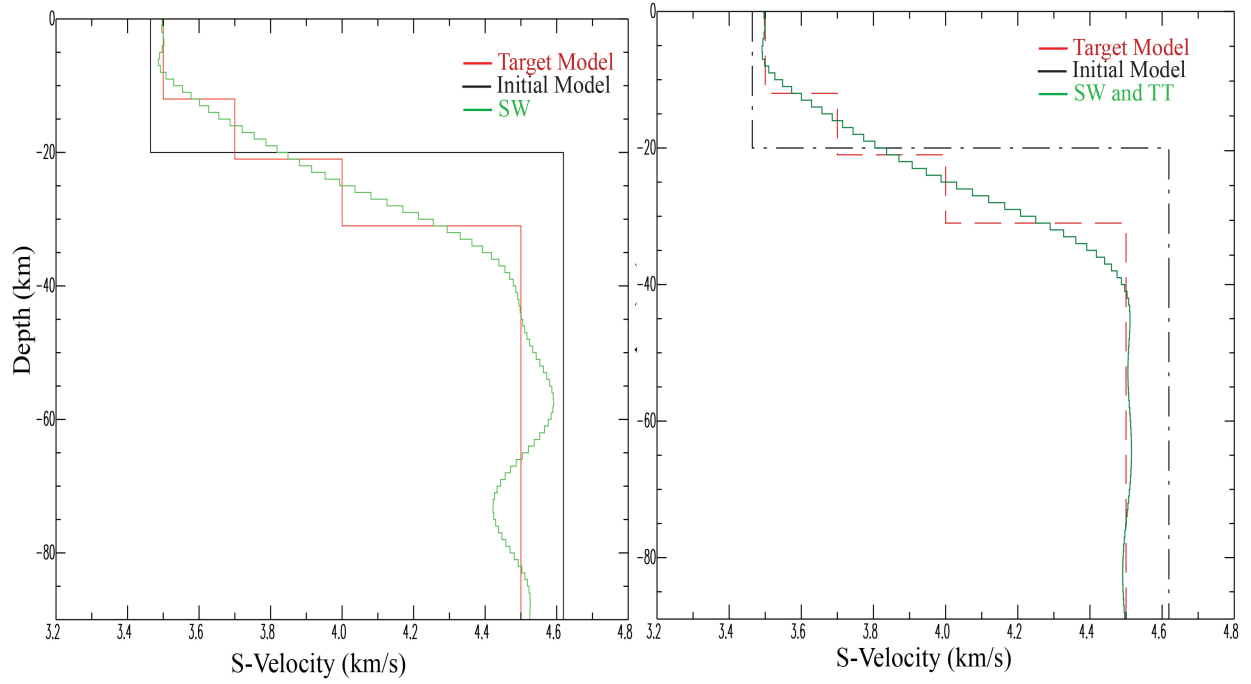


Figure 4.7: (Left) Single data inversion of surface wave dispersion measurements and synthetic rift model. (Right) joint inversion using surface wave dispersion measurements and travel times to obtain a better estimate of the rift model (target model).

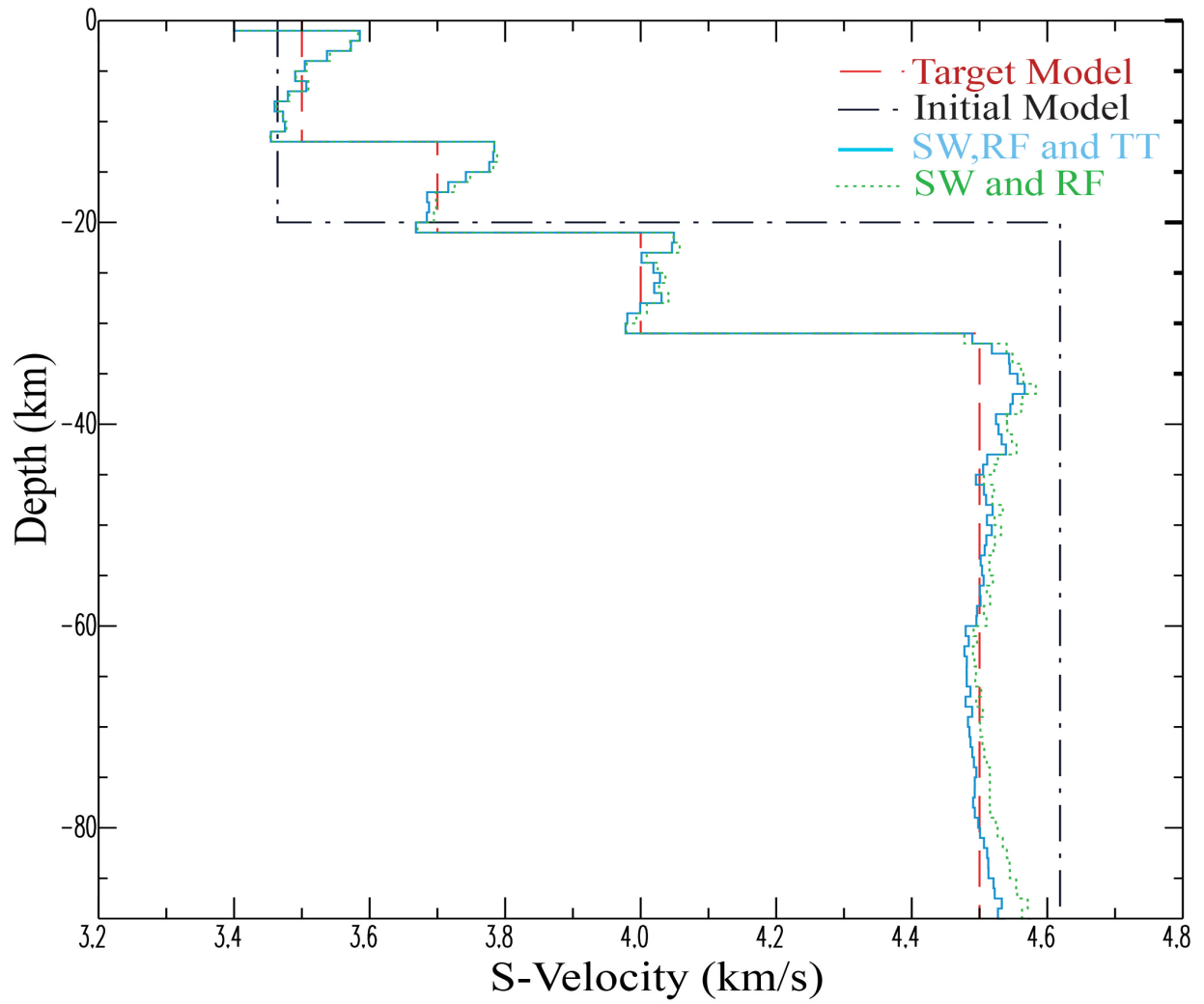


Figure 4.8: (Left) Joint inversion of surface wave dispersion, receiver functions, & travel times to obtain a better estimate of the rift model (target model).

the initial velocity test model. The synthetic rift model was the best velocity model used for this joint inversion scheme. Different synthetic Earth velocity models were used but overall, the rift model was the best. Numerous tests were performed to test the compatibility and complementary nature of the multiple geophysical datasets based from the results in figures 9-11. The algorithm used to perform the joint inversion of the multiple geophysical datasets using the multi-objective joint inversion scheme was written in FORTRAN 77 and coupled with a C code that performs the Multi-Objective Optimization method, based on the work of [121].

Joint Inversion Algorithm (Two Datasets)

1. Given an initial velocity model $(V_s)_0 = x_0$ RF observations y^{RF} , SW dispersion observations y^{SW} , and a max. number of iterations l .
2. for $\eta = 0, 0.1, 0.2, \dots$, until 1.0 do
3. Evaluate y^{RF}, y^{SW}
4. for $k = 0, 1, 2, \dots$, until l do
5. Evaluate $F(x_k), F'(x_k)$ and compute $b = F(x_k) - y - F'(x_k)x_k$
6. Solve Equation (15) by using PDIP method.
7. if $\|F(x_k) - y\| \leq \epsilon$ then
8. break
9. end if
10. Go to Step 4
11. end for
12. end for

Joint Inversion Algorithm (Three Datasets)

1. Given an initial velocity model $(V_s)_0 = x_0$ RF observations y^{RF} , SW dispersion observations y^{SW} , y^{TT} and a max. number of iterations l .

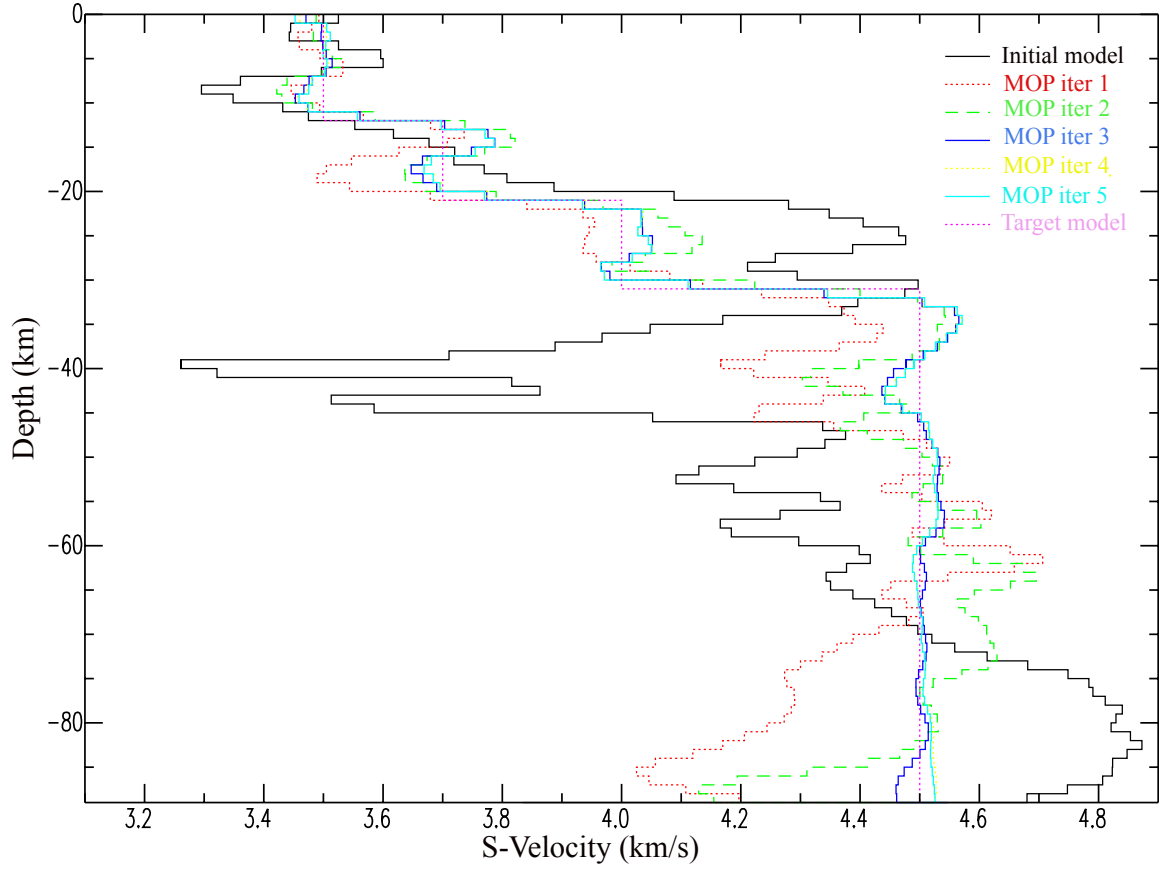


Figure 4.9: 1-D joint inversion results for a synthetic 1-D orogen velocity model. The target velocity model in purple is the rift velocity model. The black color represents the initial orogen model. The other colors represent the improvement of the 1-D shear wave velocity models matching with the target model (purple).

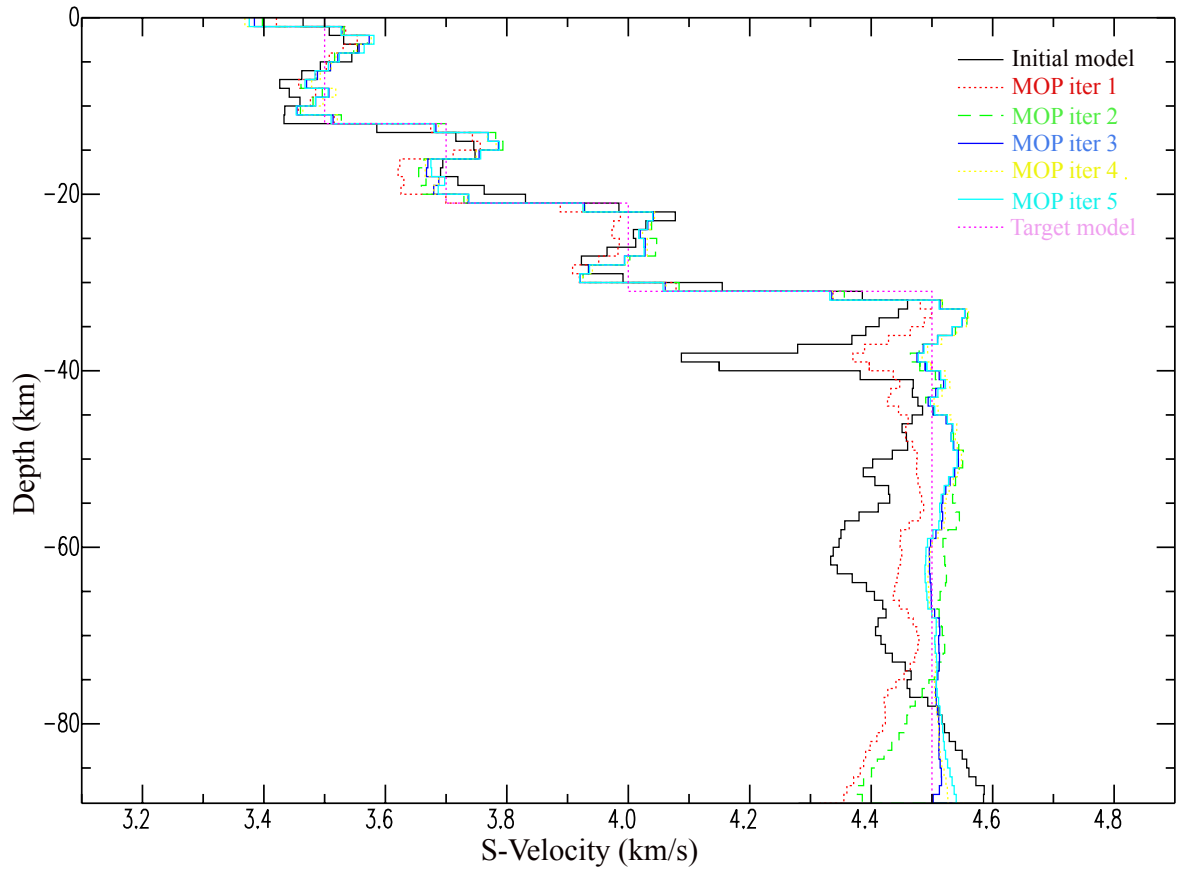


Figure 4.10: 1-D joint inversion results for a synthetic 1-D continental velocity model. The target velocity model in purple is the rift velocity model. The black color represents the initial continental model. The other colors represent the improvement of the 1-D shear wave velocity models matching with the target model (purple).

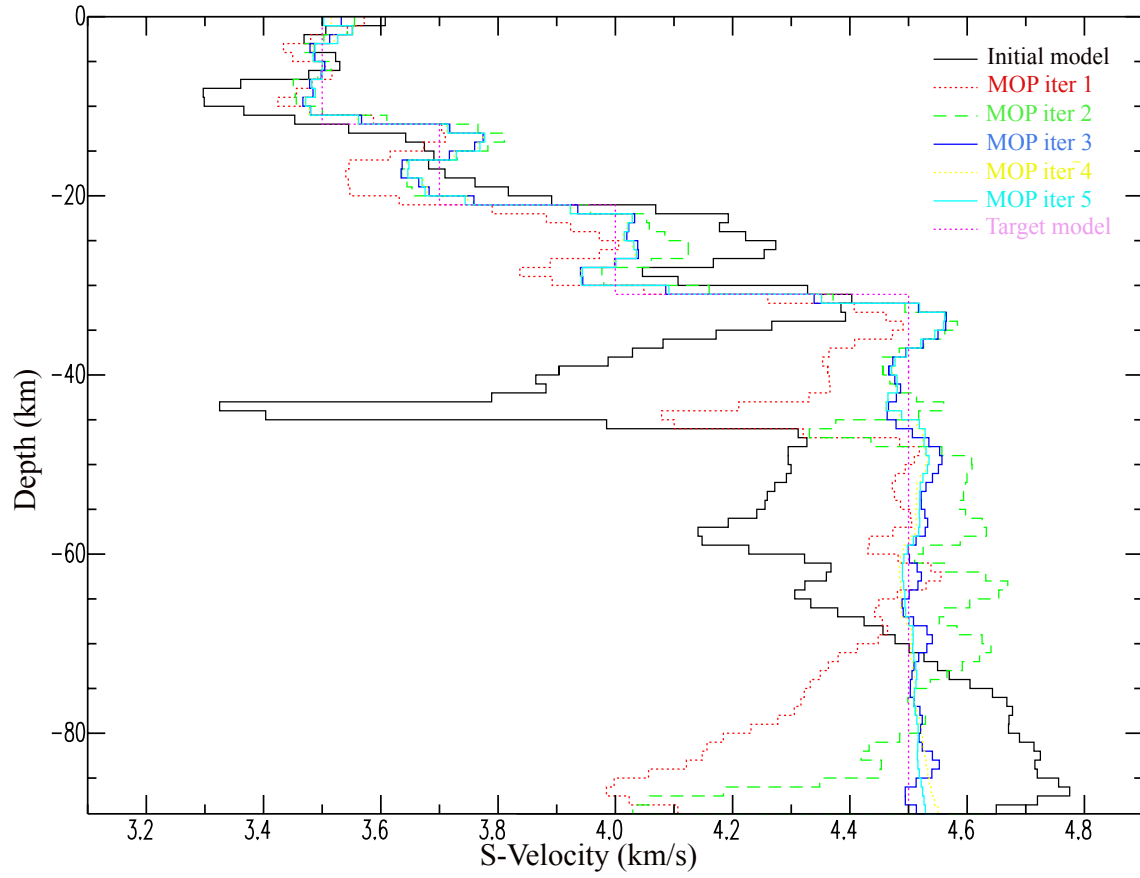


Figure 4.11: 1-D joint inversion results for a synthetic 1-D archean velocity model. The target velocity model in purple is the rift velocity model. The black color represents the initial archean model. The other colors represent the improvement of the 1-D shear wave velocity models matching with the target model (purple).

2. for $\eta = 0, 0.1, 0.2, \dots$, until 1.0 do
3. Evaluate y^{RF}, y^{SW}
4. for $k = 0, 1, 2, \dots$, until l do
5. Evaluate $F(x_k), F'(x_k)$ and compute $b = F(x_k) - y - F'(x_k)x_k$
6. Solve Equation (15) by using PDIP method.
7. if $\|F(x_k) - y\| \leq \epsilon$ then
8. break
9. end if
10. Go to Step 4
11. end for
12. end for

4.6 Conclusion

We apply an inversion scheme that expands a joint-inversion constraint least-squares (LSQ) algorithm used to characterize a one-dimensional Earth's structure. We utilize the Multi-Objective Optimization technique to perform joint inversion of multiple data sets (receiver functions, surface wave dispersion, and travel times). By jointly inverting these three geophysical data sets, we improve the model's accuracy. In the ideal situation when we know the relative accuracy of different datasets, we can formulate the joint inversion problem as a (single) least-square optimization problem. In practice, however, we only know approximate values of these accuracies; so, for inversion, we use the Multi-Objective Optimization Problem (MOP) approach. Different combination of weights were incorporated in the MOP inversion scheme in order to map the Pareto Set (Solution Space) corresponding to receiver functions, surface wave dispersion measurements, and travel times. From the Pareto Set, the MOP technique performs a direct search method that selects the most feasible solution from a set of alternative solutions from the model space [107], [108], [57]. For future work,

we plan to incorporate gravity into our inversion scheme to obtain a more constrained Earth structure and to better determine physical properties of the Earth structure.

Chapter 5

Generation of 3-D shear wave based models using 1-D constrained multiobjective optimization

Understanding Earth's tectonic processes requires determining the Earth structure. We focus on determining Texas tectonic processes by applying a joint inversion scheme for multiple geophysical datasets. Specifically, we apply an inversion scheme that expands a joint-inversion constraint least-squares (LSQ) algorithm used to characterize a one-dimensional Earth's structure. We use seismic data from regional networks in Texas and from the USArray, a dense network of permanent and portable seismographs placed systematically across the continental United States. We collect and process seismic data, namely receiver functions and surface waves. By jointly inverting these two geophysical data sets, we improve the models accuracy. In the ideal situation when we know the relative accuracy of different datasets, we can formulate the joint inversion problem as a single least-square optimization problem. In practice, however, we only know approximate values of these accuracies; so, for inversion, we use the Multi-Objective Optimization Problem (MOP) approach. We use this approach to find the 1-D models, and then we use kriging to combine these 1-D models into a single 3-D model. The resulting 3-D velocity models provide insight into the tectonic history and physical properties of the Earth structure of the Texas region. We also compare the structure of the upper mantle beneath the Texas region with other ancient and/or active rift systems in the vicinity of the Texas region. This work is based on the following publication:

Thompson, L., A. A. Velasco, V. Kreinovich, R. Romero, and A. Sosa, 2015, 3-D shear wave based models of the Texas region using 1-D constrained multi-objective optimization, *Journal of Geophysical Research*, (manuscript preparation).

5.1 Introduction

For this research, we combine multiple geophysical datasets for the purpose of better determining physical properties of the Earth structure. By simultaneously inverting multiple datasets, we obtain a better estimate of the true Earth structure model based on inversion. In general, in terms of inversion there are two reasons why the estimated Earth structure model differs from the true Earth model after inversion. The first reason is the inherent non-uniqueness of the inverse problem that causes several (usually infinitely many) models that satisfy the data. The second reason is that real geophysical data is always affected by noise, which introduces error associated with the estimation of the Earth Structure model after inversion. By jointly inverting multiple geophysical data sets, we can avoid the inherent non-uniqueness from inverting the geophysical datasets (e.g., receiver functions, surface wave dispersion measurements) individually [133]. In this research, we use receiver functions, and surface wave dispersion measurements data sets to characterize the crust and upper mantle of various tectonic provinces within the Texas region.

In general, geophysical data sets such as receiver functions are suited to constrain the depth of discontinuities and are sensitive to relative changes in S-wave velocities in different layers but they are unable to determine the velocities between discontinuities. Surface waves measurements on the other hand, constrain the absolute shear velocities between discontinuities, but they cannot determine the depth of discontinuities [117]. The complementary information provided by these datasets, reduces the inherent ambiguity or non-uniqueness of performing inversion (e.g., [40]). By jointly inverting seismic data, we can overcome the difficulties of non-uniqueness and be able to facilitate the construction of the true Earth model.

When we process a single data set (e.g., Surface Wave Dispersion), we normally use the least squares method to find the best-fit model. For multiple data sets (e.g., Surface Wave Dispersion and Receiver Functions), if we knew the variance (uncertainty of data) of the different measurements of the multiple data sets, we could still be able to use the least squares approach to finding the model space. In practice, we only have an approximate knowledge of the variances. So, instead of producing a single model, we want to generate several models corresponding to different possible variances. A technique for generating several possible models (corresponding to different optimization criteria) is known as Multi-objective Optimization (MOP). The resulting set of models is known as the Pareto Front. Once several models corresponding to different possible variances are computed, we can then proceed to select the most geophysical meaningful model from the Pareto Front. Methods for such are described selection, e.g., in ([107],[108], [57]).

5.2 Tectonic Setting

The state of Texas, one of the major continental crustal regions in western US, was generated as part of the 1.37 Ga southern granite-rhyolite provinces [95]. In the southeastern portion of the Llano Uplift province within Texas, according to [95], it was found that the Llano Uplift upper crustal structure is primarily made up of dioritic and tonalitic gneiss material, which is inferred to be a representation of the 1.33-1.30 Ga allochthonous magmatic arc [95]. This arc, according to [95], [79], is thought to have been accreted to Laurentia during the Grenville orogeny (mountain building period) at approximate 1.1 Ga due to N-dipping subduction beneath Laurentia during this period. Another tectonic province within Texas, the Balcones province, lies above the transitional crust and separates the Mesoproterozoic craton from the Pennsylvanian Ouachita orogenic lithosphere [95]. Geologists use the Balcones province as an approximate southern limit of the North American cratonic boundary and the western limit of attenuated transitional crust beneath the Texas Gulf coastal plain according to [95]. The western limit of attenuated transitional

crust beneath the Texas Gulf coastal plain is a result of the Gulf of Mexico Jurassic oceanic lithosphere opening according to [74], [110], [95]. It still remains unclear the nature of the transitional crust of the Texas Gulf coastal plain. The transitional crust is buried beneath thick sediments of the Gulf of Mexico coastal plain and according to [95], they conjectured how the crust could be composed of one or more of the following components:

- metamorphosed sediments and crust of the Paleozoic passive margin, deformed during the Ouachita orogeny [75];
- fragments of Gondwana, left behind when Gondwana separated from Laurentia in Jurassic time [99];
- juvenile mafic crust of Jurassic age [74];
- thinned crust of southern Laurentia.

Observations of high and low velocities within the mantle are usually determined through tomographic methods. For this research, we will use MOP formulation to generate 3D shear wave models to determine regions of Texas that would have high and low velocities within the mantle. Using the MOP optimization technique to generate 3D imagery that helps to determine the petrophysical nature of the Lithosphere-Asthenosphere Boundary (LAB), which is still considered a subject for discussion among the geoscience community (e.g., [59], [30], [32]). Through the development of our 3D models using multiple data sets, we are able to evaluate and analyze the Upper mantle seismic velocity structures in both the vertical and horizontal directions, which will be the key to our understanding of the overall tectonic structure of Texas. Within the 2009-2011 time period, the USArray Transportable Array (TA) deployed an array of seismic stations within the Midwestern US and through the USArray data, which we are able to use to investigate the velocity structure beneath the stable North American craton. Within the Midwestern region of the US (Texas), previous seismic work using USArray dataset found that the thick continental lithosphere is neutrally buoyant and this buoyancy effect is caused by iron depletion and

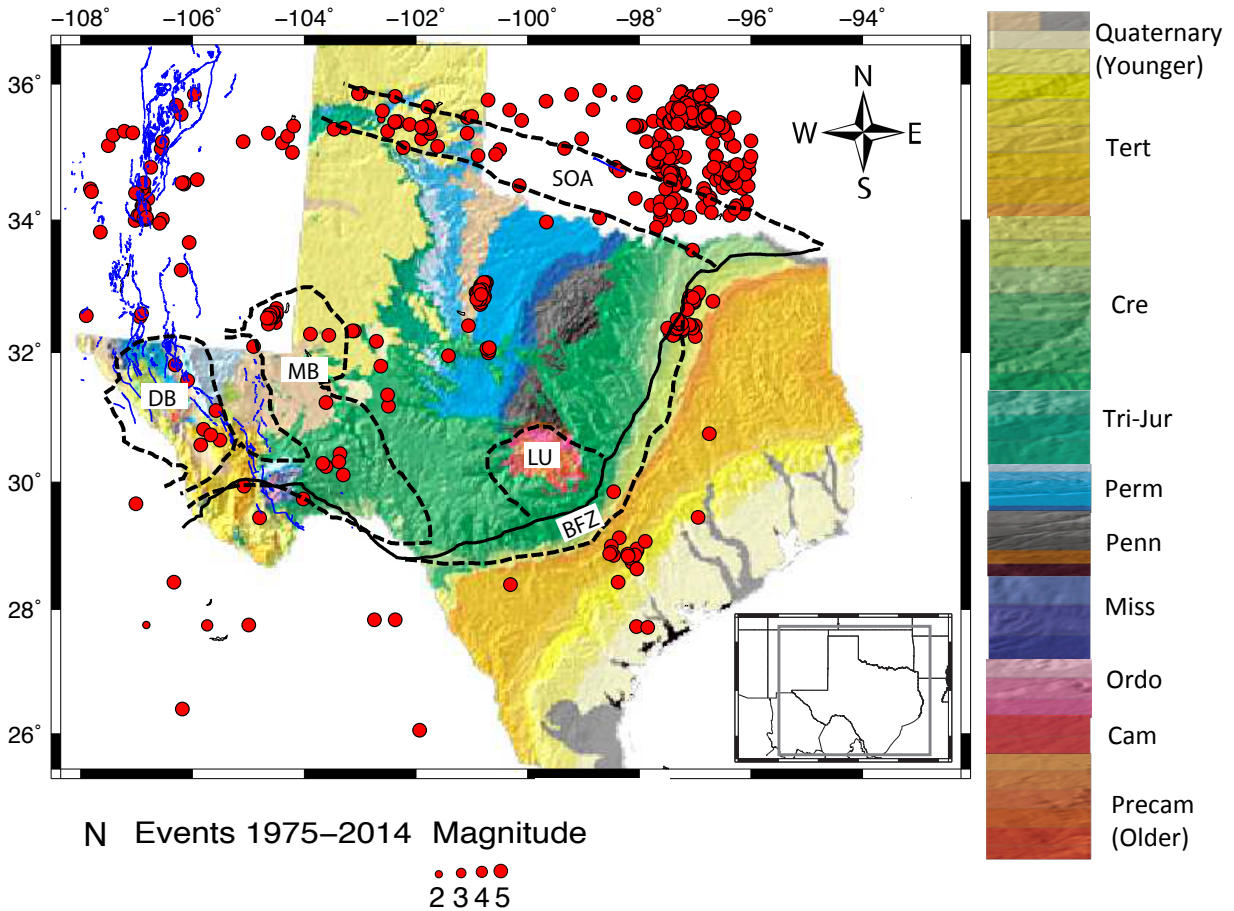


Figure 5.1: Tectonic map of Texas with dark colors that represent older rock material and lighter colors represent younger rock material. DB label is Delaware Basin, MB is Midland Basin, LU is Llano Uplift, BFZ is the Balcones Fault Zone, and SOA is Southern Oklahoma Aulacogen.

decreased buoyancy from cold temperatures in a thick thermal boundary layer ([23]). By utilizing the USArray Data for Texas, we are able to characterize the upper mantle structure beneath a large portion of the North American stable Precambrian interior based on the development of shear wave models using receiver functions, and surface wave dispersion measurements. Through shear wave models for the entire state of Texas, we might expect to find the presence of low-velocity layers on top of the 410 km discontinuity, which could be an indication of high water concentrations in the transition zone ([23]). Teleseismic imaging methods have been used in previous geophysical research using the USArray TA dataset to map upper mantle seismic structure, e.g., Body wave travel-time tomography, surface wave dispersion tomography, and receiver function analysis [18], [118]; [84], [80], [92], [20].

5.3 Receiver Functions

A receiver function maps the seismic response of the earth beneath a seismic station to an incoming, teleseismic P wave. Deconvolving the vertical component of a teleseismic earthquake seismogram from the radial component (e.g., [60]) results in a receiver function, which then allows for the identification of converted phases corresponding to strong impedance contrasts (e.g., the crustal-mantle boundary). Since we use teleseismic events that arrive to the stations with near-vertical incidence, receiver functions can also be used for imaging deep structure ([36], [42], [59], [64], [112]). Receiver functions can provide valuable information for investigating magma lenses within the crust, determining the Moho depth, other upper-mantle discontinuities [70], structure and evolution of the crust [9], and rifting extension and magmatism (e.g., [28]).

EarthScope provides a framework for broad, integrated studies in the Earth sciences (e.g., research on fault properties, earthquake process, plate boundary process, large-scale continental deformation, etc). USArray, which is part of the EarthScope experiment, is a

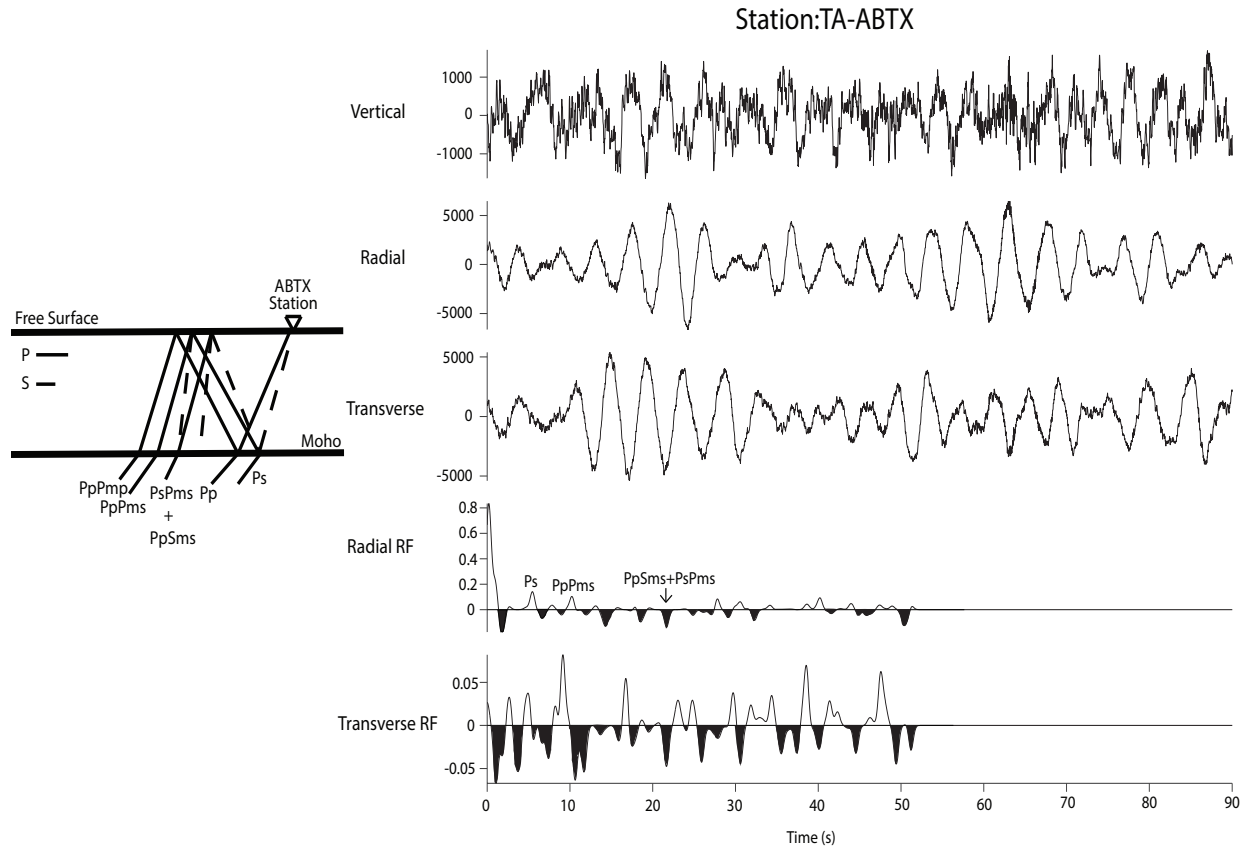


Figure 5.2: (Right) Components of the waveform for station ABTX along with the radial and transverse receiver functions. (Left) The halfspace represents usually the Moho converted phase P_s and the multiples, along with their ray paths.

15-year program to place a dense network of permanent and portable seismographs across the continental United States. The Plate Boundary Observatory (PBO) component of EarthScope is a geodetic observatory designed to study the three-dimensional strain field resulting from deformation across the active boundary zone between the Pacific and North American plates in the western United States. The San Andreas Fault Observatory at Depth (SAFOD) component of Earthscope is a borehole observatory across the San Andreas Fault to directly measure the physical conditions under which earthquakes occur. USArray Transportable Array (TA), an element of EarthScope, places a rolling, dense network of over 400 permanent and portable seismographs on a nearly uniform 70 km grid across the U.S. Each station of the TA network collects data for two years before it is re-deployed to a new location [80]. The USArray was in the region from approximately 2008 to 2010. We initially collected data from the EarthScope Automated Receiver Survey (EARS) website (<http://www.seis.sc.edu/ears/>) for the Texas region. The EARS website contains the data sets of all frequent seismic activity and a database of all the recording stations that monitor seismic events.

In Texas, parts of the Transportable Array were deployed to western Texas in 2008. Between 2009 and 2010, these Transportable Array stations in western Texas were moved to central and eastern Texas (Fig 3b). The Reference Network consists of permanent stations that are separated by 300 km spacing across the United States. In Texas, there are 10 Reference Network stations and we collected data from 8 recording stations that are located in central and eastern Texas.

We requested seismogram data from the Incorporated Research Institutes for Seismology (IRIS) Data Management Center (DMC) for teleseismic events occurring from January of 2011 to December of 2011 for Transportable Array stations. A collaborator, Keisuke Irie, acquired USArray data for 2009 and 2010 and we incorporated the data that we requested from the IRIS DMC with what Keisuke acquired for his M.S. thesis and used it for our complete receiver function analysis for Rio Grande Rift and the entire state of Texas. The

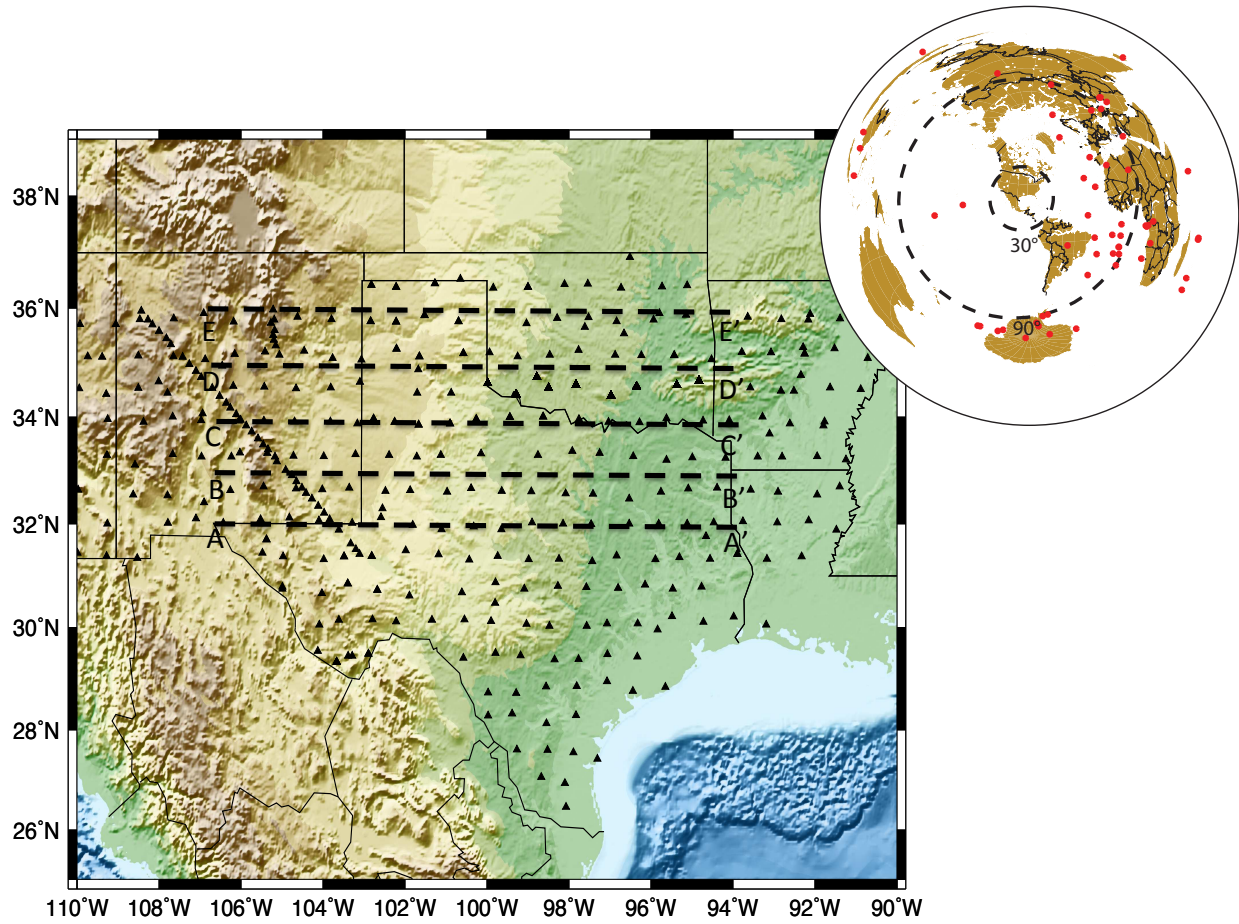


Figure 5.3: Black triangles represent all the USArray locations within the Texas region. (Topright) Represents the epicenter locations of the events used for the receiver functions.

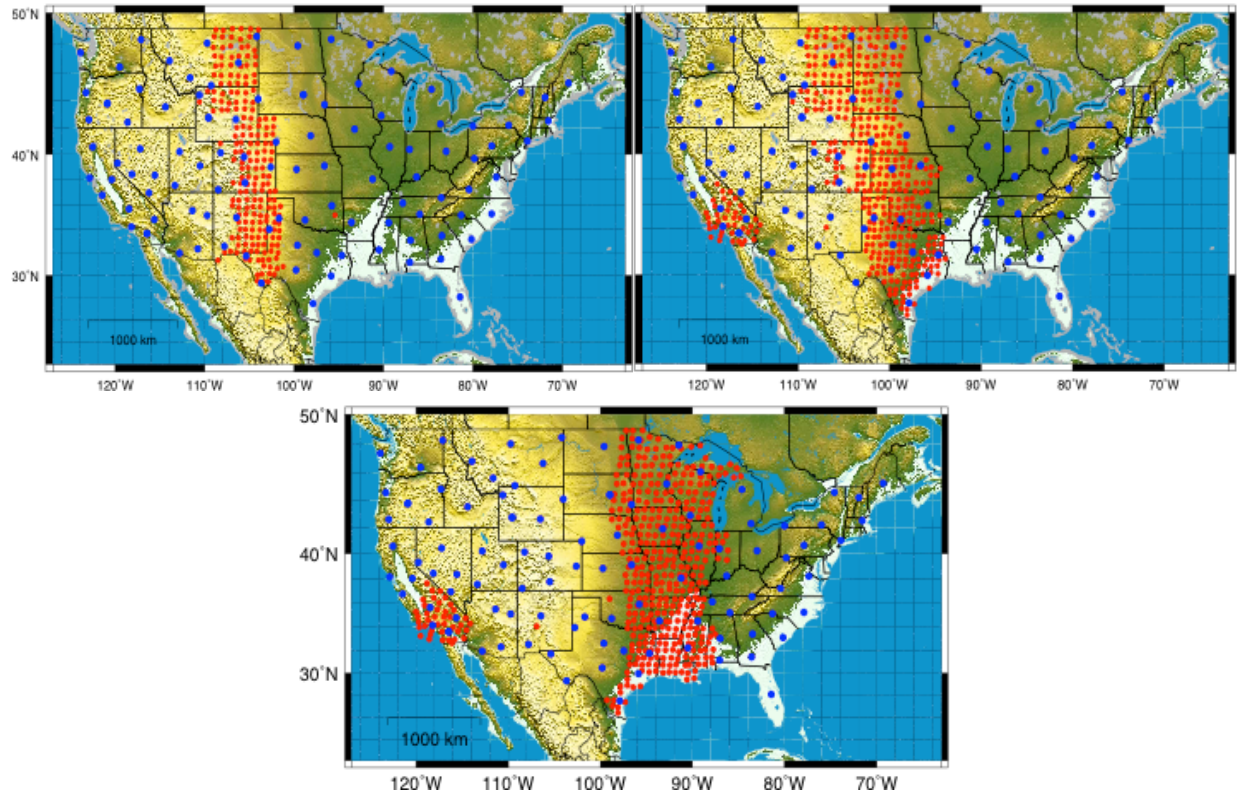


Figure 5.4: Red dots represent the Earthscope Transportable USArray stations used for this research. The figure illustrates how the Transportable Array migrated from west to east during 2008-2010.

locations of the USArray stations during 2009 are shown on Fig 3b. The blue dots represent the locations of the reference network stations and the red dots symbolize the locations of the Transportable Array stations. The USArray stations were then transported from west of Texas to central Texas in May of 2010 indicated by Fig 3b. The USArray stations were then moved to east of Texas during the time period between 2010 and 2011.

5.4 Surface Wave Dispersion

Surface waves in general differ from body waves in many respects: they travel slower, lower frequencies, larger amplitudes, and their velocities are in fact dependent on frequency [116]. The surface wave velocities vary with respect to depth being sampled by each period of the surface wave. The sampling by each period of the surface wave is known as dispersion [121]. Valuable information can be inferred by measuring surface wave dispersion because it allows better understanding of the Earth’s crustal and mantle velocity structure [84], [121]. As part of the systematic determination of earthquake moment tensors for North American earthquakes, Saint Louis University measures fundamental mode Love and Rayleigh wave spectral amplitudes and group velocities using a multiple filter analysis from local to regional earthquakes. Tomography is then used to obtain tomographic images of group velocity dispersion for North America with emphasis on the continental United States (e.g., [22], [129]; [148], [18]).

In particular, Love and Rayleigh wave group dispersion observations generally account for average velocity structure as a function of depth [51], [72]. The dispersion curves for surface waves are extracted from station records of three component seismograms for different frequencies and distances, by using reduction algorithms that rely on spectral analysis techniques. The important fact here is that, based on Rayleigh’s principle, surface wave velocities are more sensitive to S wave velocity, although they are also theoretically sensitive to P wave velocity and density. Rayleigh’s principle states that the phase velocity

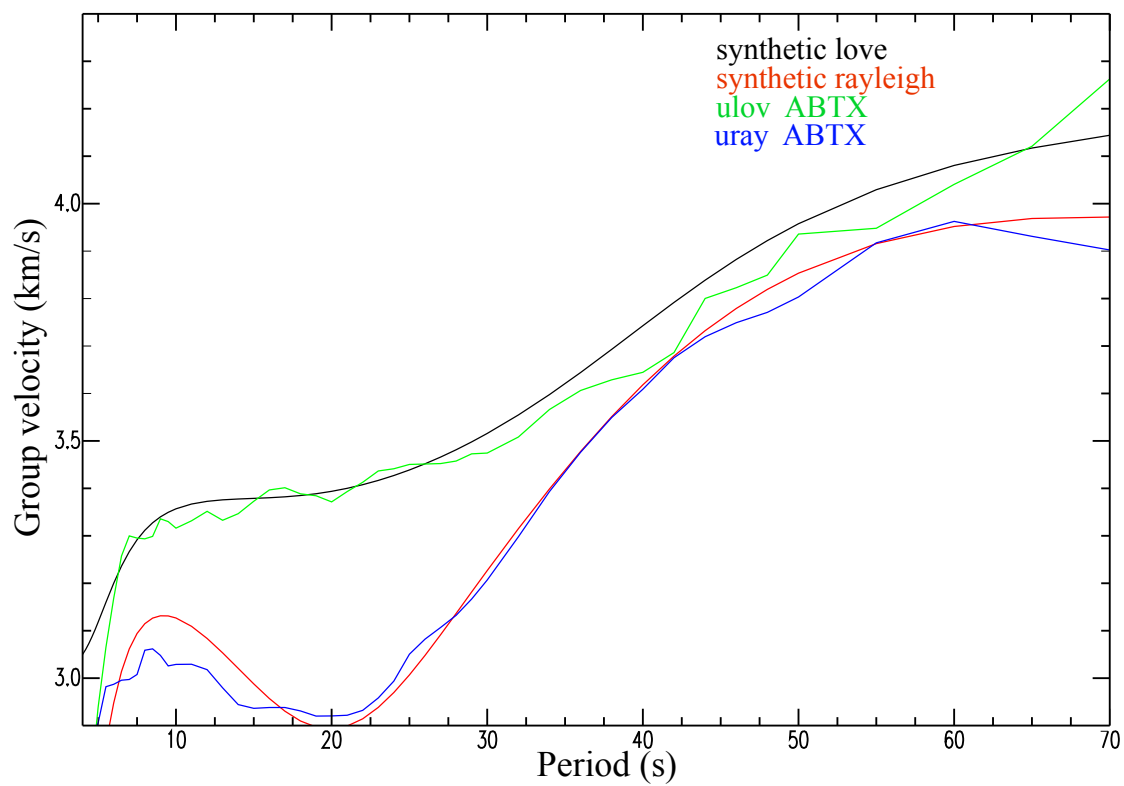


Figure 5.5: Rayleigh and Love Group Velocity Dispersion curves for station ABTX.

perturbation, denoted by $\frac{\delta c}{c}$, can be viewed as a function of $(K_\alpha, K_\beta, K_\rho)$ the sensitivity coefficients for P wave velocity, S wave velocity and density, respectively, i.e.

$$\frac{\delta c(T)}{c(T)} = \int \left(K_\alpha \frac{\delta \alpha(z)}{\alpha(z)} + K_\beta \frac{\delta \beta(z)}{\beta(z)} + K_\rho \frac{\delta \rho(z)}{\rho(z)} \right) \quad (5.4.1)$$

where T is the period and z is the depth. By investigating sensitivity function variation in depth, the relative contribution of each property to dispersion can be shown. This subject is beyond the scope of our work, thus we just mention here that such analysis allows geophysicists to show that the relative contribution of P wave velocity and density to dispersion is smaller than the one for S wave velocity [51]. Thus, surface wave dispersion is much more sensitive with respect to S wave velocity, and therefore we have established the dependence of this data set on shear wave velocity.

5.5 Joint Inversion

We want to find the parameters $x = (x_1, \dots, x_n)$ from the results $y = (y_1, \dots, y_m)$ of measuring quantities $y_i = F_i(x)$ depending on these parameters. A usual way to estimate parameters in the presence of noise is to use the least squares method, i.e., to find the values x that minimize the expression

$$S = \sum_{i=1}^m \frac{(F_i(x) - y_i)^2}{\sigma_i^2} \quad (5.5.1)$$

where σ_i is the standard deviation corresponding to the i -th measurement. When we use measurements of different types t, t' , etc, it is reasonable to assume that all the measurements i of type t have the same accuracy $\sigma_i = \sigma^t$. Then,

$$S = \sum_{i=1}^m \frac{(F_i(x) - y_i)^2}{\sigma_i^2} = \sum_{t=1}^T \sum_{i \in t} \frac{(F_i(x) - y_i)^2}{(\sigma^t)^2} = \sum_{t=1}^T \frac{1}{(\sigma^t)^2} \sum_{i \in t} (F_i(x) - y_i)^2 \quad (5.5.2)$$

This expression can be rewritten as $J(x) = \sum_{t=1}^T w_t \cdot f_t(x)$, where $w_t = \frac{1}{(\sigma^t)^2}$ and $f_t(x) = \sum_{i \in t} (F_i(x) - y_i)^2$. In practice, as we have mentioned earlier, we do not know the exact values

of σ^t . As a result, instead of a single model x , we need to consider models x corresponding to all possible combinations of the weights w_t . It is known that under some reasonable conditions, this set coincides with the Pareto set for a multi-objective optimization problem

$$f_1(x) \rightarrow \min, f_2(x) \rightarrow \min, \dots, f_T(x) \rightarrow \min, \quad (5.5.3)$$

where a Pareto set is defined as the set of all tuples x for which it is not possible to improve all criteria, i.e., for which

$$\nexists x' (\forall t (f_t(x') \leq f_t(x))) \ \& \ \exists t (f_t(x') < f_t(x)). \quad (5.5.4)$$

Once we have generated all such tuples, it is reasonable to select a tuple x which is the closest to the ideal point. Specifically, for each type t , we find the smallest f_t^{\min} and the largest f_t^{\max} of the values, $f_t(x)$, and then we find the tuple x for which

$$d^2(f^{\min}, f(x)) = \sum_{t=1}^T \left(\frac{f_t(x) - f_t^{\min}(x)}{f_t^{\max}(x) - f_t^{\min}(x)} \right)^2 \quad (5.5.5)$$

is the smallest possible [57]. To find a model x corresponding to each combination of weights w , we use an iterative algorithm. Once the k -th approximation x_k is computed, we linearize the problem, by replacing $F_i(x)$ with $F_i(x_k) + F'_i(x_k)(x - x_k)$, and then we use the primal dual interior-point method to solve the corresponding quadratic optimization problem [121].

To make sure that the resulting model is geophysically meaningful, i.e., all velocities x_i are within the bounds $a_i \leq x_i \leq b_i$ corresponding to the depth of the i -th location, we optimize the quadratic function $f_t(x)$ under the constraints $a_i \leq x_i \leq b_i$. For all the stations involved in the geophysical inversion, the initial velocity model, x_0 , corresponds to the AK-135 model of [56], starting at 10 km depth and distributed at a 2 km interval up to 70 km depth, then at a 5 km interval up to 250 km and finally at 10 km until 300 km. By incorporating explicit velocity bound constraints into the inversion model, we are able to produce a better constrained model while having more stable inversions [121].

5.6 Joint Inversion Results

Joint inversion means minimizing the weighted sum $w_1 \cdot f_1 + w_2 \cdot f_2 + \dots$ of the sums f_1, f_2, \dots corresponding to measurements of a different type; the weights w_t are inversely proportional to the standard deviations of the measurements of the corresponding type. However, we do not know the exact values of these standard deviations. We can make such crude estimates of these values and thus, of the weights. However, if the 3-D models are based on these approximate weights, then, if we see some unusual features on the 3-D model, a geophysicist would be somewhat skeptical to what extent the features are real and not caused by a specific selection of weights. It would be much more convincing if, instead of a single combination of weights, we would try different combinations and show for all these combinations, we see similar features in the 3-D models.

This is, in effect, why we see the multi-objective optimization technique: instead of finding a model that optimizes a single objective function S , we find the set of optimal models corresponding to different weighted combinations of several objective functions f_1, f_2, \dots . If the new feature appears in all these models, then a geophysicist would be much more confident that the similar features revealed in the 3-D models are real and not caused by a arbitrary selection of weights.

An additional advantage is that by comparing different models, we can see how accurately we can determine the parameters of this feature, for example, if the average size of a pluton in different models is approximately 7.5 km, and the standard deviation estimates based on different maps is 2.5 km, then we know that the actual size is 7.5 ± 2.5 .

5.7 Kriging interpolation from 1D velocity profiles

Since our ultimate goal is to create a 3D Earth structure model of the Texas region, we use the 1D S wave velocity profiles of each station as input data for a kriging interpolation algorithm [113]. In general, interpolation algorithms estimate values by using a weighted

sum of surrounding data. Kriging represents an example of a computationally efficient interpolation technique that allows the incorporation of uncertainty on the predicted values. We implement a Bayesian kriging approach that integrates variable spatial damping, a useful tool to control the kriged solution in extrapolation zones where few or no data is available [113]. In our case, the station spacing within our region represents a 2D spatial grid, with each station now having a depth varying 1-D velocity structure. We thus can estimate the unknown velocities of the 2-D grid at different depths based on the known velocities, thus creating our 3-D model.

Initially, we remove an appropriate trend prior to applying kriging [113], which, in our case, corresponds to the mean of the velocities at a certain depth. A spatially damped kriging estimator then incorporates variable damping and measurement error multiplied by a unit-normalized function, which decreases noise values to zero according to the predictions point relative distance. As a result, we obtain a smoothly damping effect over the predicted velocities that varies according to each velocity node and its surroundings. For our results, we choose the blending functions of 2° to guarantee good spatial sampling.

Interpolating the 1-D profiles by means of kriging can help us to illuminate better the Earth structure beneath each station in the Texas region. Schematically, if each station had perfect azimuthal coverage, the region below each station would have cone shaped raypaths, where at a certain depth (that depends on station spacing), the raypaths at adjacent stations begin to overlap, providing us with full subsurface structure coverage. Before this depth, we expect that the surface wave group dispersion information obtained from regional earthquakes can improve the average crustal velocity structure and also the vertical resolution [112]. In this fashion, we account for velocity structure resolution avoiding additional inversions by grouping the 1D profiles depending on azimuthal range as indicated by [6]. Generally, the upper mantle of a tectonically active region is expected to exhibit 3D heterogeneities with a length scale smaller than both the lateral resolution of surface waves and vertical resolution of receiver functions [84]. Therefore, the models obtained by using these two data sets should resolve the main features beneath the region of study.

5.8 3-D Texas S-Wave Velocity Model

We perform 1-D joint inversions using our MOP approach for 270 stations from EarthScope USArray network. In general, each independent joint inversion includes at least 3 receiver function bins created according to an average ray parameter, with a width of approximately 0.01 s/km between 0.04 s/km and 0.07 s/km. The average ray parameter was determined by taking the mean value of the maximum and minimum ray parameters for each station before being used for stacking. The number of receiver functions employed to create these stacks depends on the station, but in general is not less than 25 per ray parameter. Each receiver function consists of 820 data points for a time range from -5 to 80 seconds. We also include fundamental mode Love and Rayleigh group velocities with 50 to 65 dispersion measurements, with periods between 5 to 140 seconds. Since the station spacing of the USArray is about 70 km, we anticipate lateral resolution of that order for each individual 1-D inversion in the Texas region.

We first determined the average map and deviation by taking the arithmetic average of maps corresponding to different combination of models. A minor problem with this approach is that for abrupt discontinuous transitions (typical of the Earth structure), even when all the models show an abrupt transition but they all show it at somewhat different depths; the average model will have this transition blurred in the 3-D model. From this viewpoint, a better approach may be to take a median of several models instead of arithmetic average; in this case, there is no blurring.

We first estimated the difference between different models by taking the differences $d_1(x) - d_2(x)$ between the velocity values at the corresponding 3-D points x , and then take, e.g., the rms of these differences, i.e., the square root of the sum of all the values $(d_1(x) - d_2(x))^2$ corresponding to different spatial points x . The problem is that if we have two abrupt transitions at somewhat different depths, these models are very similar, but the sum of the squares of the difference may be huge.

We will get a more accurate description of how close the models are if we compare the

value $d_1(x)$ at each 3-D point x not only with the value $d_2(x)$ at the same spatial location x , but also with the values $d_2(y)$ for all values y in an appropriate neighborhood of x . To be more precise, instead of the sum of the values $(d_1(x) - d_2(x))^2$, we take the sum of the values $\min_y (d_1(x) - d_2(y))^2$ where minimum is taken over all points y for which the distance $d(x, y)$ does not exceed a certain threshold ϵ

Figures 5 through 9 show the resulting 3D crustal and upper mantle structure images with a combination of different weights with the constraint $W_R F + W_S W = 1$. Figure 10 and 11 show the overall average of the ten different shear wave models with varying weights. Figure 10 shows the average of all the models in a 3-D perspective for the Texas region. Figure 11 shows the resulting average shear wave depth slices in a different perspective other than a 3-D, in order to highlight any anomalies within the major tectonic provinces of the Texas region. For Figure 11, we use a color scale (maximum to minimum velocities) to highlight crustal anomalies and a different color scale in Figures 5 through 9 with a reduced color spectrum varying from 4.0 km/s to 5 km/s, similar to that used by [136] to highlight mantle anomalies.

Figures 12a-e show our 2-D velocity model cross-sections A-A', B-B', C-C', D-D', and E-E'. Figure 12a, cross section A-A', coincides with latitude 32°, and passes through Delaware Basin, Midland Basin, and ends at the Balcones Fault Zone. Cross-section B-B' (Fig 12b) coincides with latitude 33° and covers the part of the Rio Grande Rift (RGR) and the center of our study area. In figures 12c-e, coincides with latitudes of 34°-36° and covers part of the RGR, as well as, the Southern Oklahoma Aulacagen (SOA). The lithosphere-athenosphere boundary (LAB) seems to lie approximately 140 km depth for the average 2-D shear wave model cross-sections that coincides with latitudes 32°-33°. The cross-sections at 34°-36°, seem to have the LAB around 160 km near the SOA tectonic province. To further investigate how accurate are our upper mantle shear wave models, Figures 13-14, shows depth slices from 50 to 300 km depth in 50 km intervals and 3-D

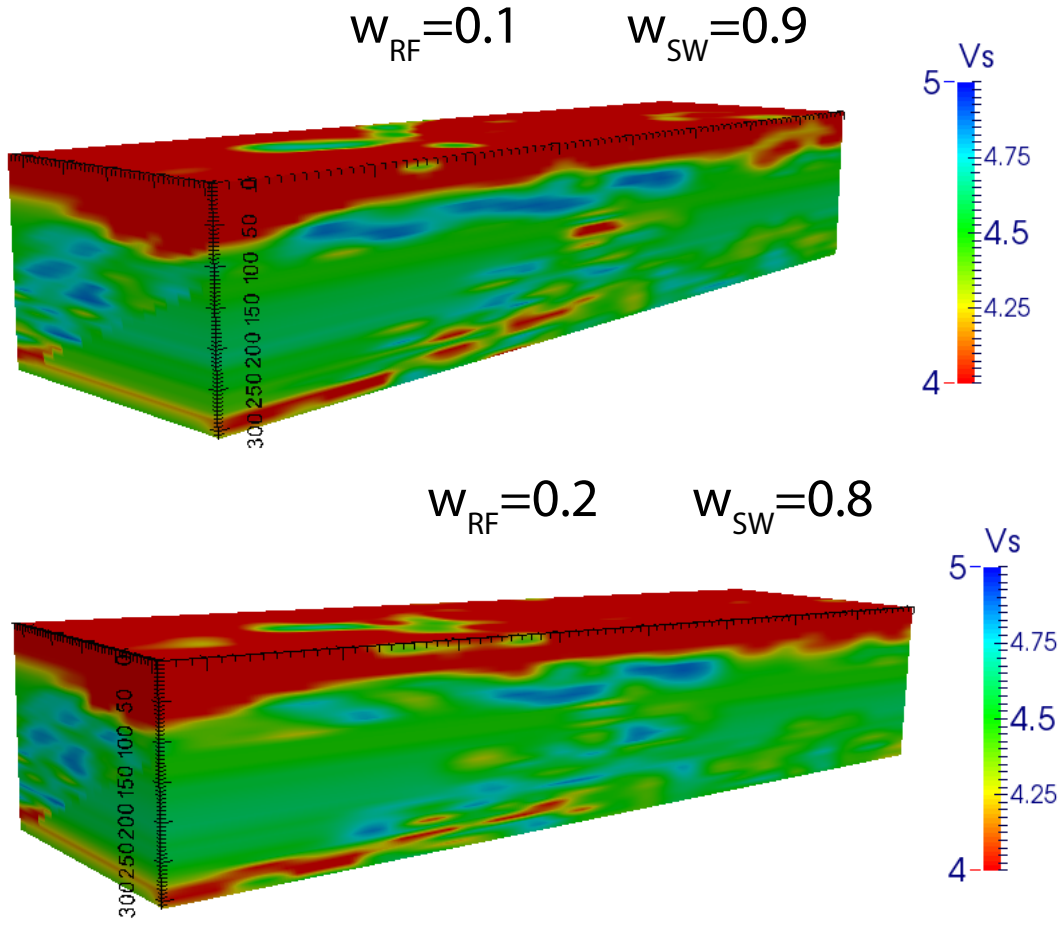


Figure 5.6: 3-D shear wave model of the Texas region from the surface to 300km depth. Red colors represent low shear wave velocities (4 km/s) and blue colors represent high shear wave velocities (5 km/s). $W_{RF}=0.1, 0.2$ and $W_{SW}=0.9, 0.8$ are the weights that were used to generate the 3-D models with the constraint $W_{RF} + W_{SW} = 1$.

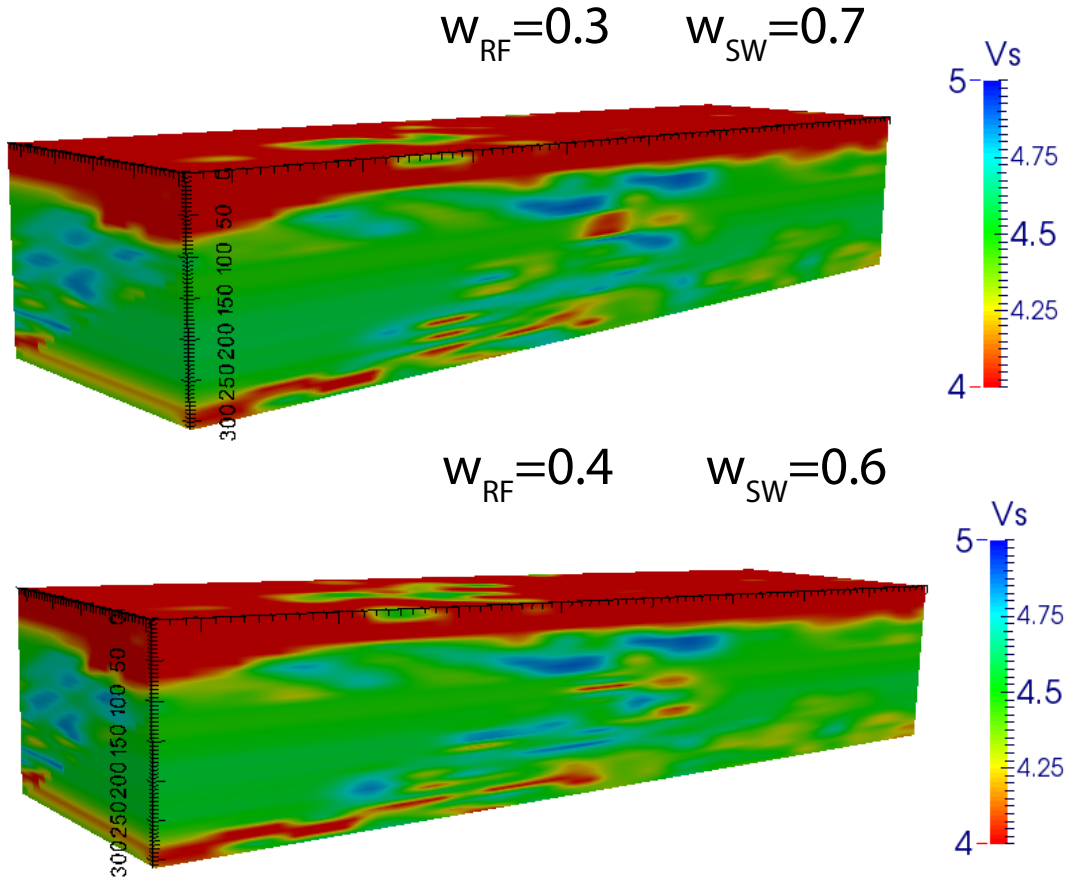


Figure 5.7: 3-D shear wave model of the Texas region from the surface to 300km depth. Red colors represent low shear wave velocities (4 km/s) and blue colors represent high shear wave velocities (5 km/s). $W_{RF}=0.3, 0.4$ and $W_{SW}=0.7, 0.6$ are the weights that were used to generate the 3-D models with the constraint $W_{RF} + W_{SW} = 1$.

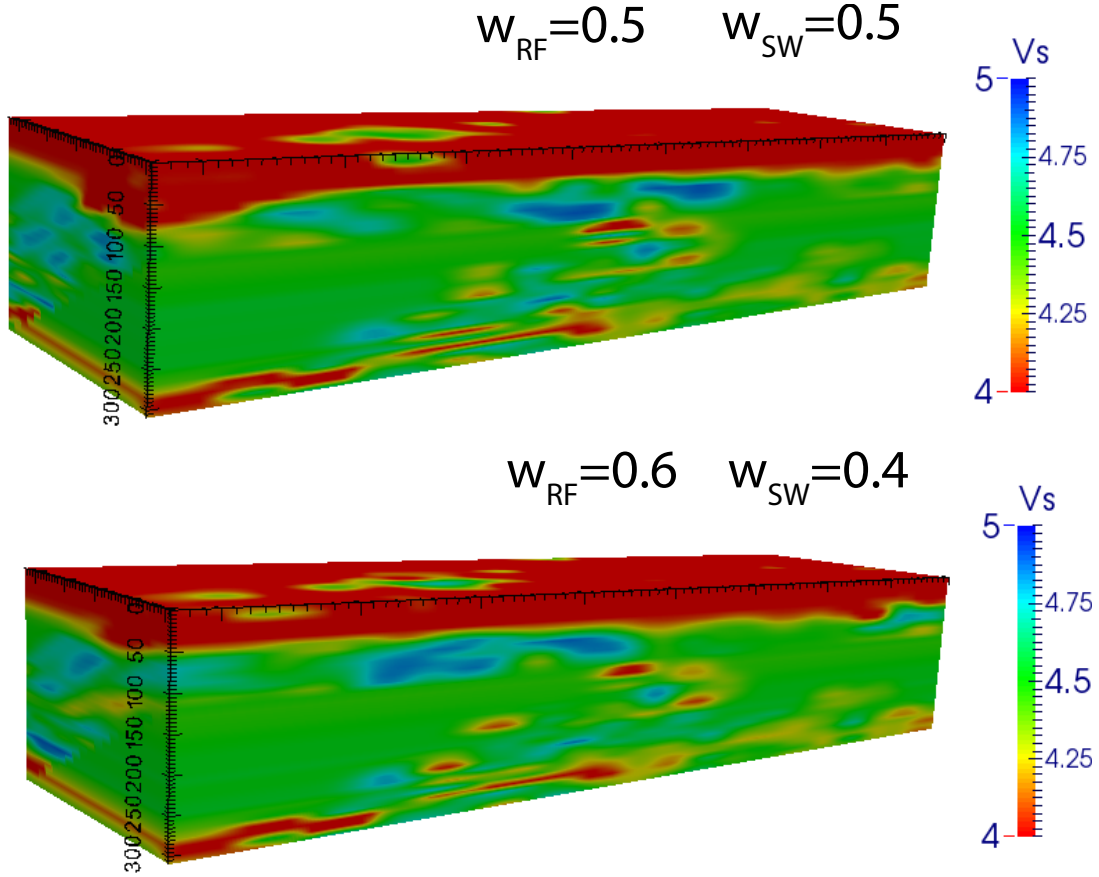


Figure 5.8: 3-D shear wave model of the Texas region from the surface to 300km depth. Red colors represent low shear wave velocities (4 km/s) and blue colors represent high shear wave velocities (5 km/s). $W_{RF}=0.5, 0.6$ and $W_{SW}=0.5, 0.4$ are the weights that were used to generate the 3-D models with the constraint $W_{RF} + W_{SW} = 1$.

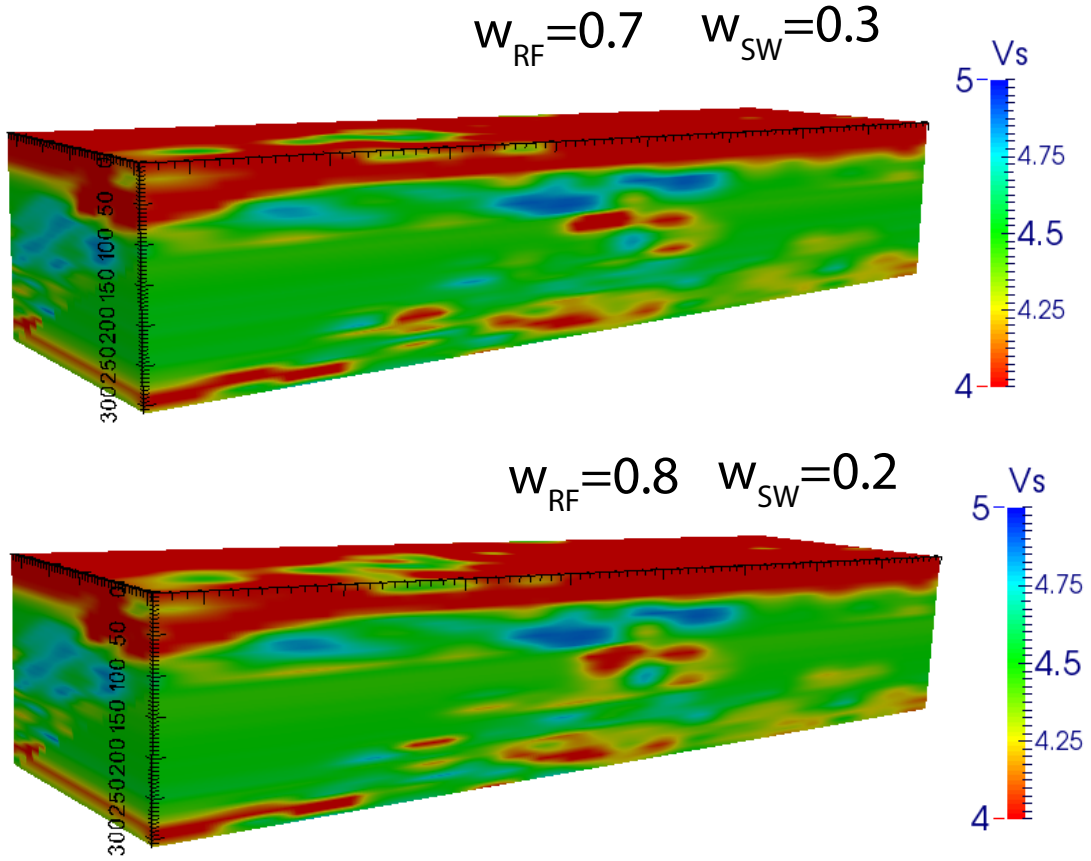


Figure 5.9: 3-D shear wave model of the Texas region from the surface to 300km depth. Red colors represent low shear wave velocities (4 km/s) and blue colors represent high shear wave velocities (5 km/s). $W_{RF}=0.7, 0.8$ and $W_{SW}=0.3, 0.2$ are the weights that were used to generate the 3-D models with the constraint $W_{RF} + W_{SW} = 1$.

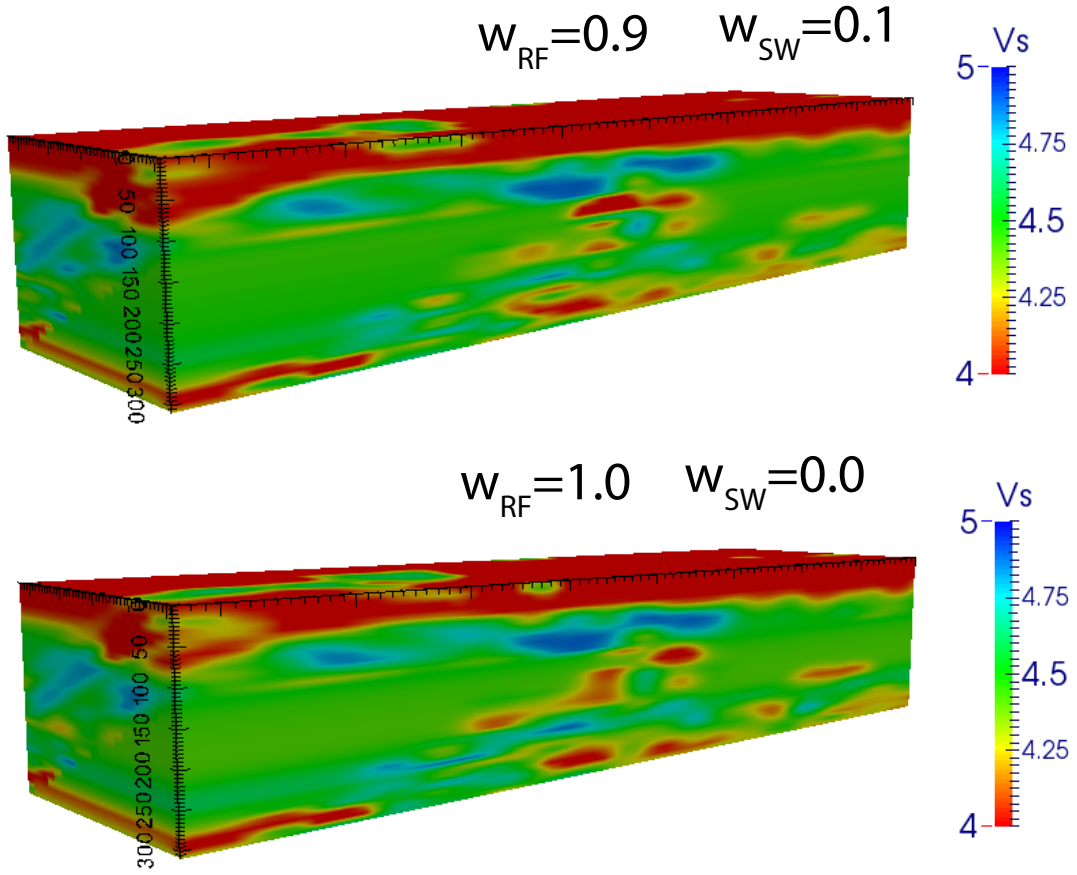


Figure 5.10: 3-D shear wave model of the Texas region from the surface to 300km depth. Red colors represent low shear wave velocities (4 km/s) and blue colors represent high shear wave velocities (5 km/s). $W_{RF}=0.9, 1.0$ and $W_{SW}=0.1, 0.0$ are the weights that were used to generate the 3-D models with the constraint $W_{RF} + W_{SW} = 1$.

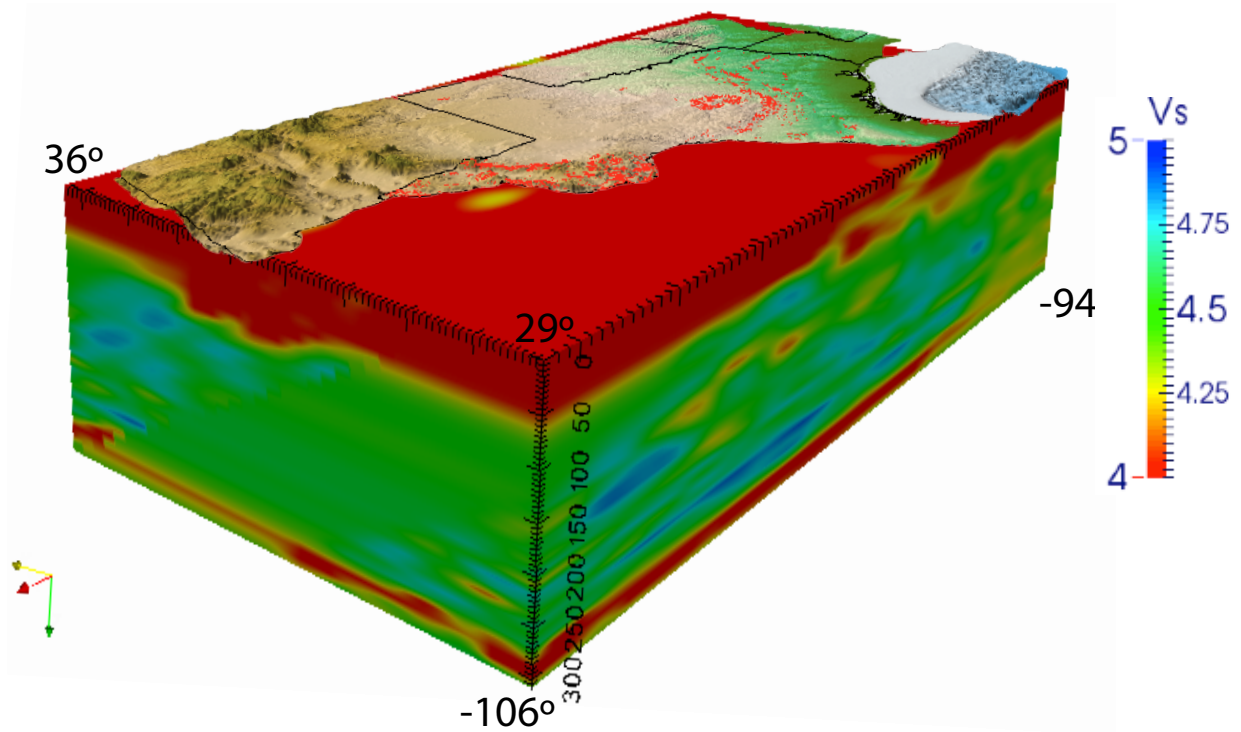


Figure 5.11: Average of the ten 3-D shear wave model of the Texas region from the surface to 300km depth. Red colors represent low shear wave velocities (4 km/s) and blue colors represent high shear wave velocities (5 km/s).

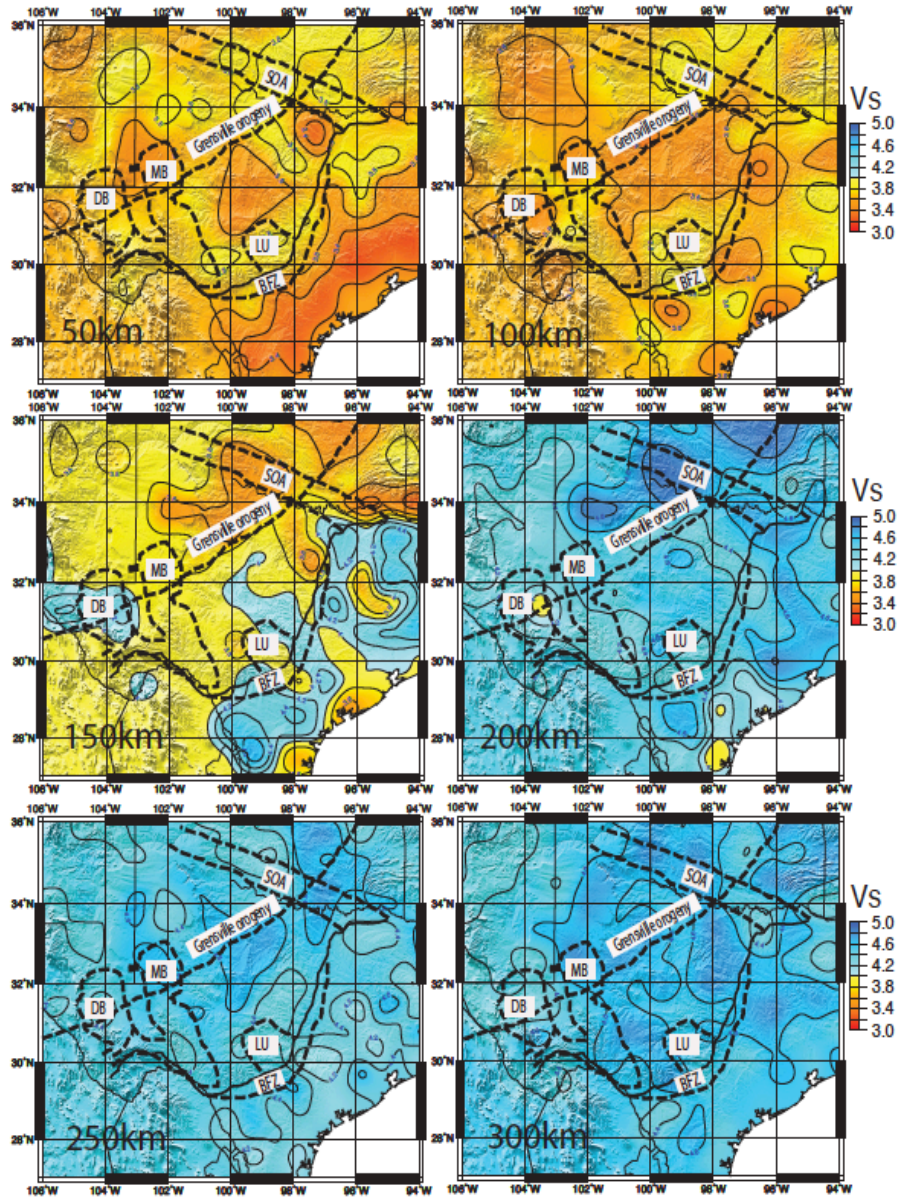
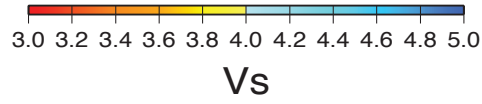


Figure 5.12: Average of the ten-shear wave model of the Texas region from the surface to 300km depth. Red colors represent low shear wave velocities (3 km/s) and blue colors represent high shear wave velocities (5 km/s).



Cross Section average 32 degrees

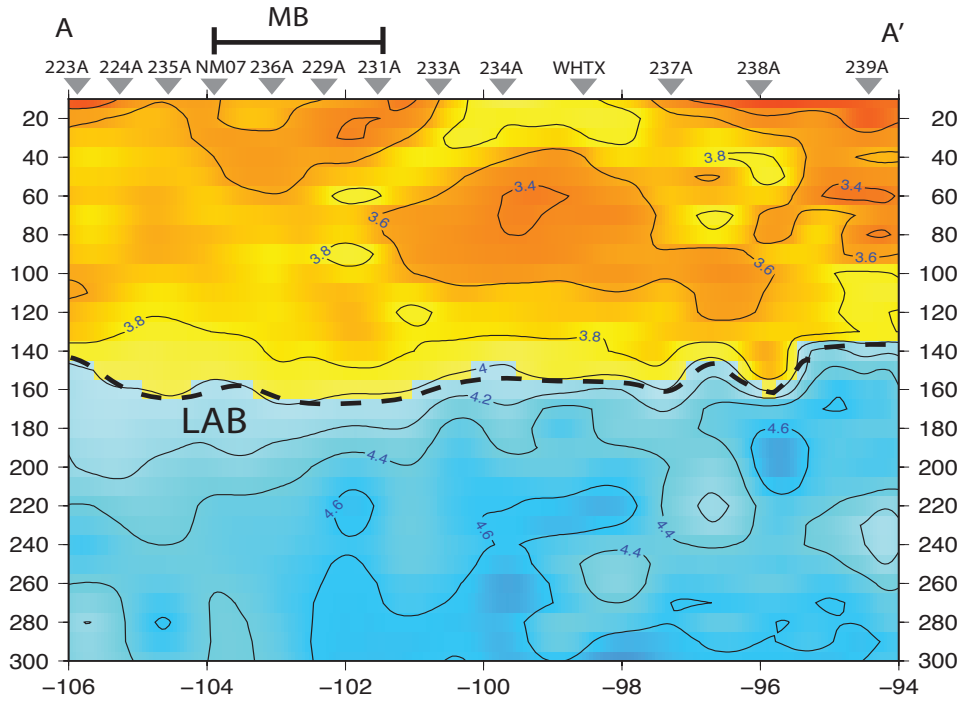


Figure 5.13: 2-D average shear wave model of ten models using different combination of weights for the Texas region. Fig 3a shows where we took the cross section within Texas. Gray inverted triangles represent USArray stations. Red colors represent low shear wave values and blue colors represent high shear wave values.

3.0 3.2 3.4 3.6 3.8 4.0 4.2 4.4 4.6 4.8 5.0

V_s

Cross Section average 33 degrees

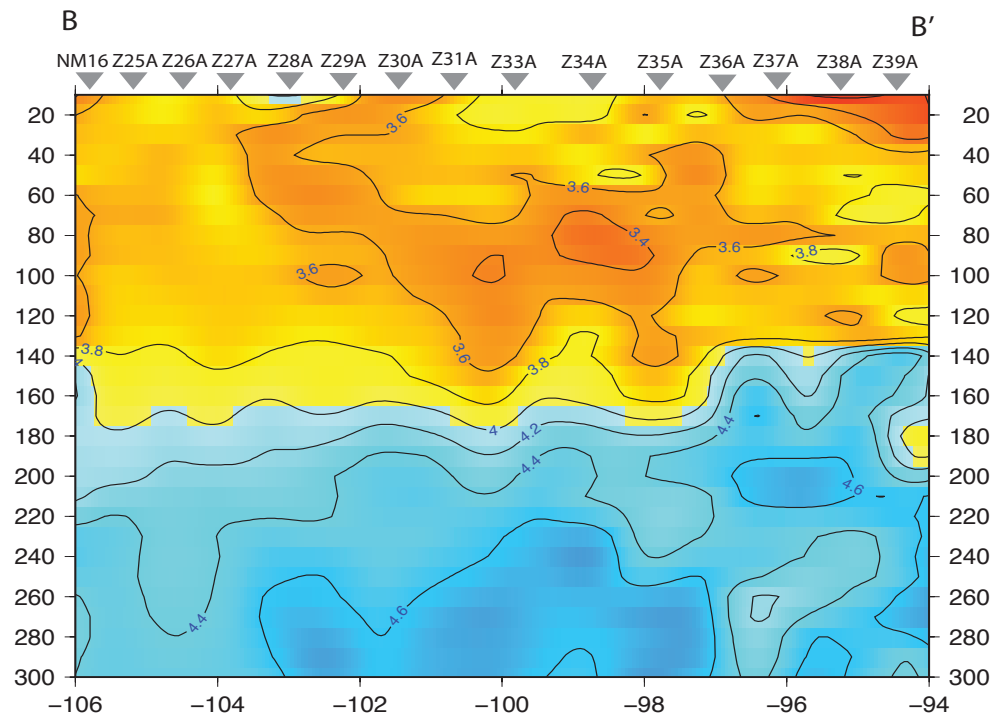
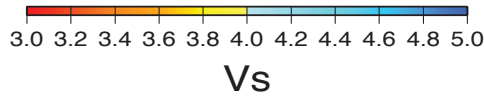


Figure 5.14: 2-D average shear wave model of ten models using different combination of weights for the Texas region. Fig 3a shows where we took the cross section within Texas. Gray inverted triangles represent USArray stations. Red colors represent low shear wave values and blue colors represent high shear wave values.



Cross Section average 34 degrees

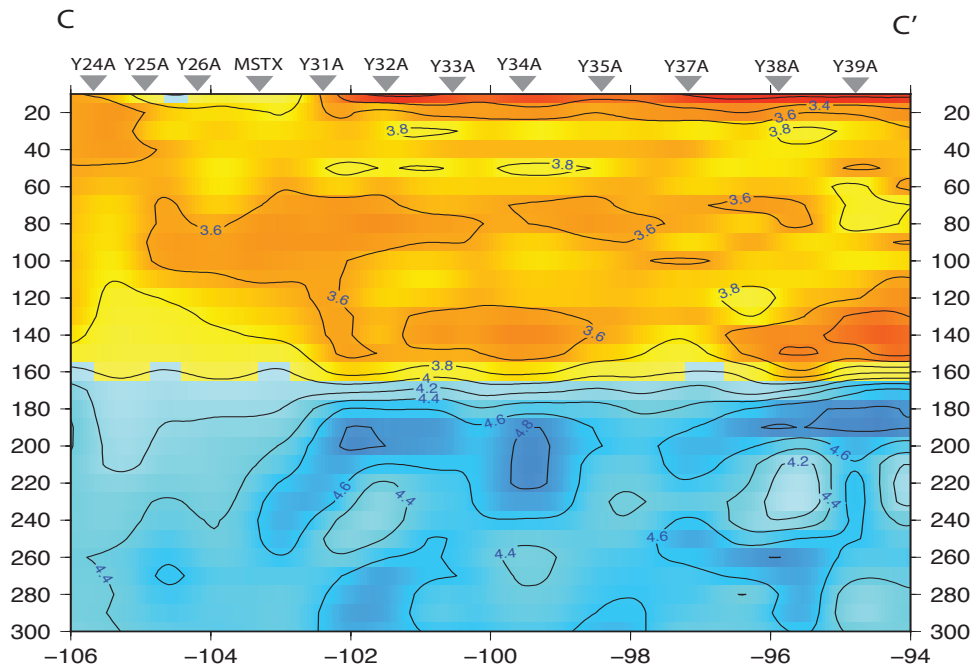


Figure 5.15: 2-D average shear wave model of ten models using different combination of weights for the Texas region. Fig 3a shows where we took the cross section within Texas. Gray inverted triangles represent USArray stations. Red colors represent low shear wave values and blue colors represent high shear wave values.

3.0 3.2 3.4 3.6 3.8 4.0 4.2 4.4 4.6 4.8 5.0

V_s

Cross Section average 35 degrees

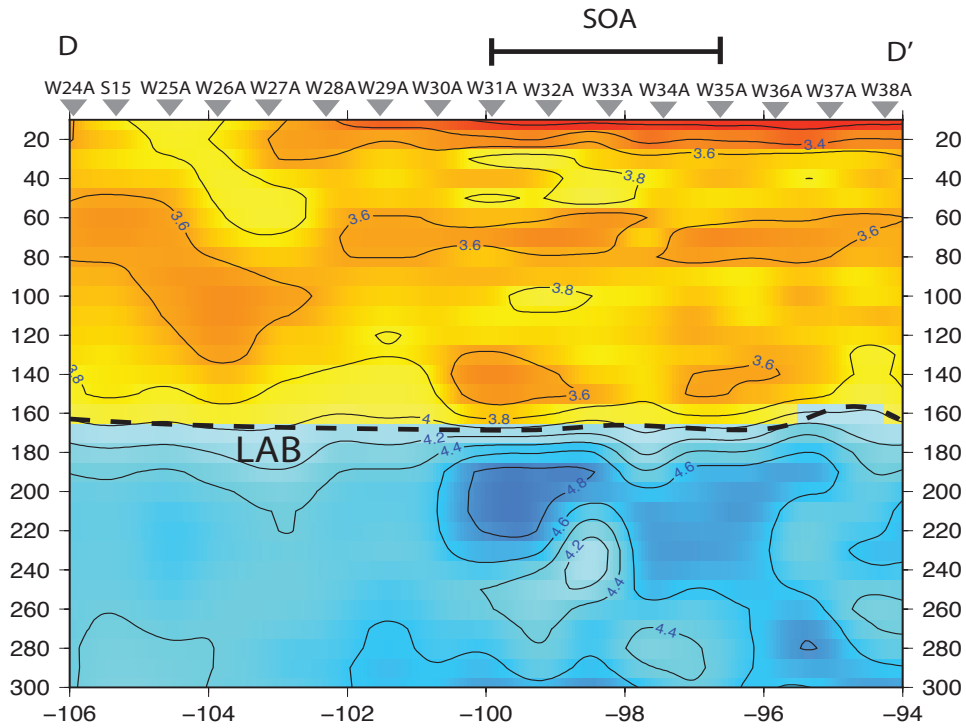
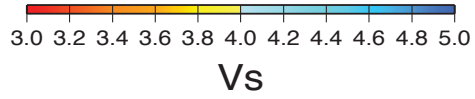


Figure 5.16: 2-D average shear wave model of ten models using different combination of weights for the Texas region. Fig 3a shows where we took the cross section within Texas. Gray inverted triangles represent USArray stations. Red colors represent low shear wave values and blue colors represent high shear wave values.



Cross Section average 36 degrees

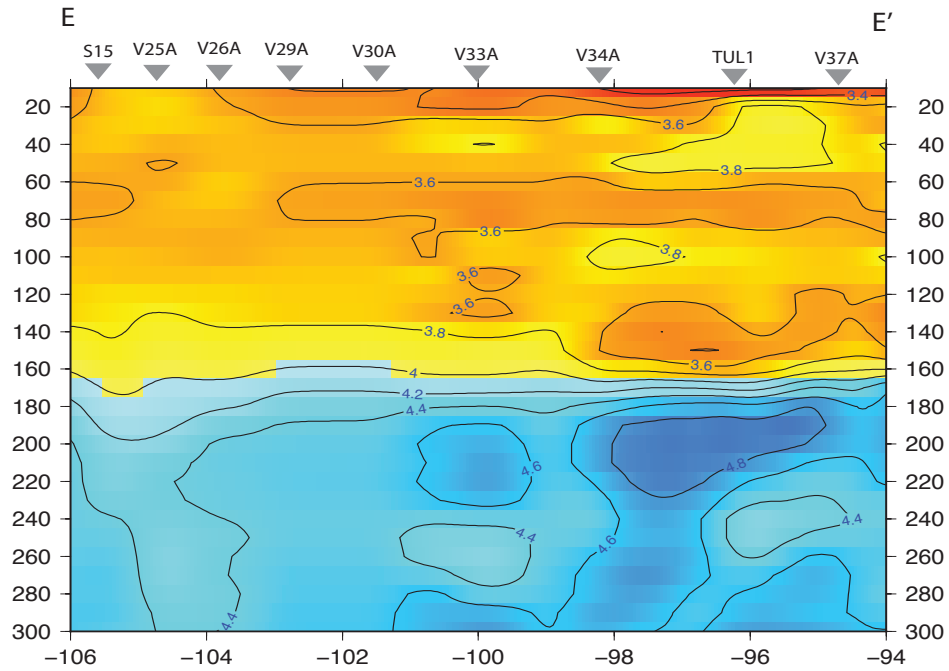


Figure 5.17: 2-D average shear wave model of ten models using different combination of weights for the Texas region. Fig 3a shows where we took the cross section within Texas. Gray inverted triangles represent USArray stations. Red colors represent low shear wave values and blue colors represent high shear wave values.

standard deviation of the ten shear wave models in a 3-D perspective for our study area.

Comparing our 3-D model directly to others can be difficult and requires complete visualization of all models in one figure (e.g., [10], [88]). Since our models were derived from 1-D inversions that included two complementary data sets and then interpolated to 3-D, we expected to find differences in our 3-D model compared to previous studies. We also focus solely in this region, while many of the other models include much larger regions ([35]; [64], [81], [119]). Thus, we show features relevant to Texas evolution, and highlight those that we believe are fully consistent with most recent and past models in the Texas region obtained by using different data sets and techniques (e.g., [93],[94], [111]). We discuss below the impact of the MOP approach, the new model for the Texas region.

5.8.1 MOP methodology

We implement a new approach for joint inversion of receiver function and surface wave group dispersion data based on constrained optimization [121]. We create independent 1D Earth velocity profiles of upper mantle velocity structure along the Texas region. In general, no smoothing or damping factors were used to stabilize the inversions. This is an advantage when compared to other standard techniques, which often require tuning of several regularization parameters. It is well known that for severely ill-posed inversions that may appear for some stations, determining the optimum regularization parameters remain as a difficult and often speculative task [121], [150]. There are several strategies to choose these regularization parameters [44], [150], and some recent advances on transdimensional inversion to include them directly as inversion parameters [14]. However, there is not yet an analytical or automatic way to find the best parameter for all particular cases [44], [78]. The MOP approach reduces the subjectivity of these selections, since the simple inclusion of a damping parameter stabilizes the inversion [121]. Some stations have inversion convergence issues mostly due to the absence of good quality RF data to identify absolute S-wave velocities and sharp discontinuities.

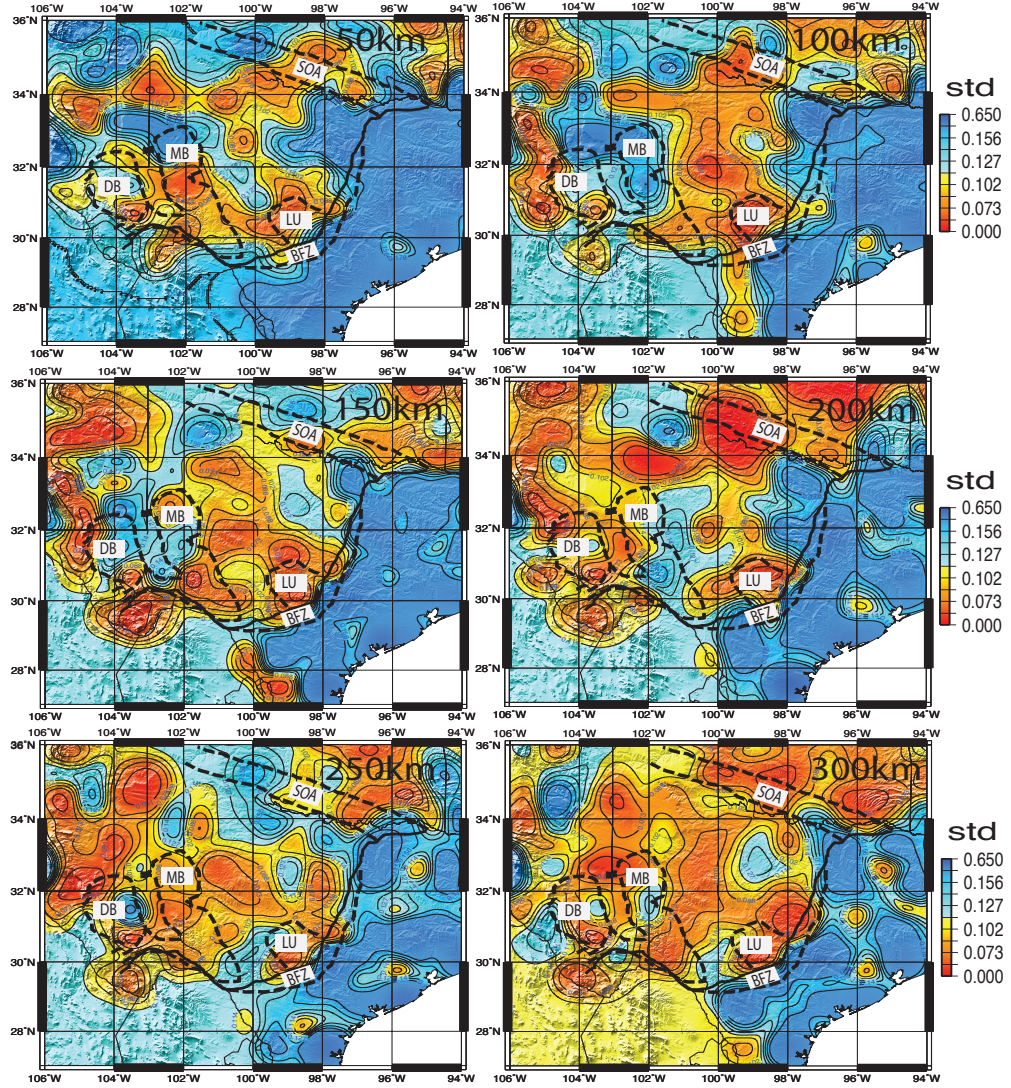


Figure 5.18: 3-D standard deviation (std) model of the Texas region using the difference of the ten 3-D shear wave models. Red colors represent low standard deviation values or error of the ten models and blue colors represent high std values of all the 3-D models.

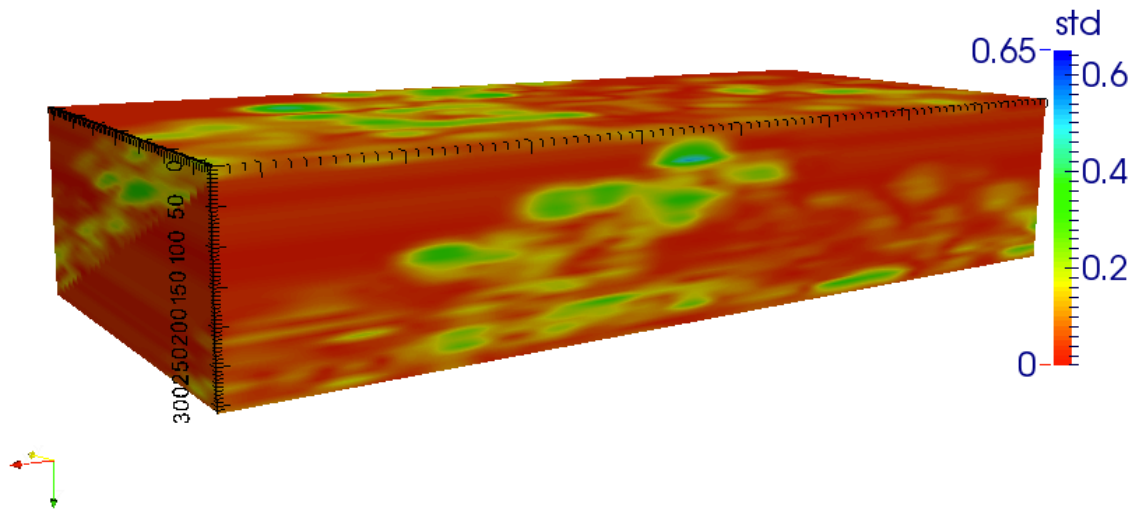


Figure 5.19: 3-D standard deviation (std) model of the Texas region using the difference of the ten 3-D shear wave models. Red colors represent low standard deviation values or error of the ten models and blue colors represent high std values of all the 3-D models.

[6] use a similar approach combining results obtained by independent joint inversion of surface wave phase velocities and receiver function information, to compute 1D S wave velocity profiles for the Colorado Plateau. By creating 3D images of Earth structure based on independent inversions that produce 1D velocity profiles, results are likely inferior compared to a full 3D parameter approach [6]. However, a 3D inversion involves a high volume of information and a great number of parameters, which may be numerically intractable. We will explore the migration of our technique to a full 3D inversion in future work using other geophysical datasets such as travel times and gravity.

We can adjust our inversion approach by tuning the kriging interpolation blending function according to the geological province where the stations are located, e.g. Delaware Basin, Llano Uplift, Balcones Fault Zone, and the Southern Oklahoma Aulagen, as an attempt to enhance the performance of the inversion algorithm. However, our approach allows us to begin with a standard initial model (ak135: [56]), and produce coherent independent inversions that could be combined for a consistent 3D structure. This represents another difference to the approach presented by [6], where they weight the contribution of each inversion separately to select a suitable initial model.

5.8.2 Interpretation

[43] argue that the Southern Oklahoma and adjacent parts of Texas contain an extensive igneous province emplaced during Early Cambrian rifting within the Southern Oklahoma rift zone. The SOA rift zone contains an extensive, largely bimodal igneous assemblage and its considered to be Early Cambrian in age but could also extend back into late Neoproterozoic for the Texas-Oklahoma border ([43]). Southern Oklahoma rift zone, contains felsic igneous rocks of ~ 1.4 Ga. The southern midcontinent of granite-rhyolite province, form the main subsurface basement units of each side of the Southern Oklahoma rift zone (e.g., [130], [97]). Based from residual gravity anomaly maps by [43], they found that the Southern Oklahoma rift is a high density-high velocity body (mafic igneous complex) that is ~ 40 km wide and extends at least to a depth of 10 km. Southern Oklahoma rift zone

is volcanic fill and has a distinct geophysical signature. Figure 11 shows a clear SW-NE seismic trend that coincides with the high density-high velocity body that [43] found using gravity. Figure 13 supports the high-velocity anomaly found because the standard deviation is very low which results in a well-constrained locality. We have taken several cross sections across 34° , 35° , and 36° , which shows consistent transition from low to high velocity contrast that lies within the Southern Oklahoma rift zone.

In general, the presence of high-velocity lithosphere in the upper mantle is known to be a ubiquitous feature for stable continental provinces like the Texas region ([146]). Under certain tectonic environments, old cratons can be reactivated and lead to extensive magmatic activity which could explain why we find such high velocities at ~ 200 km depth.

5.9 Discussion

The cross-sections at 34° - 36° , have a deep LAB signature around 160 km near the SOA tectonic province which are similar results of [59], where they found a deep LAB structure found under the border region of Texas and Oklahoma and over the Colorado Plateau, which they believe could be remnants of the Farallon slab. Previous studies (e.g., [27], [111]) have discussed that a dissected part of the Farallon slab was accreted with portions of the North American craton. [92] and [85], have found high velocities in their tomography results and interpret their observations as having subduction fragments that still remain deep in the Earth Structure. [59] found a deep LAB structure using receiver functions, and discovered that the LAB lies within the Texas-Oklahoma border region where the SOA is located and known to be a Proterozoic mid-continental rift [137], [59]. [63] have also found high velocity anomalies in their data below the Colorado Plateau, which they also derived from receiver functions and interpreted the high velocities as a result of delamination of lower parts from the lithosphere. The LAB boundary usually are not imaged underneath thick continental interiors, although the characteristics of the boundary vary in different tectonic settings, it has remained a puzzle as to why the LAB would be so variable in

nature [105].

5.10 Conclusions

We present a new model of crustal and upper mantle structure beneath the Texas region. Separate joint inversions were performed for 270 Earthscope USArray stations. We create a generic framework that connects a constrained optimization joint inversion algorithm with a Bayesian interpolation scheme for high-resolution imaging of Earth structure. Furthermore, this scheme efficiently provides a robust alternative to extend simultaneous independently created 1D S wave velocity models, to produce 3D images of Earth structure compared to full 3D inversions. Our framework generates a continuous and smooth 3D velocity model of the ancient Rift system within Texas, revealing the complexities of the tectonic provinces within our study area and helping us to better characterize its crustal and upper mantle velocity structure. We identify the LAB boundaries for the provinces of DB, MB, LU, SOA and BFZ. We have no evidence of a deep mantle plume that drives the current rifting process, and all velocity anomalies are shallower than 120 km.

References

- [1] D. L. Abt and K. M. Fischer and S. W. French and H. A. Ford and H. Yuan and B. Romanowicz, “North American lithospheric discontinuity structure imaged by Ps and Sp receiver functions,” *J. Geophys. Res.*, 2010,115,B09301,doi:10.1029/2009JB006914.
- [2] D. C. Adams, and G. R. Keller, “Precambrian Basement Geology of the Permian Basin Region of West Texas and Eastern New Mexico: A geophysical Perspective,” *AAPG Bulletin*, 1996,80,pp. 410–431.
- [3] M. J. Aldrich, “Tectonics of the Jemez Lineament in the Jemez Mountains and the Rio Grande Rift,” *J. Geophys. Res.*, 1986,91,pp. 1753–1762.
- [4] C. J. Ammon, “The Isolation of Receiver Effects from Teleseismic P Wave forms,” *Bull. Seism. Soc. Am.*, 1991,81,pp. 2504–2510.
- [5] M. G. Averill and K. C. Miller and G. R. Keller and V. Kreinovich and R. Araiza and S. A. Starks, “Using expert knowledge in solving the seismic inverse problem,” *International Journal of Approximate Reasoning*, 2007,45, pp. 564–587.
- [6] I. W. Bailey and M. S. Miller and K. Liu and A. Levander, “Vs and density structure beneath the Colorado Plateau constrained by gravity anomalies and joint inversions of receiver function and phase velocity data,” *J. Geophys. Res.*, 2012,117,B02313,doi:10.1029/2011JB008522.
- [7] S. W. Baldrige and K. H. Olsen, “The rio grande rift,” *Am Sci*, 1989,77, pp. 240–247.
- [8] V. A. Bankey and Cuevas D. Daniels and C. Finn and I. Hernandez and P. Hill and R. Kucks and W. Miles and M. Pilkington and C. Roberts and W. Roest and V. Rystrom and S. Shearer and S. Snyder and R. Sweeney and J. Velez and J. Phillips and D. Ravat, “Digital data grids for the magnetic anomaly187 map of North America,”

U.S. Geological Survey Open-File Report 02-414, U.S. Geological Survey, 2002, Denver, Colorado, USA.

- [9] L. Bashir and S. S. Gao and K. H. Liu and K. Mickus, “Crustal structure and evolution beneath the Colorado Plateau and the southern Basin and Range Province: Results from receiver function and gravity studies,” *Geochem. Geophys. Geosyst*, 2011,12,Q06008,doi:10.1029/2011GC003563.
- [10] T. W. Becker, “On recent seismic tomography for the western United States,” *Geochem. Geophys. Geosyst*, 2012,13, Q01W10, doi:10.1029/2011GC003977.
- [11] G. D. Bensen and M. H. Ritzwoller and Y. Yang, “A 3D shear velocity model of the crust and uppermost mantle beneath the United States from ambient seismic noise,” *Geophysical Journal International*, 2009,177(3), pp. 1177-1196.
- [12] H. T. Berglund and A. F. Sheehan and M. H. Murray and M. Roy and A. R. Lowry and R. S. Nerem and F. Blume, “Distributed deformation: the Rio Grande Rift, Great Plains, and Colorado Plateau,” *Geology*, 2012,40(1),pp. 2326.
- [13] P. Bird, “Continental delamination and the Colorado Plateau,” *J. Geophys. Res*, 1979,84, pp. 7561-7571.
- [14] T. Bodin and M. Sambridge and H. Tkalcic and P. Arroucau and K. Gallagher and N. Rawlinson, “Transdimensional inversion of receiver functions and surface wave dispersion,” *J. Geophys. Res*, 2012,117(B2),doi:10.1029/2011JB008560.
- [15] T. M. Brocher, “Geometry and physical properties of the Socorro, New Mexico, magma bodies,” *J. Geophys. Res*, 1981,86, pp. 9420-9432.
- [16] J. S. Buehler and P. M. Shearer, “Pn tomography of the western United States using USArray,” *J. Geophys. Res*, 2010,115,B09315,doi:10.1029/2009JB006874.

- [17] J. S. Buehler and P. M Shearer, “Localized imaging of the uppermost mantle with USArray Pn data,” *J. Geophys. Res.*, 2012,117,B09305,doi:10.1029/2012JB009433.
- [18] S. Burdick and R. D. Van der Hilst and F. L. Vernon and V. Martynov and T. Cox and C. Eakins and J. Karasu and G. H. Tyllel and J. Astiz and G. L Pavlis, “Model update January 2010: upper mantle heterogeneity beneath North America from travel time tomography with global and USArray Transportable Array data,” *Seismol. Res. Lett.*, 2010,81, pp. 689-693.
- [19] J. W. Cady, “Calculation of gravity and magnetic anomalies of finite-length right polygonal prisms,” *Geophysics*, 1980,45,pp. 1507-1512.
- [20] A. Cao and A. Levander, “Highresolution transition zone structures of the Gorda Slab beneath the western United States: Implication for deep water subduction,” *Journal of Geophysical Research: Solid Earth*, 2010,115(B7).
- [21] C. E. Chapin and S. M. Cather, “Tectonic setting of the axial basins in the northern and central Rio Grande rift, in Basins of the Rio Grande Rift: Structure, Stratigraphy, and Tectonic Setting,” eds. *G. Randy Keller, Steven M. Cather, GSA Special Paper*, 1994,291, pp. 5-25.
- [22] K. H. Cho and R. B. Herrmann and C. J. Ammon and K Lee, “Imaging the upper crust of the Korean Peninsula by surface-wave tomography,” *Bulletin of the Seismological Society of America*, 2007,97(1B), pp. 198-207.
- [23] K. H. Cho and R. B. Herrmann and C. J. Ammon and K Lee, “Upper mantle P velocity structure beneath the Midwestern United States derived from triplicated waveforms” *Geochem. Geophys. Geosyst.*, 2012,13,Q0AK04,doi:10.1029/2011GC003818.
- [24] E. W. Collins and J. A. Raney, “Tertiary and Quaternary tectonics of the Hueco Bolson, Trans-Pecos Texas and Chihuahua, Mexico, in G. R. Keller and S. M. Cather

- (eds.), Basins of the Rio Grande rift; structure, stratigraphy, and tectonic setting” *Geo. Soc. Am.*, 1994, 291, pp. 265-282.
- [25] D. Colombo and M. De Stefano, “Geophysical modeling via simultaneous joint inversion of seismic, gravity, and electromagnetic data: Application to prestack depth imaging” *The Leading Edge*, 2007, 26, pp. 326-331.
- [26] F. A. Cook and E. R. Decker and S. B. Smithson, “Preliminary Transient Heat Flow Model of The Rio Grande Rift In Southern New Mexico,” *Earth Planet. Sci. Lett.*, 1978, 40, pp. 316-326.
- [27] F. A. Cook and E. R. Decker and S. B. Smithson, “Are diamond-bearing Cretaceous kimberlites related to low-angle subduction beneath western North America?” *Earth. Planet. Sci. Lett.*, 2011, 303, pp. 59-70.
- [28] M. T. Dugda and A. A. Nyblade and J. Julia and C. A. Langston and D. J. Ammon and S. Simiyu, “Crustal structure in Ethiopia and Kenya from receiver function analysis: Implications for rift development in eastern Africa,” *J. Geophys. Res.*, 2005, 110, pp. 1-15.
- [29] Y. Dzierma and W. Rabbal, M. M. Thorwart and E. R. Flueh and M. M. Mora and G. E. Alvarado, “The steeply subducting edge of the Cocos Ridge: Evidence from receiver functions Beneath the northern Talamanca Range, south-central Costa Rica,” *Geochem. Geophys. Geosyst.*, 2011, 12, Q04S30, doi:10.1029/2010GC003477.
- [30] Eaton, D. W., F. Darbyshire, R. L. Evans, H. Grutter, A. G. Jones, and X. Yuan, “The elusive lithosphere-asthenosphere boundary (LAB) beneath cratons” *Lithos.*, 2009, 109, pp. 1-22, doi.org/10.1016/j.lithos.2008.05.009.
- [31] Y. Fialko and M. Simons, “Evidence for on-going inflation of the Socorro magma body, New Mexico, from Interferometric Synthetic Aperture Radar imaging,” *Geophys. Res. Lett.*, 2001, 28, pp. 3549-3552.

- [32] K. M. Fishcher and H. A. Ford and D. L. Abt and C. A. Rychert , “The lithosphere-asthenosphere boundary” *Annual Review of Earth and Planetary Sciences*, 2010,38, pp. 551-575,doi:10.1146/annurev-earth-040809-152438.
- [33] L. A. Gallardo and M. A. Meju, “Joint two-dimensional DC resistivity and seismic traveltime inversion with cross-gradients constraints,” *J. Geophys. Res*, 2004,109,B03311,doi:10.1029/2003JB002716.
- [34] W. Gao and S. P. Grand and W. S. Baldrige and W. S. Wilson and M. West and J. F. Ni and R. Aster, “Upper mantle convection beneath the central Rio Grande Rift imaged by P and S wave Tomography,” *J. Geophys Res*, 2004,109,B03305,doi:10.1029/2003JB002743.
- [35] H. Gilbert, “Crustal structure and signatures of recent tectonism influenced by ancient terranes in the western United States,” *Geosphere*, 2012,18, pp. 141157,doi:10.1130/GES00720.1.
- [36] H. Gilbert and A. A. Velasco and G. Zandt, “Preservation of Proterozoic terrane boundaries within the Colorado Plateau and implications for its tectonic evolution,” *Earth Planet. Sci. Lett*, 2007,258, pp. 237-248.
- [37] F. Goff and C. J. Janik, “Gas geochemistry of the Valles caldera region, New Mexico and comparisons with gases at Yellowstone, Long Valley and other geothermal systems,” *J. Volcanol. Geotherm. Res*, 2002,116, pp. 299-323.
- [38] H. Gurrola and E. G. Baker and B. J. Minster, “Simultaneous time-domain deconvolution with application to the computation of receiver functions,” *J. geophys. Int*, 1995,120, pp. 537-543.
- [39] E. H. Isaaks and R. M. Srivastava, *An Introduction to Applied Geostatistics*, Oxford Univ. Press, New York, Oxford, 1989.

- [40] E. Haber and D. Oldenburg, “Joint inversion: a structural approach” *Inverse Problems*, 1997,13, pp. 63-77.
- [41] R. I. Hackney and W. E. Featherstone, “Geodetic versus geophysical perspectives of the gravity anomaly” *Geophys. J. Int*, 2003,154, pp. 35–43.
- [42] S. M. Hansen and K. G. Dueker and J. C. Stachnik and R. C. Aster and K. E. Karlstrom, “A rootless rockies - Support and lithospheric structure of the Colorado Rocky Mountains inferred from CREST and TA seismic data,” *Geochem. Geophys. Geosyst*, 2013,14, pp. 26702695.
- [43] R. E. Hanson and R. E. Puckett and Keller Jr. and G. R. Brueseke and M. E. Bulen and C. L. Mertzman and S. A. Finegan and D. A. McCleery, “Intraplate magmatism related to opening of the southern Iapetus Ocean: Cambrian Wichita igneous province in the Southern Oklahoma rift zone,” *Lithos*, 2013,174, pp. 57-70.
- [44] P. C. Hansen, *Discrete Inverse Problems*, SIAM, 2010.
- [45] W. A. Heiskanen and H. Moritz , *Physical Geodesy*, W. H. Freeman and Company, 1967.
- [46] R. B. Herrmann and C. J. Ammon, “Computer programs in seismology,” *Saint Louis University*, 2002,v 3.30.
- [47] R. B. Herrmann and C. J. Ammon and H. Benz, “Group velocity dispersion for North America,” [http : //www.eas.slu.edu/eqc/eqc_research/NATOMO](http://www.eas.slu.edu/eqc/eqc_research/NATOMO), 2013.
- [48] J. M. Hills, “Sedimentation, tectonism, and hydro-carbon generation in Delaware Basin, west Texas and New Mexico” *American Association of Petroleum Geologists Bulletin*, 1984,68, pp .250267.
- [49] G. Jin and J. B. Gaherty, “Surface Wave Measurement Based on Cross-correlation” *Geophys. J. Int*, 2014,(Manuscript Formatting).

- [50] A. G. Jones and J. C. Alfonso and J. Fulla and F. Salajegheh, “The lithosphereasthenosphere system beneath Ireland from integrated geophysicalpetrological modeling I: Observations, 1D and 2D hypothesis testing and modelin,” *Lithos*, 2013,<http://dx.doi.org/10.1016/j.lithos.2013.10.033>.
- [51] J. Julia and C. J. Ammon and R. Herrmann and M. Correig, “Joint inversion of receiver function and surface wave dispersion observations” *Geophysics J. Int*, 2000,142, pp. 99-112.
- [52] G. R Keller, “Geophysical constraints on the crustal structure of New Mexico” in *G. H. Macks and K. A. Giles,eds.,The Geology of New Mexico,A Geologic History: New Mexico Geological Society Special Publication*, 2004,11, pp. 439-456.
- [53] G. R. Keller and W. S. Baldrige, “The Rio Grande rift: a geological and geophysical overview” *Rocky Mount Geo*, 1999,34, pp. 121-130.
- [54] G. R. Keller and P. Morgan and W. R. Seager, “Crustal structure, gravity anomalies, and heat flow in the southern Rio Grande rift and their relationship to extensional tectonics” *Tectonophysics*, 1990,174, pp. 21-37.
- [55] G. R. Keller and M. A. Khan and P. Morgan and R. F. Wendlandt and W. S. Baldrige and K. H. Olsen and C. Prodehl and L. W. Braile, “A comparative study of the Rio Grande and Kenya rifts” *Tectonophysics*, 1991,197, pp. 355-371.
- [56] Kennett, B. L. N., Engdahl, E. R., and Buland, R., “Constraints on seismic velocities in the Earth from traveltimes” *Geophys. J. Int*, 1995,122(1), pp. 108-124.
- [57] E. Kozlovskaya, “An algorithm of geophysical data inversion based on non-probabilistic presentation of a-prior information and definition of pareto-optimality” *Inverse Problems*, 2000,16, pp. 839-861.

- [58] C. Kreemer and G. Blewitt and R. A. Bennett, “Present-day motion and deformation of the Colorado Plateau” *Geophys. Res. Lett*, 2010,37,L10311,doi:10.1029/2010GL043374.
- [59] P. Kumar and P. R. Kind and X. Yuan and J. Mechie and R. Kind and X. Yuan and J. Mechie, “USArray receiver function images of the Lithosphere-Asthenosphere boundary” *Seismol. Res. Lett*, 2012,83(3), pp. 486491.
- [60] C. A. Langston, “Evidence for the subducting lithosphere under southern Vancouver Island and western Oregon from teleseismic P wave conversions” *J. Geophys. Res*, 1981,86, pp. 3857-3866.
- [61] C. A. Langston, “Crust 2.0. The Current Limits of Resolution for Surface Wave Tomography in North America” *EOS Trans AGU*, 2000,81,F897,<http://igpppublic.ucsd.edu/gabi/ftp/crust2/>.
- [62] Lees, J.M. and Vandecar, J.C., “Seismic tomography constrained by bouguer gravity anomalies: Applications in western Washington” *PAGEOPH*, 1991,135, pp. 31-52.
- [63] Levander, A., Schmandt, B, Miller, M.S., Liu, K., Karlstrom, K.E., Crow, R.S., Lee, C.-T.A., and Humphreys, E.D., “Continuing Colorado plateau uplift by delamination-style convective lithospheric downwelling” *Nature*, 2011,472, pp. 461466.
- [64] A. Levander and M. S. Miller, “Evolutionary aspects of lithosphere discontinuity structure in the western US” *Geochem. Geophys. Geosyst*, 2012,13(7),Q0AK07,doi:10.1029/2012GC004056.
- [65] J. P. Ligorria and C. J. Ammon, “Iterative deconvolution and receiver-function estimation” *Bull. Seism. Soc. Am*, 1999,85, pp. 1395-1400.
- [66] F. C. Lin and B. Schmandt and V. C. Tsai, “Joint inversion of Rayleigh wave phase velocity and ellipticity using USArray: Constraining velocity and density structure in the upper crust” *Geophys. Res. Lett*, 2012,39,L12303,doi:10.1029/2012GL052196.

- [67] Lines, L.R., Schultz, A.K. and Treitel, S., “Cooperative inversion of geophysical data” *Geophysics*, 1988,53, pp. 8-20.
- [68] L. Liu and M. Gurnis, “Dynamic subsidence and uplift of the Colorado Plateau” *Geology*, 2010,38, pp. 663666.
- [69] Liu, K., A. Levander, F. Niu, and M. S. Miller, “Imaging crustal and upper mantle structure beneath the Colorado Plateau using finite frequency Rayleigh wave tomography” *Geochem. Geophys. Geosyst*, 2011,12,Q07001,doi:10.1029/2011GC003611.
- [70] A. Lodge and G. Helffrich, “Grid search inversion of teleseismic receiver functions” *Geophys. J. Int*, 2009,178, pp. 513-523.
- [71] A. R. Lowry and M. P. Gussinye, “The role of crustal quartz in controlling Cordilleran Deformation” *Nature*, 2011,471, pp. 353-357,doi:10.1038/nature09912.
- [72] M. Maceira and C. J Ammon, “Joint inversion of surface wave velocity and gravity observations and its application to central Asian basins s-velocity structure” *J. Geophys. Res*, 2009,114,B02314,doi:10.1029/2007JB0005157.
- [73] M. N. Machette, “Contrasts between short- and long-term records of seismicity in the Rio Grande rift; important implications for seismic-hazard assessments in areas of slow extension, in W. R. Lund (ed.), Proceedings volume; Basin and Range Province seismic-hazards summit” *Misc. Pub. - Utah Geo. Survey*, 1998,98, pp. 84-95.
- [74] K. Mickus and R. J. Stern and G. R. Keller and E. Y. Anthony, “Potential field evidence for volcanic rifted margin along the Texas Gulf Coast” *Geology*, 2009,37, pp. 387-390.
- [75] K. Mickus and R. J. Stern and G. R. Keller and E. Y. Anthony, “Lithospheric structure of the south-central U.S” *Geology*, 1992,20, pp. 335-338,doi:10.1130/0091-763(1992)020<0335:LSOTSC>2.3.CO;2.

- [76] M. N. Machette, “Formation of cratonic mantle keels by arc accretion: Evidence from S receiver functions” *Geophys. Res. Lett*, 2010,37,L18305,doi:10.1029/2010GL044366.
- [77] M. Moorkamp and A. G. Jones and S. Fishwick, “Joint inversion of receiver functions, surface wave dispersion, and magnetotelluric data” *J. Geophys. Res*, 2010,115,B04318,doi:10.1029/2009JB0006369.
- [78] M. Moorkamp and B. Heincke and M. Jegen and A. W. Roberts and R. W. Hobbs, “A framework for 3D joint inversion of MT, gravity and seismic refraction data” *Geophys. J. Int*, 2011,184, pp. 477-493.
- [79] S. Mosher, “Tectonic evolution of the Southern Laurentian Grenville orogenic belt” *Geological Society of America Bulletin*, 1998,10, pp. 1357-1375,doi:10.1130/0016-7606(1998)110<1357:TEOTSL>2.3.CO;2.
- [80] M. P. Moschetti and M. H. Ritzwoller and F. C. Lin and Y. Yang, “Crustal shear wave velocity Structure of the western United States inferred from ambient seismic noise and earthquake data” *J. Geophys. Res*, 2010,115,B10306,doi:10.1029/2010JB007448.
- [81] R. Moucha and A. M. Forte, D. B. Rowley and J. X. Mitrovica and N. A. Simmons and S. P. Grand, “Mantle convection and the recent evolution of the Colorado Plateau and the Rio Grande Rift valley” *Geology*, 2008,36(6), pp. 439-442.
- [82] R. Moucha and A. M. Forte, D. B. Rowley and J. X. Mitrovica and N. A. Simmons and S. P. Grand, “Deep mantle forces and the uplift of the Colorado Plateau” *Geophys. Res. Lett*, 2009,36,L19310,doi:10.1029/2009GL039778.
- [83] J. Nocedal and S. J. Wright, *Numerical Optimization, 2nd Edition*, Springer Verlag, 2006.
- [84] M. Obrebski and S. Kiselev and L. Vinnik and J.-P. Montagner, “Anisotropic stratification beneath Africa from joint inversion of SKS and P receiver functions” *J. Geophys. Res*, 2010,115,B09313,doi:10.1029/2009JB006923.

- [85] M. Obrebski and S. Kiselev and L. Vinnik and J.-P. Montagner, “Lithosphere-asthenosphere interaction beneath the western United States from the joint inversion of body-wave traveltimes and surface-wave phase velocities” *Geophys. J. Int.*, 2011,185, pp. 103-121,doi:10.1111/j.1365-246x.2011.04990.x.
- [86] ODriscoll, L. J., Humphreys, E. D., and Schmandt, B., “Time corrections to teleseismic P delays derived from SKS splitting parameters and implications for western U.S. P-wave tomography” *Geophys. Res. Lett.*, 2011,38,L19304,doi:10.1029/2011GL049031.
- [87] T. J. Owens and H. P. Crotwell and C. Groves and P. Oliver-Paul, “SOD: Standing order for data” *Seismol. Res. Lett.*, 2004,75(4), pp. 515-520,doi:10.1785/gssrl.75.4.515-a.
- [88] G. L. Pavlis and K. Sigloch and S. Burdick and M. J. Fouch and F. L. Vernon, “Unraveling the geometry of the Farallon plate: synthesis of three-dimensional imaging results from USArray” *Tectonophysics*, 2012,532, pp. 82102.
- [89] Pearce, J. and Fialko, Y., “Mechanics of active magmatic intraplate in the Rio Grande Rift near Socorro, New Mexico” *J. Geophys. Res.*, 2010,115,B07413,doi:10.1029/2009JB006592.
- [90] L. B. Pederson, “Interpretation of potential field data A generalized inverse approach” *Geophysical Prospecting*, 1977,25, pp. 199-230.
- [91] D. Plouff, “Preliminary documentation for a FORTRAN program to compute gravity terrain corrections based on a geographic grid” *U.S. Geological Survey Open-File Report*, 1977,45, pp. 77-535.
- [92] F. F. Pollitz and J. A. Snoke, “Rayleigh-wave phase-velocity maps and three-dimensional shear velocity structure of the western US from local non-plane surface wave tomography” *Geophys. J. Int.*, 2010,180, pp. 1153-1169,doi:10.1111/j.1365-246X.2009.04441.x.

- [93] N. Rawlinson and A. M. Reading and B. L. N. Kennett, "Lithospheric structure of Tasmania from a novel form of teleseismic tomography" *Journal of Geophysical Research*, 2006,111, doi: 10.1029/2005JB003803.
- [94] N. Rawlinso and B. L. N. Kennett and E. Vanacore and R. A. Glenn, "The structure of the upper mantle beneath the Delamerian and Lachlan, orogens from simultaneous inversion of multiple teleseismic datasets" *Gondwana Res*, 2011,20, doi:10.1016/j.gr.2010.11.001.
- [95] Raye U., Anthony E.Y., Stern R.J., Kimura J., Ren M.,Qing C. and Tani K., "Composition of the mantle lithosphere beneath south-central Laurentia: Evidence from peridotite xenoliths, Knippa, Texas" *Geosphere*, 2011,7, pp. 1-14.
- [96] M. Reiter and R. M. Chamberlin, "Alternative perspective of crustal and upper mantle phenomena along the Rio Grande Rift" *GSA Today*, 2011,21, pp. 4-9.
- [97] C. R. Rohs and W. R. Van Schmus, "Isotopic connections between basement rocks exposed in the St. Francois Mountains and the Arbuckle Mountains, southern mid-continent, North America" *International Journal of Earth Sciences*, 2007,96, pp. 599-611.
- [98] M. Reiter and R. M. Chamberlin, "Joint inversion of long-period magnetotelluric data and surface-wave dispersion curves for anisotropic structure: Application to data from Central Germany" *Geophys. Res. Lett*, 2011,38,L05304,doi:10.1029/2010GL046358.
- [99] Rowy, D. B., and Pindell, J. L., "End Paleozoic-Early Mesozoic western Pangean Reconstruction and its implications for the distribution of Precambrian and Paleozoic Rocks around Meso-America" *Precambrian Research*, 1989,42, pp. 411-444,doi:10.1016/0301-9268(89)90022-3.
- [100] M. Roy and J. K. MacCarthy and J. Selverstone, "Upper mantle structure beneath the eastern Colorado Plateau and Rio Grande Rift revealed by Bouguer

- gravity, seismic velocities, and xenolith data” *Geochem. Geophys. Geosyst.*, 2005,6, Q10007,doi:10.1029/2005GC001008.
- [101] M. Roy and T. H. Jordan and J. Pederson, “Colorado Plateau magmatism and uplift by warming of heterogeneous lithosphere” *Nature*, 2009,459, pp. 978-982.
- [102] C. Ruhl and S. L. Bilek and J. Stankova-Pursley, “Relocation and characterization of the August 2009 microearthquake swarm above the Socorro magma body in the central Rio Grande Rift” *Geophys. Res. Lett.*, 2010,37,L23304,doi:10.1029/2010GL045162.
- [103] E. M. Rumpfhuber and G. R. Keller and E. Sandvol and A. A. Velasco and D. C. Wilson, “Rocky Mountain evolution: Tying continental dynamics of the Rocky Mountains and Deep Probe seismic experiments with receiverfunctions” *J. Geophys. Res.*, 2009,114, B08301,doi:10.1029/2008JB005726.
- [104] Rye, C. J., I. C. Willis, N. S. Arnold, and J. Kohler, “On the need for automated Multiobjective optimization and uncertainty estimation of glacier mass balance models” *J. Geophys. Res.*, 2012,117,F02005,doi:10.1029/2011JF002184.
- [105] C. A. Rychert and P. M. Shearer, “A global view of the lithosphere-asthenosphere boundary” *Science*, 2009,324(324), pp. 495-498.
- [106] B. B. Sageman and M. H. Gardner and J. M. Armentrout and A. E. Murphy, “Stratigraphic Hierarchy of organic carbon-rich siltstones in deep-water facies, Brushy Canyon Formation (Guadalupian), Delaware Basin, West Texas” *Geology*, 1998,26, pp. 451-454,doi:10.1130/0091-7613.
- [107] M. Sambridge, “Geophysical inversion with a neighborhood algorithm I: searching a parameter space” *Geophys. J. Int.*, 1999a,138, pp. 479-494.
- [108] M. Sambridge, “Geophysical inversion with a neighborhood algorithm II: appraising the ensemble” *Geophys. J. Int.*, 1999b,138, pp. 727-746.

- [109] J. C. Savage and M. Lisowski and W. H. Prescott and A. R. Sanford, “Geodetic measurements of horizontal deformation across the Rio Grande rift near Socorro, New Mexico” *J. Geophys. Res.*, 1980,85, pp. 7215-7220.
- [110] J. C. Savage and M. Lisowski and W. H. Prescott and A. R. Sanford, “The crust under the Gulf of Mexico Basin” *The Gulf of Mexico Basin: Geological Society of America, The Geology of North America*, 1991,pp. 53-72.
- [111] J. Schmedes and J. B. Roth and R. L. Saltzer and G. M. Leahy, “Complex subduction and small-scale convection revealed by body-wave tomography of the western United States upper mantle” *Earth planet. Sci. Lett.*, 2010,297, pp. 435-445.
- [112] J. Schmedes and J. B. Roth and R. L. Saltzer and G. M. Leahy, “Imaging the shallow crust using teleseismic tomography” *Bulletin of the Seismological Society of America*, 2012,102(3), pp. 1276-122.
- [113] C. A. Schultz and S. C. Myers and J. Hipp and C. J. Young, “Nonstationary bayesian kriging: a predictive technique to generate spatial corrections for seismic detection, location, and identification” *Bull. Seism. Soc. Am.*, 1999,88, pp. 1275-1288.
- [114] Seager, W. R., and Mack, G. H., “Laramide paleotectonics in southern New Mexico. In: J. A. Peterson (Editor), Paleotectonics and Sedimentation in the Rocky Mountain Region” *Am. Assoc. Pet. Geol.*, 1986,41, pp. 669-685.
- [115] L. D. Serpa and B. deVoogd and L. A. Wright and J. Willemin and J. E. Oliver and E. C. Hauser and B. W. Troxel, “Structure of the Central Death Valley Pull-Apart Basin and Vicinity from COCORP Profiles in the Southern Great Basin” *Geol. Soc. Am. Bull.*, 1988,100, pp. 1437-1450.
- [116] P. M. Shearer, *Introduction to seismology*, Cambridge University Press, 2009.
- [117] W. Shen and M. H. Ritzwoller and V. Schulte-Pelkum, “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” *J. Geophys. Res. Solid Earth*, 2013, 118, doi:10.1029/2012JB009602.
- [118] K. Sigloch, “Mantle provinces under North America from multifrequency P wave tomography” *Geochem. Geophys. Geosyst*, 2011,12,Q02W08,doi:10.1029/2010GC003421.
- [119] Sine, C. R., Wilson, D., Gao, W., Grand, S. P., Aster, R., Ni, J., and Baldrige, W. S., “Mantle structure beneath the western edge of the Colorado Plateau” *Geophysical Research Letters*, 2008,35(10),L10303.
- [120] T. Song and D. V. Helmberger, “A depleted, destabilized continental lithosphere Near the Rio Grande rift” *Earth Planet. Sci. Lett*, 2007,262, pp. 175-184.
- [121] A. Sosa and L. Thompson and A. A. Velasco and R. Romero and R. Hermann, “3D Structure of the Southern Rio Grande Rift from 1D Constrained Joint Inversion of Receiver Functions and Surface Wave Dispersion” *Earth Planet. Sci. Lett*, 2013,402, pp. 127-137,doi:10.1016/j.epsl.2014.06.002.
- [122] W. Spence and R. S. Gross, “A tomographic glimpse of the upper mantle source of magmas of the Jemez lineament, New Mexico” *J. Geophys. Res*, 1990,95(B7), pp. 10,829-10,849,doi:10.1029/JB095iB07p10829.
- [123] L. K. Steck and M. L. Begnaud and S. Phillips and R. Stead, “Tomography of crustal P and S travel times across the western United States” *J. Geophys. Res*, 2011,116,B11304,doi:10.1029/2011JB008260.
- [124] Stein, S., and Wysession, M., *An introduction to seismology, earthquakes, and earth structure*, Wiley. com, 2009.
- [125] L. Thompson and A. A. Velasco and M. Hussein, “Geophysical Constraints on the crustal structure of the Southern Rio Grande Rift” *Bulletin of the Seismological Society of America (Submitted)*, Wiley. com, 2009.

- [126] M. Talwani and J. L. Worzel and M. Landisman, “Rapid gravity computations for two-dimensional bodies with application to the Mendocino submarine fracture zone” *J. Geophys. Res.*, 1959,64, pp. 49-59.
- [127] Y. Tian and Y. Zhou and K. Sigloch and G. Nolet and G. Laske, “Structure of North American mantle constrained by simultaneous inversion of multiple-frequency SH, SS, and Love waves” *J. Geophys. Res.*, 2011,116,B02307,doi:10.1029/2010JB007704.
- [128] A. N. Tikhonov and V. Y. Arsenin, *Solution of Ill-Posed Problems*, VH Winston & Sons, Washington, D.C, 1977.
- [129] K. Vozoff and D. L. B. Jupp, “Upper-mantle S-velocity structure of North America” *J. Geophys. Res.*, 1997,102, pp. 815-822.
- [130] W. R. Van Schmus and M. E. Bickford and A. Turek, “Proterozoic geology of the east-central Midcontinent basement, in: van der Pluijm, B.A., Catacosinos, P.A.(Eds), Basement and Basins of Eastern North America” *Geological Society of America Special Paper*, 1996,308, pp. 7-32.
- [131] V. Wijk and J. V. Hunen and S. Goes, “Small-scale convection during continental rifting: Evidence from the Rio Grande rift.” *Geology*, 2008,36, pp. 575-578.
- [132] C. R. Vogel, *Computational Methods for Inverse Problems*, SIAM, 2002.
- [133] K. Vozoff and D. L. B. Jupp, “Joint inversion of geophysical data” *Geophys. J. Roy Astr. Soc.*, 1975,42, pp. 977–991.
- [134] M. Webring, “A Fortran Gridding Program Based on Minimum Curvature” *U.S. Geological Survey Open-File Report*, 1982,41, pp. 81-1224.
- [135] M. West and J. Ni and W. S. Baldrige and D. Wilson and R. Aster and W. Gao and S. Grand, “Crust and upper mantle shear wave structure of the southwest United

- States: Implications for rifting and support for high elevation” *J. Geophys. Res.*, 2004,109,B03309,doi:10.1029/2003JB00.
- [136] M. West and W. Gao and S. Grand, “A simple approach to the joint inversion of seismic body and surface waves applied to the southwest U.S” *Geophys. Res. Lett.*, 2004,31, L15615,doi:10.1029/2004GL020373.
- [137] S. J. Whitmeyer and K. E. Karlstrom, “Tectonic model for the Proterozoic evolution of North America” *Geosphere*, 2007,3, pp. 220-259.
- [138] C. K. Wilson and J. Shragge and B. Artman, “Teleseismic imaging with wave equation Migration” *Stanford Exploration Project*, <http://sepwww.stanford.edu/sep/prof/>, *Report*, 2004,115, pp. 399-411.
- [139] D. Wilson, “Imagining crust and upper mantle seismic structure in the southwestern United States using teleseismic receiver function” *Leading Edge* , 2003,22, pp. 232-237.
- [140] Wilson, D., and R. Aster, “Seismic imaging of the crust and upper mantle using Regularized joint receiver functions, frequency-wave number filtering, and Multimode Kirchhoff migration” *J. Geophys. Res.* , 2005,B05305,doi:10.1029/2004JB003430.
- [141] D. Wilson and R. Aster and J. Ni and S. Grand and M. West and W. Gao, W. S. Baldrige and S. Semken, “Imaging the structure of the crust and upper mantle beneath the Great Plains, Rio Grande Rift, and Colorado Plateau using receiver functions” *J. Geophys. Res.*, 2005,110,B05306,doi:10.1029/2004JB003492.
- [142] D. C. Wilson and R. Aster and S. Grand and J. Ni and W. S. Baldrige, “High-resolution receiver Function imaging reveals Colorado Plateau lithospheric architecture and mantle-Supported topography” *Geophys. Res. Lett.*, 2010,37,L20313,doi:10.1029/2010GL044799.
- [143] D. Wilson, “Lithospheric structure of the Rio Grande Rift” *Nature*, 2005,433, pp. 851-854.

- [144] L. A. Woodward, "Rate of crustal extension across rio-grande rift near Albuquerque, New-Mexico" *Geology*, 1977,5, pp. 269-272.
- [145] Xue, M., and Allen, R. M., "Mantle structure beneath the western United States and its implications for convection processes" *J. Geophys. Res*, 2010,115,B07303,doi:10.1029/2008JB006079.
- [146] P. Xu and D. Zhao, "Upper-mantle velocity structure beneath the North China craton: Implications for lithospheric thinning" *Geophys. J. Int*, 2009,177, pp. 1279-1283,doi:10.1111/j.1365-246X.2009.04120x.
- [147] Y. Yang and M. H. Ritzwoller and F. C. Lin and M. P. Moschetti and N. M. Shapiro, " Structure of the crust and uppermost mantle beneath the western United States revealed by ambient noise and earthquake tomography" *J. Geophys. Res*, 2008,113,B12310,doi:10.1029/2008JB005833.
- [148] H. Yuan and B. Romanowicz, "Lithospheric layering in the North American Craton" *Nature*, 2010,466, pp. 10631068.
- [149] H. Yuan and B. Romanowicz and K. M. Fischer and D. Abt, "3-D shear wave radially and azimuthally anisotropic velocity model of the North American upper mantle" *Geophys. J. Int*, 2011,184, pp. 12371260.
- [150] M. S. Zhdanov, *Geophysical inverse theory and regularization problems*, v36, Elsevier, 2002.
- [151] L. Zhu and H. Kanamori, "Moho depth variation in southern California from tele-seismic receiver functions" *J. Geophys. Res*, 2000,105, pp. 2969-2980.

Curriculum Vitae

Lennox Emmanuel Thompson was born on June 06, 1985. The first son of the Johnson family who graduated from Carver Voc Tech High School, Baltimore, MD in spring 2003. He entered Coppin State University in the fall of 2003, and, started his computer science undergraduate program. After graduating in 2008, he entered The University of Texas at El Paso Masters program in Geophysics. After two years of classes and research, graduated in 2010 to started the Ph.D. geophysics program in 2011 at University the Texas at El Paso. During the Ph.D. program, he was a teaching assistant, research assistant, and GK-12 fellow. He published a series of papers using several geophysical techniques to generate 3-D visualization of the Earth structure within the New Mexico & Texas region. Attended several workshops, tutorial, & seismic training courses. Presented oral & poster presentations at professional meetings around the world and collaborated with scientists with the U.S. and South America region.

Permanent address: 200 N. Mesa Hills Drive Apt 1508

El Paso, Texas 79912-4927