



uOttawa

Passive seismic imaging of lithosphere
structure at active tectonic margins in
Canada and New Zealand

by

Jeremy Gosselin

B.Sc., University of Victoria, 2014

M.Sc., University of Victoria, 2017

A thesis submitted to the University of Ottawa
in partial fulfillment of the thesis requirement for the degree of

Doctor of Philosophy

in

Earth and Environmental Sciences

Faculty of Science

© Jeremy Gosselin, Ottawa, Canada, 2021

Examining Committee

The following served on the Examining Committee for this thesis.

External Examiner: Sergei Lebedev
 Professor of Geophysics, Dept. of Earth Sciences,
 University of Cambridge
 Adjunct Professor, Dublin Institute for Advanced Studies

Internal Members: Glenn Milne
 Professor, Dept. of Earth and Environmental Sciences,
 University of Ottawa

 Claire Perry
 Seismologist, Dept. of Natural Resources Canada
 Adjunct Professor, Dept. of Earth and Environmental Sciences,
 University of Ottawa

 Dariush Motazedian
 Professor, Dept. of Earth Sciences, Carleton University

Supervisor: Pascal Audet
 Professor, Dept. of Earth and Environmental Sciences,
 University of Ottawa

Declaration of Authorship

I hereby certify that the submitted thesis represents results from my own ideas and investigations. This work was completed while I was registered as a candidate for the degree of Doctor of Philosophy at the University of Ottawa, under the supervision of Pascal Audet. Furthermore, I hereby certify that I have not obtained a degree elsewhere on the basis of the research presented in this submitted work.

Jeremy Gosselin

September, 2021

Contemporary doctoral research in Earth sciences rarely occurs alone. This is especially true for a thesis constructed from a set of disparate journal articles, such as this submitted work. Five chapters in this thesis represent journal articles with varying co-author contributions. I discuss my own specific contributions below.

· **Chapter 1** includes excerpts from

A review of inverse methods in seismic site characterization

by

Jeremy M. Gosselin, Stan E. Dosso, Aysegul Askan, Marc Wathelet, Alexandros Savvaidis, and John F. Cassidy.

The Consortium of Organizations for Strong Motion Observation Systems (COSMOS) consisting of the U.S. Geological Survey, the Geological Survey of Canada, and a group of North American power utility companies (Southern California Edison and Pacific Gas and Electric) invited me to write a review of methods for geophysical inversion applied to small-scale geotechnical assessment of seismic hazard. Although this paper does not include original research, it did represent a significant time commitment, effort, and accomplishment during my doctoral program. Furthermore,

much of the theory and literature that I reviewed is relevant to chapter 2 and chapter 3 of this thesis. Hence, I adopt excerpts as a useful introductory chapter to part I of this thesis. As of November 2021, this paper is in press for publication in a special issue of the *Journal of Seismology*.

- **Chapter 2** is a journal article titled

Azimuthal anisotropy in Bayesian surface wave tomography: application to northern Cascadia and Haida Gwaii, British Columbia

by

Jeremy M. Gosselin, Pascal Audet, Andrew J. Schaeffer, Fiona A. Darbyshire, and Clément Estève

I conceptualized this study, as well as developed and tested software coded in the FORTRAN and PYTHON programming languages. I downloaded and processed surface-wave dispersion data. I applied the novel tomography method to synthetic experiments and real data, as well as wrote the article with input from co-authors. This paper is published in *Geophysical Journal International*.

- **Chapter 3** is a journal article titled

Circular statistics in Bayesian inversion for two-station phase-velocity dispersion estimation

by

Jeremy M. Gosselin, Pascal Audet, Andrew J. Schaeffer, and Clément Estève

I conceptualized this study, as well as developed and tested software coded in the FORTRAN and PYTHON programming languages. I applied the novel inversion method to real data, which I downloaded and processed. I also wrote the article with input from co-authors. As of November 2021, this paper is in preparation for submission to *Geophysical Journal International*.

- **Chapter 5** is a journal article titled

Seismic evidence for megathrust fault-valve behavior during episodic tremor and slip

by

Jeremy M. Gosselin, Pascal Audet, Clément Estève, Morgan McLellan, Stephen G. Mosher, and Andrew J. Schaeffer.

This paper represents a collaboration of the seismology group at the University of Ottawa (at the time this work was done). I conceived the hypothesis and scope of the project, which I discussed in my doctoral thesis proposal. Specifically, I postulated a change in seismic velocity structure in the vicinity of the plate interface in subduction zones during episodic tremor and slow slip (ETS) as a consequence of fluid pressure fluctuations. I worked with my thesis advisor (Pascal Audet) and proposed stacking receiver function data relative to the nearest ETS event (critical for this study). I interpreted the results and wrote the paper with Pascal Audet, as well as with input from other co-authors. This paper is published in *Science Advances*.

- **Chapter 6** is a journal article titled

Seismic constraint on heterogeneous deformation and stress state in the forearc of the Hikurangi subduction zone, New Zealand

by

Jeremy M. Gosselin, Pascal Audet, Bill Fry, and Emily Warren-Smith

This paper describes research conducted while I was a visiting student at the Geological Survey of New Zealand (GNS Science) for several months during my doctoral program. I conceived this study. I downloaded teleseismic data and processed receiver functions. I also interpreted the results and wrote the paper with input from co-authors. This paper is published in *The Seismic Record*.

Abstract

Our knowledge of Earth's structure and dynamics is owed, in large part, to techniques that allow us to infer the physical properties of rocks based on observations made near the surface. In particular, passive seismic imaging relies on natural sources of elastic wave energy (typically earthquakes) to illuminate Earth's interior. The frequency-dependent dispersive characteristics of surface waves from earthquakes provides a valuable constraint on the depth heterogeneity beneath the surface, which can be used to infer structure of the lithosphere (i.e., tectonics). The first part of this thesis develops novel tools for passive seismic imaging considering surface-wave dispersion. Specifically, Bayesian (probabilistic) methods are developed that provide rigorous uncertainty quantification. The ability to estimate the directional dependence of surface-wave speeds (i.e., seismic anisotropy) is demonstrated. Furthermore, a general approach for considering circular (wrapped) data, such as surface-wave phase measurements, is developed and applied to estimate the average dispersion between pairs of seismic stations. These ideas are applied to data recorded at seismic stations over British Columbia, Canada, to produce a large volume of data products that will help improve our understanding of the tectonics throughout the region. The second part of this thesis investigates the structural and mechanical conditions in subduction zones, where tectonic plates collide and one plate is thrust beneath the other. Specifically, a type of passive seismic imaging based on recordings of body waves from distant earthquakes (known as receiver functions) is used to infer subduction zone structure in relation to the coupling between tectonic plates. Receiver function data calculated for stations over the Cascadia subduction zone (southwestern Canada) suggest that episodes of slow slip on the plate interface occur in tandem with changes in mechanical conditions. Receiver function data calculated for stations over the Hikurangi subduction zone (New Zealand) suggest that stress and deformation along the margin are spatially linked to plate coupling. These results improve our understanding of the dynamics of these tectonic systems, including the seismic hazards that they pose.

Acknowledgements

I would not have been able to complete this thesis without the assistance and encouragement from many people.

I wish to thank my supervisor, Pascal Audet, for persuading me to come to Ottawa in the first place. You have supported me while also providing me autonomy and freedom to pursue ideas (some of which are good, some of which are not so good). We have shared some incredible experiences over these four years (i.e., Svalbard, New Zealand). You motivate me to broaden my mind and approach big problems.

I also wish to thank my unofficial mentor, Andrew Schaeffer, as well as my office mates Clément Estève and Steve Mosher. We have had great times together at UOttawa, and travelling to conferences (I will never forget playing rocket league in our “gritty” airBNB in Washington, DC). I am grateful for the part you all played during my time in Ottawa, and for our lasting friendships. Thank you also to the other members of our geophysics group: Phil, Élyse, Morgan, Taylor, and Hao.

Thank you to the many people who have collaborated on research with me throughout my doctoral program. This includes the people listed above as well as Fiona Darbyshire. In particular, I would like to thank Bill Fry and Emily Warren-Smith who welcomed me to New Zealand for several months.

Thank you to Stan Dosso, John Cassidy, and Jan Dettmer for ongoing advice and wisdom.

Thank you to my family, and to my friends in Ottawa: Jeff, Hester, Petra, Gwendolyn, Brett, and Nicole. You helped to keep my life balanced.

Last, but certainly not least, I thank my wife, Iris, for loving and supporting me. You followed me to Ottawa, away from friends and family. For that am I forever grateful. You are a constant source of motivation in all aspects of life, and I could not have finished this thesis without you.

To my wife Iris.

Table of Contents

List of Tables	xiii
List of Figures	xiv
Thesis introduction	1
I Bayesian tools for surface-wave imaging	8
1 A brief review of geophysical inverse theory	9
1.1 Theoretical overview	10
1.1.1 Models and data	10
1.1.2 The forward problem	11
1.1.3 Errors and misfit	13
1.1.4 Parameterization	17
1.1.5 Other considerations	18
1.2 Linearized inversion	20
1.3 Nonlinear optimization	24
1.3.1 Downhill simplex	25

1.3.2	Simulated annealing	26
1.3.3	Genetic algorithms	27
1.4	Bayesian Inversion	28
1.4.1	McMC sampling	29
1.4.2	Model selection	31
1.4.3	Error model and likelihood function	33
1.5	Tomography	34
2	Azimuthal anisotropy in Bayesian surface-wave tomography: application to northern Cascadia and Haida Gwaii, British Columbia	39
2.1	Introduction	39
2.2	Tectonic Setting	43
2.3	Surface-wave dispersion data	45
2.4	Inversion methodology	50
2.4.1	Azimuthal anisotropy	50
2.4.2	Model parameterization	51
2.4.3	Bayesian inversion	54
2.4.4	Practical considerations	57
2.5	Synthetic inversion tests	60
2.6	Anisotropic group velocity over northern Cascadia and Haida Gwaii	64
2.7	Discussion and Conclusions	72
3	Circular statistics in Bayesian inversion for two-station phase-velocity dispersion estimation	75
3.1	Introduction	75

3.2	Methodology	77
3.2.1	Bayesian inversion	77
3.2.2	Error model for circular data	79
3.2.3	Teleseismic two-station phase-velocity dispersion	81
3.3	Surface-wave dispersion over British Columbia, Canada	82
3.3.1	Surface-wave phase spectra	83
3.3.2	Path-averaged Earth structure and surface-wave dispersion	90
3.4	Conclusions	95

II Spatial and temporal variability of subduction zone structure from receiver function analysis 96

4 A brief review of receiver functions 97

4.1	Teleseismic recordings	97
4.2	Deconvolution	101
4.3	Interpretation	103
4.4	Receiver functions at subduction zones	108

5 Seismic evidence for megathrust fault-valve behavior during episodic tremor and slip 113

5.1	Introduction	113
5.2	Results	116
5.3	Discussion	118
5.4	Materials and methods	121
5.4.1	Receiver function analysis	121

5.4.2	Harmonic decomposition	122
5.4.3	Interface velocity contrast	123
5.4.4	ETS times from GPS data	124
5.4.5	Temporal stacking	125
5.4.6	Bootstrap analysis	126
6	Seismic constraint on heterogeneous deformation and stress state in the forearc of the Hikurangi subduction zone, New Zealand	127
6.1	Introduction	127
6.2	Data and Methods	129
6.3	Results	134
6.4	Conclusions	139
	Thesis conclusions	140
	Final thoughts	145
	References	146
	APPENDICES	175
A	Supplementary materials for: Azimuthal anisotropy in Bayesian surface-wave tomography	175
B	Supplementary materials for: Circular statistics in Bayesian inversion for two-station phase-velocity dispersion estimation	186

C	Supplementary materials for: Seismic evidence for megathrust fault-valve behavior during episodic tremor and slip	190
D	Supplementary materials for: Seismic constraint on heterogeneous deformation and stress state in the forearc of the Hikurangi subduction zone, New Zealand	201

List of Tables

3.1	Prior bounds for model parameters in Bayesian inversion.	79
4.1	1D Earth model with a dipping layer.	105
4.2	1D Earth model with an anisotropic layer.	105
A.1	Prior bounds for model parameters in Bayesian tomography.	183
A.2	Simple crustal model.	184
C.1	Strike and dip angles of the plate interface beneath each station location. .	198
C.2	Estimated start, end and mid-point dates of ETS events from 1994-2019. .	199
C.3	Statistical test p -values.	200
D.1	Results of harmonic decomposition of receiver function data.	206

List of Figures

i	Illustration of the megathrust earthquake cycle.	5
1.1	Forward vs. inverse problems.	12
1.2	Hypothetical misfit surfaces in a 2D model space.	16
1.3	Hypothetical tradeoff between data misfit and regularization.	22
1.4	Illustration of Bayesian inference.	30
1.5	Example of surface-wave ray path deflections.	36
2.1	Tectonic setting of coastal British Columbia.	44
2.2	Depth sensitivity of fundamental-mode Rayleigh wave group velocity.	47
2.3	Surface wave path coverage from regional earthquakes.	48
2.4	Examples of Voronoi and Delaunay interpolations.	53
2.5	Synthetic Bayesian surface-wave tomography tests.	62
2.6	Inversion hyper-parameters from synthetic tests.	63
2.7	Isotropic structure of British Columbia from Bayesian tomography.	65
2.8	Azimuthal anisotropy over British Columbia from Bayesian tomography.	67
2.9	Isotropic velocity uncertainty from Bayesian tomography.	69
2.10	Anisotropy uncertainty from Bayesian tomography.	70

2.11	Inversion hyper-parameters from Bayesian tomography.	71
3.1	Station coverage over British Columbia, Canada.	84
3.2	Earthquake locations for two-station phase spectra calculation.	86
3.3	Earthquake recordings at stations CRAG and R33M.	87
3.4	Surface-wave phase spectra between stations CRAG and R33M.	88
3.5	Bayesian inversion results for station pair CRAG and R33M.	91
3.6	Bayesian inversion results for all station pairs over British Columbia. . . .	93
4.1	Earthquake phase travel times.	98
4.2	Receiver function illustration.	100
4.3	Receiver functions for Earth model with a dipping layer.	106
4.4	Receiver functions for Earth model with an anisotropic layer.	107
4.5	Receiver functions for structure with a dipping low-velocity layer.	110
5.1	The study region in the forearc of the Cascadia subduction zone.	115
5.2	ETS-related changes in velocity structure beneath station PGC.	117
5.3	Conceptual model of ETS-related fluctuations in P_f in Cascadia.	120
6.1	Tectonic setting and seismic station distribution over the Hikurangi subduc- tion zone.	128
6.2	Receiver function data at station MRHZ.	131
6.3	Receiver function harmonic decomposition orientations and magnitudes. . .	135
6.4	Along-strike heterogeneity in forearc deformation in the Hikurangi subduc- tion zone.	136

A.1	Example of surface wave dispersion processing from a regional earthquake recording.	176
A.2	Surface wave path coverage.	177
A.3	Surface wave azimuthal path coverage.	178
A.4	Synthetic Bayesian tomography tests under varying noise conditions. . . .	179
A.5	Inversion hyper-parameters from Bayesian tomography under varying noise conditions.	180
A.6	Uncertainty of azimuthal anisotropy magnitude.	181
A.7	Inversion hyper-parameters after first iteration.	182
B.1	Depth sensitivity of fundamental-mode Rayleigh wave phase velocity. . . .	187
B.2	Histograms of all estimated dispersion curves.	188
B.3	Histograms of all estimated dispersion curve uncertainties.	189
C.1	S_V receiver functions for station PGC.	191
C.2	Harmonic components for station PGC.	192
C.3	Cross-correlation coefficients between observed and predicted receiver functions.	193
C.4	Example calculation of strike and dip angles at each station location. . . .	194
C.5	Estimated S-wave velocity contrast across the top and bottom of the LVL. . .	195
C.6	GPS data used in the determination of ETS event times.	196
C.7	Changes in δ_{V_S} across the top and bottom boundaries of the LVL.	197
D.1	Predicted P-S delay times.	202
D.2	Harmonic components of receiver function data for all stations.	203
D.3	Comparison of P and PP receiver functions.	205

Thesis introduction

Earth's outermost solid layer is partitioned into a set of rigid lithospheric (tectonic) plates that move and interact, manifesting the surficial expression of mantle convection due to Earth cooling (Davies, 2001). The relative motion of tectonic plates leads to convergent, divergent, and transform plate boundaries. Active margins characterize the boundary between tectonic plates along the edges of continents, and are typically convergent. Furthermore, active margins often involve a variety of dynamic processes including orogenesis (mountain building), volcanism, and the subduction of oceanic lithosphere as it is recycled into the mantle. Subduction zones play an important role in the study of Earth geodynamics by providing a window into the processes of mantle convection and the chemistry of Earth's interior. These regions are of economic significance for hosting processes that form precious ore-bodies in the crust. Lastly, active margins and subduction zones are of societal concern as they can host a variety of natural hazards that pose a significant threat to human life and property. These include volcanic eruptions as well as large earthquakes and associated tsunamis.

Geophysical imaging plays an important role in Earth sciences across a spectrum of spatial scales including uncovering the structure of Earth's core, exploring for natural resources, and evaluating the soil conditions of a site for infrastructure construction. Our knowledge of the structure and dynamics of active tectonic margins is owed, in large part, to minimally invasive techniques that estimate the physical properties of rocks (and rock processes) based on observations made near Earth's surface. In particular, seismic imaging techniques rely on measurements of elastic waves that propagate through (and interact

with) rocks in the subsurface to gain insight on the distribution of elastic (and anelastic) rock properties. The source of these elastic waves may be man-made or naturally occurring, leading to active and passive methods, respectively.

Seismology encompasses a broad spectrum of research. My own scientific interests are similarly broad. My doctoral research spans a variety of problems both methodologically and geographically. The focus of this thesis is on the development of passive seismic imaging methods, as well as the application of existing passive seismic imaging methods to address important questions related to tectonic structure and dynamics at active margins in Canada and New Zealand. This thesis consists of a collection of disparate journal articles (presented as chapters here) that are broadly separated in two parts.

· Part I :

There is a fundamental natural causality between seismic wave phenomena and the mechanical properties of rocks in the Earth. For example, the speed at which seismic waves travel through rocks is a function of the elastic moduli of the rocks, derived from a general Hooke's law (Shearer, 2019). The *cause* is the elastic moduli of the rocks, the *effect* is the amount of time required for seismic waves to travel through the rocks (which seismologists can measure). Natural processes are often referred to as forward problems, and are mathematically approximated based on physical principles (i.e., starting with the cause and calculating the effect). Studies on tectonic structure and dynamics must consider Earth models that are 10s to 100s of km deep, where direct observation and measurement of rock properties is generally not possible. Instead, seismologists rely on observations of seismic wave phenomena recorded near the surface. Consequently, the goal of seismic imaging is to estimate the cause (rock mechanical properties) from the effect (recordings of seismic waves). This is referred to as an inverse problem. Inverse problems are important in many scientific disciplines since they allow for inference on quantities where direct observation is not possible.

A common measurement used in seismic imaging is based on recordings of seismic waves that propagate along Earth's surface. Surface waves with different frequen-

cies are sensitive to rock elastic properties over different depth ranges, with lower-frequency (longer-period) signals generally sensitive to deeper structure (Shearer, 2019). Because Earth is vertically (radially) heterogeneous, surface waves of varying frequency will *disperse* as they propagate. Consequently, the measurement of surface-wave dispersion can be used to infer the spatial distribution of rock elastic properties beneath Earth’s surface.

Uncertainty is pervasive throughout seismology, including seismic imaging and inverse problems, though it has historically often been ignored. There is uncertainty at every step in passive seismic imaging. There is uncertainty due to seismic instrumentation, and in measurements due to competing seismic signals. There is uncertainty due to the various approximations seismologists use to model Earth structure and the physical processes that produce seismic wave phenomena. Understanding and quantifying the uncertainties in passive seismic imaging is important for robust interpretation, and is an active area of research. Uncertainty estimation in passive seismic imaging is challenging, and many methods are approximate and/or qualitative.

The focus of part I of this thesis is on inverse problems. Specifically, Bayesian (probabilistic) approaches that provide rigorous uncertainty quantification are developed and applied to passive seismic imaging using surface-wave measurements.

- **Chapter 1** of this thesis provides a brief review of common inverse methods applied in seismology (and geophysics in general) in order to provide context for the subsequent thesis chapters.
- **Chapter 2** of this thesis develops a Bayesian approach to estimate anisotropic maps of surface-wave group velocities, with associated uncertainties. The method is applied to surface-wave dispersion measured from regional earthquakes recorded at seismic stations throughout British Columbia, Canada.
- **Chapter 3** of this thesis develops a Bayesian approach to estimate the average seismic velocity structure between two stations in order to quantify the average surface-wave dispersion, with associated uncertainties. The method is applied

to teleseismic surface-wave measurements from distant earthquakes recorded at stations throughout British Columbia, Canada.

· Part II :

Subduction zones are regions where two tectonic plates converge and interact along a large fault known as a megathrust. The megathrust can remain locked for decades or centuries while accumulating elastic strain, which is ultimately released during an earthquake. Ruptures of large portions of a megathrust generate giant and damaging earthquakes and associated tsunamis like the 2011 magnitude 9.0 Tohoku earthquake in Japan (Mori et al., 2011). In a simplified model of the megathrust earthquake cycle (fig. i), the convergence of the two tectonic plates (during the inter-event time period) causes shortening in the overriding continental plate. The seaward edge of the continental plate is forced downward along the margin, and there is an upward deformation of the continent further inland (i.e., a flexural bulge). During a megathrust earthquake, this deformation reverses as the inland part of the continent collapses downward and the seaward edge springs upward, which typically causes a tsunami as the water column is displaced. This pattern of deformation is observable on sensitive Global Navigation Satellite Systems (GNSS) instruments, and provides a useful constraint on the inferred size of the locked portion of the megathrust in subduction zones (e.g., Mazzotti et al., 2003; Wang et al., 2003).

The locked portion of the megathrust can also be considered within a thermal context. The shallowest (up-dip) and deepest (down-dip) extents of the locked zone are often defined by the behaviour of rocks at different temperatures (e.g. Hyndman & Wang, 1993). The shallowest section of the plate interface is typically beneath an accretionary prism of sediments, as opposed to the continental crust further inland. Up to a depth of a few km, sediments are unconsolidated and likely incapable of accumulating any significant elastic strain. At greater depths, the lithified sedimentary rocks undergo alterations as they are subjected to higher temperatures and pressures (Underwood, 2007). These rocks have sufficient strength to exhibit seismic

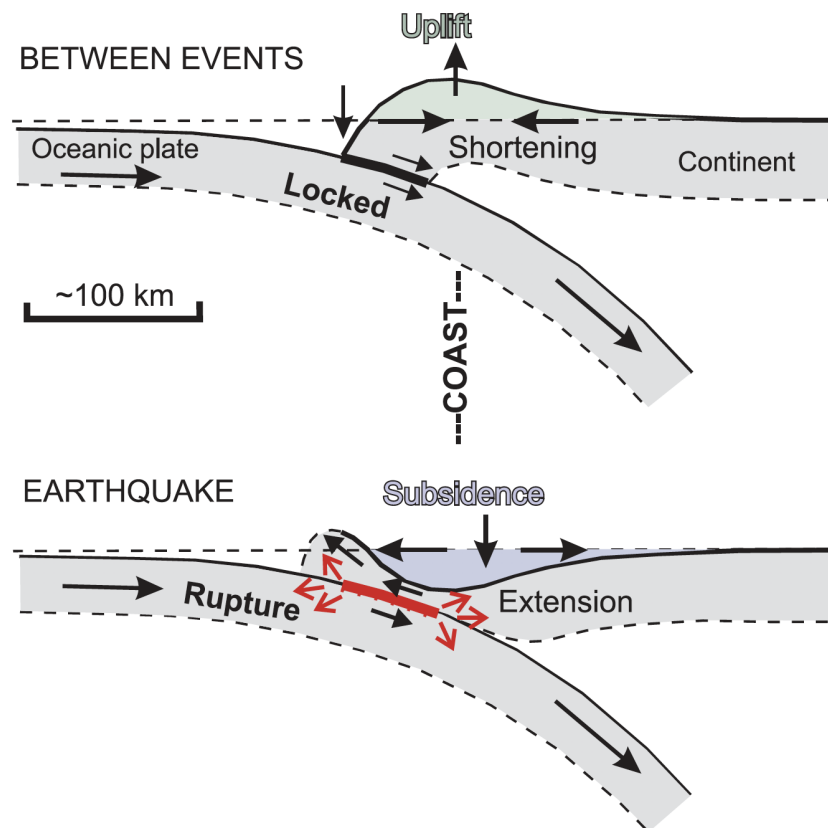


Figure i: Illustration of the megathrust earthquake cycle (adapted from Hyndman & Rogers, 2010).

behaviour once they reach 100-150°C, which defines the up-dip extent of the locked megathrust zone (Hyndman & Wang, 1993; Harris & Wang, 2002). This boundary has implications for tsunamigenesis, as it impacts the extent of seafloor vertical displacement during a megathrust rupture (e.g., Gao et al., 2018). Temperatures continue to increase with depth along the plate interface. Eventually, temperatures are high enough that the rocks on the fault no longer behave elastically. The brittle-ductile transition temperature of quartz (the most ductile major mineral in granite) is $\sim 350^\circ\text{C}$ (Scholz, 1988, 2019), and is often used to define the down-dip limit of the locked zone (e.g., Hyndman & Wang, 1993). Although, in cooler subduction zones, the down-dip limit may be related to aseismic forearc mantle serpentinite (e.g., Hyndman & Peacock, 2003). Mechanically, the locked zone refers to the frictionally unstable region within which a megathrust rupture may nucleate (Scholz, 1998). The brittle-ductile transition temperature of feldspars (the most ductile major mineral in basalt) is $\sim 450^\circ\text{C}$ (Scholz, 1988, 2019). At higher temperatures, rocks are dominated by ductile behaviour. Consequently, the temperature range between 350-450°C defines a transition region that is frictionally conditionally-stable. Earthquakes are not expected to nucleate in this region, though a megathrust rupture originating in the locked zone may potentially propagate through this region (Hyndman & Wang, 1993).

Periodic slow fault motion that can last for days to months (or even years) has been observed at many subduction zones around the world, and is one of the most important discoveries in seismology in recent years (Dragert et al., 2001). See recent reviews by Bürgmann (2018) and Behr & Bürgmann (2021) for useful discussions on the current state of knowledge on this phenomenon. These slow slip events do not produce traditional seismic signals (Beroza & Ide, 2011), yet in some cases they produce fault motion that would be equivalent to a magnitude 7.2 earthquake without damaging ground shaking (Bürgmann, 2018). Some speculate that even larger slow slip events are possible (Meade & Loveless, 2009). Slow slip is effectively imperceptible to people, and can only be detected using sensitive GNSS instruments or strainmeters (e.g., Wang et al., 2008). These events occur in the transition region

between the sections of megathrust faults that are locked and sections that are continuously freely slipping (i.e., creeping), which is generally deeper than the locked zone (Lay et al., 2012). Observations of slow slip often correlate spatially and temporally with observations of a noise-like seismic signal called tremor (Rogers & Dragert, 2003). The combined observation is referred to as episodic tremor and slip (ETS). It has been suggested that slow slip could play a potential role in triggering large earthquakes by transferring stresses along the megathrust (Obara & Kato, 2016).

Knowledge of the structural and mechanical conditions of currently locked, creeping, and ETS segments of megathrusts is important for understanding subduction zone dynamics and the seismic hazard posed by these tectonic features. The focus of part II of this thesis is on temporal and spatial variability of subduction zone structure, and its relation to deformation, plate interface coupling, and ETS. Specifically, properties of the Cascadia subduction zone (Canada) and the Hikurangi subduction zone (New Zealand) are inferred from a type of passive seismic data known as receiver functions (RFs), which are based on recordings of body waves from distant earthquakes.

- **Chapter 4** of this thesis provides a brief review of RF estimation and interpretation in order to provide context for the subsequent thesis chapters.
- **Chapter 5** of this thesis presents RF data that suggest temporal changes in mechanical properties near the plate interface in subduction zones occur in tandem with ETS.
- **Chapter 6** of this thesis presents RF data that suggest deformation and stress state vary along the margin of the Hikurangi subduction zone, and correlate with estimates of plate coupling along the megathrust.

Part I

Bayesian tools for surface-wave imaging

Chapter 1

A brief review of geophysical inverse theory

The general topic of geophysical inversion is considered in several texts (e.g., Parker, 1994; Tarantola, 2005; Menke, 2018; Aster et al., 2018) and reviews in the literature (e.g., Jackson, 1972; Parker, 1977; Treitel & Lines, 2001; Sambridge & Mosegaard, 2002). This chapter presents a review of common inversion approaches for estimating geophysical models from seismic data. The chapter is structured roughly in order of increasing complexity of the approaches reviewed, concluding with probabilistic (Bayesian) inversion and tomography, which are considered further in chapters 2 and 3. Many of the same inversion approaches reviewed here have been applied to estimate Earth structure from local to global scales. Furthermore, these approaches have been applied to many other types of geophysical data (e.g., gravity, magnetic, electromagnetic). It is not intended, nor is it possible, for this thesis chapter to provide an exhaustive account of the geophysical inversion literature. The intent of this chapter is to provide a reference upon which subsequent thesis chapters will build.

Here I review the theory, assumptions, and limitations of common inversion methods applied in seismology (and geophysics in general). Many factors in the inverse problem can influence the recovered model, and ultimately impact our interpretation and understanding

of subsurface structure. The range of issues considered here is broad and varies with respect to seismic data types, algorithmic complexity, computational expense, physical dimension, and the ability to quantitatively estimate the uncertainty in the inverse solution. Given the range of methods discussed here, this chapter first provides a brief overview of some general aspects of inversion (applicable to all problems) in the following section before discussing specific inversion approaches in more detail. Although chapters 2 and 3 focus on Bayesian inversion, this chapter reviews a progression of inverse methods of increasing complexity for context. Most approaches reviewed here are based on recovering an optimal (best-fit) model, typically of the subsurface. Linearized and fully nonlinear approaches to finding optimal models are reviewed in sections 1.2 and 1.3, respectively. As an alternative to methods that find a single optimal solution, Bayesian inversion, which provides a probabilistic result, is discussed in Section 1.4. Lastly, seismic tomography, which is the subject of Chapter 2 (and the future goal of the work presented in chapter 3), is discussed in classical terms in section 1.5.

1.1 Theoretical overview

1.1.1 Models and data

Inversion can be defined as the estimation of the parameters of a postulated model that represents a physical system of interest, using observations of some process that interacts with the system (i.e., data). The model is said to be postulated because it is based on our assumptions of the physical processes, and the complexity of structure, beneath Earth's surface. Here we make the distinction between the model (which is proposed) and the model parameters (which are estimated during inversion). In the case of seismology, the physical system is the seismic velocity (or attenuation) structure of the subsurface, and the data are observations of seismic waves that interact with this structure. It is important to recognize that the model always represents an idealization of the actual physical system. For example, a model consisting of a 1D profile of seismic velocity assumes the subsurface is laterally homogeneous (and typically isotropic). An important issue in all inverse

problems is whether the postulated model adequately represents the physical system. For example, assuming lateral homogeneity when the subsurface is actually laterally heterogeneous introduces modelling errors that can bias or preclude meaningful results. Such cases may necessitate more complex 2D or 3D models of the system. The statistician George E. P. Box summarized this best when he stated that “*all models are wrong, but some are useful*”; a principle that applies to all scientific models, not just in inverse problems.

In the case of a 1D profile, model parameters commonly represent the shear-wave velocity (V_S) and compressional-wave velocity (V_P) within discrete layers. Other parameters such as density (ρ) and attenuation are also considered. In 2D or 3D models, parameters typically represent seismic velocities of discrete cells in a spatial grid. The number and sizes of layers or cells in the model are also issues in model selection. Depending on the particular approach, these spatial properties may also be considered as model parameters estimated in the inversion (e.g., the number and thicknesses of layers in a 1D scenario).

As mentioned above, the data considered to constrain subsurface structure represent observations of seismic wave phenomenology. Common seismic data include measurements of the dispersive (variation of phase or group velocity with frequency) nature of surface waves. Although both chapters 2 and 3 consider surface wave dispersion, neither of the inversion implementations in these studies formally consider dispersion as data. In chapter 2, the data represent the travel times of surface waves (at specific periods) from regional earthquakes. In chapter 3, the data are the phase spectrum of the cross-correlation of surface-wave recordings from station pairs. This is analogous to a phase-wrapped dispersion curve of unknown initial phase (as will be discussed in chapter 3). Dispersion measurements can be extracted from ambient seismic noise, earthquakes (regional or teleseismic), or even controlled sources (in near-surface, geotechnical applications). Specifics of dispersion processing are discussed in chapters 2 and 3.

1.1.2 The forward problem

In order to consider an inverse problem, a solution must be available for the corresponding forward problem (also called the forward model or direct problem). For a given set

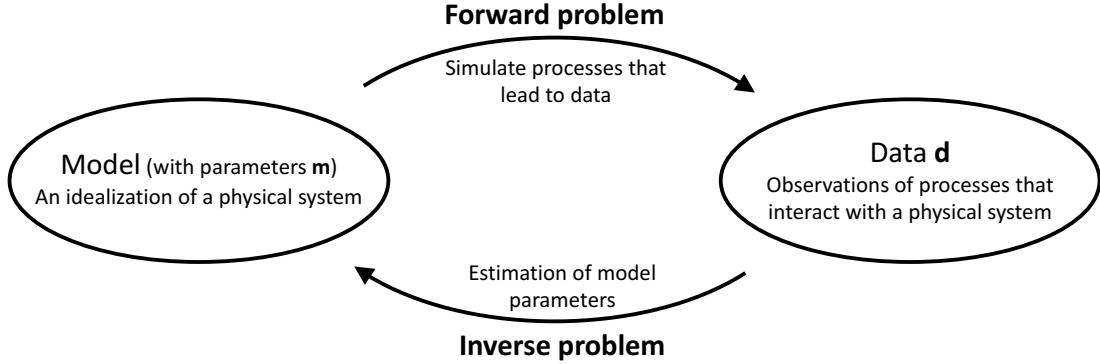


Figure 1.1: Forward vs. inverse problems.

of model parameters, the forward problem computes (predicts) the data that would be observed for this representation of the physical system (i.e., the forward problem simulates the physical processes that lead to the data). Mathematically, the forward problem represents a mapping from the model-parameter space to the data space, while the inverse problem represents the reverse mapping (fig. 1.1). Geophysical forward problems (predicting data for models) are typically deterministic and stable in the sense that small changes to the model cause only small changes to the resulting (predicted) data. In contrast, the inverse problem (estimating model parameters from data) is typically non-unique (more than one model can fit the data) and can be unstable (i.e., small changes to the data, such as errors, can cause large changes in the recovered model). Hence, solutions to inverse problems should always be appraised critically, and quantitative uncertainty estimation, when possible, is valuable.

To briefly consider the forward problem for the data types mentioned above, predicting

surface-wave dispersion for laterally homogeneous models usually employs efficient numerical methods such as the Thomson-Haskell propagator-matrix (Thomson, 1950; Haskell, 1953; Knopoff, 1964; Dunkin, 1965; Gilbert & Backus, 1966). Most 1D forward models are based on discrete layers with uniform geophysical properties within each layer. Model parameterizations that involve gradients (e.g., a power-law profile) can be parameterized in terms of multiple uniform sub-layers. For tomographic inversion, the forward problem must determine the wavefront evolution or ray path for a particular seismic phase (in this case, surface waves) to accurately predict travel times (e.g., Rawlinson & Sambridge, 2004). All inverse methods incorporate an underlying approach to solving the corresponding forward problem, with associated assumptions that can impact the model solution. The specific forward problems in this thesis are discussed in chapters 2 and 3.

1.1.3 Errors and misfit

The goal of most inversion approaches is to determine the set of model parameters for which the predicted data best fit the observed data (or fit the observed data to a specified level). A perfect fit is generally not possible (or desirable) because of data errors, which include measurement errors on the observed data (e.g., due to instrument effects and competing seismic vibrations in the environment) and theory errors on the predicted data (due to simplified model parameterizations and approximate physics of the forward problem). A common measurement error when considering surface-wave dispersion is the misidentification, or interaction, of dominant modes in the data (O'Neill & Matsuoka, 2005). This is predominantly an issue for short-period, geotechnical-scale surface-wave studies. Sources of measurement error unique to each problem in this thesis are discussed in subsequent thesis chapters. Data errors (sum of measurement and theory errors) can be considered to represent all factors that cannot be modelled or accounted for in the inverse problem. In some cases, measurement errors can be characterized statistically from repeated sets of observed data. However, this approach does not apply to theory errors, the statistics of which are usually poorly known.

In many inversion approaches, the data misfit function is formulated based on knowl-

edge or assumptions on the error processes. Considering the differences between observed and predicted data (called data residuals) to represent the errors, the misfit function can be derived by assuming a particular statistical distribution of the residuals given the model, which is then interpreted as the likelihood of the model. For example, the common assumption of Gaussian-distributed errors leads to an L_2 misfit function (negative log-likelihood function) consisting of the sum of squared residuals, weighted by the inverse error covariance matrix. The Gaussian assumption is supported by the Central Limit Theorem, which states that the sum of a number of error processes tends to a Gaussian distribution regardless of the distributions of the individual processes. Further, L_2 misfit minimization corresponds to least-squares methods, for which analytic solutions exist for linear and linearized problems. However, the assumption of Gaussian errors and least-squares misfit can be inappropriate for data sets containing outliers (data with improbably large errors). In such cases, the assumption of Laplacian errors (i.e., a two-sided exponential distribution) is more robust. This assumption leads to an L_1 misfit function based on the sum of absolute values of residuals, weighted by the square-root of the inverse covariance matrix. No analytic solution to L_1 misfit minimization exists.

The L_1 and L_2 misfit functions mentioned above formally require knowledge of the error covariance matrix, which is often not available. However, under the assumption of independent, identically distributed (IID) errors (i.e., a diagonal covariance matrix with constant variance σ^2), the problem of misfit minimization is independent of the covariance matrix and unweighted misfit functions apply. This approach is commonly used for misfit minimization; however, it does not apply to parameter uncertainty estimation where error quantification is required. Further, it should be noted that if the errors are not IID (common for geophysical data), the approach can lead to suboptimal (biased) solutions. In such cases, it may be preferable to estimate the error statistics from the data as part of the inverse problem. This requires variance estimation if the errors can be considered independent, or full covariance estimation for inter-dependent (e.g., serially correlated) errors. Variance/covariance can be estimated using non-parametric approaches where the residuals from an initial unweighted inversion are used to compute a diagonal or Toeplitz (band-diagonal) covariance matrix for a subsequent inversion. Alternatively, parametric

approaches can be applied in nonlinear (numerical) inversions by including the parameters for a model of the covariance matrix in the inversion.

Misfit functions can also be defined without a statistical foundation to serve the same purpose of quantifying the difference between observed and predicted data. For example, the optimal-transport metric has gained popularity as a misfit function in several seismological inverse problems (e.g., Métivier et al., 2016; Hedjazian et al., 2019). This misfit metric enhances certain desirable properties of the problem such as linearity of the relationship between data and model parameters, as well as the uniqueness of the problem. It is worth noting that, generally, the nonlinear inversion strategies reviewed in this chapter need not be based on specific assumptions on data errors.

Regardless of how the misfit function is defined, it can be interpreted as a hyper-surface in the model space (i.e., a multi-dimensional function over the model parameters). The goal for many inverse methods is to estimate the optimal set of model parameters that represent the global minimum of the misfit surface. An important distinction is whether the inverse problem is linear, weakly nonlinear, or strongly nonlinear. Linear inverse problems often can have simple misfit surfaces (fig. 1.2a). In particular, for a linear inverse problem with Gaussian-distributed errors, the L_2 misfit function represents the (negative) logarithm of a Gaussian likelihood function over the parameter space, and an analytic expression for this solution exists.

Most inverse problems in seismology (and geophysics in general) are nonlinear. A rare example of a linear problem is the inversion of measured surface-wave attenuation curves to estimate compressional-wave and/or shear-wave attenuation (described by dimensionless quality factor Q) where, for known velocity structure, the relationship between surface-wave attenuation and body-wave attenuation is linear (e.g., Xia et al., 2002). In some cases, the nonlinearity of an inverse problem may be weak, such that the problem can be solved via linearization, iteratively stepping down the local misfit gradient to the minimum. For strongly nonlinear problems, the misfit surface can be complex, potentially including multiple disconnected regions with low misfit (i.e., local minima) in the parameter space (fig. 1.2b). Linearized inversions can fail (converge to local minima or diverge) for strongly

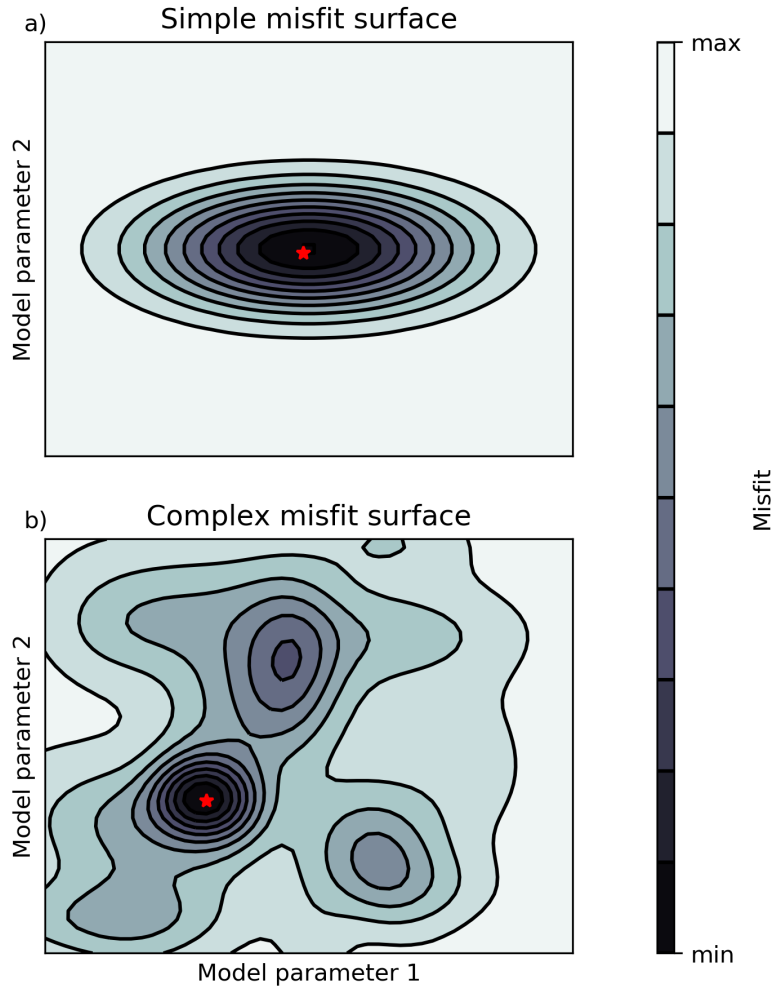


Figure 1.2: Hypothetical misfit surfaces in a 2D model space. A simple misfit surface for an idealized linear problem is shown in (a). A complex misfit surface, with multiple local minima, for a nonlinear problem is shown in (b). The locations of the global minima of the surfaces are shown by red stars.

nonlinear problems, and nonlinear approaches may be required.

Nonlinear inversion in seismology comprises a diverse collection of methods. The most-common approaches include the downhill simplex method (DHS), which moves roughly down the misfit gradient through geometric moves without computing derivatives of the hyper-surface. Additionally, the use of global search methods such as simulated annealing (SA) and genetic algorithms (GA) is widespread in geophysical literature. These approaches apply a directed random search based on optimization processes found in nature. An alternative to inversion methods that seek the best-fit model solution (a point estimate in the parameter space), Bayesian inversion is based on probabilistic (numerical) sampling over the parameter space to estimate properties of the posterior probability density (PPD), providing parameter estimates together with quantitative (nonlinear) uncertainty analysis.

1.1.4 Parameterization

The approach to parameterizing the model is another important issue in inversion. Adopting too few parameters (e.g., layers in a 1D problem or grid cells in 2D or 3D) can underfit the observed data, leaving model structure unresolved and biasing parameter estimates. Conversely, adopting too many parameters can overfit the data (i.e., fit the errors on the data), resulting in models with spurious (unconstrained) structural features. Furthermore, more model parameters typically means greater computation cost.

The inversion also depends on the form of the model represented by the parameterization. For example, as mentioned above, most 1D problems consider the subsurface to be represented by a discrete stack of uniform layers, with discontinuities in geophysical properties at layer boundaries. However, some recent inversions consider gradient (smooth) structures represented by linear or power-law functions (e.g., Molnar et al., 2010) or, more generally, by polynomial basis functions (Gosselin et al., 2017). In near-surface seismic applications, these gradient-based parameterizations have been shown to characterize the structure of unconsolidated soils and sediments better than discrete layered models, although this is not universal for all sites.

Some approaches to parameterization, particularly for linearized inversions, include only the geophysical properties for a fixed discretization of the subsurface model, since solving for spatial parameters (e.g., layer thicknesses or cell sizes) can significantly increase the nonlinearity of the problem. In this approach, the model is typically over-parameterized with many layers/grid cells that are below the spatial resolution of the data to provide flexibility for the solution. The resulting under-determined inversion can be constrained by regularization (described in a later section) to explicitly control the data misfit and structure of the solution.

An alternative parameterization approach is to solve for the spatial properties of the model (individual layer thicknesses or cell sizes) as part of the problem. This is more often applied in nonlinear approaches, where the increased nonlinearity is of less concern. This approach typically assumes a small number of layers/cells to constrain model structure, but the issue of over- or under-parameterization is sometimes given little attention beyond qualitative practices such as progressively increasing the number of layers until the data misfit stops decreasing significantly. However, quantitative approaches to parameterization based on formal information criteria or by sampling probabilistically over the number of parameters have recently gained popularity in seismology, and are discussed further in section 1.4.

1.1.5 Other considerations

The inversion approaches discussed in this chapter have various advantages and disadvantages. For instance, linearized inversions often require the fewest data predictions (forward solutions) and are the least computationally demanding. Bayesian inversions numerically sample the model probability over the parameter space, and can require thousands (or millions) of data predictions. Consequently, Bayesian methods are typically the most computationally demanding, but also provide the most informative solutions. Nonlinear optimizations generally fall somewhere between these extremes in terms of computational demand. Computational cost is of less concern for 1D problems, where forward solutions are quite efficient such that inversions can generally be solved in a matter of seconds (for

linearized methods) to minutes (for Bayesian methods) with serial algorithms on a desktop computer. For tomographic studies, which are 2D or 3D by design, the computational demand of the forward problem is much greater and most approaches rely on linearized inversion. Bayesian methods have only recently been applied to large-scale seismic inversions at significant computational expense (e.g., Bodin & Sambridge, 2009; Bodin et al., 2012). This is discussed further in chapter 2.

Inverse methods are sometimes treated as a “black box”, with data as input and a model of earth structure as output, but with little consideration of the underlying processes. However, the output model can be nonphysical or meaningless if the data are of poor quality or if the error assumptions or model parameterization do not apply. The observed data should be inspected (visually or otherwise) for quality control and to assess the reliability of model solutions. For all inverse problems, the agreement between the observed and predicted data should be examined to ensure a meaningful fit (but note that good data fit is a necessary but not a comprehensive condition for a meaningful model solution). Examining the data residuals after an inversion can also provide useful insight into the data-error statistics.

Due to the non-uniqueness and potential instability of inverse problems, it is important to assess the uncertainty of the model solution, when possible. Understanding and quantifying the uncertainties in estimated subsurface structure is of significant value for robust interpretation. However, uncertainty estimation can be challenging for inverse problems, and many methods are approximate and/or qualitative. For instance, linearization errors as well as regularization schemes generally preclude quantitative uncertainty estimation for linearized inversions. Global-search methods are designed to locate the optimal solution, but not estimate uncertainties. However, Bayesian inversion methods can provide rigorous uncertainty quantification for nonlinear problems (for known/assumed/estimated data errors). Further, the estimated inversion uncertainties can be propagated into subsequent analyses (e.g., thermal or geodynamic modelling, seismic response of sedimentary basins, etc.).

1.2 Linearized inversion

As mentioned in the previous section, although most inverse problems in seismology are functionally nonlinear, in some cases the analytic theory for linear inversion can be applied via local linearization and iteration. This section reviews common linearized inversion techniques as a foundation for discussion on more complex inverse methods.

Consider a nonlinear inverse problem with vectors of model parameters $\mathbf{m}$, observed data $\mathbf{d}$, and predicted data $\mathbf{d}(\mathbf{m})$. Equating the observed data to those predicted for the model sought, and expanding the data functional (forward problem) about a starting model $\mathbf{m}_0$ to first order leads to

$$\begin{aligned}\mathbf{d} &= \mathbf{d}(\mathbf{m}) \\ &\approx \mathbf{d}(\mathbf{m}_0) + \mathbf{A}(\mathbf{m} - \mathbf{m}_0).\end{aligned}\tag{1.1}$$

Here, $\mathbf{A}$ is the sensitivity or Jacobian matrix of partial derivatives, $A_{ij} = \partial d_i(\mathbf{m}_0)/\partial m_{0j}$. Neglecting higher-order terms in the expansion linearizes the inverse problem about the initial model. Rearranging equation (1.1) and defining data and model perturbations $\delta\mathbf{d} = \mathbf{d} - \mathbf{d}(\mathbf{m}_0)$ and $\delta\mathbf{m} = \mathbf{m} - \mathbf{m}_0$, respectively, leads to a linear relation:

$$\delta\mathbf{d} = \mathbf{A} \delta\mathbf{m}.\tag{1.2}$$

This is the fundamental relationship between changes in a proposed model and resulting changes in the forward-modelled data, which can be used to refine the initial model.

For some inverse problems, analytic expressions for the required partial derivatives are available, but in other cases they must be estimated numerically (e.g., via finite differences). The sensitivity matrix encapsulates the physics/geometry for the (linearized) problem, and can provide useful insight. For example, Xia et al. (1999) examined the sensitivity matrix for 1D linearized inversion of surface-wave dispersion data, and concluded quantitatively that dispersion data are significantly more sensitive to V_S than V_P or ρ . Furthermore, Xia et al. (2003) determined that, for a given frequency, higher-order modes are more sensitive than the fundamental mode to deeper structure, and thereby provide greater information for, and constraint on, such structure. Note that this defines the sensitivity matrix within

the context of a discrete, parameterized model. The continuous (and analytic) equivalents to the sensitivity matrix are often called the Fréchet derivatives (McGillivray & Oldenburg, 1990).

As mentioned previously, the assumption of Gaussian-distributed errors for a linear inverse problem leads to an analytic solution. For equation (1.2), this assumption leads to a likelihood function

$$L(\delta\mathbf{m}) = \frac{1}{(2\pi)^{N/2}|\mathbf{C}|^{1/2}} \exp \left[-\frac{\Phi(\delta\mathbf{m})}{2} \right], \quad (1.3)$$

where

$$\Phi(\delta\mathbf{m}) = (\delta\mathbf{d} - \mathbf{A}\delta\mathbf{m})^T \mathbf{C}^{-1} (\delta\mathbf{d} - \mathbf{A}\delta\mathbf{m}) \quad (1.4)$$

is the misfit (negative log-likelihood) function and $\mathbf{C}$ is the data error covariance matrix. The best-fit model perturbation $\delta\mathbf{m}$ (for the linearized problem) can be found by maximizing the likelihood or, equivalently, minimizing the misfit: setting $\partial\Phi(\delta\mathbf{m})/\partial\delta\mathbf{m} = 0$ leads to

$$\delta\mathbf{m} = [\mathbf{A}^T \mathbf{C}^{-1} \mathbf{A}]^{-1} \mathbf{A}^T \mathbf{C}^{-1} \delta\mathbf{d}. \quad (1.5)$$

The model solution is given by $\mathbf{m}_1 = \mathbf{m}_0 + \delta\mathbf{m}$. Since higher-order terms were neglected, this may not represent a satisfactory solution, but the procedure can be repeated iteratively until the data are fit appropriately and/or the parameters no longer change between iterations.

The linearized solution, equation (1.5), requires the matrix $\mathbf{A}^T \mathbf{C}^{-1} \mathbf{A}$ to be well-conditioned. In practice, depending on the choice of parameterization, inverse problems in seismology are often ill-conditioned (nearly singular), which leads to unstable inversions. The inversion can be stabilized using singular-value decomposition (e.g., Parolai et al., 2006) and/or by incorporating additional constraints on the model parameters independent of the data.

A common strategy to incorporate constraints for stability is regularization (fig. 1.3). Rather than minimizing the data misfit, equation (1.4), regularization considers a more general objective function that augments the data misfit with a model misfit term

$$\begin{aligned} \Psi(\delta\mathbf{m}, \beta) = & (\delta\mathbf{d} - \mathbf{A}\delta\mathbf{m})^T \mathbf{C}^{-1} (\delta\mathbf{d} - \mathbf{A}\delta\mathbf{m}) + \\ & \beta(\delta\mathbf{m} - \delta\hat{\mathbf{m}})^T \mathbf{R}^T \mathbf{R} (\delta\mathbf{m} - \delta\hat{\mathbf{m}}), \end{aligned} \quad (1.6)$$

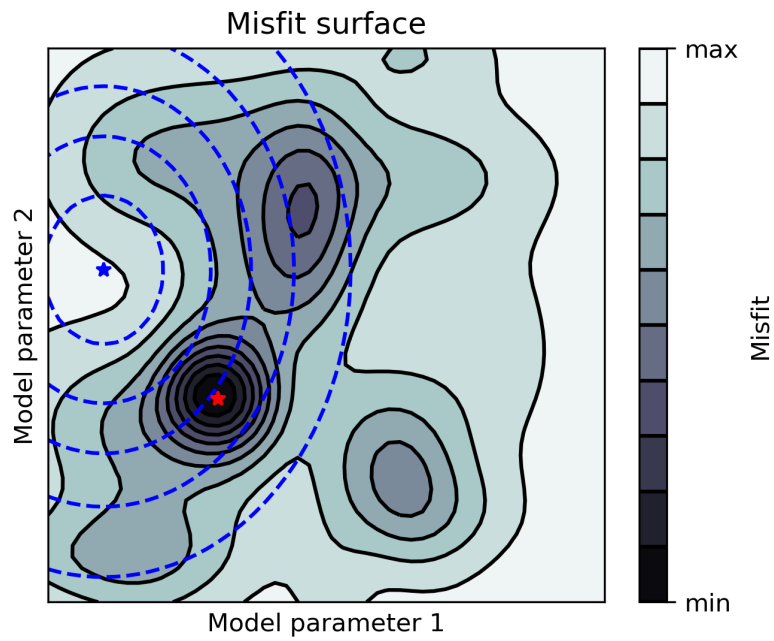


Figure 1.3: Tradeoff between data misfit and regularization. Solid contours show the misfit surface with its minimum at the red star. Dashed contours show a hypothetical regularization surface with its minimum at the blue star (which does not fit the data). The regularized solution balances minimizing both functions. Adapted from Sambridge & Mosegaard (2002).

where $\delta\hat{\mathbf{m}}$ represents a preferred value for $\delta\mathbf{m}$, $\mathbf{R}$ represents a weighting matrix (the regularization matrix), and β is a trade-off parameter determining the relative importance of the two terms. Minimizing equation (1.6) with respect to $\delta\mathbf{m}$ leads to the regularized solution which may be expressed by

$$\delta\mathbf{m} = \delta\hat{\mathbf{m}} + [\mathbf{A}^T\mathbf{C}^{-1}\mathbf{A} + \beta\mathbf{R}^T\mathbf{R}]^{-1} \mathbf{A}^T\mathbf{C}^{-1} [\delta\mathbf{d} - \mathbf{A}\delta\hat{\mathbf{m}}]. \quad (1.7)$$

In equation (1.7), the term $\beta\mathbf{R}^T\mathbf{R}$ stabilizes the matrix inversion, given appropriate choices of β and $\mathbf{R}$. The role of β as a trade-off parameter is clear: in the limit $\beta \rightarrow 0$ equation (1.7) simplifies to equation (1.5) which minimizes data misfit alone, while the limit $\beta \rightarrow \infty$ leads to $\delta\mathbf{m} = \delta\hat{\mathbf{m}}$ which minimizes model misfit. The goal is to determine an appropriate value for β that provides an acceptable data misfit while stabilizing the inversion. If suitable knowledge of the data error covariance is available, β can be chosen to fit the data according to a statistical criterion (e.g., χ^2 test). If not, a more subjective approach is to plot the data misfit versus the model misfit to determine a balance near the inflection point of this curve (the L-curve method, e.g., Hansen, 1992).

One of the most common forms of regularization used in seismology sets $\mathbf{R} = \mathbf{I}$ (the identity matrix) and $\delta\hat{\mathbf{m}} = \mathbf{0}$ (i.e., a preference for small linearized step size), such that equation (1.7) becomes

$$\delta\mathbf{m} = [\mathbf{A}^T\mathbf{C}^{-1}\mathbf{A} + \beta\mathbf{I}]^{-1} \mathbf{A}^T\mathbf{C}^{-1}\delta\mathbf{d}. \quad (1.8)$$

This solution adds β to the main diagonal of $\mathbf{A}^T\mathbf{C}^{-1}\mathbf{A}$ to overcome ill-conditioning. For IID errors (i.e., $\mathbf{C} = \sigma^2\mathbf{I}$), this method is often referred to as damped least squares. This regularization, favoring small $\delta\mathbf{m}$, is consistent with local linearization that may only apply in a small region around the initial model.

Levenberg (1944) proposed a strategy for assigning β based on the number of linearized iterations such that β is larger for initial iterations (consistent with the local linear assumption) but gradually decreases for later iterations in order to better fit the observed data near convergence. Marquardt (1963) improved the inversion by specifying $\mathbf{R}^T\mathbf{R}$ to be the diagonal components of $\mathbf{A}^T\mathbf{C}^{-1}\mathbf{A}$ such that the relative weighting of model parameters

in the regularization is defined by information in the sensitivity matrix, resulting in the Levenburg-Marquart approach.

Alternatively, regularizations can be defined to represent first- or second-order spatial derivative operators applied to the parameters to minimize model gradients or roughness, respectively (Constable et al., 1987; Aster et al., 2018). These regularizations minimize complex structure, producing simple (flat or smooth) models. For these regularizations, the linearized problem, equation (1.2), is typically recast as

$$\delta \mathbf{d} + \mathbf{A} \mathbf{m}_0 = \mathbf{A} \mathbf{m}, \quad (1.9)$$

such that the inversion and regularization are formulated for the updated model rather than the model perturbation. Minimum-structure regularized inversions are often applied in tomographic studies, discussed later.

Linearized inversion may fail by diverging, or converging to a local (rather than global) minimum, if the nonlinearity of the problem is strong and/or the initial model is poor. Nonlinear methods, discussed in the following section, are designed to overcome these problems. Compared to nonlinear methods, linearized inversion typically requires fewer forward operations (data predictions) since they are based on exploiting misfit gradient information rather than employing directed random searches. Some studies have considered multi-step hybrid inversions that initially apply a nonlinear approach to estimate an accurate initial model for a subsequent linearized inversion (e.g., Picozzi & Albarello, 2007).

In linear inverse theory, the quality of the model solution can be assessed through the calculation of resolution and model covariance matrices. However, for linearized inversions, these measures suffer from linearization error, and regularization can preclude meaningful results.

1.3 Nonlinear optimization

Linearized inversion methods considered in the previous section are sometimes referred to as local searches since, although they move efficiently downhill based on misfit gradient

information, they typically remain close to the starting model and are prone to becoming trapped in local minima. As an alternative, nonlinear search (optimization) methods are designed to widely search the space and (ideally) avoid sub-optimal solutions. Although optimization is not explicitly considered in this thesis, it represents the logical bridge between classical (linear/linearized) inverse theory and the Bayesian methods applied in this thesis. The goal of these methods is to determine the set of model parameter values that minimizes the data misfit via numerical optimization. Specifically, the various methods all solve the same mathematical problem, but apply different optimization schemes.

A variety of nonlinear optimization methods have been applied to geophysical inversion, and only a few (historically significant) approaches are described in this section including DHS, as well as global search methods of SA and GA. The greatest advantage of global-search algorithms (i.e., SA and GA) is that they are relatively insensitive to the starting/initial model, unlike linearized inversions or DHS. No global-search algorithm is guaranteed to converge to the global minimum in misfit within a finite number of steps, although they can be much more effective than linearization (at increased computational cost). Furthermore, there is no general approach to determine whether the solution obtained actually represents the global minimum, although confidence is increased if repeated runs of the algorithm (with different random initializations) produce similar results.

For a more complete and technical discussion on some of the techniques reviewed in this section (for general applications in geophysics), as well as for further discussion on exploitation vs. exploration of the misfit hyper-surface in geophysical inverse methods, see Sambridge & Mosegaard (2002).

1.3.1 Downhill simplex

The DHS method (Nelder & Mead, 1965) is an optimization method based on a geometric scheme for moving “downhill” in parameter space without calculating partial derivatives. DHS operates on a simplex (convex hull) of $M + 1$ models in an M -dimensional model parameter space. Computing the misfit for each model of the simplex provides (limited) information on the local misfit gradient without derivative computations. DHS moves

progressively downhill in misfit without relying on linearization, but as it has limited ability to move uphill except potentially via multiple contractions (a geometric operation), it is prone to becoming trapped in local minima in the parameter space. To improve the confidence in finding the global minimum, it is recommended to run the procedure several times, initiated from different starting models. Because the DHS method is prone to converge to local minima in the parameter space, it has been combined with global search methods to exploit the advantages of each (e.g., García-Jerez et al., 2019).

1.3.2 Simulated annealing

SA is based on an analogy to the natural optimization process of thermodynamic annealing, by which crystals are grown and metals hardened (Van Laarhoven & Aarts, 1987). The optimization algorithm consists of a series of iterations involving random perturbations of the unknown model parameters (representing the thermodynamic system) of $\mathbf{m} \rightarrow \mathbf{m}'$, with a resulting change to the data misfit function (analogous to free energy of the system) of $\Phi(\mathbf{m}) \rightarrow \Phi(\mathbf{m}')$. Perturbations that decrease the misfit are always accepted, while perturbations that increase misfit are sometimes accepted with an acceptance probability given by the Gibbs distribution of statistical mechanics

$$A(\mathbf{m}'|\mathbf{m}) = \exp(-\Delta\Phi/T), \quad (1.10)$$

where $\Delta\Phi = \Phi(\mathbf{m}') - \Phi(\mathbf{m})$ represents the increase in misfit due to the perturbation and T is a control parameter (analogous to absolute temperature). Over the process of many such iterations the temperature T is gradually reduced from an initial high value (cooling/annealing the system in thermodynamic terms). Accepting some perturbations that increase Φ allows the algorithm to escape from local minima in search of a better solution. At early iterations (high T), the algorithm searches the parameter space in an essentially random manner. As T decreases, accepting increases in Φ becomes increasingly improbable, and the algorithm spends more time searching regions of low Φ , eventually converging to a solution that should approximate the global minimum.

The starting temperature, rate of reducing T , and the number and type of perturbations define the annealing schedule, which controls the efficiency and effectiveness of the algorithm. Adopting an annealing schedule that is too fast (i.e., decreases T too quickly or allows too few perturbations) can lead to sub-optimal solutions. Alternatively, adopting an annealing schedule that is overly cautious wastes computation time. Determining an appropriate annealing schedule is problem specific and generally requires some experimentation and familiarity with the inverse problem. SA can be related to Markov-chain Monte Carlo (MCMC) methods (described for Bayesian inversion in section 1.4), with the goal of optimization based on non-convergent sampling with decreasing T rather than probability estimation based on sampling to convergence at unit temperature.

1.3.3 Genetic algorithms

GA is based on an analogy to biological evolution according to the concept of “survival of the fittest”. GA simulates the genetic evolution of a collection (population) of models through many iterations (generations) to minimize their misfit. This is analogous to maximizing the population’s fitness to a specific ecological niche (Fogel et al., 1966; Holland et al., 1992; Sambridge & Mosegaard, 2002). Each generation of models acts as parents for the next generation of (offspring) models through processes designed to mimic selection (pairing of parent models), crossover (recombination of parent genetic information in offspring), and mutation (random variations) in a manner that probabilistically favors models with lower misfits. A large variety of approaches to the selection, crossover, and mutation steps exist, which will not be discussed here. Like the annealing schedule for SA, these GA steps (and the tradeoff between their efficiency and robustness) are specific to the problem. GA has been widely used in seismic studies that invert surface-wave dispersion data (e.g., Yamanaka & Ishida, 1996; Parolai et al., 2006).

1.4 Bayesian Inversion

Bayesian inference approaches to geophysical inversion are based on quantifying the PPD of Earth models given observed data and prior information. To quantify model probability over the multi-dimensional parameter space, Bayesian inversions generally seek to compute statistical properties of the PPD. PPD properties of interest include parameter estimates, such as most-probable and mean values. More significantly, Bayesian inversion can quantify parameter uncertainties, which can be expressed as variances/covariances, credibility intervals, and (most notably) marginal probability densities. Since rigorous uncertainty analysis, rather than point estimation, is central to Bayesian inference, nonlinear inversion and rigorous model selection or of greater significance (than in other approaches). In the common case where appropriate Earth and error models are not known *a priori*, they can be estimated from the data, as part of the inverse problem.

Generally, nonlinear parameter and uncertainty estimation is carried out numerically using MCMC methods, as discussed in Section 1.4.1. MCMC draws random samples of the parameters from the PPD, such that statistical parameter/uncertainty estimates can be computed from the ensemble of samples. Efficient sampling strategies are key for practical inversion algorithms.

In terms of model selection, determining an appropriate Earth-model parameterization is an important aspect of quantitative inversion. As stated previously, over-parameterizing the model under-constrains parameters and can lead to spurious structure and to over-estimating parameter uncertainties. Conversely, under-parameterization can leave structure unresolved, biasing parameter estimates and under-estimating uncertainties. An objective approach to model selection is to choose the simplest parameterization that explains the data as quantified by the Bayesian information criterion (BIC), which is discussed in Section 1.4.2. Another approach is to marginalize over a set of possible parameterizations. Since this involves probabilistic sampling over models with different numbers of parameters (model dimensions), this approach is referred to as trans-dimensional (trans-D) inversion. Trans-D inversion, described in Section 1.4.2, has the advantage that the uncertainty in the parameterization is included in the parameter uncertainty estimates. Both of these model-

selection approaches avoid subjective regularizations in the inversion, which can preclude meaningful uncertainty estimation.

Defining the data error model is another important component of rigorous uncertainty estimation. In many practical cases, the error distribution, including both measurement and theory errors, is not well known. The lack of specific information suggests that a simple distribution be assumed, with statistical parameters estimated from the data.

An important advantage of the Bayesian approach is that it is straightforward to propagate the uncertainty analysis for the inversion result directly into uncertainties in subsequent analyses. This cannot be accomplished with other approaches reviewed in this chapter.

1.4.1 McMC sampling

To briefly describe nonlinear Bayesian inversion, let $\mathcal{M}$ represent the choice of model, with $\mathbf{m}$ the corresponding set of M unknown model parameters, and let $\mathbf{d}$ be N observed data. Assuming data and parameters to be random variables, they are related by Bayes' theorem

$$P(\mathbf{m}|\mathbf{d}, \mathcal{M}) = \frac{P(\mathbf{m}|\mathcal{M}) P(\mathbf{d}|\mathbf{m}, \mathcal{M})}{P(\mathbf{d}|\mathcal{M})}. \quad (1.11)$$

In equation (1.11), $P(\mathbf{m}|\mathcal{M})$ is the prior probability, representing available information for the model parameters (given a choice of model), independent of the data. $P(\mathbf{d}|\mathbf{m}, \mathcal{M})$, the conditional probability of $\mathbf{d}$ given $\mathbf{m}$ and $\mathcal{M}$, defines the data information. Interpreted as a function of $\mathbf{d}$, this represents the residual error distribution. However, when $\mathbf{d}$ is considered fixed (once data are observed), the term is interpreted as the likelihood of the parameters, $\mathcal{L}(\mathbf{m}|\mathcal{M})$. The normalization term, $P(\mathbf{d}|\mathcal{M})$, is referred to as Bayesian evidence. $P(\mathbf{m}|\mathbf{d}, \mathcal{M})$ defines the PPD, representing the state of parameter information given the data (via the likelihood) and prior (as illustrated in fig. 1.4). Henceforth, the dependence on model $\mathcal{M}$ is suppressed for simplicity when not required. In practice, many studies in geophysics that use Bayesian inversion use a bounded uniform prior (as an approximation of an uninformative prior), such that the shape of the posterior is equivalent

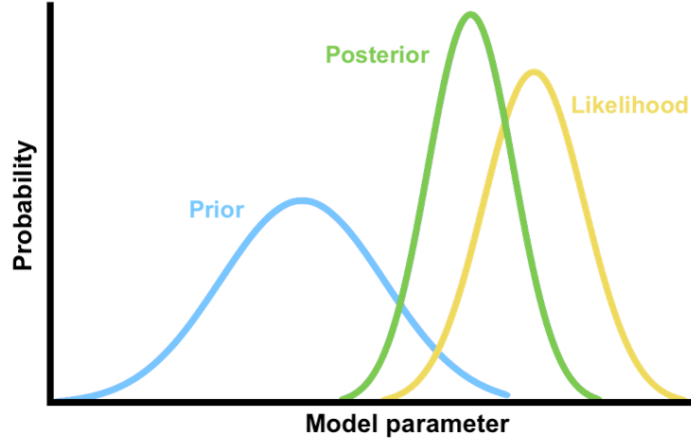


Figure 1.4: Illustration of Bayesian inference. The posterior probability (state of knowledge) is informed by the prior probability and data information (via the likelihood).

to the shape of the likelihood. However, Bayesian inversion also allows for quantitative incorporation of meaningful prior knowledge within the inverse problem.

Nonlinear Bayesian inversion is typically based on using MCMC methods to draw (dependent) random samples from the PPD while satisfying the requirement for reversibility of the Markov chain (Gilks et al., 1996; Sambridge & Mosegaard, 2002). In particular, Metropolis-Hastings sampling constructs a Markov chain by applying a proposal density function $Q(\mathbf{m}'|\mathbf{m})$ to generate new parameters $\mathbf{m}'$ based only on the current values $\mathbf{m}$, and accepting the proposed parameters as the next step in the chain with probability

$$A(\mathbf{m}'|\mathbf{m}) = \min \left[1, \frac{Q(\mathbf{m}|\mathbf{m}')}{Q(\mathbf{m}'|\mathbf{m})} \frac{P(\mathbf{m}')}{P(\mathbf{m})} \frac{\mathcal{L}(\mathbf{m}')}{\mathcal{L}(\mathbf{m})} \right]. \quad (1.12)$$

The acceptance criterion is applied by drawing a uniform random number ξ on $[0, 1]$ and accepting the new parameters if $\xi \leq A(\mathbf{m}'|\mathbf{m})$. If the proposed step is rejected, another copy of the current model is included in the Markov chain. Once convergent sampling is achieved, parameter estimates and uncertainties can be computed statistically from

the ensemble of samples (omitting an initial burn-in stage while stationary sampling is established).

The choice of proposal density $Q(\mathbf{m}'|\mathbf{m})$ controls the efficiency of MCMC sampling. The goal is to achieve a well-mixed Markov chain that efficiently samples the parameter space, avoiding both small, ineffectual perturbations and high rejection rates. Dosso & Wilmut (2008) introduced efficient proposal densities based on principal-component (PC) decomposition of the parameter covariance matrix. The PC decomposition provides both directions and length scales for effective parameter updates. Perturbations are applied in a rotated parameter space where the axes align with the dominant correlation directions of the model parameters (i.e., PC parameters are uncorrelated). Many studies now consider the techniques of parallel tempering in order to achieve wide sampling over a potentially multi-modal parameter space (Earl & Deem, 2005; Dosso et al., 2012). This is based on a series of interacting Markov chains with successively relaxed likelihoods raised to powers $1/T \leq 1$, where T is referred to as the sampling temperature (similar to the temperature parameter in SA, discussed in Section 1.3.2). While only samples collected at $T = 1$ are unbiased and retained for analysis, probabilistic interchanges between chains provides a robust and efficient sampling of the parameter space.

1.4.2 Model selection

As mentioned earlier, determining an appropriate model parameterization is an important aspect of Bayesian inference. In equation (1.11), the Bayesian evidence $P(\mathbf{d}|\mathcal{M})$, which formally represents the conditional probability of the data for a particular model, can be considered the likelihood of the model given the data. Hence, a natural approach to model selection is to seek the model that maximizes the evidence. Unfortunately, the integral defining evidence is particularly challenging to evaluate numerically to sufficient precision (Chib, 1995). Commonly, an asymptotic point estimate of evidence, the BIC, is applied, defined by (Schwarz, 1978)

$$\begin{aligned} \text{BIC}(\mathcal{M}) &\approx -2 \log_e P(\mathbf{d}|\mathcal{M}) \\ &= -2 \log_e \mathcal{L}(\hat{\mathbf{m}}|\mathcal{M}) + M \log_e N, \end{aligned} \tag{1.13}$$

where $\hat{\mathbf{m}}$ is the maximum-likelihood parameter estimate. Since the BIC approximates the negative of evidence, the model with the smallest BIC over a set of possible models is selected as the most appropriate choice. The first term on the right of equation (1.13) favors models with high likelihood (low data misfit); however, this is balanced by the second term which applies a penalty for additional parameters. Hence, minimizing the BIC provides the simplest model parameterization consistent with the resolving power of the data.

Trans-D inversion addresses model selection by sampling probabilistically over models with differing numbers of parameters (Green, 1995; Malinverno, 2002; Sambridge et al., 2006; Dettmer et al., 2010; Dosso et al., 2014). Let k index the choice from $\mathcal{K}$ possible models; Bayes' theorem for hierarchical models can be written (Green, 1995)

$$P(k, \mathbf{m}_k | \mathbf{d}) = \frac{P(k) P(\mathbf{m}_k | k) P(\mathbf{d} | k, \mathbf{m}_k)}{\sum_{k' \in \mathcal{K}} \int P(k') P(\mathbf{m}'_{k'} | k') P(\mathbf{d} | k', \mathbf{m}'_{k'}) d\mathbf{m}'_{k'}}. \quad (1.14)$$

In equation (1.14), $P(k)P(\mathbf{m}_k | k)$ is the prior probability of the state $(k, \mathbf{m}_k)$, $P(\mathbf{d} | k, \mathbf{m}_k)$ is interpreted as the likelihood $\mathcal{L}(k, \mathbf{m}_k)$, and $P(k, \mathbf{m}_k | \mathbf{d})$ is the trans-D PPD. The PPD can be sampled numerically using the reversible-jump Markov-chain Monte Carlo (rjMCMC) algorithm, which accepts a proposed transition between the current state $(k, \mathbf{m}_k)$ and a proposed state $(k', \mathbf{m}'_{k'})$ with a probability given by the Metropolis-Hastings-Green criterion (Green, 1995)

$$A(k', \mathbf{m}'_{k'} | k, \mathbf{m}_k) = \min \left[1, \frac{P(k', \mathbf{m}'_{k'})}{P(k, \mathbf{m}_k)} \frac{L(k', \mathbf{m}'_{k'})}{L(k, \mathbf{m}_k)} \frac{Q(k, \mathbf{m}_k | k', \mathbf{m}'_{k'})}{Q(k', \mathbf{m}'_{k'} | k, \mathbf{m}_k)} |\mathbf{J}| \right], \quad (1.15)$$

where $Q(k', \mathbf{m}'_{k'} | k, \mathbf{m}_k)$ is the proposal probability density and $|\mathbf{J}|$ is the Jacobian determinant for the transformation between parameter spaces. Most studies only consider model parameter perturbations such that $|\mathbf{J}| = 1$. Note that the index k is omitted in subsequent discussion (including subsequent chapters) for simplicity.

1.4.3 Error model and likelihood function

Defining the error model requires specifying the statistical distribution of residual errors, which is often not well known. The lack of specific information suggests that a simple distribution be assumed, with statistical parameters estimated from the data. Assuming Gaussian-distributed errors with an unknown error covariance matrix $\mathbf{C}$, the likelihood is

$$\mathcal{L}(\mathbf{m}, \mathbf{C}) = \frac{1}{(2\pi)^{N/2} |\mathbf{C}|^{1/2}} \exp \left[-\frac{1}{2} \mathbf{r}(\mathbf{m})^T \mathbf{C}^{-1} \mathbf{r}(\mathbf{m}) \right], \quad (1.16)$$

where $\mathbf{r}(\mathbf{m}) = \mathbf{d} - \mathbf{d}(\mathbf{m})$ are data residuals. Errors are often assumed to be IID random variables. However, significant error correlations can occur, and neglecting these can bias parameter estimates and under-estimate uncertainties.

A variety of approaches have been applied to address error covariance in Bayesian inversion. Molnar et al. (2010, 2013) applied a non-parametric approach to estimate $\mathbf{C}$ based on residuals from the optimal model of an initial inversion assuming IID errors. In this approach, the residuals are considered a realization of the error process from which statistical quantities can be estimated. Assuming the residuals represent an ergodic random process, a Toeplitz (diagonally-banded) covariance matrix can be estimated from the residual autocovariance (Dosso et al., 2006); this covariance-matrix estimate is then used in a subsequent Bayesian inversion. Alternatively, parametric approaches to error covariance can be used by considering data residuals to be a first-order auto-regressive process, which is equivalent to a Toeplitz covariance matrix with exponentially-diminishing off-diagonal (covariance) terms (e.g., Dettmer et al., 2012; Gosselin et al., 2017, 2018). In this case, the standard deviation of the data errors and the autoregressive parameter are sampled in the inversion to account for error variance and covariance. Both non-parametric and parametric methods can be extended to consider error statistics that vary over the data set by dividing the data into segments over which error statistics are assumed constant.

1.5 Tomography

Seismic tomography has been an important tool for imaging heterogeneous Earth structure over the last several decades, and has been applied over a wide range of spatial scales considering a variety of seismic phases (wave types). The topic is vast and well developed, with several texts and reviews available (e.g., Nolet, 2012; Rawlinson et al., 2014). This section does not attempt to provide a general review of the subject. Rather, this section discusses the extension of surface-wave dispersion measurements to constrain laterally heterogeneous structure via surface-wave travel-time tomography, which is relevant to subsequent chapters. Here, the classical tomography problem is reviewed.

Because seismic surface waves propagate along the surface of Earth (rather than propagating in depth), the associated tomographic problem can be (and often is) formulated in two steps, involving two different inverse problems. The first step is the actual “tomography”, whereby 2D maps of phase or group velocity (depending on the dispersion measurement methods) at various frequencies are constructed on a cellular grid, based on the spatial distribution of surface-wave travel paths. The second step is to form dispersion curves at specific locations throughout the study area (by combining phase/group-velocity maps at various frequencies) and perform a series of 1D inversions for structure directly beneath these locations. These 1D inversion results are then interpolated to form a pseudo-3D model. This second inversion step (estimating 1D structure) is typically solved using any of the methods described in previous sections of this chapter. This section focuses on the tomographic aspect of the problem (i.e., the first inversion step).

Consider an array of seismometers that provides a set of surface-wave travel-time measurements between N station pairs (assuming cross-correlation-based dispersion measurements). Using the seismic-ray assumption, the surface-wave travel time $t_i(f)$, at frequency f , between the i -th station pair is given by the integral of phase or group slowness $s(l, f)$ over the ray path l_i along which the wave propagates,

$$t_i(f) = \int_{l_i} s(l', f) dl', \quad i = 1, \dots, N. \quad (1.17)$$

Assuming a 2D model of slowness discretized into M cells of uniform slowness, equa-

tion (1.17) can be expressed as

$$t_i(f) = \sum_{j=1}^M s_j(f) \Delta l_{ij}, \quad i = 1, \dots, N, \quad (1.18)$$

where $s_j(f)$ is the slowness of the j -th cell and Δl_{ij} is the path length of the i -th ray through this cell ($\Delta l_{ij} = 0$ for cells that are not along the i -th ray path). This is the forward problem in classical tomography. For a homogeneous medium, a seismic ray follows the great-circle path (GCP) connecting the two stations. However, for a heterogeneous medium, the ray path is itself a function of the spatial distribution of slowness, leading to a nonlinear problem.

Tomography is often linearized by assuming a GCP geometry or by calculating ray paths for a starting (reference) model and assuming stationarity. Within a linearized formulation, it is clear that the partial derivatives of the data (travel times) with respect to the model parameters (cell slownesses) that form the sensitivity matrix $\mathbf{A}$ (cf. Eqs. (1.1) and (1.18)) are simply the ray-path length segments (i.e., $A_{ij} = \Delta l_{ij}$). Linear inverse methods (as discussed in Section 1.2) can then be applied to solve for the slowness values in the discretized 2D map. Regularization schemes defined to represent first- or second-order spatial derivative operators applied to the parameters to minimize model gradients or roughness are often employed in tomography to overcome the ill-posedness of the matrix inversion, and to produce simple, minimum-structure models (Constable et al., 1987; Aster et al., 2018). As discussed in Section 1.2, after an initial linearized inversion, the problem can then be linearized about the resulting model (updating ray path geometry accordingly), and the inversion procedure repeated iteratively until convergence. In reality, seismic waves exhibit off-path sensitivity where their true sensitivity is to geophysical properties in a volume about the ray path. For surface waves, such volumes can be approximated in 2D by a zone (ellipse) around the ray path (e.g. Yoshizawa & Kennett, 2002; Lebedev & Van Der Hilst, 2008). However, for tomography problems at certain scales, the use of such sensitivity zones often produce comparable results to the ray assumption (e.g., Darbyshire et al., 2018).

The extent/significance of ray path deflections (from straight paths or GCPs) depends

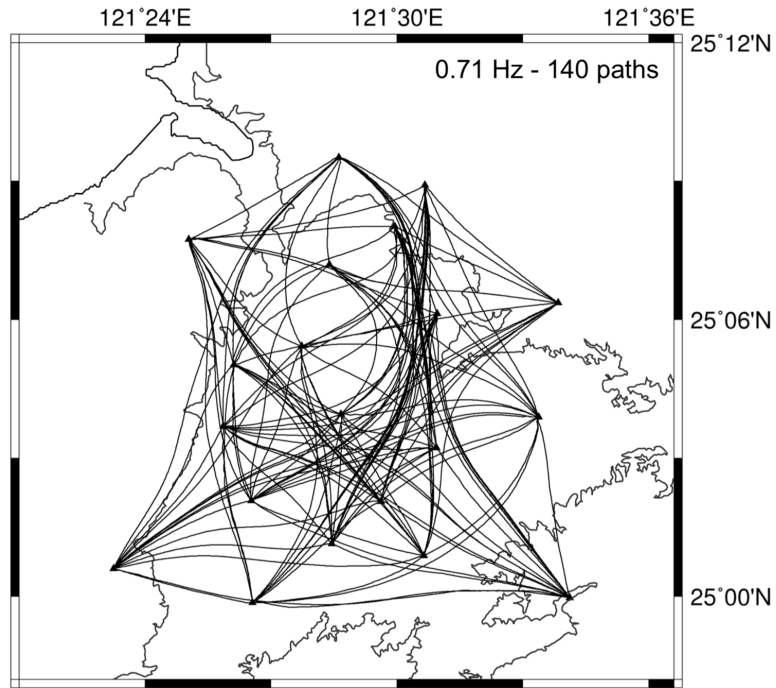


Figure 1.5: Example of surface-wave ray paths (at a frequency of 0.71 Hz) estimated for a model of the Taipei Basin of Taiwan. Lateral heterogeneity in the seismic velocity structure of the basin causes significant deflections from great-circle ray paths (modified from Fang et al., 2015).

on propagation distance and the magnitude of lateral variability in velocity structure. In the shallow subsurface, where lateral variability in structure can be significant (e.g., due to sedimentary basins), actual ray paths can deviate significantly from straight-line or GCP assumptions (e.g., fig. 1.5). These assumptions can lead to theory errors that bias inversion results. However, at larger scales considering longer-period data, surface waves are sensitive to structure at greater depths where lateral variability in structure is generally small in magnitude. If Earth structure along the surface wave propagation path does not include significant heterogeneity (e.g., ocean-continent margins), Ritzwoller & Levshin (1998) show that deviations from GCP do not significantly affect tomographic inversion results for path lengths less than approximately 5000 km for periods greater than approximately 30 s. Similarly, Bensen et al. (2008) suggest that these effects are negligible in surface wave tomography at periods greater than 20 s and path lengths less than 1000 km. Alternatively, some studies use sophisticated wavefront tracking methods to accurately update ray path geometry over multiple linearized inversion steps (e.g., Fang et al., 2015).

There are associated advantages and disadvantages to performing surface-wave tomography in two independent steps. An advantage of the two-step inversion approach to estimating 3D structure is that it decouples the 1D depth sensitivity of the dispersion data from the 2D ray path sensitivity (i.e., the tomographic problem). Furthermore, tomography is typically performed individually for each frequency, reducing computation cost and complexity. However, 2D phase or group velocity maps at closely-spaced frequencies are generally expected to show similar structure. Some studies have inverted all frequencies simultaneously and applied an additional regularization term for inter-frequency smoothness to impose consistent structure between 2D maps at adjacent frequencies (e.g. Hannemann et al., 2014). However, this requires the numerical inversion of a significantly larger matrix. Once the 2D maps are estimated, a series of computationally inexpensive inversions can be performed to estimate 1D structure beneath each point in the study area.

Direct inversion for 3D models from surface-wave dispersion for specific paths has also been considered (e.g., Fang et al., 2015; Inzunza et al., 2019). This has the advantage of skipping the intermediate step of estimating 2D phase or group velocity maps. Fur-

thermore, data errors are propagated directly into the final inversion result, and avoid (potential) distortion by the intermediate inversion step. However, it is more challenging to define the sensitivity of the data to model parameters in the discretized 3D volume.

Finally, the model parameterization can also have a significant effect on tomographic inverse problems. Equation (1.18) formulates the tomographic problem with model parameters representing the slowness of discrete grid cells, but other approaches are possible. For example, Fang et al. (2015) transform the model from a regular grid into a sparse wavelet-basis domain, where the parameters are wavelet coefficients. The advantage of this approach is that an L_1 damping regularization applied to the wavelet coefficients implicitly creates a minimum-structure model that (in theory) only allows detailed structure in the model where required by the data, as opposed to having two explicit regularization terms that apply smoothing and damping over the entire 2D model (in physical/geographic space). This is particularly attractive in tomographic inversion, where uneven path coverage inherently leads to a multi-scale problem (i.e., some regions of the model are better resolved than others). Consequently, Bayesian trans-D inversion has also recently been applied to tomography for large-scale problems (e.g., Bodin & Sambridge, 2009; Bodin et al., 2012), providing an adaptive multi-scale parameterization that is estimated as part of the inversion. This is the topic of the next chapter in this thesis.

Chapter 2

Azimuthal anisotropy in Bayesian surface-wave tomography: application to northern Cascadia and Haida Gwaii, British Columbia

2.1 Introduction

Seismic anisotropy is the dependence of seismic wave speeds on the direction of propagation, and manifests the coherent alignment of rock fabrics via deformation in the crust and upper mantle. As such, measuring seismic anisotropy is valuable for understanding Earth's evolution, dynamics, and structure. Deformation can result in lattice preferred orientation (sometimes called crystallographic orientation) of intrinsically anisotropic minerals, and shape preferred orientation of discontinuous and heterogeneous Earth structure (Montagner & Guillot, 2002). Seismically-observable anisotropy in the crust is generally attributed to rock micro-fractures, the alignment of anisotropic minerals with major foliations, and layered sequences of sedimentary rocks (Babuska & Cara, 1991). Anisotropy in the upper mantle may be explained by the alignment of anisotropic olivine minerals (Montagner

& Anderson, 1989; Babuska & Cara, 1991). When subjected to a simple shear force, olivine aggregates develop a lattice preferred orientation via dislocation creep, with a fast-axis direction parallel to the direction of shear (Karato et al., 2008). For these reasons, seismic anisotropy reflects the record of deformation within the Earth and provides useful constraint on the dynamics of the crust and mantle.

Seismic anisotropy can be estimated from single station techniques to much larger spatial scales. Observations of shear wave splitting from refracted teleseismic body-wave (e.g., SKS) phases offer a relatively simple method for constraining 1D anisotropy beneath a seismograph station (e.g., Vinnik et al., 1989, 1992; Savage, 1999). Shear-wave splitting is also commonly considered using local earthquake data in order to constrain crustal anisotropy (e.g., Balfour et al., 2012). However, since these measurements represent the depth-integrated effect of anisotropy beneath the station, they are typically interpreted as the result of a single layer of anisotropy, with a trade-off between the thickness of this layer and the anisotropic intensity. The effects of multiple anisotropic layers beneath a station are challenging to interpret. Single-station azimuthal dependence has also been observed in teleseismic receiver function studies and has been attributed to anisotropic layering structure (e.g., Savage, 1998; Audet, 2015). For such methods, the lateral resolution of anisotropy depends on the spatial distribution of seismograph stations that record body waves.

Azimuthal anisotropy has been constrained using surface waves via tomographic inversions at regional (e.g., Darbyshire & Lebedev, 2009; Darbyshire et al., 2018; McLellan et al., 2018) and global (e.g., Schaeffer et al., 2016) scales. Such tomographic studies constrain both the lateral and depth variations in anisotropy, since different periods of surface waves are sensitive to structure at different depth ranges in the Earth. To date, these studies utilize linearized and regularized inversions, where the resulting model depends on the choice and level of regularization. Such methods must simultaneously solve for the distribution of isotropic seismic velocities, in addition to seismic anisotropy, with the caveat that anisotropic structures are not as well resolved as isotropic features. As a result, anisotropy within these models must be regularized more strongly. Such choices make quantification of uncertainty of the various isotropic and anisotropic features, and

their interpretation, difficult. Furthermore, regularization of anisotropy within a model is challenging. As discussed in Schaeffer et al. (2016), this is particularly important in polar regions where standard smoothing and gradient damping methods may produce model artifacts (where anisotropic structures appear to circle the poles or point towards them), due to the significant changes in ray azimuth direction (relative to north) along travel paths at these latitudes.

Understanding and quantifying the uncertainties in estimated anisotropy allows for more reliable interpretation of these structures. In recent years, Bayesian methods have become popular for solving seismic tomography problems in studies considering surface wave (e.g., Bodin & Sambridge, 2009; Bodin et al., 2012; Zulfakriza et al., 2014; Galetti et al., 2017) and body wave (e.g., Piana Agostinetti et al., 2015) travel-time data. Bayesian inversion allows for rigorous uncertainty quantification in geophysical inverse problems. Bayesian inversion is a probabilistic approach in which model parameters (e.g. the distribution of isotropic seismic velocities, and anisotropy) along with the observed data are treated as random variables. These model parameters are constrained by data as well as prior information, which may be informative (thereby providing additional constraint) or uninformative (allowing the solution to be entirely constrained by data information content). The posterior probability density (PPD) of the model parameters is the solution to the inverse problem. Typically, the PPD is numerically estimated using Markov-chain Monte Carlo (MCMC) methods (Mosegaard & Tarantola, 1995; Brooks et al., 2011). A preferred model and associated uncertainty can be determined from features of the PPD.

Studies that have considered Bayesian seismic tomography have typically applied an adaptive model parameterization that allows the spatial complexity of the model to be determined by the data information content. In such a framework, the number of model parameters is treated as an unknown within the inversion. These methods are commonly referred to as trans-dimensional, as the PPD spans multiple possible model dimensions (i.e., a different number of model parameters). In this case, the generalized reversible-jump Markov-chain Monte Carlo (rjMCMC) algorithm is used to numerically estimate the PPD (Green, 1995, 2003). Trans-dimensional inversion allows for increased model complexity where it is required by the data, and estimates of model uncertainty include

the uncertainty in the model complexity/parameterization. Consequently, this method avoids subjective regularization choices that can preclude meaningful uncertainty analysis, as well as avoiding user-based choices of a particular model complexity.

This work aims to incorporate azimuthal anisotropy in a trans-dimensional Bayesian tomographic inversion of surface wave dispersion data in order to avoid subjective regularization procedures, as well as to quantify the uncertainty in the orientation of seismically-detectable rock fabrics. Our work is motivated by the modelling challenges posed by data collected along the west coast of Canada. The complex interactions between the North American (NA), Pacific (PA), and Juan de Fuca (JdF) plates along the coast of British Columbia produces significant seismicity, as well as a poorly understood transition between the Cascadia subduction zone and the transpressive Haida Gwaii region (fig. 2.1). Resolving seismic velocity structure (isotropic and anisotropic) in this region using standard single-station passive seismological techniques is complicated by the sparsity of station coverage to the west of the plate boundaries, with a large gap between Vancouver Island and Haida Gwaii. Body-wave derived velocity models (i.e., tomography, receiver functions) suffer from insufficient lateral sampling of the 3D volume. However, surface waves can efficiently sample volumes where station coverage is sparse. We demonstrate our novel inversion approach through synthetic experiments, and apply this method to resolve anisotropic surface wave velocities across northern Cascadia and Haida Gwaii. We consider surface wave dispersion measurements obtained from recordings of regional earthquakes over several decades. These measurements allow for robust estimation of anisotropic Rayleigh-wave group velocity maps at periods between 15 and 50 s. Our results demonstrate the value of this novel inversion technique and provide an initial step towards resolving a full, 3D, seismic velocity model in future work, with the incorporation of additional datasets. Furthermore, this work provides broad structural context for future seismic station deployments (both onshore and offshore) to study the tectonic structure and dynamics of the region, with the aim of improving our understanding of seismic and tsunami hazards.

2.2 Tectonic Setting

The west coast of British Columbia is the most seismically active region in Canada, with a history of large earthquakes and tsunamis and the potential for hosting future large events. The region comprises a number of active tectonic boundaries including the northernmost extent of the Cascadia subduction zone, which has historically produced large megathrust earthquakes and associated tsunamis (Nelson et al., 1995; Satake et al., 1996, 2003). North of Cascadia, the termination of subduction involves a number of triple junctions that are poorly understood (fig. 2.1). This produces significant seismicity and a complex system of transform faulting (Hutchinson et al., 2019), as well as the breakup and possible capture of the Explorer (Ex) microplate (formerly part of the JdF plate) by the NA plate (Braunmiller & Nábělek, 2002; Audet et al., 2008; Savard et al., 2020). Near Haida Gwaii, along the northernmost coast of British Columbia, the transpressive interaction between the PA and the NA plates produces significant seismicity, including Canada’s two largest instrumentally-recorded earthquakes (Bostwick, 1984; Cassidy et al., 2014). Margin-parallel motion between the PA and NA plates is predominantly accommodated by the strike-slip Queen Charlotte Fault (QCF). North of Haida Gwaii, the QCF continues into Alaska as the Fairweather fault. The orientation of the QCF with respect to the direction of plate motion requires that a component of convergence of approximately 15 mm/yr exists across the margin (Kreemer et al., 2003). However, the nature of how plate convergence is accommodated across Haida Gwaii is unsettled.

A significant body of geophysical evidence including seismic velocity structure, as well as thermal and gravity modelling suggests convergence is accommodated by subduction, or underthrusting, of PA beneath NA (Hyndman, 2015; Gosselin et al., 2015, and references therein). On 28 October 2012, a moment-magnitude 7.8 thrust earthquake occurred off the west coast of Haida Gwaii producing a large local tsunami (Leonard & Bednarski, 2014; Cassidy et al., 2014). The 2012 earthquake occurred on a shallow northeast-dipping thrust fault, confirming the presence of some form of convergent strain accumulation offshore of Haida Gwaii. Plate kinematic models place the leading edge of the subducting PA plate east of Haida Gwaii (fig. 2.1), assuming subduction initiation between 3.9 and 8 Ma, and

a location of convergence initiation between northern Vancouver Island and Haida Gwaii (Smith et al., 2003). Despite the geophysical evidence and the occurrence of the 2012 thrust earthquake, the model of oblique subduction beneath Haida Gwaii does not explain all observations in the region. Microseismicity along the west coast of Haida Gwaii occurs predominantly in the vicinity of the QCF (Bérubé et al., 1989; Bird et al., 1997). In the northern Haida Gwaii region, earthquake focal mechanisms are predominantly strike slip and consistent with the QCF (Ristau et al., 2007). In the southern Haida Gwaii region, most earthquakes have thrust mechanisms, though they do not reside on the same fault as that of the 2012 event. There is no clear evidence of Wadati-Benioff seismicity in the region, and all of the observed aftershocks from the 2012 event were located offshore beneath the Queen Charlotte Terrace, and none beneath Haida Gwaii (Cassidy et al., 2014). Furthermore, recent submarine geomorphological mapping suggests plate convergence rates across Haida Gwaii may be significantly less than previously reported, and may be insufficient for subduction initiation (Brothers et al., 2020). Understanding the extent (or existence) of subduction has implications for the tectonic history, structure, and dynamics of the region including the size of major seismogenic zones along the margin. Constraining the transition from pure subduction beneath southern Vancouver Island to transpressive deformation near Haida Gwaii is an important objective for understanding the hazards posed by large earthquakes along the west coast of Canada.

2.3 Surface-wave dispersion data

Regional earthquakes typically release surface wave energy over period ranges that are sensitive to crustal and uppermost mantle depths, making the dispersion measurements obtained from regional earthquakes useful for studies of tectonic structure (fig. 2.2). In general, group velocity dispersion is simpler to measure than phase velocity dispersion, as no information on the earthquake source-time function is required. Since the source group time (the earthquake effect on surface wave travel time) is effectively negligible for earthquakes shallower than approximately 25 kilometres, for periods less than approximately 75 s (Levshin et al., 1999), a wide range of earthquakes can be considered (including those without

centroid moment tensor solutions). In contrast, accurate information of the source phase is required for single-station phase velocity measurements from earthquakes. Phase velocity (and group velocity) dispersion measurements can also be obtained via multi-station interferometric approaches (e.g., Bensen et al., 2007), where source phase information is factored out of the processing procedure. These techniques require concurrently operating seismic stations. Unfortunately, many of the broadband seismic stations in western Canada considered in this work were temporary deployments that predominantly did not run simultaneously (fig. 2.1). Measurement errors for group velocity dispersion are typically larger than for phase velocity measurements (e.g., Bensen et al., 2008). The interpretation of group velocity maps is also more complex (due to the negative depth sensitivity even in fundamental-mode sensitivity kernels). Nonetheless, group velocity dispersion is a valuable measurement that leads to improved data path coverage in tomographic inversions, which is particularly valuable in this study as station coverage along the west coast of Canada is sparse (fig. 2.3).

We used recordings of regional earthquakes from over 300 broadband seismic stations from both permanent and temporary campaign networks with varying deployment times. For each station, we used events with magnitude greater than 4.5 between 2000 and 2019. For some stations with long deployment times (i.e., several decades), we analyzed over 400 events. The majority of the earthquakes considered here are associated with the JdF ridge and QCF (fig. 2.1b). We extract 30-minute vertical-component seismograms for each event at each available station. Next, we remove the instrument response from each seismogram and apply a zero-phase bandpass filter for periods 5-100 s to suppress microseismic noise and long-period instabilities. Assuming the vertical-component recording of the surface-wave coda is dominated by the fundamental-mode Rayleigh wave, basic group velocity dispersion measurements can be obtained from the maximum amplitude of the wave-train envelope as a function of time and period. For a known source-station distance, this time-dependence can be easily converted to (group) velocity. In this study, we measure Rayleigh wave group velocity from regional earthquakes using the multiple-filter analysis technique implemented by Herrmann (2013), which allows for manual selection of the velocity-period envelope range for each seismogram (fig. A.1). This is important as each earthquake is

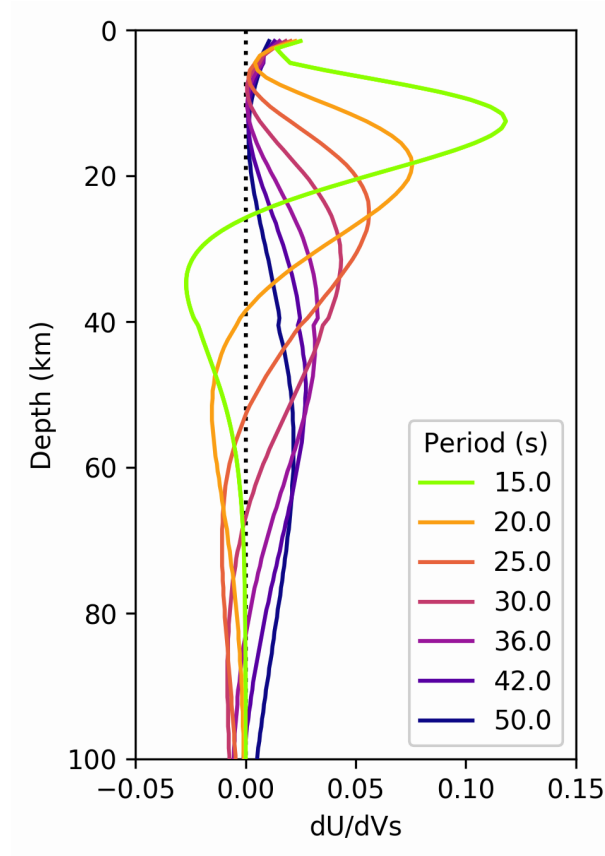


Figure 2.2: Fundamental-mode Rayleigh wave group-velocity depth sensitivity for periods considered in this work, calculated using a simple 40 kilometer thick crustal model (available in the supplementary material).

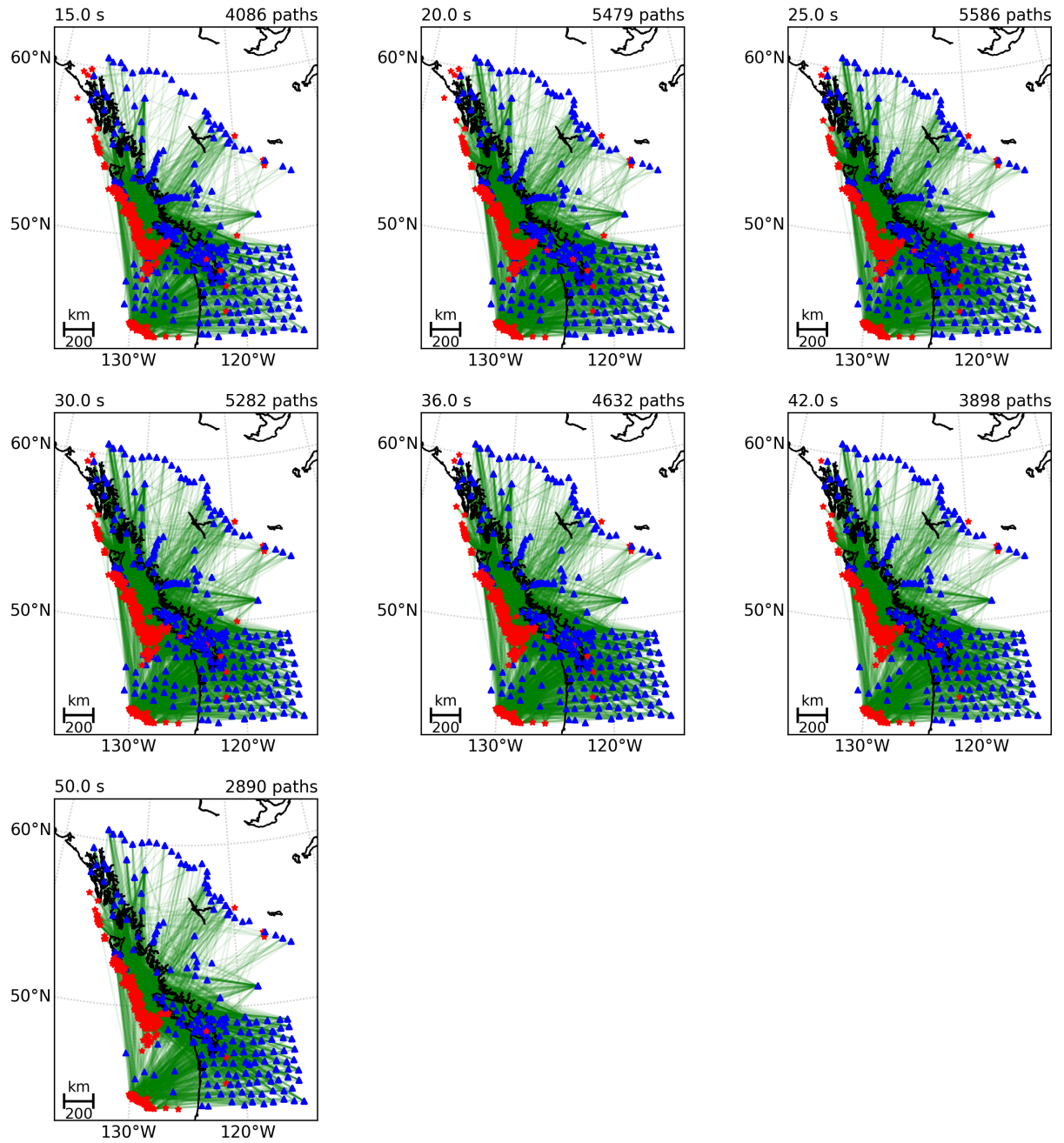


Figure 2.3: Surface wave path coverage at periods considered in this work. Earthquake locations and recording stations are shown by red stars and blue triangles, respectively. The total number of data considered in the inversion at each period is shown above each map.

unique, and the period range of surface wave energy recorded at each station can vary. Furthermore, significant quality control is required when considering dispersion data from earthquake recordings, which can be accomplished via visual inspection and manual selection. In total, over 30,000 seismograms were analyzed, of which over 8000 yielded usable dispersion curves. Although data processing involves manual selection of dispersion curves in group velocity space, the data are considered as travel-time measurements in the tomographic inversions. This is particularly important as surface wave travel paths are updated within the inversion (as discussed in the next section), and the estimated group velocities from the manual selection assumes surface waves travel along minimum distance great-circle paths.

A common practice in tomographic studies that consider regional earthquakes is to bin and average measurements from earthquakes that are in close proximity to each other. This technique is used to reduce the number of data that must be considered in linearized tomographic inversions as well as to reduce the effect of path bias, where the data are dominated by paths over a particular region. However, such binning procedures make characterization of data errors more challenging, and can lead to incorrect data weighting if not accounted for. In an effort to remove outliers from the dispersion dataset without altering the measurements themselves, we implement a modified binning strategy. We examine the measurements within each source bin (with a radius of 100 kilometres) recorded at each station individually and discard outlier data using a median absolute deviation threshold of 2.5. However, we do not average the data within each source bin but rather retain each individual measurement. Regardless of this procedure, the final dataset may still contain outliers if some source bins do not contain a sufficient number of events (less than 3). Additional treatment of outliers is considered within the inversion methodology (discussed in the next section). We also discard data that have a propagation path length shorter than five wavelengths. The final number of paths considered at each period after this procedure is shown in figure 2.3. We made group velocity dispersion measurements over the period range between 10 and 80 seconds. However, only periods between 15 and 50 s produced adequate path coverage for tomographic inversions with the aim of resolving anisotropy. Specifically, paths over a wide range of directions are required in order

to resolve azimuthal anisotropy in tomographic inversions. In this work, azimuthal path coverage is best over the coastal region of British Columbia including Vancouver Island, Haida Gwaii, and the continental shelf (figs. A.2 and A.3).

2.4 Inversion methodology

We consider the contributions from isotropic and azimuthally anisotropic components in the inversion of surface-wave dispersion data. The parameters that describe the group velocity maps are adaptively defined over a distribution of nodes, where the number and location of the nodes is determined as part of the inversion. Furthermore, we treat the errors in the dispersion data as unknown within the inversion. We solve for all model parameters (including hyper-parameters describing the model parametrization and data error distribution) within a Bayesian framework in order to reduce overall model dimensionality (compared to linearized approaches) and avoid subjective regularization procedures. Each one of these steps is detailed in the sections below. Bayesian inversion and associated numerical methods are discussed in Tarantola (2005) and Brooks et al. (2011). The specific application of Bayesian inversion to surface-wave tomography is discussed in detail in Bodin & Sambridge (2009) and Bodin et al. (2012). We note that the inclusion of azimuthal anisotropy within Bayesian surface-wave tomography represents a novel extension of existing methods.

2.4.1 Azimuthal anisotropy

Seismic anisotropy, in general terms, is the dependence of seismic wave speeds on the direction of propagation. For simplification, studies of seismic anisotropy typically assume hexagonal symmetry and define two types of anisotropy depending on the orientation of the axis of symmetry. Radial anisotropy assumes a vertical axis of symmetry within the Earth, whereas azimuthal anisotropy assumes a horizontal axis of symmetry. Consequently, such definitions of anisotropy can be described by a magnitude and direction. This work

considers azimuthal anisotropy, and any reference to seismic anisotropy henceforth is with respect to this type. Under the assumption of hexagonal symmetry with a horizontal axis of symmetry, Smith & Dahlen (1973) approximate the velocity of a surface wave propagating through an anisotropic medium at period T and azimuth ψ to first order by

$$C(T, \psi) = C_0(T) + A_1(T)\cos(2\psi) + B_1(T)\sin(2\psi) + A_2(T)\cos(4\psi) + B_2(T)\sin(4\psi), \quad (2.1)$$

where $C_0(T)$ is the isotropic velocity of the medium. The 2ψ and 4ψ terms represent the velocity variations with π and $\pi/2$ periodicity, respectively. Though equation (2.1) was originally defined for surface wave phase velocities, it may be applied to model group velocity in a similar way (e.g., Pawlak et al., 2012; Darbyshire et al., 2018; Yeck et al., 2017). It is recognized that Rayleigh wave data have low sensitivity to anisotropy with $\pi/2$ periodicity, particularly in the mantle where anisotropic signatures are expected to be simpler and lower in magnitude. In the crust, complex interactions of multiple fabrics could produce seismically-observable patterns with $\pi/2$ periodicity. However, for computational simplification, we only consider anisotropy with π periodicity and neglect the 4ψ terms in equation (2.1). The spatial distributions of three quantities ($C_0(T)$, $A_1(T)$, and $B_1(T)$) are considered in the inversion.

2.4.2 Model parameterization

In any inversion, a choice must be made in order to define (or parameterize) the model to be solved for. Using a grid with large spacing can result in under-fitting the data, where significant resolvable structure is ignored. Conversely, too small of a grid spacing can result in over-fitting data, where non-physical signatures in the data (i.e., noise) are fitted, causing spurious model structures. Furthermore, small grid spacing can result in significant additional computational expense. Irregular grids can be applied in tomographic inversions as a geometric approach to impose uniform resolution of model parameters or to estimate model resolution. For example, Debayle & Sambridge (2004) use Voronoi diagrams to qualitatively estimate the length scale of structures that can be resolved in global tomographic inversion by optimizing the Voronoi parameterization to satisfy an

appropriate quality criterion. Such a parameterization is defined by a set of nodes with assigned coordinates and values. The model value at any point in space is defined by the assigned value of the nearest node (i.e., nearest-neighbour interpolation), generating a set of discrete geometric cells (fig. 2.4). In this case, the size of the individual Voronoi cells is proportional to the resolution length scale for the model parameters. In Bayesian inversion, Bodin & Sambridge (2009) introduced an adaptive model parameterization to seismic tomography also using Voronoi cells. The advantage of this parameterization is that regions in space where the data require additional complexity will be modelled with additional nodes, and vice versa. In this way, the Voronoi cell parameterization is a simple and intuitive means of spatial model adaptation based on data information content.

In practice, however, Voronoi cells possess several undesirable qualities for application to seismic tomography. Foremost, the discrete nature of nearest-neighbour interpolation leads to individual models that appear non-physical. Hawkins et al. (2019) recently introduced an alternative means of adaptive model parameterization in geophysical inversion. This approach is similar to Voronoi cells in that it is defined using a distribution of nodes in space. Instead of nearest-neighbour interpolation, a Delaunay triangulation is applied to the nodes and the model quantity (in our case, $C_0(T)$, $A_1(T)$, and $B_1(T)$) at any point in space is linearly interpolated using the velocity values assigned to the nodes of the confining triangle (fig. 2.4). The Delaunay triangulation avoids narrow triangles by maximizing the minimum angle of the triangulation. Linear interpolation within the triangles is easily calculated using Barycentric coordinates (Sambridge et al., 1995; Hawkins et al., 2019). Here we apply the Delaunay triangulation with linear interpolation, which naturally produces more realistic models than Voronoi cells and are preferable in tomographic inversion. Each node is assigned five parameters in total: isotropic group velocity ($C_0(T)$), two parameters describing anisotropy with π periodicity ($A_1(T)$, and $B_1(T)$), and two coordinate values. Similar to Bodin & Sambridge (2009) and Bodin et al. (2012) we solve for the number of nodes in the parameterization as part of the inversion, allowing the overall complexity of the model to be determined by the data information content.

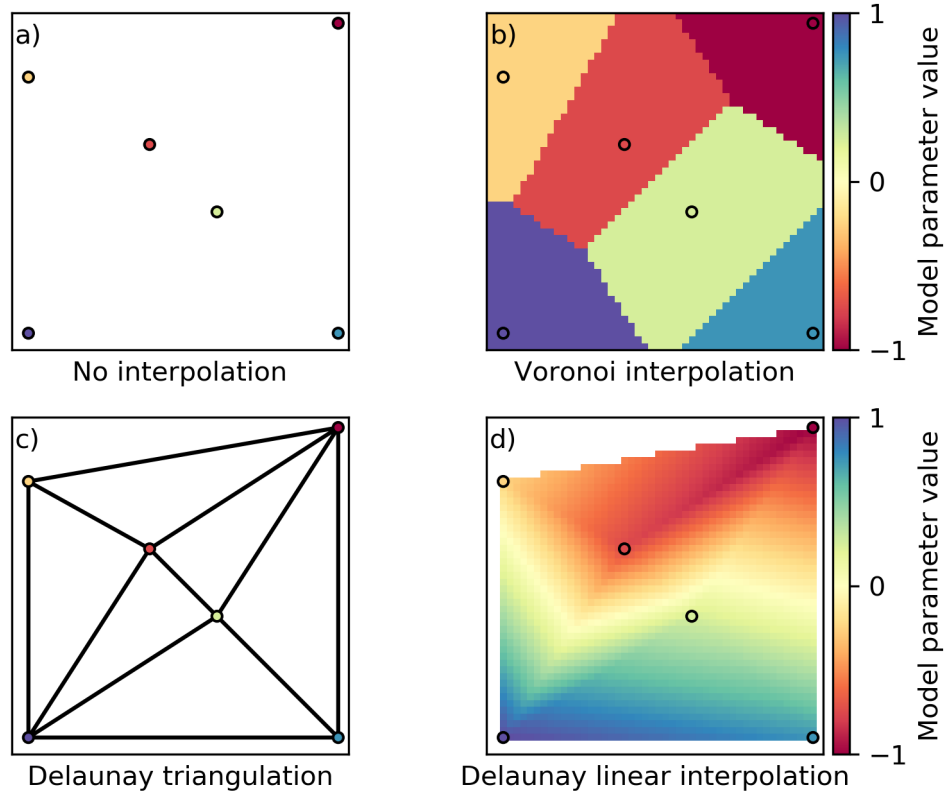


Figure 2.4: Examples of Voronoi and Delaunay interpolations. A distribution of nodes with randomly assigned parameter values is shown in a). The corresponding Voronoi interpolation is shown in b). The corresponding Delaunay triangulation over the nodes is shown in c). The linear interpolation using Barycentric coordinates within each triangle of the Delaunay triangulation is shown in d).

2.4.3 Bayesian inversion

Bayes' theorem can be written as

$$P(\mathbf{m}|\mathbf{d}, \mathcal{M}) = \frac{P(\mathbf{m}|\mathcal{M}) P(\mathbf{d}|\mathbf{m}, \mathcal{M})}{P(\mathbf{d}|\mathcal{M})} \quad (2.2)$$

where $\mathbf{d}$ is a random variable (of length N) representing the observed data (in this case, surface-wave travel times at a particular period of interest) and $\mathbf{m}$ is a random variable (of length M) representing model parameters that describe an anisotropic group-velocity map. Measured data are a fixed realization of $\mathbf{d}$ in practice. Bayes' theorem describes the relationship between conditional and joint probabilities of $\mathbf{d}$ and $\mathbf{m}$, where $\mathcal{M}$ is the specific model (e.g., the parameterization or model dimension of the Rayleigh-wave group velocity map). $P(\mathbf{m}|\mathcal{M})$ represents the prior knowledge of the model parameters before the inversion. $P(\mathbf{d}|\mathbf{m}, \mathcal{M})$ represents the likelihood of $\mathbf{m}$, or $\mathcal{L}(\mathbf{m})$ (for model $\mathcal{M}$, since $\mathbf{d}$ is fixed). $P(\mathbf{m}|\mathbf{d}, \mathcal{M})$ is the solution to the inverse problem (the PPD) since it is the probability of the model parameters given a specific model $\mathcal{M}$, prior information, and observed data. $P(\mathbf{d}|\mathcal{M})$ is a normalization term often referred to as Bayesian evidence.

As discussed above, we parameterize the anisotropic Rayleigh-wave group velocity map using Delaunay triangulation with linear interpolation across a set of nodes with associated properties (isotropic velocity, and two anisotropic parameters). The number of nodes in the model is unknown within the inversion procedure, and the level of model complexity is thereby defined by data information content. Consequently, the number of model parameters required to describe the group velocity map is not fixed. Such inversion methods are commonly referred to as trans-dimensional. The number of nodes in the model is a hyper-parameter within the inversion that indexes over a countable set of possible model parameterizations (i.e., model $\mathcal{M}$ in eq. 2.2). In this case, the PPD represents the complete solution to the inverse problem over this set of models and model dimensions. Typically, an analytic solution for equation (2.2) does not exist and numerical methods are used to approximate the solution. Specifically, we use the reversible-jump Markov-chain Monte Carlo (rjMCMC) algorithm to draw a series of dependent (on the previous sample) and asymptotically-unbiased samples from the PPD (Green, 1995, 2003). This algorithm

is a generalization of the Metropolis-Hastings sampler, which is a standard Markov-chain Monte Carlo (McMC) technique (Metropolis et al., 1953; Hastings, 1970). The rjMcMC algorithm involves randomly perturbing a current set of model parameters $\mathbf{m}$ (i.e., the number/position of nodes, and their associated properties) to a new state $\mathbf{m}'$, and accepting this perturbation with probability

$$A(\mathbf{m}'|\mathbf{m}) = \min \left[1, \frac{Q(\mathbf{m}|\mathbf{m}')}{Q(\mathbf{m}'|\mathbf{m})} \frac{P(\mathbf{m}')}{P(\mathbf{m})} \frac{\mathcal{L}(\mathbf{m}')}{\mathcal{L}(\mathbf{m})} |\mathbf{J}| \right], \quad (2.3)$$

where $Q(\mathbf{m}'|\mathbf{m})$ is the probability of proposing the new state from the current state, and $|\mathbf{J}|$ is the determinant of the Jacobian matrix for the transition between different model states. For model perturbations that do not change the dimension of the model (i.e., moving node positions and changing isotropic or anisotropic parameters), the determinant of the Jacobian is unity. Careful choices of prior and proposal densities also result in the determinant of the Jacobian being unity for perturbations that change model dimensionality (i.e., adding or removing nodes) (Bodin & Sambridge, 2009; Dosso et al., 2014). Specifically, bounded uniform prior distributions are used such that $P(\mathbf{m}) = P(\mathbf{m}')$. Prior bound widths are listed in appendix A, and are selected in order to constrain model parameters to realistic values while still allowing the solution to be predominantly determined by the data. When perturbing a model parameter, a new value is drawn from a Gaussian distribution centered on the current value. In this way, proposal densities are symmetric ($Q(\mathbf{m}'|\mathbf{m}) = Q(\mathbf{m}|\mathbf{m}')$). In the case of adding a new node to the model, the values of all the parameters associated to this node are drawn from the prior. With the chosen prior and proposal densities, it can be shown that equation (2.3) simplifies to a ratio of likelihoods for all types of model perturbations (Bodin & Sambridge, 2009; Dosso et al., 2014). Given a sufficiently large number of samples, the rjMcMC algorithm provides a good approximation to the PPD.

Within the inversion, the likelihood function depends on the assumed distribution of the data errors, which is typically unknown. The Central Limit Theorem suggests that data errors may follow a multi-variate Gaussian distribution, and a likelihood function defined by this distribution is valid. However, in the presence of data outliers, a Gaussian likelihood may be biased. A Laplace distribution is similar to the Gaussian distribution, but with

higher probability at greater distance from the mean (i.e., larger tails). Alternatively, the Gaussian distribution can be viewed as the likelihood associated with the L_2 misfit norm, whereas the Laplace distribution is the likelihood associated with the L_1 misfit norm. Consequently, the Laplace distribution is more tolerant of data outliers. As discussed in the previous section, earthquake data from regions with limited seismicity may elude the source binning procedure and introduce outliers into the dispersion dataset. For this reason, we apply the Laplace likelihood function in this work, which is given by

$$\mathcal{L}(\mathbf{m}) = \frac{1}{\prod_{i=1}^N 2s_i} \exp \left[- \sum_{i=1}^N \frac{|r_i|}{s_i} \right], \quad (2.4)$$

where

$$r_i = d_i^{obs} - d_i(\mathbf{m}) \quad (2.5)$$

are the data residuals, and s_i is the scale parameter that controls the diversity (i.e., width) of the Laplace distribution for each datum (similar to the standard deviation of a Gaussian distribution). The multiple-filter analysis technique used in this work does not provide useful estimates of the errors on the measured dispersion data. Furthermore, qualitatively assigning relative errors within the dispersion dataset based on uncertainties in earthquake source properties (i.e., location and timing) is not possible as this information is only available for a small number of events in our dataset. For these reasons, we implement a parameterized error model within the inversion framework. Similar to Bodin et al. (2012), we consider the errors on surface wave travel times as unknown and model them using a linear dependence with source-to-station distance Δ_i (for path i) by

$$s_i = a\Delta_i + b. \quad (2.6)$$

In this way, each travel time is assigned a different error according to its associated propagation path length. Only two additional hyper-parameters (a and b) are considered within the inversion.

2.4.4 Practical considerations

One of the great challenges with rjMcMC is inefficient sampling over the parameter space, particularly between spaces of different dimensions (in this case, adding or removing model nodes). We attempt to address this issue here by using parallel interacting chains within rjMCMC sampling, often called parallel tempering (Dosso et al., 2012; Sambridge, 2014). In this method, the acceptance criterion, equation (2.3), in each chain is relaxed by raising the likelihood to powers $1/T$ (where T is often called the sampling temperature and is greater than unity). Chains with higher temperatures have greater probability of accepting models that fit the data poorly (low-likelihood), preventing them from becoming trapped in local minima. Probabilistic swapping between the chains allows chains with lower temperatures to efficiently search all regions of the model parameter space. The temperature values for each chain are chosen such that approximately 20 percent of proposed chain swaps are accepted. This is unique to every problem, and is chosen via trial-and-error. Setting these parameters is subjective. However, in theory, the ultimate choice of these parameters has no effect on the overall inversion solution but only on the efficiency of the algorithm. It is computationally cheap to attempt a swap between chains since data prediction (i.e., forward modelling) is not required. Only chains with temperature set to unity provide unbiased sampling of the PPD. Typically one chain is set at $T = 1$ and only the samples drawn from this chain are kept in the final analysis. Parallel tempering is a robust sampling technique despite the additional computation burden of running multiple chains. For such a large inverse problem with so many model parameters, employing schemes like parallel tempering is valuable for avoiding wasted computation time. This is not only useful for increasing the acceptance rate of adding or removing model nodes to the model, but also allows the sampler to act as a highly effective global optimizer (Sambridge, 2014). This greatly reduces the number of initial burn-in samples (dependent on the random initial starting model) required before the algorithm locates the high-likelihood part of the model space.

Previous works that have considered trans-dimensional Bayesian tomography have implemented an iteratively nonlinear inversion scheme, where surface-wave travel paths are

updated after each inversion procedure in order to provide more accurate forward modelling for the following inversion iteration (Bodin & Sambridge, 2009; Bodin et al., 2012). More recently, Bayesian tomography has been successfully implemented in a fully non-linear inversion where new surface wave travel paths are computed for each new model perturbation in the rjMcMC algorithm, at great computational expense (Galetti et al., 2017). Since the data considered in this work are measured from recordings of regional earthquakes (as opposed to ambient noise cross-correlations as in other studies) there are significant uncertainties in the timing and locations of sources that are likely to have a greater effect on the data than the accuracy of the surface wave travel paths (i.e., forward model). Ritzwoller & Levshin (1998) suggest that, other than in extreme cases, off-great circle path propagation can be ignored at periods above approximately 30 s for paths with distances less than approximately 5000 kilometres. Similarly, Bensen et al. (2008) suggest that these effects are negligible in surface wave tomography at periods greater than 20 s and path lengths less than 1000 kilometres, which is generally applicable to our group velocity dataset over northern Cascadia and Haida Gwaii. It is worth noting that previous studies on trans-dimensional Bayesian surface-wave tomography considered much shorter period dispersion data as well as longer propagation paths. However, our study region includes an oceanic-continental transition with potentially large lateral gradients in velocity structure that may produce significant deviations to great-circle ray paths, despite the long periods and short path lengths considered. For this reason (and for the sake of generality of our inversion methodology), we adopt the iteratively nonlinear procedure of Bodin & Sambridge (2009) as an optimal balance between accurate forward modelling and computational efficiency. We assume that surface waves travel along great circle paths during the initial iteration. For subsequent inversion iterations, the ray paths are calculated using the inversion solution from the previous iteration. This process is repeated four times to produce the final models.

A consequence of the iteratively nonlinear inversion scheme is that data error hyperparameters within the inversion (i.e., a and b values) will be estimated to be larger during initial inversion iterations due to larger theory (forward modelling) errors (Bodin et al., 2012). For all inversions in this work, we perform four iterations of ray path updating

using a computationally efficient, pseudo ray-bending procedure for piecewise travel-time minimization (Um & Thurber, 1987). In order to efficiently perform forward modelling by integrating the travel time over the surface wave travel paths, we map the Delaunay triangulation with linear interpolation model (discussed above) to a regular, finely-sampled, integration grid. We attempted several inversion tests with varying scales of integration grid and found that results do not change significantly once the integration grid spacing is finer than 30 kilometres. At coarser spacing, the linear interpolations within Delaunay triangles become inaccurate. At finer spacing, the forward modelling becomes slower. For all subsequent inversions in this work, we use an integration grid with 30 kilometre spacing. Coincidentally, 30 kilometres is approximately the minimum wavelength of the data considered for northern Cascadia and Haida Gwaii, and represents a physical limitation on the scale of resolvable structure from the data.

The rjMcMC algorithm is used to draw samples from the PPD. Given a sufficient number of samples, a better approximation of the PPD is achieved. It is important to collect enough samples so that the inversion solution becomes stationary (with respect to properties of the PPD). Rigorous quantification of algorithm convergence is an emerging and challenging topic in the field of Bayesian inversion (Brooks et al., 2011). It is especially difficult to assess convergence for trans-dimensional inversions, where the PPD spans multiple model dimensions, and a specific parameter in one dimension does not necessarily represent the same structure in a different model dimension (Bodin & Sambridge, 2009). We use a similar approach to Bodin & Sambridge (2009) to assess converge of the algorithm by examining C_0 , A_1 , and B_1 at given locations in the group velocity map to assess if these quantities are stationary over the rjMcMC samples. Furthermore, we ensure that the distribution over the number of nodes in the model is stationary and provides a smooth distribution. In all inversions discussed here, we ran 20 parallel rjMcMC samplers at varying temperatures to generate 5 million samples each (keeping only those sampled at $T=1$). The choice of running the inversion with 20 parallel tempering chains stems from the limitations in available computing resources. Inversion tests with fewer parallel chains produce similar results, though they require significantly more samples. Of the retained samples, the first half are discarded as burn-in samples dependent on the (random) initial starting

model. Final results for each inversion are represented by the remaining 2.5 million model samples (for the final iteration of updated ray path geometry).

2.5 Synthetic inversion tests

Performing tomographic inversions on simulated data (for a known true model) is useful for studying the resolving power of observed data through so-called checkerboard or resolution tests. In this case, synthetic inversions are also valuable for studying and testing new methodologies. In theory, solving for the PPD in Bayesian tomography precludes resolution tests. However, in practice, these tests can help illuminate possible deficiencies in the data, or artifacts in the final model solution. This section demonstrates the Bayesian surface-wave tomography with azimuthal anisotropy, and illustrates the spatial resolution for the northern Cascadia and Haida Gwaii dataset. The synthetic inversion test presented here is for the source-station path coverage of the dataset at a period of 15 s. Tests at other periods show similar results, as the spatial and azimuthal path coverage is similar for the periods considered in this work, even though the total number of data at each period differs (fig. 2.3). The true model consists of an alternating checkerboard pattern of slow and fast group velocity anomalies that are ± 25 percent of the average velocity in the 15 s period dataset (3.0 km/s), with a length scale of 360 kilometres. Azimuthal anisotropy in the true model consists of 5 percent anisotropy (relative to the isotropic velocity) with an alternating opposing pattern of fast-axis directions. The ray paths for the true model are estimated using piecewise travel-time minimization (Um & Thurber, 1987), and are used to calculate the simulated data. Gaussian-distributed errors with a standard deviation of 1 second are added to the simulated travel-time data. Furthermore, Gaussian-distributed errors with a standard deviation of 20 seconds are added to 1 percent of the dataset to simulate outliers. Inversions were carried out using the Bayesian methodology described in the previous section (with four iterations of ray path updating).

Figure 2.5 summarizes the results of the synthetic inversion test. The mean values of isotropic and azimuthally anisotropic group velocity extracted from the ensemble of

rjMcMC samples at each location in the 2D map are shown. In general, the inversion test recovers the pattern of isotropic group velocities and azimuthal anisotropy. However, the anisotropic components of the true model are less well recovered, and only resolved in regions with greater azimuthal path coverage. The magnitudes of isotropic anomalies are accurately recovered, especially in regions with dense path coverage. Figure 2.5 also shows the estimated model uncertainty for the isotropic group velocity map, and the estimated uncertainty in the orientation of the fast-axis direction of anisotropy. Isotropic group velocity uncertainty is estimated from the standard deviation of the ensemble of rjMcMC samples at each location in the 2D map. Since the fast-axis direction is a polar quantity estimated from A_1 and B_1 , we propagate the errors in A_1 and B_1 through the simple equation $\frac{1}{2}\arctan(B_1/A_1)$ rather than taking the standard deviation of fast-axis directions themselves. These figures reveal that the estimated model uncertainty depends on the data path coverage as well as the underlying group velocity model. Uncertainty is greater where path coverage is poor. Furthermore, estimated uncertainty is greater near lateral heterogeneity in the velocity model, as the inversion is not able to perfectly recover the location of discontinuous structure.

Figure 2.6 shows the distribution of the estimated inversion hyper-parameters for the synthetic test. Since the simulated errors on the data were simply Gaussian-distributed, the inversion test correctly estimates no linear relationship between the level of data errors and the length of the data travel paths. The inversion also correctly estimates the constant term for the data errors (b in eq. 2.6). The slight under-estimation of the constant error terms is likely the result of the discrepancy between the simulated error distribution (Gaussian) and the modelled likelihood (Laplace), where the error diversity (modelled by a and b parameters) is smaller than the error standard deviation by a factor of $2^{1/2}$. The results appear insensitive to outliers. However, the effect of outliers within the inversion may be significant if outliers are present in the data in regions with sparse path coverage. Ultimately, the inversion performs well for estimating the unknown level of data errors. Additional synthetic tests were performed using anomalies with a smaller length scale (240 kilometres) under varying noise conditions with Gaussian-distributed errors with a standard deviation of 1, 2, and 5 seconds added to the simulated travel-time data with

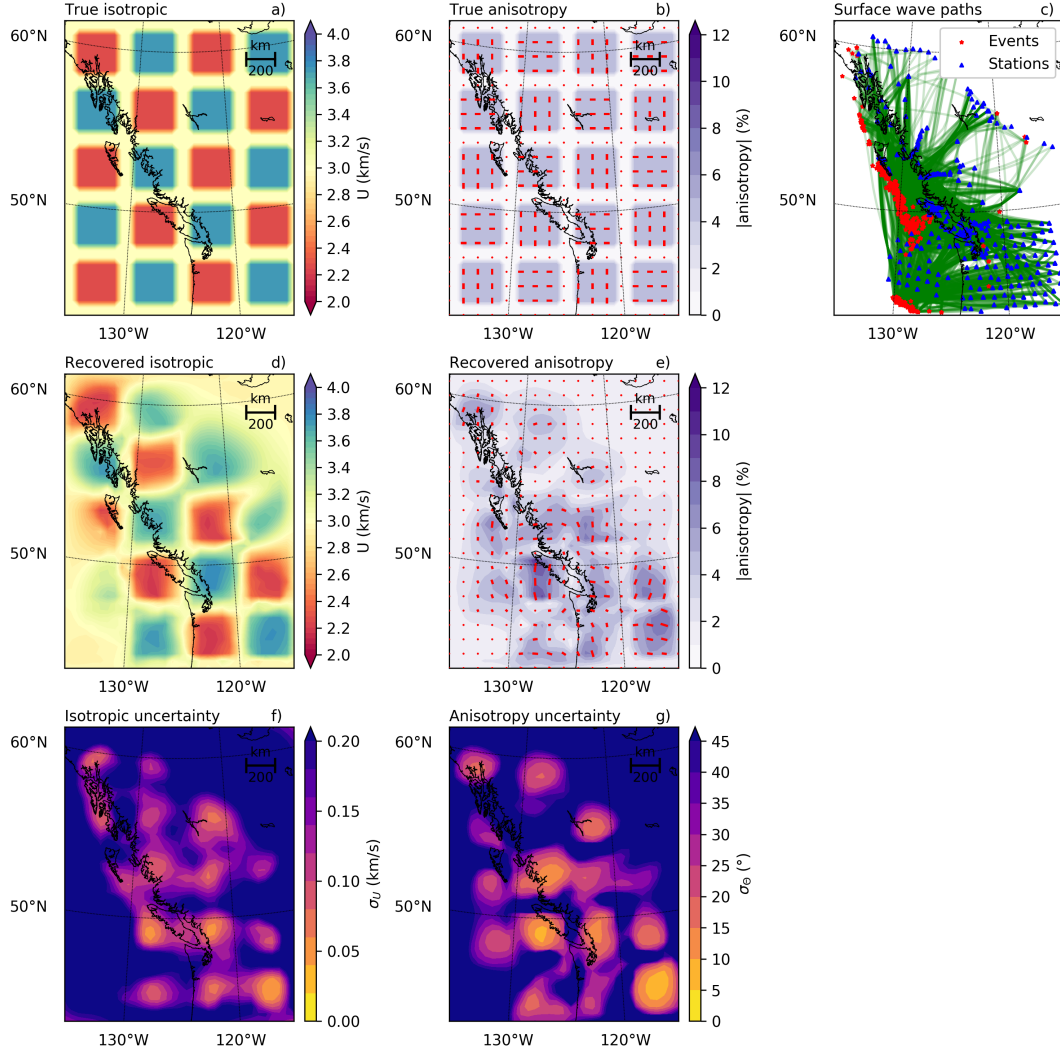


Figure 2.5: Results of the synthetic Bayesian inversion test. The isotropic and anisotropic components of the true model are shown in a) and b), respectively. The surface wave travel paths used to calculate the simulated dataset are shown in c). Mean isotropic velocities estimated from the ensemble of rjMcMC samples for the inversion are shown in d). Similarly, the recovered fast-axis orientations and magnitudes of anisotropy are shown in e). The standard deviations of isotropic velocities estimated from the ensemble of rjMcMC samples are shown in f). The uncertainty of anisotropy fast-axis directions are shown in g). Red lines delineate the fast direction of anisotropy.

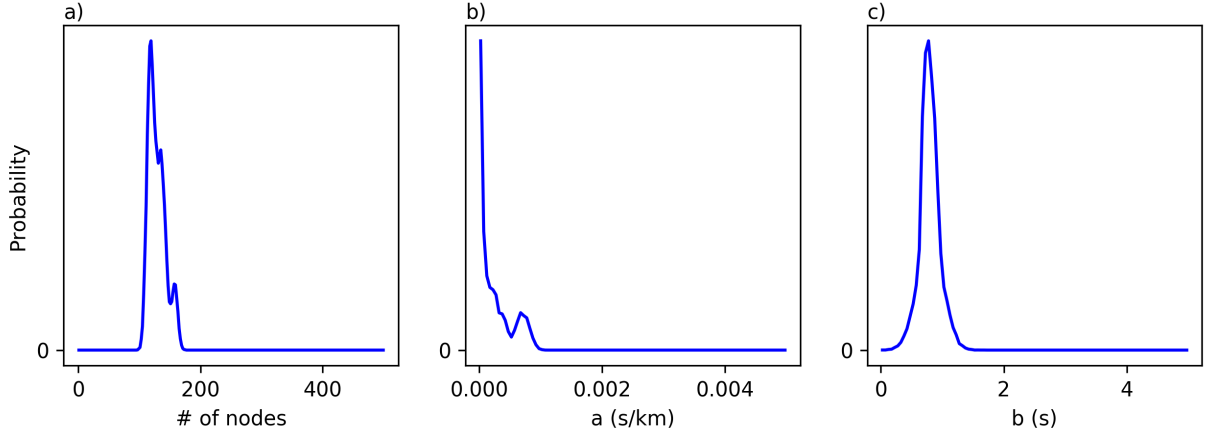


Figure 2.6: Hyper-parameter estimates from the synthetic Bayesian inversion test. The number of nodes in the Delaunay triangulation with linear interpolation is shown in a). The linear coefficient between unknown data errors and travel path length is shown in b). The constant term for unknown data error is shown in c).

outliers as described above (figs. A.4 and A.5). These tests are shown in the supplement and reveal that the estimation of azimuthal anisotropy in surface wave tomography depends on the tradeoff between the length scale (or magnitude) of anisotropy and the level of data errors. In some cases, the data can be adequately fit without anisotropy when data errors are large (and anisotropy signal is small). Ultimately, the inversion will not recover anisotropy if it is not required by the data, as the mean of the prior distributions for the anisotropic parameters is zero (i.e., no anisotropy). The Bayesian inversion approach allows for detailed examination of the data information content and how this affects the resolution of isotropic group velocity and azimuthal anisotropy.

2.6 Anisotropic group velocity over northern Cascadia and Haida Gwaii

In this section we present results for Bayesian tomographic inversion of surface wave dispersion data collected over northern Cascadia and Haida Gwaii. It is challenging to make detailed interpretations from group velocity maps without inverting for a pseudo 3D velocity model (by estimating 1D velocity profiles at each location in the study region, then combining them). Interpretation of group velocity maps in our study region is further complicated due to the combination of onshore-offshore structure, where water depth and thick accretionary sediments can have a significant effect on surface wave velocities (e.g., Darbyshire et al., 2018; Yeck et al., 2017; Janiszewski et al., 2019). As such, we provide only a synoptic interpretation of the isotropic and anisotropic group velocity maps and reserve more detailed investigation for future work. For the periods considered here, surface waves on land are sensitive from mid crustal depths to shallow lithospheric mantle depths (fig. 2.2).

Figure 2.7 shows variations in isotropic group velocity as a function of period. At short periods, isotropic group velocity anomalies derive predominantly from compositional variations in the crust, as well as from the presence of sedimentary basins, volcanic provinces, and mountain ranges. At longer periods, surface waves are increasingly sensitive to crustal thickness as well as temperature anomalies in the lower crust and uppermost mantle (Moschetti et al., 2007). At periods between 15 and 25 s we observe a significant change in isotropic group velocity along the continental margin that delineates the boundary between oceanic and continental lithosphere (fig. 2.7). At 15 s, we observe a narrow band of very slow group velocity, which is likely attributed to water-laden sediments, and sedimentary rocks, along the continental shelf (AC in fig. 2.7). This slow velocity anomaly also extends onshore through the Olympic Peninsula south of Vancouver Island, consistent with the location of the onshore Cascadia accretionary complex (Brandon et al., 1998). At periods between 25 and 36 s we observe a margin-parallel band of relatively slow group velocity, with higher velocities further inland, throughout the Cascadia margin (CR and

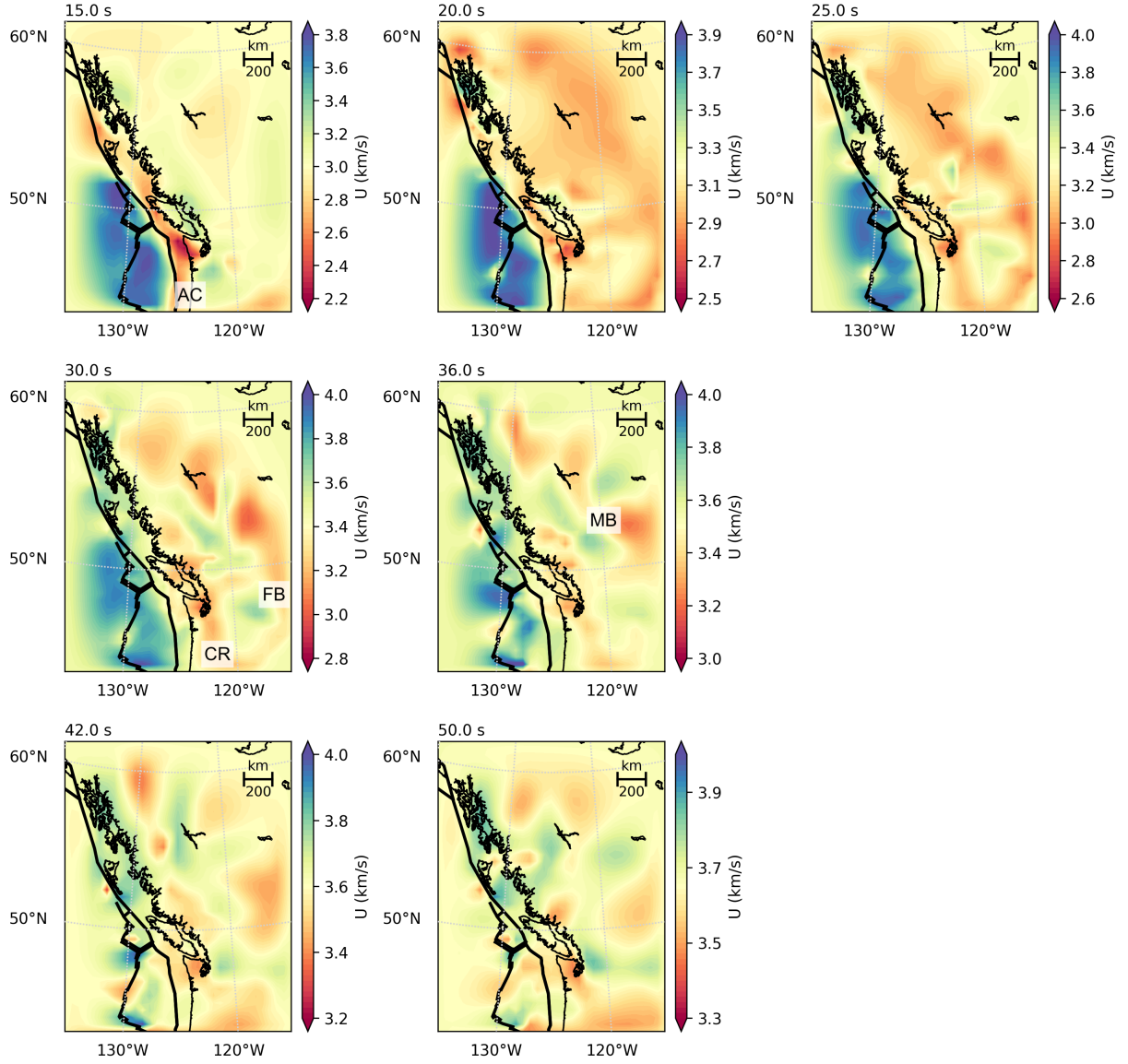


Figure 2.7: Isotropic group velocity over northern Cascadia and Haida Gwaii. The mean isotropic velocities estimated from the ensemble of rjMcMC samples for inversions at periods 15-50 s are shown. AC, CR, FB, and MB stand for accretionary complex, Cascade range, flood basalts, and morphological belts, respectively, and are discussed in the text.

FB in fig. 2.7). This pattern is consistent with results from other surface wave studies in Cascadia and reflects the presence of the Cascade range (slow) and the Columbia river flood basalts (fast) further inland (e.g., Moschetti et al., 2007; Janiszewski et al., 2019). At periods between 25 and 42 s we observe a margin-parallel band of high group velocity over 200 kilometres east of the coast (MB in fig. 2.7). The geology of British Columbia comprises a series of margin-parallel morphological belts that contain various accreted terranes (Nelson et al., 2013). The observed velocity anomalies likely reflect compositional changes between these belts/terranes or possibly variations in crustal thickness beneath these belts/terranes (Clowes et al., 1995). Further detailed examination of the structure of inland British Columbia may be challenging as group velocities are less well resolved here, and may be biased due to the lack of resolvable anisotropy. At the longest periods considered in this work (42 and 50 s), velocity anomalies are generally small in magnitude (note the narrower colour scale range with increasing period) but are generally large in spatial extent, reflecting the large integrated effect of long wavelength surface wave data.

Figure 2.8 shows the mean magnitude and directions of anisotropy estimated from the Bayesian inversion. Our anisotropy results show consistent directions across all periods from Vancouver Island to Haida Gwaii, as expected from the overlap in depth sensitivity for the periods considered here (fig. 2.2). The slight change in anisotropy direction and magnitude between northern Vancouver Island and Haida Gwaii (most clearly seen at 42 and 50 s in fig. 2.8) correlates with a change in isotropic group velocity (42 and 50 s in fig. 2.7) and may be a result of the transitional tectonic regime along the coast. At periods considered in this work, seismically-observable anisotropy is expected to be attributed to the local compressive stress field, or major foliations within the crust. Local horizontal compressive stress can result in aligned fluid-filled cracks, which produce fast anisotropy directions parallel to the stress direction (Crampin, 1994). Throughout northern Cascadia, local shear-wave splitting measurements show margin-parallel fast-axis directions (Currie et al., 2001; Balfour et al., 2012), which are consistent with horizontal compressive stress directions estimated from earthquake moment tensor solutions (Balfour et al., 2011). The complicated and low-amplitude pattern of anisotropy recovered over southern Vancouver Island (most notable at 15-25 s) is a robust feature, due to excellent path coverage over

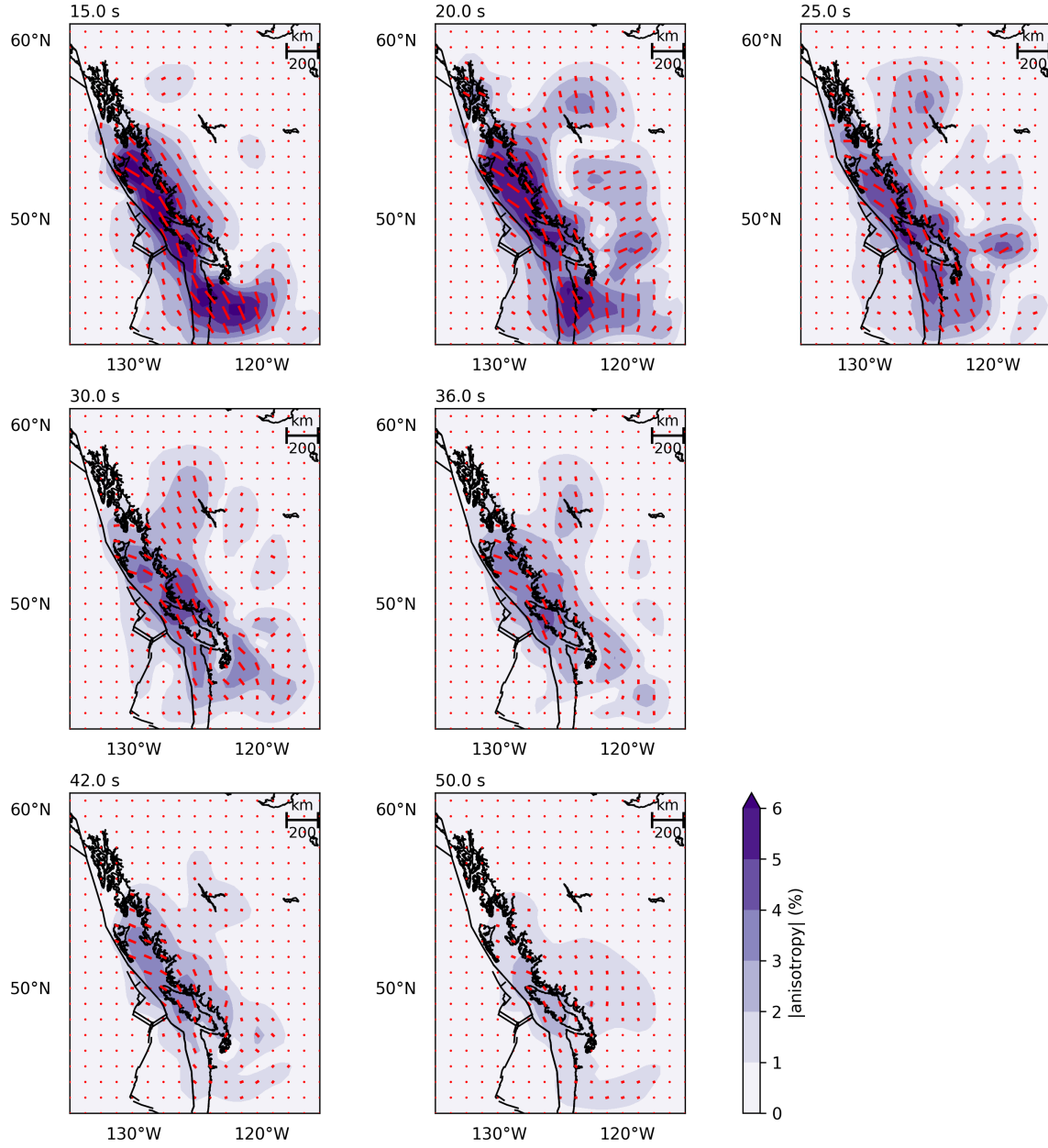


Figure 2.8: Azimuthal anisotropy in group velocity over northern Cascadia and Haida Gwaii. The mean magnitude and directions of anisotropy estimated from the ensemble of rjMcMC samples for inversions at periods 15-50 s are shown.

this region. It may be the result of superimposed structural anisotropy as some of the stations in this region are located near margin-perpendicular crustal faults with strong anisotropic signatures (Bostock & Christensen, 2012). Further north, near Haida Gwaii, limited local shear-wave splitting measurements also show margin-parallel fast-axis directions (Cao et al., 2017). However, these are perpendicular to local compressive stress directions (Ristau et al., 2007), suggesting that crustal anisotropy near Haida Gwaii is the result of structural foliations, possibly related to the QCF. Recent work by Eilon & Forsyth (2020) suggests that shallow offshore anisotropy within and beneath the JdF crust is predominantly margin-perpendicular, though weak in magnitude. Offshore path coverage in our model is not as dense as the continental shelf region, and we suspect such weak offshore anisotropy is likely not resolvable with our dataset. In general, our estimates of fast-axis directions in this region appear consistent with local shear-wave splitting measurements.

Figures 2.9 and 2.10 show the uncertainties in isotropic group velocity and fast-axis direction of azimuthal anisotropy estimated from the standard deviations of the ensemble of rjMCMC samples from the Bayesian inversion. Since the fast-axis direction is a polar quantity estimated from A_1 and B_1 , we propagate the errors in A_1 and B_1 through the simple equation $\frac{1}{2}\arctan(B_1/A_1)$ rather than taking the standard deviation of fast-axis directions themselves. Uncertainty in isotropic group velocity is low throughout Vancouver Island and the continental shelf northward towards Haida Gwaii (fig. 2.9). Uncertainties are also low throughout the northwestern United States likely due to the high seismic station density throughout this region (fig. 2.3). Uncertainties in the fast-axis direction of azimuthal anisotropy are generally low across all periods for the continental shelf region including Vancouver Island and Haida Gwaii (fig. 2.10). However, uncertainties throughout the northwestern United States are higher despite the abundant station coverage. This is likely the result of the lack of azimuthal coverage throughout this region (i.e., all of the ray paths are pointing in a similar direction). This may also be the result of weaker or more complex anisotropy in this region. In summary, only the continental shelf region including Vancouver Island and Haida Gwaii presents low uncertainties in both isotropic group velocity and the direction of azimuthal anisotropy. As such, only results from this region (including associated uncertainties) may be considered for future analysis. Uncertainties

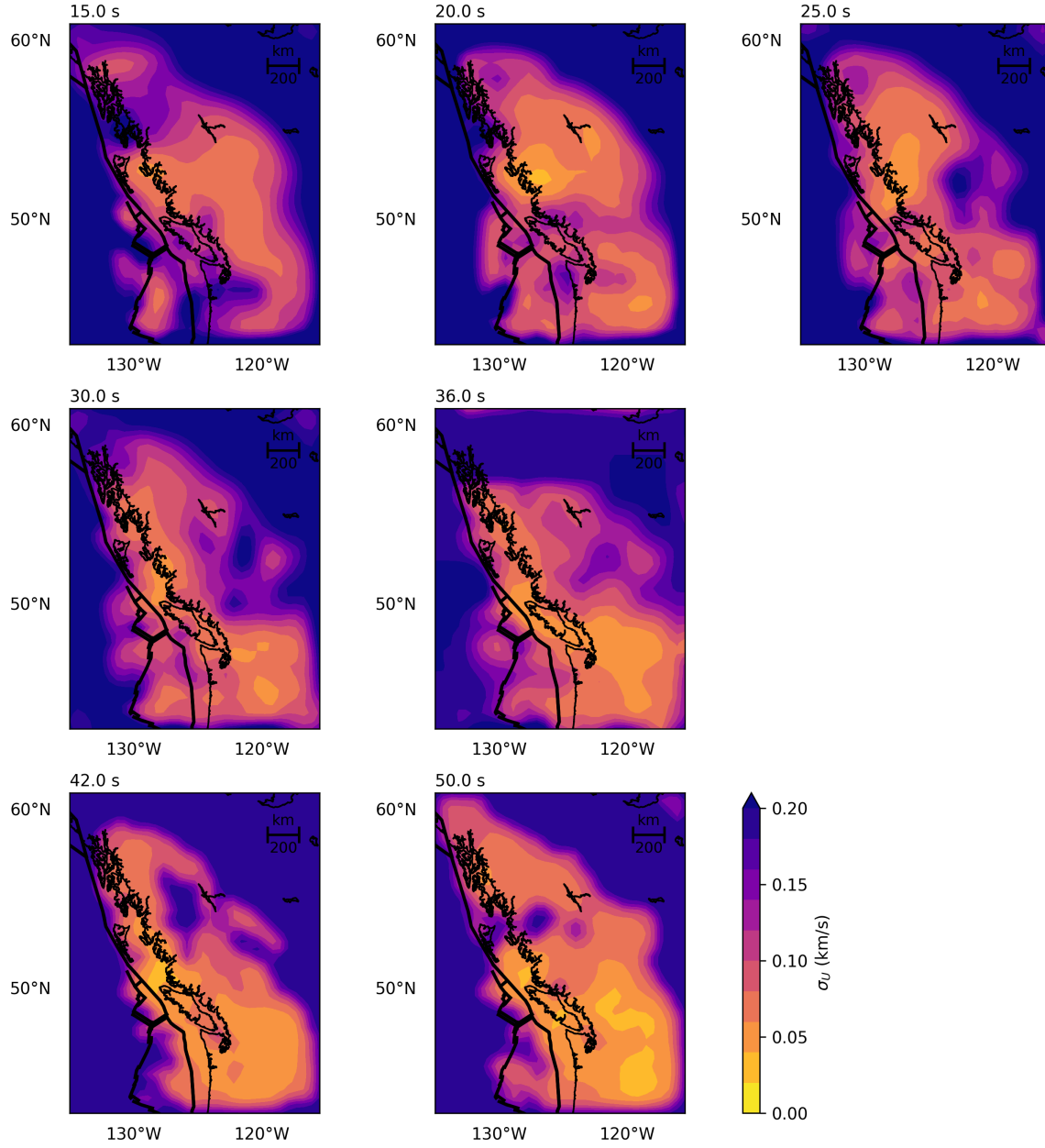


Figure 2.9: Uncertainty of isotropic group velocity over northern Cascadia and Haida Gwaii. The standard deviations of isotropic velocities estimated from the ensemble of rjMcMC samples for inversions at periods 15-50 s are shown.

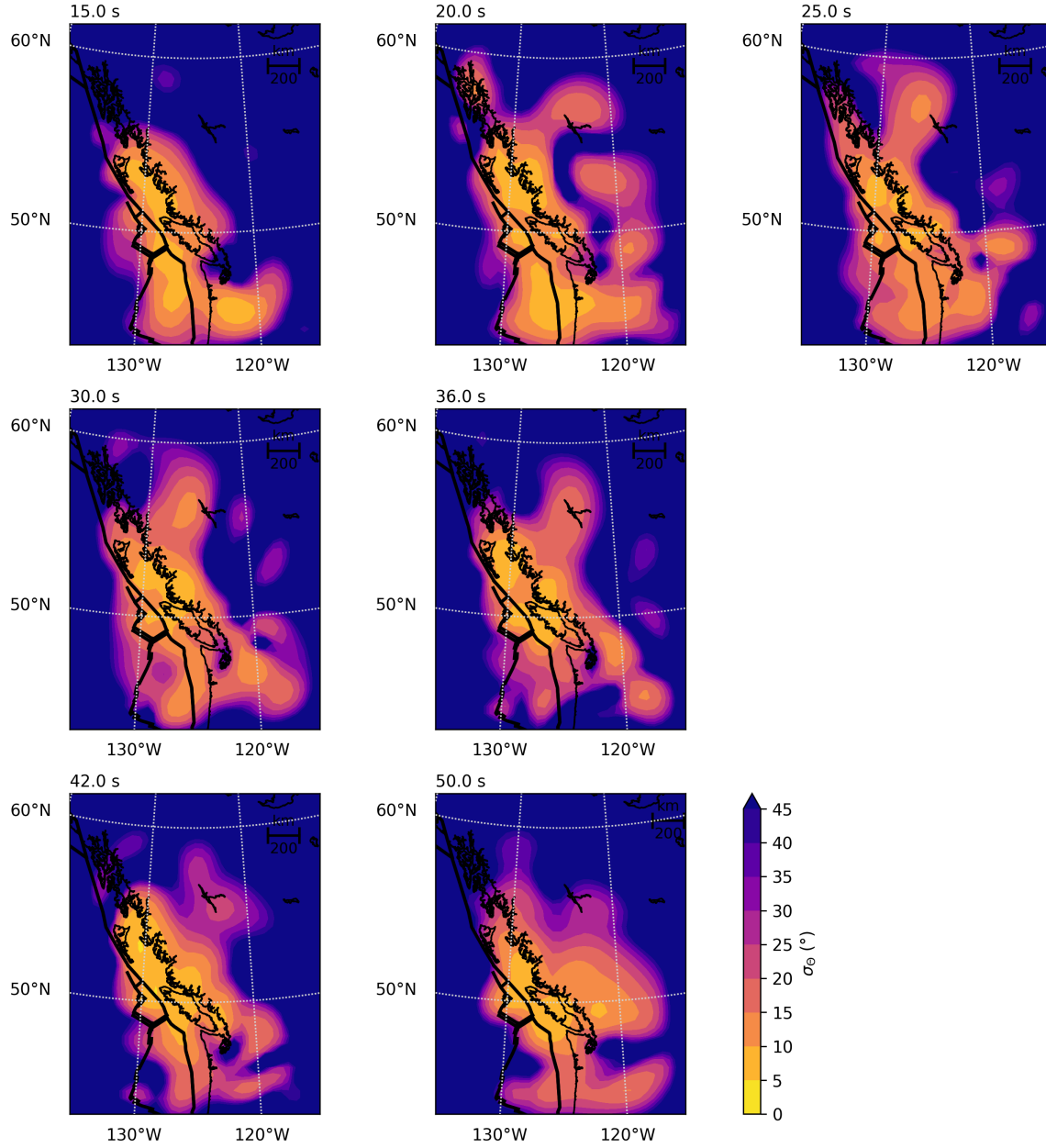


Figure 2.10: Uncertainty of fast-axis direction of azimuthal anisotropy over northern Cascadia and Haida Gwaii, for inversions at periods 15-50 s.

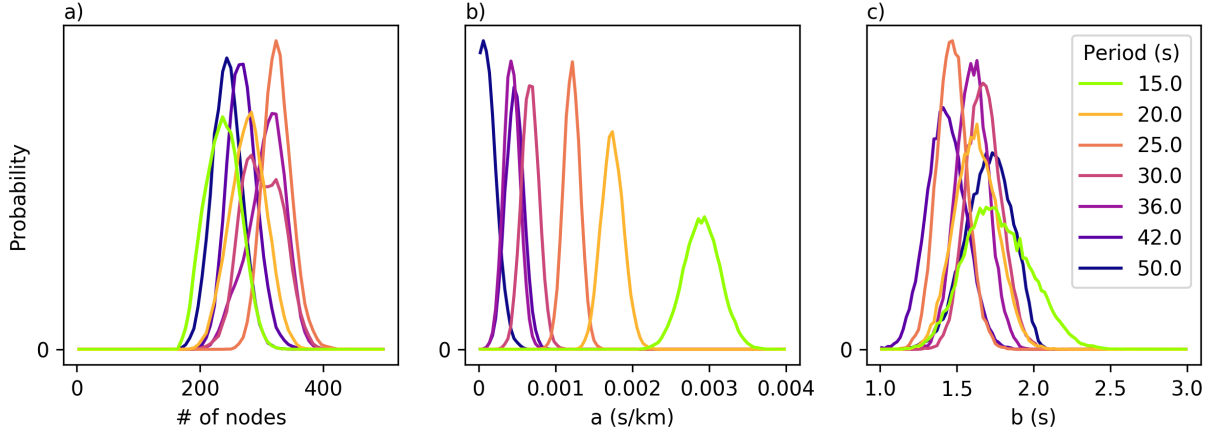


Figure 2.11: Hyper-parameter estimates from Bayesian inversions. The number of nodes in the Delaunay triangulation with linear interpolation is shown in a). The linear coefficient between unknown data errors and travel path length is shown in b). The constant term for unknown data error is shown in c).

in the magnitude of anisotropy are shown in the supplementary material, and are of the order of 1 percent where recovery is best (fig. A.6). For both the isotropic and anisotropic structure, estimated uncertainties correlate with the recovered isotropic and anisotropic maps, where larger uncertainty is estimated near lateral heterogeneity as the inversion is not able to perfectly recover the location of discontinuous structure.

Figure 2.11 shows the distribution of the estimated inversion hyper-parameters at each period. In general, all inversions at all periods recover a similar number of nodes in the inversion. The equivalent figure for the distribution of the estimated inversion hyper-parameters after 1 inversion iteration (i.e., assuming great circle ray paths) is shown in the supplement (fig. A.7). Comparison between the inversion iterations suggests that, other than for 15 s period data, updating the ray path geometry had little effect on the estimated data errors, as might be expected when considering long period data over short path lengths. It is reasonable then that the inversion of 15 s data should initially recover the smallest number of nodes, since the inversion estimates the largest errors (a and b parameters) for these data, suggesting that they have reduced information content and are

not able to resolve more complex structure. After four iterations of ray path updating, the inversion estimates similar constant terms for data errors, b values (fig. 2.11c), for all periods. This suggests some consistent, period-independent, source of errors in the data that may be related to errors in source location or onset time. This may be significant as many of the events considered in this work are located offshore and are likely poorly located (and often have analyst-fixed depths). Unfortunately, information on location uncertainty is not available for the majority of the events considered in this work. Figure 2.11b shows an interesting trend between period and the estimated linear dependence of data errors with travel path length. At long periods, the dependence is estimated to be low (small a value), and vice versa. This suggests that surface wave travel time errors are a function of wavelength, as well as the number of wavelengths between source and receiver. These wavelength-dependent data errors may be attributed to unaccounted-for scattering effects.

2.7 Discussion and Conclusions

In this paper we consider tomographic inversion of surface-wave dispersion data using a novel Bayesian approach to quantify azimuthal anisotropy and associated uncertainty while avoiding subjective regularization procedures. We implement an alternative model parameterization compared to previous work that considered Bayesian tomography in order to produce continuous (more realistic) models. Like previous studies, we consider the inverse problem within a trans-dimensional Bayesian framework, and allowed the model dimensionality (complexity) to be determined by the data information content. Like previous works, nonlinearity in the tomographic inverse problem (i.e., ray path geometry) is considered via an iterative process in which paths are updated after each inversion run. It is well known that isotropic structures are better resolved than anisotropic structures using surface wave data. Future work will consider decoupling the parameterizations of isotropic and anisotropic structure. This is similar to Bodin et al. (2016) who considered anisotropy in 1D shear-wave velocity profiles constrained by data from single-station seismic methods.

Our inversion method was applied to surface wave dispersion data at periods between

15 and 50 s measured from recordings of earthquakes over northern Cascadia and Haida Gwaii, a region of complex active tectonics but limited station coverage. Group velocity dispersion measurements from regional earthquakes have greater errors than dispersion data obtained from other methods. We consider data errors in a robust way via manual dispersion selection for each seismogram, outlier removal procedures, and the use of a Laplace Likelihood distribution (with unknown diversity estimated within the inversion). Our results for isotropic group velocities are consistent with previous studies further south over Cascadia. Our results for azimuthal anisotropic fast-axis directions are consistent with, and supplement, the limited shear-wave splitting measurements over the region. In general, our inversion recovers margin-parallel fast-axis directions between Vancouver Island and Haida Gwaii, with a small change in direction and magnitude along the margin. These results are broadly consistent with azimuthal anisotropy from teleseismic shear-wave splitting measurements (Cao et al., 2017; Mosher et al., 2014), which presumably sample the lithospheric and/or sub-lithospheric mantle, possibly pointing to vertically-coherent deformation.

Our method estimates the unknown level of data errors, and reveals a correlation between surface wave period and the dependence of data errors on travel path length. For periods greater than 15 s, we find that estimated data errors do not change significantly between inversion iterations (i.e., ray path updating) suggesting that for such long period data, over such short propagation path lengths, deviations from the great circle paths are not the dominant source of data errors. The consistency between periods for the estimated constant term of the data errors suggests a systematic source, potentially related to poor event locations (particularly for offshore seismicity). The inversion method provides robust results in regions with sufficient azimuthal path coverage to recover anisotropy. Otherwise, estimated anisotropy is equivalent to the prior distribution (with a mean of zero anisotropy) and recovered isotropic velocities may be biased. Our results demonstrate the value of this novel inversion technique for robust determination of anisotropy from surface wave dispersion data. Our results over the continental shelf, including Vancouver Island and Haida Gwaii, provide an initial step towards resolving a full, 3D, seismic velocity model in future work. This work provides broad structural context for future seismic station deployments

(both onshore and offshore) to study the tectonic structure and dynamics of the region, with the aim of improving our understanding of seismic and tsunami hazards.

Lastly, one of the great advantages of Bayesian inversion methods is the ability to incorporate meaningful prior information into the problem. Future work, particularly with respect to structural studies of northern Cascadia and Haida Gwaii, will consider additional model constraints through the use of informative prior information and other datasets.

Chapter 3

Circular statistics in Bayesian inversion for two-station phase-velocity dispersion estimation

3.1 Introduction

Circular data commonly occur in the Earth sciences, including geophysics and seismology. Such quantities are wrapped, and are generally classified as either directional or periodic. For example, directional measurements of paleomagnetic orientations or bedding orientations are important data in tectonic plate kinematics and structural geology, respectively (e.g., Darot & Bouchez, 1976; Van Alstine, 1980; Lark et al., 2014). Alternatively, temporal observations that wrap over the corresponding measurement period (e.g., hours in a day) result in circular data. Circular quantities that commonly appear in seismology include the orientation of earthquake focal mechanisms, stress fields, seismic anisotropy, as well as the phase of seismograms in the Fourier domain.

The circular nature of such quantities creates challenges in applying traditional statistical methods. A classic example is the arithmetic mean of angular measurements of 5° and

355°, which is 180°. This is clearly not a suitable measure of the location of the data, where a better choice would be 0°. Other linear statistical measures (e.g., median, variance, etc.) are likewise often not appropriate for circular data. The statistical analysis of circular data is an active field of research. Many techniques have been developed that respect the unique topology of circular quantities, yet parallel standard statistical analysis techniques. Useful details on this topic can be found in textbooks (e.g., Fisher, 1995; Mardia & Jupp, 2009) as well as recent reviews in literature (e.g., Lee, 2010). Many geophysical inverse problems assume an L_2 misfit function (or Gaussian-distributed data errors), as supported by the Central Limit Theorem. However, such assumptions are generally not valid for circular data. This work aims to incorporate a meaningful data error model for Bayesian geophysical inversion of circular data.

Bayesian methods have become increasingly popular for solving geophysical inverse problems in recent years as they allow for rigorous uncertainty quantification. Bayesian inversion is a probabilistic approach in which the parameters of a postulated model, along with the observed data, are treated as random variables. The solution to the inverse problem is described by properties of a posterior probability density (PPD) constrained by data information and prior knowledge. Such priors can be uninformative, thereby allowing the solution to be predominantly determined by the data information content. Alternatively, meaningful prior information can be quantitatively introduced to the inverse problem, which is an advantage of Bayesian methods. Typically, Markov-chain Monte Carlo (MCMC) methods are used to estimate the PPD numerically (Mosegaard & Tarantola, 1995; Malinverno, 2002; Tarantola, 2005). Bayesian (probabilistic) inversion requires knowledge or estimation of the statistical distribution of data errors. In this work, we apply the circular von Mises distribution to model data errors (Mardia & Jupp, 2009).

This work is motivated by the application of automating the estimation of surface-wave phase-velocity dispersion for the teleseismic two-station method, which is commonly measured manually. Specifically, we apply Bayesian inversion to measurements of surface-wave phase spectra in order to estimate 1D, path-averaged Earth structure between station pairs. The dispersion curve and associated uncertainties are additional results of the inversion, which can then be used as input for subsequent analysis (e.g., tomography). Alternatively,

this work presents a foundation for the inversion of surface-wave phase spectra to estimate 3D Earth structure directly. The technique is applied to surface-wave recordings from long-running stations over British Columbia, Canada. This work is further motivated by the disconnection between existing surface-wave studies in the region (e.g., Bensen et al., 2008; Dalton et al., 2011; Kao et al., 2013; Bao et al., 2016; Gosselin et al., 2021), which do not consider the entire region or incorporate all available data, and (mostly) consider structure further inland. Ultimately, this work produces data products (i.e., dispersion) that will be valuable for studying the complex (and poorly understood) interaction between the northern Cascadia subduction zone, the transpressive Haida Gwaii (HG) region, and the southern Canadian Cordillera

3.2 Methodology

This section discusses the application of Bayesian inversion to circular data for phase-velocity dispersion estimation. The dispersion is an additional result of an inversion for 1D path-averaged Earth structure. This is similar to recent work by Hawkins & Sambridge (2019), who consider the real component of ambient noise cross-correlation functions in order to solve for path-averaged Earth structure and, by extension, inter-station phase-velocity dispersion. This work provides an alternative to existing automated approaches to inter-station dispersion estimation. More importantly, we consider the inversion of 1D path-averaged Earth structure as it represents a simple example/application of the inversion of phase measurements (as opposed to the direct inversion for 2D or 3D structure). We discuss logical extensions of the ideas presented below, which represent directions for future work. Bayesian inversion and associated numerical methods are discussed in Tarantola (2005) and Brooks et al. (2011), and we provide only a brief summary below.

3.2.1 Bayesian inversion

Consider a model $\mathcal{M}$ with M unknown model parameters $\mathbf{m}$. In this case, the model is an approximation of Earth structure along a surface-wave propagation path between two

seismic stations. Also consider a set of N observed data $\mathbf{d}$ of some process that interacts with the system. Specifically, surface-wave phase measurements at various periods. In a Bayesian framework, $\mathbf{m}$ and $\mathbf{d}$ are treated as random variables and are related via Bayes' theorem

$$P(\mathbf{m}|\mathbf{d}, \mathcal{M}) = \frac{P(\mathbf{m}|\mathcal{M}) P(\mathbf{d}|\mathbf{m}, \mathcal{M})}{P(\mathbf{d}|\mathcal{M})}, \quad (3.1)$$

where $P(\mathbf{m}|\mathbf{d}, \mathcal{M})$ defines the PPD, which is the inversion solution that represents the state of model parameter knowledge given the prior information and observations. $P(\mathbf{m}|\mathcal{M})$ is the prior probability of the model parameters independent of the data, which represents initial knowledge. $P(\mathbf{d}|\mathbf{m}, \mathcal{M})$ defines the data information distribution given a model. In practice, this is interpreted as the likelihood of the model parameters $\mathcal{L}(\mathbf{m})$ since $\mathbf{d}$ is considered fixed once data are observed. $P(\mathbf{d}|\mathcal{M})$ is a normalization over the model parameter space.

As mentioned above, we invert for 1D path-averaged Earth structure. We parameterize the model with an unknown number of layers with unknown shear-wave velocity V_S . Consequently, the model complexity is controlled by the data information content. Such an inversion strategy is typically referred to as trans-dimensional because the number of model parameters required to describe the 1D Earth structure is not fixed. The number of layers in the model is a hyper-parameter within the inversion that indexes over a countable set of possible model parameterizations (i.e., model $\mathcal{M}$ in eq. 3.1). The use of $\mathcal{M}$ as well as the index of the number of model parameters are suppressed henceforth for simplicity.

Analytic solutions for equation (3.1) typically do not exist, but can be estimated numerically using Monte Carlo methods. Specifically, we apply the reversible-jump Markov-chain Monte Carlo (rjMCMC) algorithm to draw a series of dependent (on the previous sample) and asymptotically-unbiased samples from the PPD (Green, 1995, 2003). The algorithm consists of randomly perturbing a current set of model parameters $\mathbf{m}$ to a new state $\mathbf{m}'$, based on a proposal distribution $Q(\mathbf{m}'|\mathbf{m})$. In this case, the model parameters are the depths of layer interfaces as well as the V_S values within each layer. The proposed set of model parameters $\mathbf{m}'$ is then accepted with probability

$$A(\mathbf{m}'|\mathbf{m}) = \min \left[1, \frac{Q(\mathbf{m}|\mathbf{m}')}{Q(\mathbf{m}'|\mathbf{m})} \frac{P(\mathbf{m}')}{P(\mathbf{m})} \frac{\mathcal{L}(\mathbf{m}')}{\mathcal{L}(\mathbf{m})} |\mathbf{J}| \right], \quad (3.2)$$

Table 3.1: Prior bounds for model parameters in Bayesian inversion.

Parameter	Min	Max
V_S at surface (km/s)	1.0	5.0
V_S at max depth (km/s)	AK135 - 20%	AK135 + 20%
Interface depth (km)	0	400
Number of model layers	3	20

where $|\mathbf{J}|$ is the determinant of the Jacobian matrix for the transition between different model states. We apply Gaussian model parameter perturbations (centered on the current value) such that $Q(\mathbf{m}'|\mathbf{m}) = Q(\mathbf{m}|\mathbf{m}')$. Furthermore, we apply bounded uniform prior distributions on model parameters such that $P(\mathbf{m}) = P(\mathbf{m}')$. When proposing a new layer to the model, the depth and V_S of the layer are drawn from the prior. Given these choices of prior and proposal densities, it can be shown that $|\mathbf{J}| = 1$ and equation (3.2) simplifies to a ratio of likelihoods for all types of model perturbations (Bodin & Sambridge, 2009; Dosso et al., 2014). Prior bound widths are listed in table 3.1 and are selected in order to allow the solution to be predominantly determined by the data, while still constraining model parameters to realistic values. Specifically, we apply a depth-dependent prior bound on V_S such that a wide range of velocity values are acceptable at shallow depths, but a narrower (and more realistic) range of velocities are acceptable at greater depths. Prior bounds on V_S at the greatest depth considered in the model are defined by AK135 velocities (Kennett et al., 1995) $\pm 20\%$. Given a sufficiently large number of samples, the rjMCMC algorithm provides a good approximation to the PPD. For all inversions in this work, we use the algorithm to generate 500000 samples but only keep the last 250000 samples (independent of the initial random model) for subsequent analysis.

3.2.2 Error model for circular data

Assuming data errors are represented by the difference between observed and predicted data (i.e., residuals), the likelihood function defines the data error distribution within the

inversion. Typically, this distribution is unknown and a simple distribution is assumed with statistical parameters estimated from the data. Gaussian-distributed errors are commonly assumed, as supported by the Central Limit Theorem. Alternatively, a Laplace distribution may be more applicable in the presence of data outliers. However, both Gaussian and Laplace distributions are defined for linear data, and are not suitable for data where the residual errors are distributed on the unit circle. Data misfit has previously been considered for complex data in the argand plane (e.g., Dosso & Wilmut, 2006, 2012), and has been applied to Bayesian trans-dimensional inversion (e.g., Ray et al., 2016). Although this represents a means of considering circular data, consideration of signal amplitude is not necessary in our application (as discussed further below). Instead, here we explore the use of a probability distribution for real-valued, circular quantities.

Consider a random angular variable $\theta \in (-\pi, \pi]$ (modulo 2π). The continuous probability distribution of θ can be expressed on the unit circle via the von Mises distribution, which is similar to a wrapped normal distribution (Mardia & Jupp, 2009). The general form of the von Mises distribution is given by

$$P(\theta|\mu, \kappa) = \frac{1}{2\pi I_0(\kappa)} \exp[\kappa \cos(\theta - \mu)], \quad (3.3)$$

where μ is a measure of the distribution location (analogous to circular mean) and κ is the concentration of the distribution (i.e., the reciprocal of the dispersion of a probability distribution). Specifically, $1/\kappa$ is analogous to variance. For small κ values, the distribution is nearly uniform on the circle (and perfectly uniform for $\kappa = 0$). As κ increases, the distribution becomes more concentrated and approaches a normal distribution with mean μ and variance $1/\kappa$. The normalization in equation (3.3) is the modified Bessel function of first kind and order zero, given by

$$I_0(\kappa) = \frac{1}{2\pi} \int_0^{2\pi} \exp[\kappa \cos(\theta)] d\theta. \quad (3.4)$$

The rjMcMC algorithm is implemented in this work using a likelihood based on equation (3.3), which is a suitable residual error distribution for circular data such as surface-wave phase measurements. For simplicity, data errors are assumed to be independent

between periods, such that the general likelihood is given by the product of von Mises distributions for each period. However, data errors may be correlated between periods depending on their sources, and the assumption of independence between periods may underestimate uncertainties. The locations μ and concentrations κ of the general von Mises likelihood are estimated from the distribution of phase measurements for all earthquakes recorded by a given station pair (as will be discussed below). For this application, the calculation of equation (3.4) is unnecessary as κ values are fixed within the inversion and the rjMCMC algorithm requires only the calculation of ratios of likelihoods.

3.2.3 Teleseismic two-station phase-velocity dispersion

The measurement of the dispersive characteristics of surface waves based on comparison of seismograms of the same earthquake recorded at two stations has a long history in seismology. Useful early overviews on the topic are given by Knopoff (1972) and Kovach (1978), while relatively recent discussions are provided by Meier et al. (2004) and Magrini et al. (2020). Consider an earthquake approximately aligned with the great-circle path between two seismic stations. Given recordings of this earthquake at the two stations, the average phase velocity $c(\omega)$ of the fundamental-mode Rayleigh wave along the propagation path between the two stations can be estimated from the cross-correlation of the vertical-component (or radial-component) seismograms by

$$c(\omega) = \frac{\omega(x_2 - x_1)}{\phi(\omega) + 2n\pi}, \quad (3.5)$$

where x_1 and x_2 are the epicentral distances between the event and the two stations, $\phi(\omega)$ is the phase spectrum of the cross-correlation, and ω is angular frequency.

Note that the phase delay suffers from an intrinsic circular ambiguity of 2π (for integer n). This ambiguity is particularly prominent in phase-velocity space for short periods and large station separations. Typically, the ambiguity is accounted for by determining the dispersion curve branch (for integer n) that best matches a reference dispersion curve either through manual selection (e.g. Meier et al., 2004) or automation based on specific

tuning criteria (e.g. Soomro et al., 2016). By inverting the phase spectra directly, we avoid the selection of the correct dispersion branch (n value). Furthermore, we do not depend on a reference model other than to define relatively wide uniform prior bounds on V_S (as discussed above). To calculate predicted phase spectra, we start by using the Computer Programs in Seismology package of Herrmann (2013) to calculate predicted dispersion curves, then use equation (3.5) to convert the dispersion into a wrapped phase spectrum (i.e., such that $\phi(\omega) \in (-\pi, \pi]$).

One obvious problem with applying equation (3.3) as the likelihood is that the misfit (negative log-likelihood) surface will be infinitely multi-modal. Fortunately, surface-wave phase-velocity sensitivity is an integrated property with significant overlap in depth sensitivity between periods (see Appendix). Consequently, dispersion curves are inherently smooth, and most phase-velocity dispersion curves are monotonic. By enforcing realistic prior bounds on the deepest structure of the model, the long-period dispersion is restricted to the correct mode of the misfit surface (or 2π branch) via the physics of the forward modelling. The overlapping nature of depth sensitivity between dispersion at adjacent periods propagates this restriction (to some extent) to shallower depths.

3.3 Surface-wave dispersion over British Columbia, Canada

This section discusses the application of our method to estimate 1D, path-averaged, Earth structure and surface-wave dispersion from phase measurements collected from stations over British Columbia, Canada (fig. 3.1). The region consists of multiple active tectonic boundaries with histories of large earthquakes and associated tsunamis (e.g., Nelson et al., 1995; Cassidy et al., 2014). Furthermore, British Columbia spans the entire region from the coast, through the Canadian Cordillera, to the edge of the North American craton. Our study is motivated by attempting to improve our understanding of the transitions between these different tectonic modes, and their impact on Cordilleran evolution. Furthermore, we note the general sparse coverage of seismic stations throughout the interior of the

province, which makes it challenging to study this entire region by single-station-type seismic observations (e.g., receiver functions or shear-wave splitting). Dispersion estimated from this work represents valuable data for studying the tectonics of this region.

Southwestern British Columbia, including Vancouver Island (VI), sit above the northern terminus of the Cascadia subduction zone, where the oceanic Juan de Fuca (JdF) plate subducts beneath North America (NA). Historically, Cascadia has produced large megathrust earthquakes and associated tsunamis (Nelson et al., 1995; Satake et al., 1996, 2003), and is capable of future large events. North of Cascadia, the HG archipelago sits above the transpressive boundary between the Pacific (PA) and the NA plates, which produces significant seismicity including Canada’s two largest instrumentally recorded earthquakes (Bostwick, 1984; Cassidy et al., 2014). The region between HG and VI consists of a complex triple junction, transform faulting, as well as the breakup and possible capture (by NA) of the Explorer (Ex) microplate (Braunmiller & Nábelek, 2002; Audet et al., 2008; Hutchinson et al., 2019; Savard et al., 2020). Further inland, the Canadian Cordillera consists of a series of (approximately) margin-parallel morphological belts containing various exotic terranes that accreted to the North American continent (Nelson et al., 2013). The lithosphere of inland British Columbia records a long (~ 750 My.) history of deformation and evolution, is enigmatically hot and thin (Hyndman, 2017), and supports high topography throughout the province (fig. 3.1).

3.3.1 Surface-wave phase spectra

Surface-wave dispersion measurements obtained from recordings of large earthquakes are useful for studies of tectonic structure since the surface wave energy is released over period ranges that are sensitive to structure at crustal and uppermost mantle depths (see Appendix). Typically, single-station measurements of surface-wave phase velocities necessitate information on the earthquake source-time function. This challenge is easily overcome by the analysis of the cross-correlation from two seismograms of the same event, assuming the event is aligned with the great-circle path between the stations, since the station that is nearest to the event acts as a source relative to the other station. In this work, we analyse

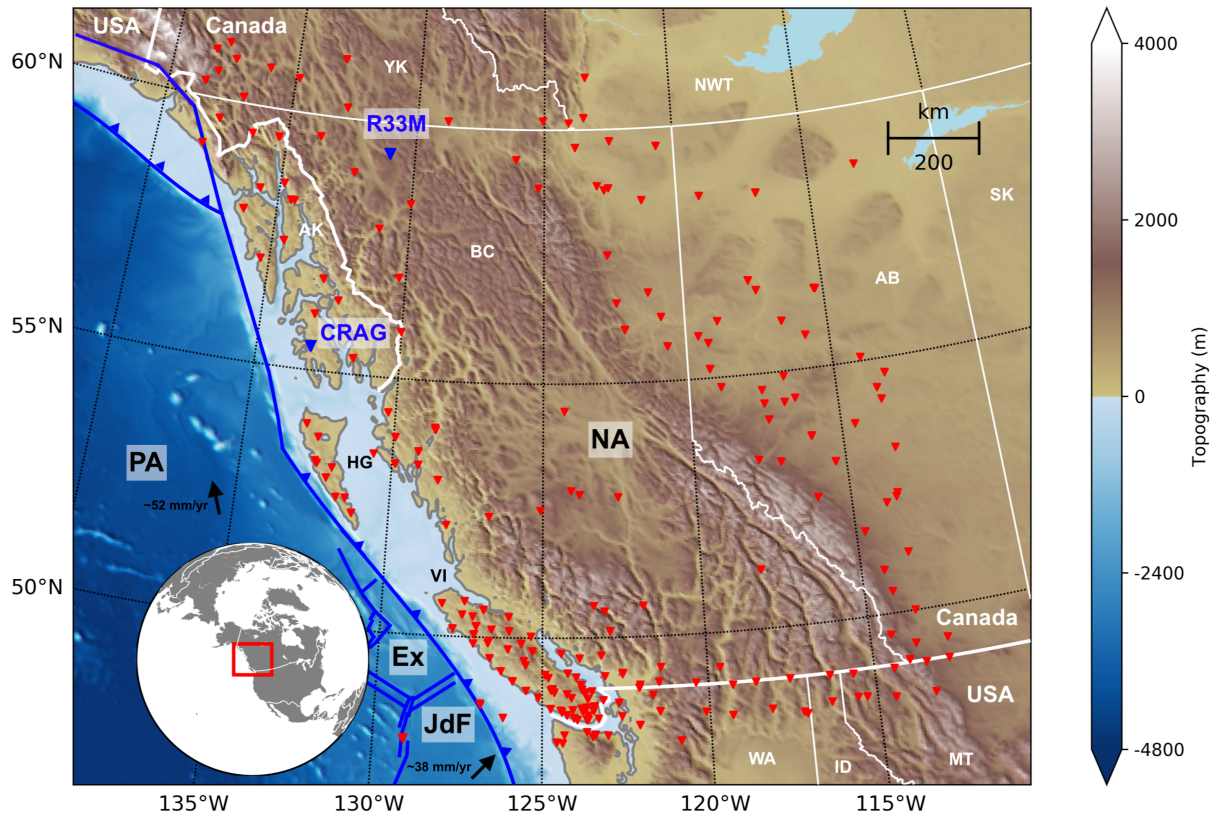


Figure 3.1: Seismic station coverage over British Columbia (BC), Canada. The locations of broadband seismic stations considered in this work are shown as inverted red triangles. The locations of stations CRAG and R33M are shown by inverted blue triangles. Tectonic boundaries are shown in blue. PA, JdF, and NA represent the Pacific, Juan de Fuca, and North American tectonic plates, respectively. Ex represent the Explorer microplate. Tectonic plate motions are shown by black arrows and are relative to NA (Kreemer et al., 2003). HG and VI represent Haida Gwaii and Vancouver Island, respectively. Political boundaries are shown in white. YK, NWT, AK, AB, SK, WA, ID, and MT represent Yukon, Northwest Territories, Alaska, Alberta, Saskatchewan, Washington, Idaho, and Montana, respectively.

recordings from over 9000 teleseismic earthquakes, between January 2000 and December 2020, recorded at over 300 broadband stations (fig. 3.1). We only consider station pairs with at least 2 years of concurrent recordings in order to ensure recording of a sufficient number of earthquakes.

We consider events with $M > 5.5$, epicentral distances $> 20^\circ$ such that the surface waves can be approximated as planar wavefronts, and epicentral distances $< 150^\circ$ such that the surface waves arrive sufficiently before the complementary waves travelling along the far side of the Earth (fig. 3.2). For each station pair, we consider events with back-azimuths for both stations within 8° of the great-circle path azimuth between the stations. For comparison, Meier et al. (2004) and Soomro et al. (2016) consider a similar threshold of 7° , whereas Magrini et al. (2020) use a threshold of 5° . We extract vertical-component seismograms for each event beginning 10 minutes before the event and ending 10 minutes after a theoretical slow surface wave travelling at 2 km/s. Next, we apply a zero-phase bandpass filter for periods 1-200 s to suppress high-frequency noise and long-period instabilities, and remove the instrument response from each seismogram. Ultimately, the bandpass filter is purely for visualization purposes as additional bandpass filters are applied at unique frequencies for phase measurement. We define the approximate surface-wave time window based on a conservative range of velocities (2.5-4.5 km/s), and calculate the signal-to-noise ratio (SNR) for each seismogram for this window relative to pre-event noise at 10 unique periods (discussed below). We only consider events with $\text{SNR} > 2$ dB on both seismograms. Examples of surface-wave recordings at stations CRAG and R33M are shown in figure 3.3. The locations of these stations are shown in figure 3.1.

As previously mentioned, we process the data at 10 unique periods by applying a series of narrow band-pass filters. Assuming the vertical-component recording of the surface-wave arrival is dominated by the fundamental-mode Rayleigh wave, a single surface-wave phase delay measurement (for a given earthquake at a given period) is taken as the phase of the cross-correlation of the two filtered seismograms at the filtered period. Previous work consider time-domain windowing in order to isolate the dominant (fundamental) mode arrival (e.g. Meier et al., 2004). Testing the use of such time windows produces similar results as processing without time windows (in this application). We do not apply time

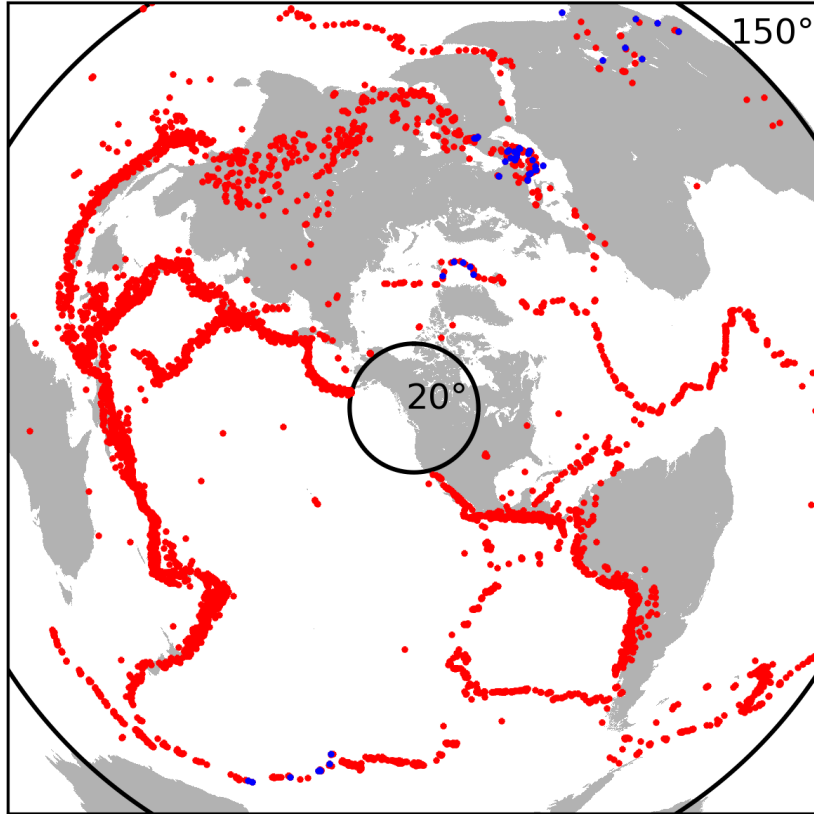


Figure 3.2: Earthquake locations for two-station phase spectra calculation. The locations of all events considered in this work are shown in red. The locations of events considered for phase spectra calculation between CRAG and R33M are shown in blue.

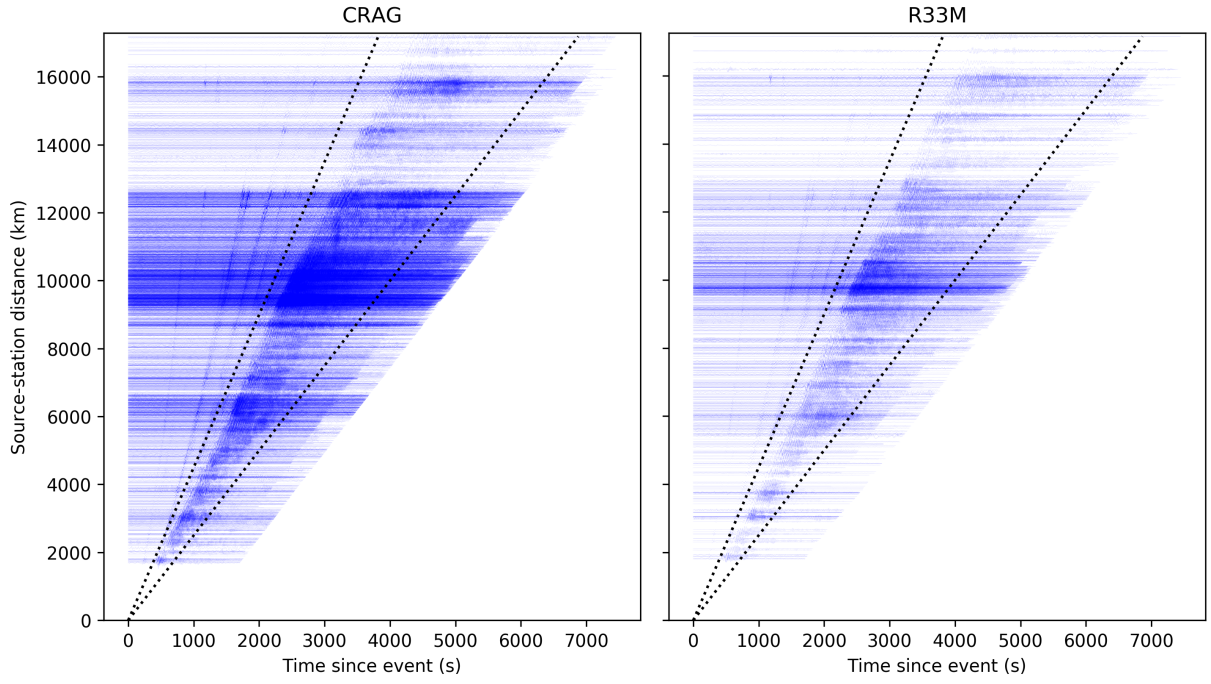


Figure 3.3: Earthquake recordings at stations CRAG and R33M. Note that only a subset of events that are approximately aligned with the great-circle path between CRAG and R33M are analyzed. All recorded events are shown for illustration purposes. The surface wave arrival window is approximated by minimum and maximum surface-wave velocities of 2.5 and 4.5 km/s, respectively (illustrated by black dotted lines).

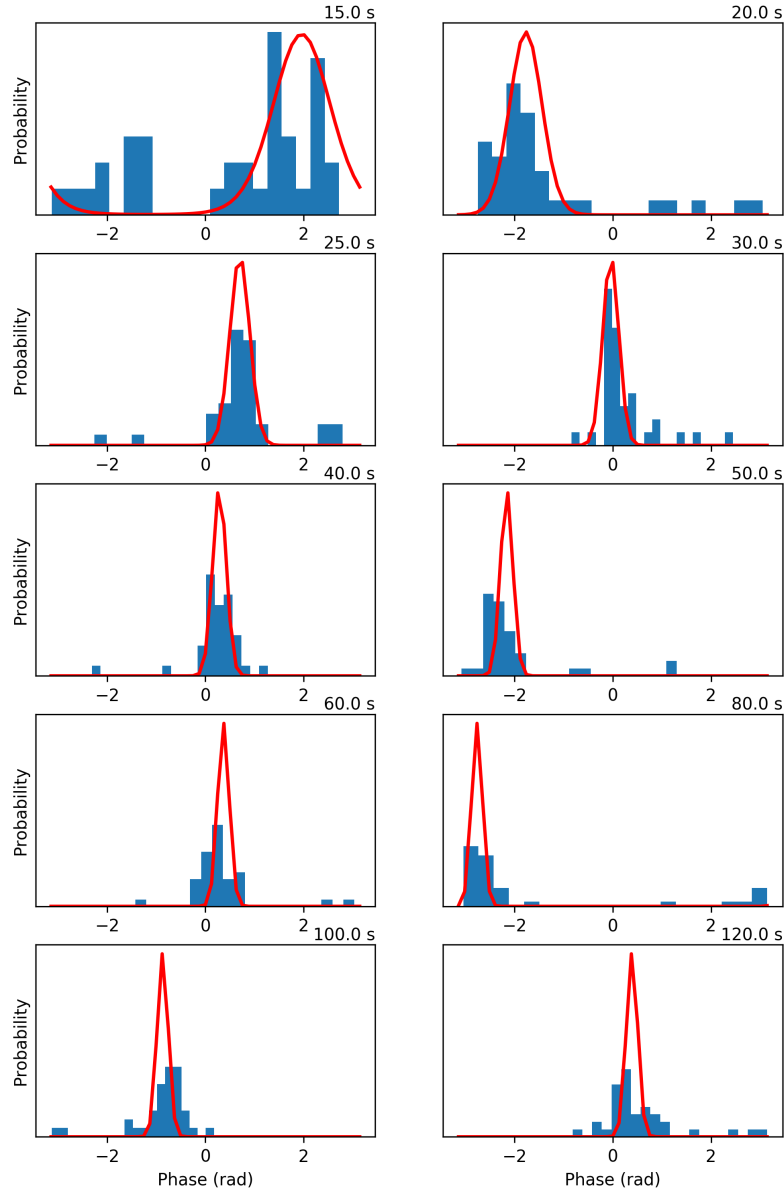


Figure 3.4: Surface-wave phase spectra between stations CRAG and R33M. Blue histograms represent the distribution of surface-wave phase measurements at ten periods (shown above each panel) from all earthquakes for this station pair. The red line represents the distribution of predicted phase spectra from the ensemble of rjMCMC samples.

windows in our final processing. The period range of surface wave energy recorded at each station can vary due to many factors including event depth, magnitude, and focal mechanism. Furthermore, the event distance and specific propagation path can introduce attenuation and scattering that affect the surface-wave spectra. As with the 2π ambiguity from equation (3.5) discussed above, the valid period range of the dispersion for a specific event is typically selected manually (e.g. Meier et al., 2004) or automatically based on specific tuning criteria (e.g. Soomro et al., 2016). We consider every period independently and base the valid period range of the dispersion on the SNR of the seismograms (as discussed above) in order to avoid this selection process. Phase measurements between stations CRAG and R33M are shown in figure 3.4.

The surface-wave propagation path between a given station pair is unique for each earthquake since the events are not perfectly aligned with the stations (i.e., x_1 and x_2 in equation (3.5) are unique for each event). Furthermore, the velocity structure of the Earth is not homogeneous along the propagation path of the surface wave, which causes deviations from the great-circle path assumption. Magrini et al. (2020) showed that the correct surface-wave arrival azimuth can be estimated from the horizontal-component rotation angle that optimizes the $\pi/2$ phase shift between radial and vertical seismograms. For most stations, we find that horizontal component seismograms have significantly lower SNR than the vertical component. The reduced SNR significantly limits the number of events that this azimuth correction can be applied to for each station pair. Consequently, we do not consider horizontal component data in this work. In order to estimate properties of the circular data error model, we assume that the propagation path length between station pairs is the same for all events (i.e., $x_2 - x_1 = \text{inter-station distance}$). Although this assumption introduces bias due to the unknown surface-wave arrival angle, recent studies have shown that dynamic scattering and diffraction effects (leading to non-planar wave propagation) likely introduce errors that are several times larger than those due to simple geometrical assumptions (e.g. Kolínský et al., 2021, and references therein).

The Bayesian inversion applied in this work requires knowledge of the statistical distribution of data errors. For every station pair and every period, we use the maximum-likelihood estimates of μ and κ of a von Mises distribution based on the set of phase

measurements from all suitable events as the input to the inversion. Specifically, the estimated μ values represent observed data. The distribution concentrations are then fixed for the inversion process. In our implementation, predicted phase measurements according to equation (3.5) assume the difference in propagation path lengths between stations is equivalent to the inter-station distance (as mentioned above). An alternative strategy would be to treat the measurements from each event independently (with different propagation path lengths and accounting for arrival azimuth). However, properties of the statistical distribution of the data errors (i.e., κ) could not be estimated from the data in this case, and we choose to adopt the former strategy.

The issues of propagation distance and arrival azimuth may be overcome by formulating the inverse problem to solve directly for 3D structure from measured phase spectra. This assumes the use of an appropriate forward model. Given the computational demand of such forward modelling, Bayesian approaches that generally require thousands (or millions) of data predictions may not be suitable.

3.3.2 Path-averaged Earth structure and surface-wave dispersion

The results of the Bayesian inversion for path-averaged Earth structure between stations CRAG and R33M are shown in figure 3.5. The inversion is trans-dimensional in that the number of layers in the Earth model is treated as an unknown. Based on these results, relatively simple structure (i.e., a smaller number of layers) is required to fit the observed phase data (fig. 3.5a). Although we model V_S structure to 400 km depth, our results indicate that structure at depths > 200 km are effectively unconstrained. This is expected given the periods considered in this work. However, the inversion does recover low near-surface V_S values, as well as potentially discontinuous structure at ~ 30 km depth (fig. 3.5b). However, it is well known that surface-wave dispersion data are insensitive to discontinuous structure (e.g. Lebedev et al., 2013), and the estimated uncertainty on such structure from our results is large. Because this represents path-averaged Earth structure, it is not surprising that the recovered model is relatively simple. Consequently, we do not attempt any further interpretation of structure from these results. Additional analysis

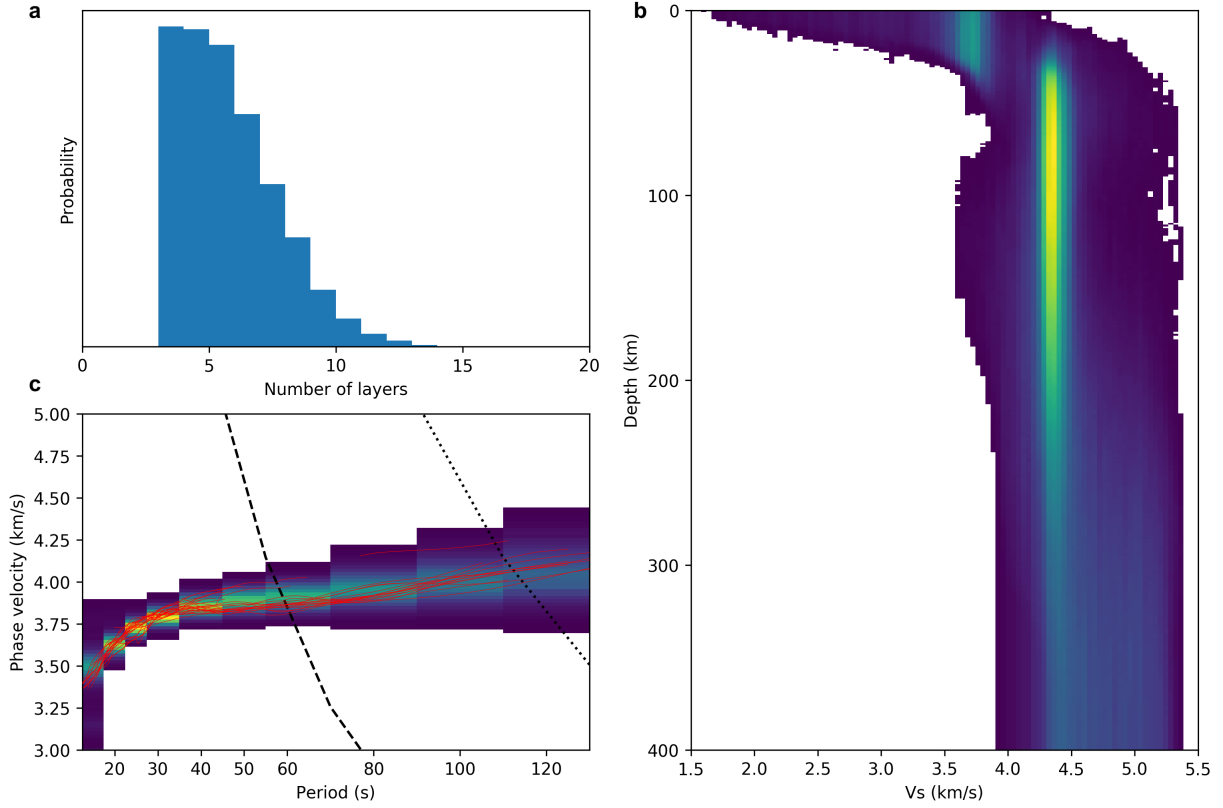


Figure 3.5: Bayesian inversion results for station pair CRAG and R33M. The distribution for the number of layers in the model from the rjMcMC samples is shown in (a). The marginal V_S density profile is shown in (b), with yellow and blue representing high and low probability, respectively. The predicted dispersion for the rjMcMC samples are shown in (c). Manually-selected dispersion curves (using the method of Meier et al., 2004) for individual earthquakes for this station pair are shown in red for comparison. The black dotted and dashed lines represent 1 and 2 wavelength contours, respectively.

(e.g., tomography) is required to fully address the tectonic implications of our results, which is beyond the scope of this work.

The predicted dispersion between stations CRAG and R33M from the *rjMcMC* sampling is shown in figure 3.5c. For comparison, we also apply the method of Meier et al. (2004) to manually select dispersion curves for this station pair. Manually selected dispersion curves for each event are shown in red. We note the reasonable agreement between the manually selected dispersion and the predicted dispersion, which is an additional result of our inversion. The uncertainty in the predicted dispersion can be similarly quantified from the ensemble of *rjMcMC* samples, and propagated into subsequent analysis. We note that the number of events (and period-band for each event) is not identical between the manually selected dispersion approach and our probabilistic approach. Furthermore, we note that the manually-selected dispersion curves include significantly finer sampling in period space (compared to the 10 periods considered in this work), and qualitatively resolve more detail in the dispersion curve. This difference could be overcome by considering surface wave phase measurements at additional (finely-spaced) periods in our inversion procedure. The fit to the observed phase measurements is shown in figure 3.4.

Reliable dispersion curve estimation from time-domain cross-correlations traditionally requires propagation distances greater than 2-3 wavelengths for the desired period, based on far-field assumptions. Hence, there is typically a maximum reliable period for dispersion estimation for each station pair based on the distance between the pair of stations. These thresholds can be represented as lines in the phase-velocity dispersion space (fig. 3.5c). We note that the relatively short station separation (456 km) between CRAG and R33M results in relatively large uncertainties for long-period dispersion. Recent studies show that reliable dispersion can be estimated for station separations of ~ 1 wavelength (e.g., Luo et al., 2015). However, this applies to dispersion from ambient noise cross-correlations rather than cross-correlations of recordings of distant earthquakes. Although dispersion measurements may not be as reliable up to the ~ 1 wavelength threshold, these data may still be incorporated into tomographic inversions in order to improve path coverage since they may be appropriately weighted according to their larger uncertainties.

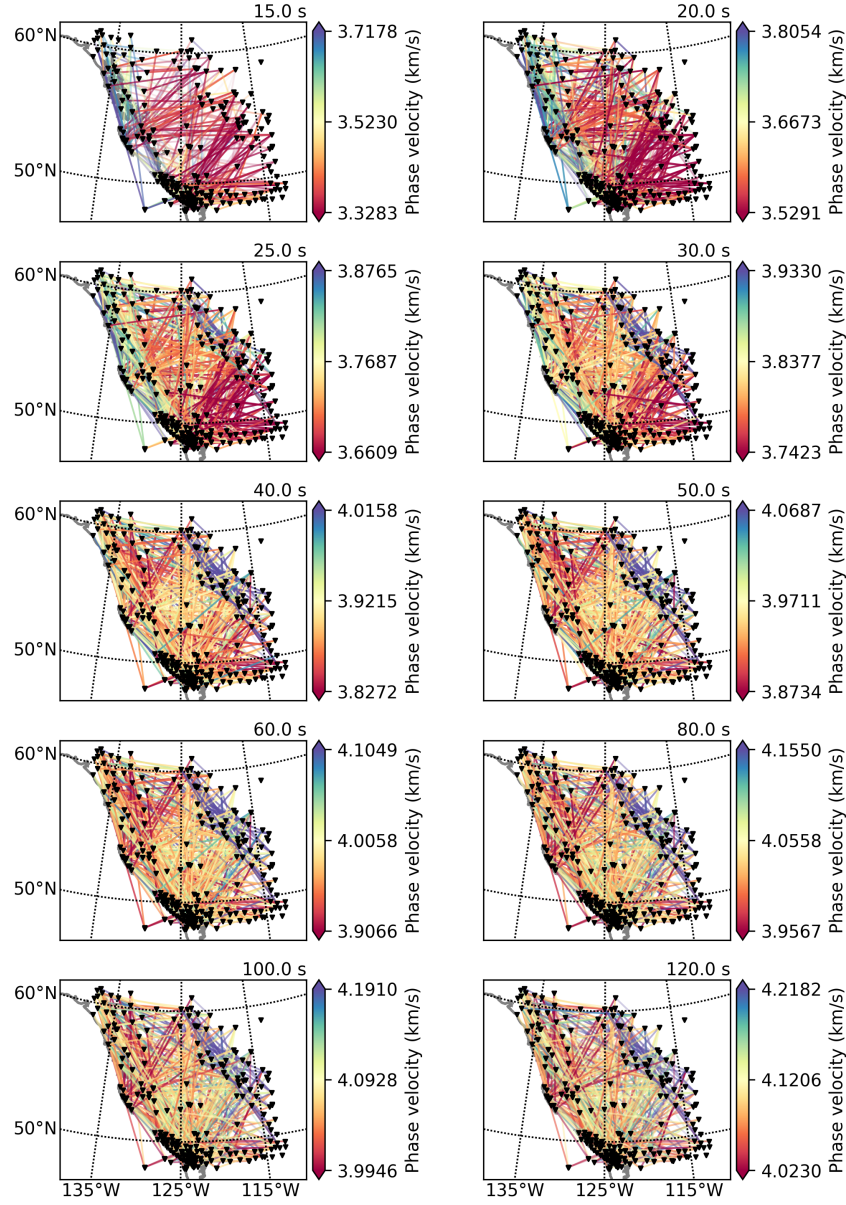


Figure 3.6: Bayesian inversion results for all station pairs over British Columbia. Paths are coloured according to the estimated phase velocity at the corresponding period for each station pair. Only a random subset of 5 % of paths are shown at each period, for illustration purposes. Station locations are shown by black triangles.

We use our Bayesian inversion approach to estimate path-averaged Earth structure (and dispersion) for all available station pairs. In total, measured phase spectra from ~ 10000 station pairs over British Columbia are inverted to estimate dispersion curves and associated uncertainties (fig. 3.6). For each station pair, we estimate the dispersion and associated uncertainty as the mean and standard deviation of the predicted surface-wave dispersion from the ensemble of rjMcMC samples. Distributions of estimated phase velocities and uncertainties for all station pairs (at each period) are shown in the Appendix.

As previously mentioned, it is challenging to address tectonic implications without further analysis, which is beyond the scope of this work. However, we point out two interesting observations from figure 3.6. First, at shorter periods (which are sensitive to shallower structure) there is a distinction between higher velocities along the coast compared to inland. This may be attributed to the thick sedimentary sequences over the western Canada sedimentary basin (the low topography at the eastern edge of the study region; fig. 3.1), since paths associated with stations in this region generally show lower short-period phase velocity. Alternatively, this may be attributed with variations in Moho depth across the province and/or the possible subduction of PA beneath NA along the transpressive HG region (e.g., Hyndman, 2015; Gosselin et al., 2015, and references therein). This observation is consistent with recent surface-wave group-velocity tomography over the region (Gosselin et al., 2021) at slightly longer periods (30-40 s). This is likely because fundamental-mode surface-wave phase velocities have positive V_S sensitivity to greater depths (for the same period). The second observation we make is that paths along the eastern edge of the study region generally show the highest phase velocities at the longest periods considered in this work. This may be attributed to the juxtaposition between the thin lithosphere of the Canadian Cordillera and the thick lithosphere of the NA craton to the east (Schaeffer & Lebedev, 2014). These observations provide confidence that our approach recovers tectonic structure.

3.4 Conclusions

The analysis of circular measurements requires special consideration. This is particularly important in geophysical inversion, which commonly assumes Gaussian-distributed data errors. Bayesian inversion is a probabilistic approach to inverse problems that allows for rigorous uncertainty quantification, and requires knowledge or estimation of the statistical distribution of data errors. In this work, we incorporate a meaningful data error model for Bayesian trans-dimensional inversion of circular data. Furthermore, we estimate the locations and concentrations of the distributions from the data.

We apply this idea to the inversion of two-station surface-wave phase delay measurements to estimate 1D path-averaged Earth structure. Surface-wave dispersion and associated uncertainties are useful additional results of the inversion. Consequently, our approach provides a means to objectively automate phase-velocity dispersion estimation, which is often estimated by (subjective) manual selection or automation with complicated measurement criteria. The method is applied to data from long-running stations over British Columbia, Canada, which is a region of varying and complex tectonic modes. We demonstrate that the approach estimates surface-wave dispersion that is consistent with manually selected measurements, while requiring less analysis time and without dependence on a reference model. We observe spatial patterns in the estimated dispersion from all station pairs that may have tectonic implications, further validating the approach. The surface-wave dispersion estimated from this work represents valuable data (for consideration in future work) for studying the tectonics of this region.

We note shortcomings with our application to two-station dispersion estimation in the potential bias due to the unknown propagation distance and arrival azimuth, and present several alternative strategies that represent future areas of research. Specifically, the data error model applied in this work does not necessitate the assumption of 1D path-averaged structure. Lastly, the ideas presented in this work are not limited to the inversion of surface-wave phase spectra from teleseismic earthquakes, but are easily applicable to phase spectra derived from ambient noise. Furthermore, these ideas can be considered for Bayesian geophysical inversion of any circular quantities.

Part II

Spatial and temporal variability of subduction zone structure from receiver function analysis

Chapter 4

A brief review of receiver functions

Receiver function (RF) is a generic term appropriated by seismologists to refer to an approximation of the Earth's impulse response (reflectivity structure) beneath a seismic station from a teleseismic body-wave source. Beginning with the works by Phinney (1964), Burdick & Langston (1977), and Langston (1979), the calculation of RFs generated from land-based teleseismic recordings has been one of the most important methods to infer/image crustal and upper-mantle structure over the past several decades. In this section, I review the basic theory of RF estimation and structural inference/interpretation. RFs have been calculated for various wave types. However, below I focus on the original theory applicable to direct compressional-wave (P) phases. Furthermore, because chapters 5 and 6 focus on tectonic (rather than technical) questions, below I review important observations of RFs at subductions zones.

4.1 Teleseismic recordings

Distant earthquakes produce many phases that are observed on teleseismic recordings, depending on the distance between the event and the recording station (fig. 4.1). At distances greater than $\sim 30^\circ$, a direct P phase travels deep through the Earth's mantle and

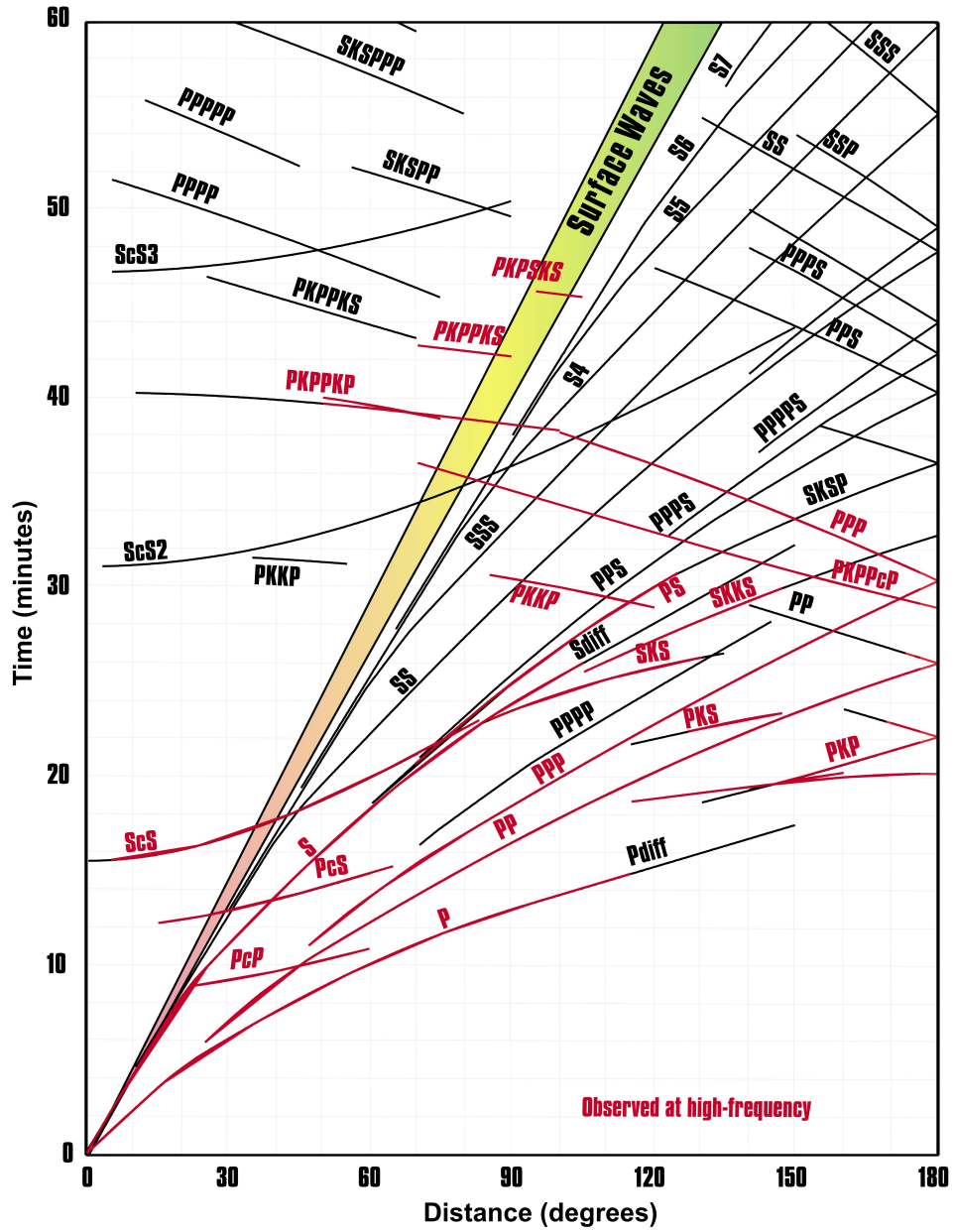


Figure 4.1: Iasp91 global Earth model (Kennett & Engdahl, 1991) earthquake phase travel-time predictions (figure adapted from Astiz et al., 1996).

arrives nearly vertically beneath a seismograph station (and can be approximated by a planar wavefront). Furthermore, the direct P phase (and subsequent coda) arrives several minutes before any other phases, and can be observed and analysed without interference. At distances greater than $\sim 100^\circ$, the direct P phase is typically no longer observed as the station is in a “shadow zone” relative to the earthquake due to refraction of the wave at the core-mantle boundary.

Typically, only moderate to large magnitude ($M > 5$) earthquakes are considered for RF analysis, since smaller events have a poor signal-to-noise ratio (SNR) at the required distances. However, larger events involve greater rupture area and duration, leading to complicated source-time functions (and ultimately more complicated teleseismic P-wave recordings). The Earth’s structure is not homogeneous near the earthquake, as well as along the propagation path through the Earth. This causes refractions and scattering that alters the waveform. Consequently, such source and path effects represent unwanted contributions to the teleseismic recording for RF analysis, which aims to isolate the scattering structure near (beneath) the seismograph station (fig. 4.2a).

Assuming the P wave (containing source and path effects) arrives beneath the station as a nearly vertically propagating planar wavefront, some energy will be converted to shear (S) waves at seismic velocity discontinuities beneath the station (fig. 4.2b). RF calculation involves isolating the amplitude and timing of these P-to-S converted waves (fig. 4.2c). The amplitudes of these phases depend on the incidence angle of the approaching P wave (which depends on earthquake depth and distance) as well as the magnitude of the seismic velocity contrast that generated the conversion. Similarly, the arrival times of these phases depend on the depth of the seismic velocity contrast, the seismic velocity structure overlying the contrast, as well as the incidence angle of the approaching P wave (which is often represented by the ray parameter or slowness). Lastly, the frequency content of these arrivals depend on the spatial nature of the seismic velocity contrast (i.e., distributed broadly vs. sharp), as well as the frequency content of the incident P wave.

These locally generated S waves propagate nearly vertically, and are observed on horizontal component recordings (at 3-component seismograph stations). Approximating the

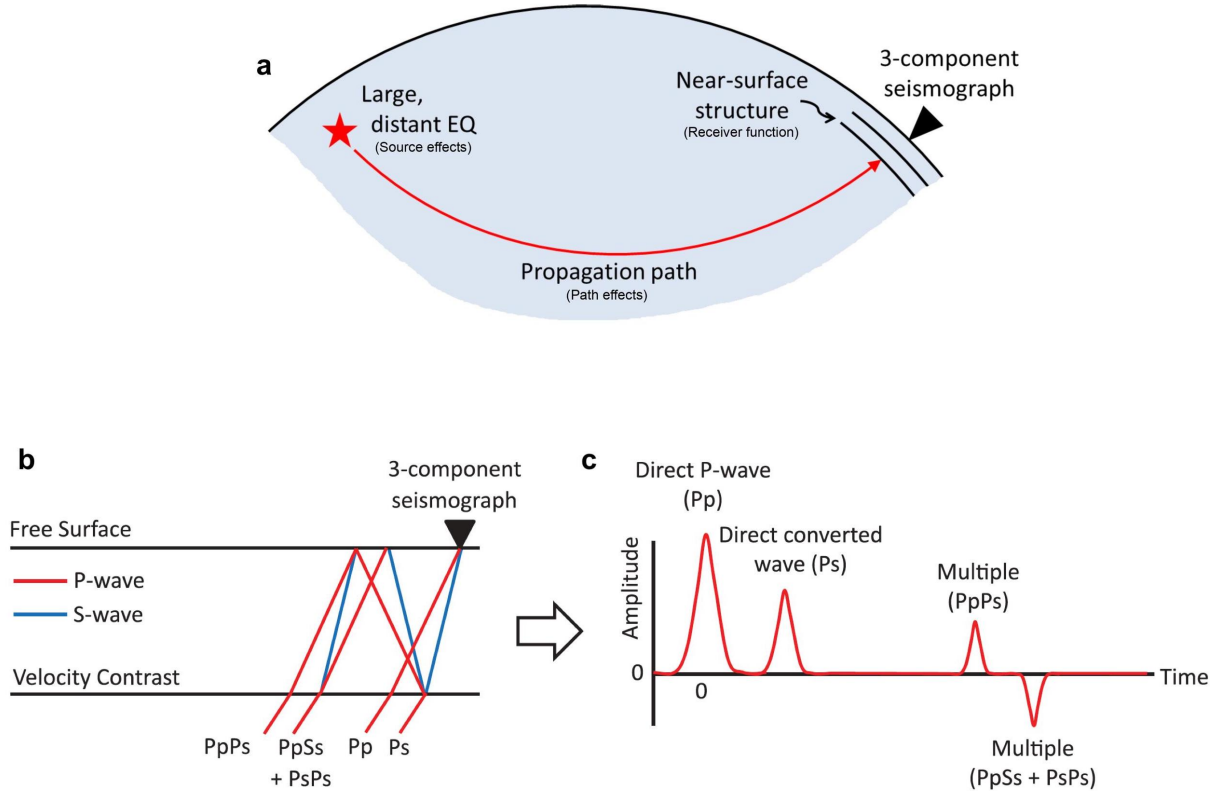


Figure 4.2: Receiver function overview. Illustration of the contributions, and travel path, of a teleseismic P wave (a). The P wave arrives nearly vertically beneath the seismograph station. The incident planar P wavefront is converted to S waves at velocity contrasts beneath the station (b). A receiver function is estimated by isolating the amplitude and timing of the locally generated S waves (c).

vertical component as a source wavelet incident on the receiver-side structure allows redundant information about earthquake rupture and propagation path effects (recorded on multiple components) to be exploited and corrected for. Deconvolving the vertical component from the horizontal components equalizes these effects, and isolates the response for local P-to-S conversions.

4.2 Deconvolution

Consider a teleseismic P-wave recording with vertical and radial (i.e., horizontal component in-line with the distant earthquake) time series given by $z(t)$ and $r(t)$, respectively. In theory, approximating the vertical component as a source wavelet and deconvolving it from the horizontal component leads to

$$RF_R(\omega) = \frac{R(\omega)}{Z(\omega)}, \quad (4.1)$$

where $Z(\omega)$ and $R(\omega)$ are the Fourier transforms of the vertical and radial recordings, respectively. $RF_R(\omega)$ is the Fourier transform of the radial RF, and ω is angular frequency. The same theory applies to the transverse horizontal component.

In practice, in the presence of seismic noise, $Z(\omega)$ can be small and the deconvolution unstable. The most common approach to stabilize the deconvolution was introduced by Langston (1979), who applied the water-level technique of Clayton & Wiggins (1976). With this technique, the deconvolution for the RF is given by

$$RF_R(\omega) = \frac{R(\omega)\overline{Z(\omega)}}{\phi(\omega)}G(\omega), \quad (4.2)$$

where $\overline{Z(\omega)}$ is the complex conjugate of $Z(\omega)$,

$$G(\omega) = \xi \exp \left[\frac{-\omega^2}{4a^2} \right] \quad (4.3)$$

is a Gaussian filter with normalization ξ , and

$$\phi(\omega) = \max \left[|Z(\omega)|^2, c \max [|Z(\omega)|^2] \right]. \quad (4.4)$$

The parameter c is the “water level” that controls the minimum allowable value in the denominator of the deconvolution in equation (4.2). The parameter a in equation (4.3) controls the width of the low-pass Gaussian filter, and consequently the frequency content of the RF. The choices for c and a are subjective, and are typically based on visual inspection of the RF.

The water-level deconvolution technique implicitly assumes that the seismic noise spectrum is flat (i.e., white noise). This is not true in practice, particularly at the frequencies of interest for RF studies (~ 0.1 to ~ 1 Hz for teleseismic P waves) due to microseism peaks. Furthermore, the seismic noise spectrum is not constant in time, which can result in biased RF estimation for constant water level parameter choices. Instead, both chapter 5 and chapter 6 of this thesis apply a Wiener deconvolution approach (Audet, 2010), in order to minimize the effect of seismic noise. In this implementation of Wiener deconvolution, the stabilization during the spectral division is estimated from the pre-event noise spectrum based on the covariance of noise on multiple channels. With this technique, the RF is given by

$$RF_R(\omega) = \frac{R(\omega)\overline{Z(\omega)}}{\phi(\omega)}. \quad (4.5)$$

In this case, the denominator $\phi(\omega)$ is given by

$$\phi(\omega) = |Z(\omega)|^2 + \left[\frac{1}{4}|N_Z(\omega)|^2 + \frac{1}{4}|N_R(\omega)|^2 + \frac{1}{2}\overline{N_Z(\omega)}N_R(\omega) \right], \quad (4.6)$$

where $N_Z(\omega)$ and $N_R(\omega)$ are the Fourier transforms of the pre-event noise recordings on the vertical and radial channels, respectively. The RF is subsequently filtered for an appropriate frequency band.

There exist many other techniques to solve the deconvolution problem for RF analysis. I will not exhaustively review them here. Rather, it is worth mentioning that some techniques consider the problem entirely in the time domain and treat the deconvolution as an inverse problem. The iterative time-domain technique of Kikuchi & Kanamori (1982), introduced to RF estimation by Ligorria & Ammon (1999), attempts to estimate the RF as the summation of many Gaussian pulses of varying time offset, width, and amplitude.

The quality of the RF is then assessed by the fit between the observed horizontal component and the predicted horizontal component (from the convolution between the vertical component and the estimated RF). As its name suggests, the technique initially fits the dominant arrivals in the seismogram then iteratively refines the details in the waveform. Similarly, Kolb & Lekić (2014) consider the deconvolution problem as an inverse problem. Rather than adding Gaussian pulses to estimate the RF waveform through an iterative procedure, they apply trans-dimensional Bayesian inversion (discussed in Part I of this thesis) to estimate the RF and associated uncertainties as properties of a posterior probability density.

4.3 Interpretation

The RF is an approximation of the impulse response (i.e., the response to a Dirac delta function) of Earth structure beneath a seismic station. In reality, the RF has finite bandwidth, and the vertical component is only an approximation of the incident source wavelet. The phases in the RF represent impulses of local P-to-S conversions beneath the seismic station. As was discussed above, the amplitudes of these phases are sensitive to the magnitude of velocity contrasts that produce P-to-S conversions, and the arrival times of these phases are sensitive to the depth of the velocity contrast as well as the travel-time average velocity of the overlying structure. Consequently, RFs provide constraint on the discontinuous seismic velocity structure beneath a seismic station. For simple structure, inferences can be made by visually inspecting the RFs.

It can be challenging to identify the individual direct converted (Ps) and multiple (PpPs and PpSs; fig. 4.2c) phases in the presence of noise and competing signals. A common practice is to stack (average) multiple RFs computed for different earthquakes recorded at the same station such that incoherent noise destructively interferes, which improves the SNR of the stacked RF. RFs are typically stacked in bins according to earthquake location. However, RFs are not dependent on earthquake direction (only weakly dependent on depth and distance) for isotropic seismic velocity structure with horizontal discontinuities. Zhu

& Kanamori (2000) proposed a technique to estimate crustal thickness H and bulk crustal V_P/V_S ratio k beneath a seismic station, assuming simple structure (i.e., isotropic with flat layering) with a single dominant velocity contrast (i.e., the Moho). Rather than stacking RFs in the time domain, Zhu & Kanamori (2000) proposed stacking RFs in the H - k domain. The differential travel times of the direct conversion and multiples, relative to the direct P arrival, suffer a tradeoff between crustal thickness and crustal seismic velocities. However, the tradeoffs are unique for each phase in the H - k domain and stack coherently for the optimal values of H and k . The H - k method, and similar variants, are widely used. However, these techniques assume relatively simple Earth structure.

For more complicated structure with many seismic velocity discontinuities, other approaches are typically required. A common approach for inferring Earth structure from RF data is to estimate parameters of a 1D Earth model that predicts RF waveforms that match observed RF waveforms via inversion. The topic of inverse theory has been reviewed in chapter 1 of this thesis, and will not be reviewed again here. However, it is worth noting that a wide range of inverse methods have been applied to estimate Earth structure from RF data ranging from linearized techniques (e.g., Owens et al., 1984) to trans-dimensional Bayesian sampling (e.g., Piana Agostinetti & Malinverno, 2010). Many studies now simultaneously consider RF data and surface-wave dispersion (discussed in chapters 2 and 3) through joint inversions (e.g., Bodin et al., 2012), as the two data types are complementary. Constraint of dipping interfaces and anisotropy via inversion necessitates many additional model parameters (of which many are often poorly resolved), as well as (relatively) computationally expensive generalized ray-tracing forward models (e.g., Frederiksen & Bostock, 2000). Consequently, the majority of inverse problems considering RF data assume isotropic velocity structure with flat layering. Studies that consider inversion of RF data to resolve dipping structure or anisotropy often target only a portion (e.g., one layer) of a 1D Earth model (e.g., Porter et al., 2011).

Teleseismic P-wave energy is confined to the radial plane for isotropic structure with horizontal discontinuities. Consequently, local P-to-S converted waves appear only on the radial component, and not on the transverse component. In the presence of material anisotropy or dipping structure, some of the teleseismic P-waveform energy can be refracted

Table 4.1: 1D Earth model with a dipping layer (from Porter et al., 2011).

Thickness (km)	V_P (km/s)	V_S (km/s)	Strike ($^{\circ}$)	Dip ($^{\circ}$)
20	6.4	3.66	0	0
5	5.8	3.314	90	20
-	7.8	4.48	90	20

Table 4.2: 1D Earth model with an anisotropic layer (from Porter et al., 2011).

Thickness (km)	V_P (km/s)	V_S (km/s)	Anisotropy (%)	Trend ($^{\circ}$)	Plunge ($^{\circ}$)
20	6.4	3.66	0	0	0
5	6.2	3.54	-20	180	45
-	7.8	4.48	0	0	0

out of the radial plane and onto the transverse plane. The observation of coherent phases on transverse-component RFs is indicative of anisotropy and/or dipping structure beneath a seismic station. For example, table 4.1 lists parameters of a simple 1D Earth model with a dipping low-velocity layer (LVL). The corresponding RFs (without multiples) predicted for this model for various earthquake directions (back-azimuths; BAZ) are shown in figure 4.3. Similarly, table 4.2 lists parameters of a 1D Earth model with an anisotropic layer (from Porter et al., 2011), with the corresponding predicted RFs shown in figure 4.4. Both models have uniform density of 2800 kg/m^3 . Coherent phases are visible on transverse RFs for these examples.

Figures 4.3 and 4.4 illustrate that the amplitude and (to a lesser extent) the timing of phases in both the radial and transverse RFs vary as a function of BAZ in the presence of material anisotropy or dipping structure. Specifically, the phase amplitude and timing depend on the orientation/geometry of the structure (i.e., strike of dipping structure or trend of anisotropy) relative to the earthquake BAZ. Forward modelling of RFs for various Earth models with dipping structure and anisotropy suggests that anisotropy with a horizontal

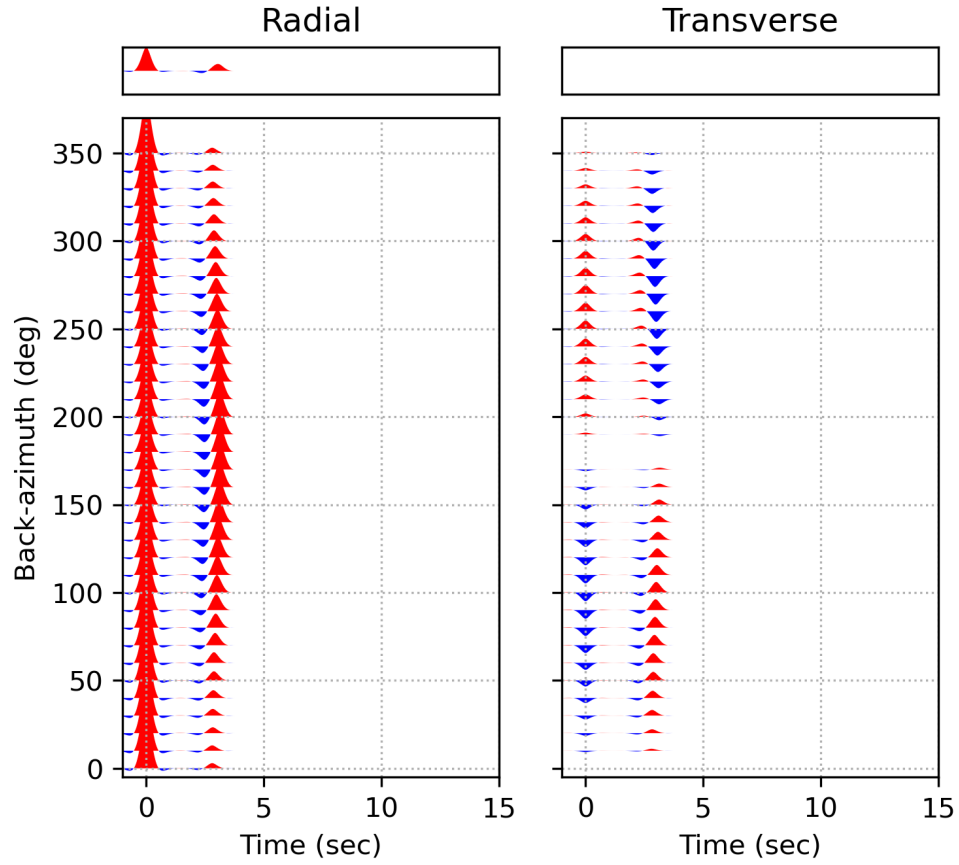


Figure 4.3: Radial and transverse receiver functions for Earth model with a dipping layer. The topmost panels are the stacks of radial and transverse receiver functions from all back azimuths. Note that only direct P-to-S conversions (not multiples) are modelled for simplicity.

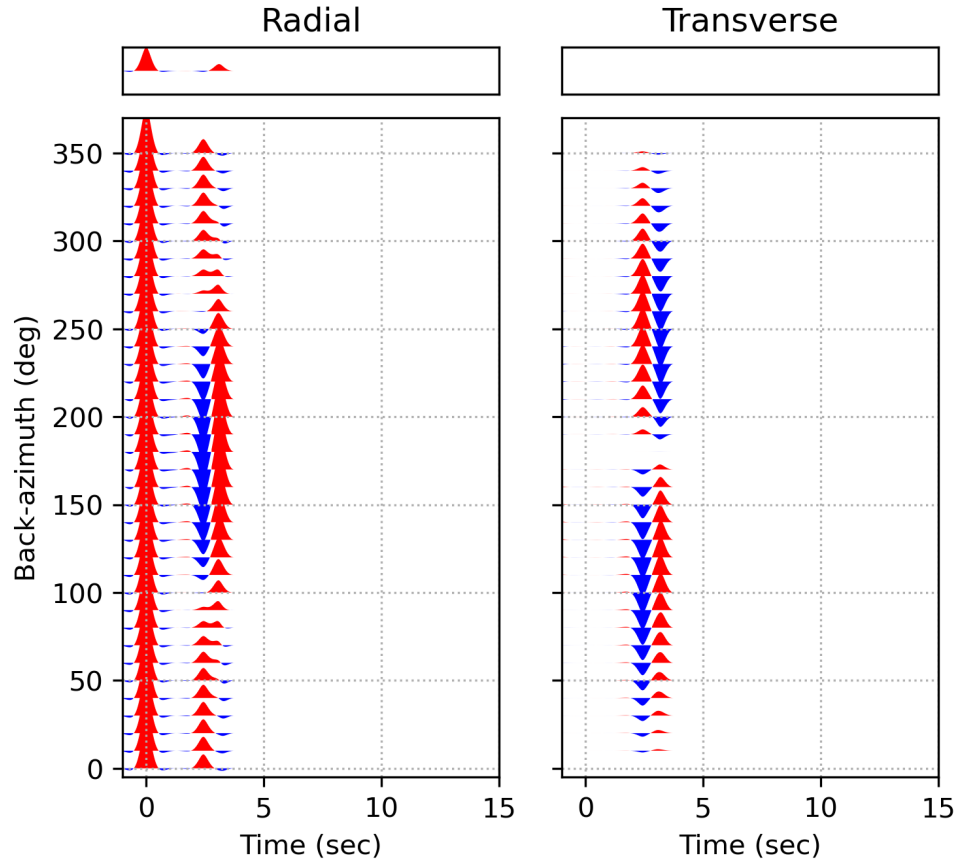


Figure 4.4: Radial and transverse receiver functions for Earth model with an anisotropic layer. The topmost panels are the stacks of radial and transverse receiver functions from all back azimuths. Note that only direct P-to-S conversions (not multiples) are modelled for simplicity.

axis of hexagonal symmetry (i.e., azimuthal anisotropy) produces phases that display a 180-degree periodicity of amplitude and timing with respect to BAZ (e.g. Cassidy, 1992; Porter et al., 2011; Schulte-Pelkum & Mahan, 2014). However, anisotropy with a plunging axis of symmetry produce phases that display a 360-degree periodicity with respect to BAZ (fig. 4.4), as does structure with dipping velocity contrasts (fig. 4.3). Consequently, the orientation of structure beneath a station can be inferred from the BAZ-periodicity of RF phases. A quantitative, though heuristic, alternative to inverting RF waveforms is to decompose RF data at a given station into a set of harmonics with respect to BAZ (Bianchi et al., 2010; Schulte-Pelkum & Mahan, 2014; Audet, 2015). This technique is applied to RF data considered in the following chapters. As this methodology is reviewed in both chapter 5 and chapter 6, I do not review it again here.

4.4 Receiver functions at subduction zones

Subduction zones are boundaries where two tectonic plates converge, and one plate is thrust beneath the other. The geometry of these boundaries necessitates dipping tectonic structures, and the nature of active tectonism can often manifest seismically detectable anisotropy. Consequently, the analysis of RFs at subduction zones can be especially challenging. However, observations of RFs have been important for developing our understanding of the structure and dynamics of subduction zones, which ultimately improves our understanding of the geohazards posed by these regions.

Geophysics (including RF analysis) has been one of the best tools for illuminating the structure and dynamics in subduction zones. There exists an abundance of structural geophysical studies at subduction zones, which I do not attempt to review here. Rather, I present a brief compilation of the most significant observations and interpretations within the context of RF analysis. There exists a general self-consistency between geophysical observations of subduction zones. Yet, a universal interpretation of these observations and their implications for the dynamic processes at subduction zones does not exist. The interpretation of these observations has implications for the inferred geometry of megathrust

faults (plate interfaces) in subduction zones. This is used as input for thermal and geodetic modelling, which are the primary constraints on the size and location of the portions of megathrusts that can rupture in large earthquakes (i.e., locked zones). The size of these faults and their proximity to major urban areas ultimately have significant consequences for seismic hazards.

The LITHOPROBE experiment during the 1980s consisted of an onshore-offshore deep multi-channel active-source seismic reflection survey with the intent of defining the geometry of the Cascadia subduction zone beneath Vancouver Island, British Columbia. The summary of this experiment (including results) is provided by Hyndman (1995). Data from this experiment clearly define what is believed to be the top of the oceanic crust as well as several frontal thrust faults in the overriding plate. The megathrust fault is believed to be located near the top of the oceanic crust since the frontal thrusts extend to this depth. Beneath the continental shelf offshore, and beneath Vancouver Island, there is a thick layer of strong seismic reflectors that is landward-dipping from ~ 15 km off the coast to ~ 35 km at its landward extent, where it disappears. This layer was originally referred to as the “E-layer” throughout the literature (due to an alphabetical naming convention for observations in the original LITHOPROBE data). The physical explanation for this layer has indirectly become an important topic in subduction zone science.

Cassidy & Ellis (1993) computed RFs using a small linear array of broadband seismometers across central Vancouver Island and the lower mainland of British Columbia to investigate the structure of the subduction zone. They identified a series of phases that suggest the presence of a landward dipping LVL (fig. 4.5) with a velocity contrast of ~ 1 km/s at depths consistent with the E-layer. Audet et al. (2009) considered RFs computed for data collected during the POLARIS experiment in the early 2000s (also across Vancouver Island) and similarly identified phases associated with a dipping LVL within the subduction zone. By measuring the differences in delay times between the phases associated with the top and bottom of the LVZ, they characterized the average material properties between the two boundaries (i.e., the properties of the LVL). Their results indicate extremely high values of Poisson’s ratio within the LVL, which can not be explained by temperature or composition. Instead, they interpret the high Poisson’s ratio as a con-

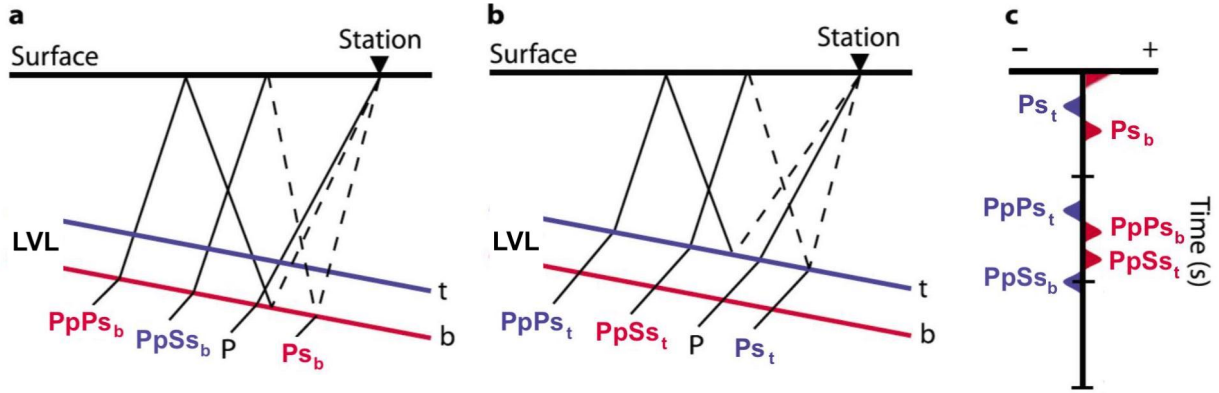


Figure 4.5: Receiver functions for structure with a dipping low-velocity layer (LVL). The bottom (a) and top (b) of a dipping LVL each produce a set of direct and multiple P-to-S converted phases that appear as a diagnostic pattern in the receiver function (c). Solid and dashed black lines represent P-wave and S-wave paths, respectively. Note that the polarity of phases generated by the top of the layer are reversed due to the velocity reduction across this discontinuity (modified from Audet et al., 2010).

sequence of near-lithostatic pore fluid pressures. Coincidentally, multiple magnetotelluric studies over Cascadia reveal a high-conductivity region that correlates with the location of the LVL/E-layer (e.g., Kurtz et al., 1990; Soyer & Unsworth, 2006), suggesting different geophysical expressions (seismic vs. conductive) are believed to be manifestations of the same structural feature that likely involves fluid processes. Henceforth, I refer to this feature as the LVL, as the term E-layer no longer appears in current literature.

The LVL is a subduction-related structural feature that is observed at most subduction zones on Earth (see Bostock, 2013, and references therein). However, the interpretation of what this layer represents is not universal. Audet et al. (2009) interpret the LVL as subducting oceanic crust. In this model, the plate interface is a narrow region of very low permeability, which explains the observed high Poisson's ratio. The nature of the low-permeability seal on the plate boundary is believed to change further down-dip at the onset of subduction-related volume changes associated with crustal eclogitization and serpentinization of the forearc mantle. Hansen et al. (2012) and Bostock (2013) adopt

a similar interpretation, with the only difference being that the LVL represents only the upper crust as opposed to the entire oceanic crust. The compositional change between upper and lower oceanic crust is believed to confine high pore fluid pressures as well as produce a significant positive seismic velocity contrast. Specifically, the lower positive phase in RFs is identified as the oceanic Moho by Audet et al. (2009), but is identified as the mid-crustal layer by Hansen et al. (2012).

A number of recent studies support an alternative interpretation of this layer. Nedimović et al. (2003) show that this layer experiences a significant change in thickness from ~ 2 km thick beneath the continental shelf to ~ 7 km thick at greater depths beneath Vancouver Island. They interpret this as an expression of the change in frictional properties on the megathrust from fully locked offshore to transitional beneath western Vancouver Island. In this model, the LVL is interpreted as a thick shear zone. The plate boundary is considered to be a 2-7 km thick fault system (as opposed to a narrow or near-planar interface) that accommodates deformation due to plate convergence. It is possible that the shear zone has elevated porosity which would explain the observed high conductivity. More recently, Audet & Schaeffer (2018) computed RFs from onshore/offshore data and found that the signature of the LVL does not extend into the locked zone offshore. This finding further supports the interpretation that the LVL is likely an expression of changing frictional conditions along the megathrust. Although they interpret the LVL as subducting material with high pore-fluid pressure confined by a low-permeability cap, they note that the increasing thickness of the LVL down-dip may instead represent the progressive development of a shear zone, similar to the hypothesis of Nedimović et al. (2003).

There are more complicated models of subduction zone structure in Cascadia (and elsewhere) than those described above. For example, Calvert (2004) interprets this layer as imbricated crustal rocks bounded by two shear zones. Combined, these form a plate-interface region that is over 16 km thick. The original RF interpretations by Cassidy & Ellis (1993) placed the subducting slab much deeper than the LVL in order to correlate with observed intra-slab seismicity as well as a deeper reflector identified in the original LITHOPROBE data (unsurprisingly, this deeper reflector was named the “F” reflector). McCrory et al. (2012) and McCrory & Wilson (2013) define the geometry of the subduction

zone from limited intra-slab seismicity, which places the plate interface well above the LVL. This has led to the interpretation that the layer beneath Vancouver Island represents a detached fragment of Eocene oceanic crust from the ancient subducted Resurrection plate, and that the layer beneath Puget Sound (further south along the Cascadia margin) represents hydrated lower forearc material. This interpretation requires a unique explanation for this layer for every region of (nearly) every subduction zone on Earth.

The majority of the examples discussed above are applicable to the Cascadia subduction zone, which is discussed in chapter 5. Cascadia is arguably the most well-studied subduction zone on Earth (other than subduction zones in Japan) for many reasons such as instrumentation and politics (among others). Consequently, the literature on this topic is regionally biased. The Hikurangi subduction zone is the convergent tectonic boundary along the eastern coast of the North Island of New Zealand, and is discussed in chapter 6. A dipping LVL has been identified near and/or above the plate interface in the Hikurangi subduction zone based predominantly on tomographic studies (e.g., Eberhart-Phillips & Reyners, 2012) as well as limited RF analysis (e.g., Bannister et al., 2007). However, the Hikurangi subduction zone is unique in many ways, with a complex history of deformation as well as neo-tectonics (as will be discussed in chapter 6). Such deformation structure can create RF phases that mask the signature of the LVL (if it is ubiquitous in this region). The overriding plate (Australian plate) is heavily deformed and previous studies in the region have shown that RFs are complicated and challenging to interpret (e.g., Savage et al., 2007). For this reason, chapter 6 focuses only on early arrivals in the RF data that correspond to seismic velocity structure within the overriding plate, rather than the entire subduction zone system.

Chapter 5

Seismic evidence for megathrust fault-valve behavior during episodic tremor and slip

5.1 Introduction

Periodic phenomena of spatially and temporally correlated seismic tremor and aseismic slow slip are collectively referred to as episodic tremor and slip (ETS), and have been observed in several subduction zones (e.g., Rogers & Dragert, 2003; Obara et al., 2004; Wallace & Eberhart-Phillips, 2013). It is generally believed that ETS occurs along the subduction zone megathrust down dip of the seismogenic zone, and that interaction between segments of the megathrust may be important for understanding increases in earthquake triggering (Obara & Kato, 2016). In most subduction zones, geophysical images reveal a landward-dipping seismically low-velocity layer (LVL; e.g., Bostock, 2013; Audet & Kim, 2016) with low electrical resistivity (Wannamaker et al., 2014) in the vicinity of ETS that is interpreted as hydrated upper oceanic crust of the subducting plate and/or shear zone material. These studies, along with thermal and petrological models (Hyndman & Peacock, 2003), provide evidence consistent with ETS occurrence in fluid-rich environments.

Investigations on the mechanical properties of the LVL in subduction zones suggest near lithostatic pore-fluid pressures (P_f) facilitated by low permeability along the plate interface (e.g., Audet et al., 2009; Peacock et al., 2011; Audet & Kim, 2016), which act to lower effective normal stress. This is supported by frictional modelling (e.g., Liu & Rice, 2007), as well as observations that suggest tremor sensitivity to small dynamic stress perturbations (e.g., Rubinstein et al., 2007, 2008). At ETS depths (30-40 km), most fluids derive from prograde metamorphic dehydration reactions of subducting material (Hyndman & Peacock, 2003). Studies of subduction zone forearcs suggest fluids eventually migrate along the megathrust and into the overlying crust, near the mantle-wedge corner (e.g., Audet & Bürgmann, 2014; Hyndman et al., 2015; Savard et al., 2018). It has been suggested that these fluids may not only play a passive role in ETS by weakening faults via high P_f , but that ETS may also be coupled with cyclic fluid processes (Hyndman et al., 2015). However, observations of such cyclic processes are limited and indirect (Nakajima & Uchida, 2018; Warren-Smith et al., 2019). Characterizing temporal variations in physical properties in subduction zones in relation to ETS may help in understanding fluid processes and their effect on slip behaviour along the plate interface.

Receiver functions (RFs), a type of teleseismic scattering data, have been used to investigate the seismic velocity structure of subduction zone forearcs in the ETS source region (Audet & Kim, 2016). In this study, we investigate temporal variations in RF data that may be associated with fluid-related P_f cycling during ETS. These data are sensitive to discontinuities in seismic shear-wave velocity (V_S) at vertical scale lengths of 1-10 km. RF data reveal the signature of the LVL characterized by double-polarity pulses of direct P-to-S converted (Ps) and back-scattered (PpPs and PpSs) waves from the top and bottom boundaries of this layer (Audet et al., 2009). We use high-quality broadband seismic data from >1300 teleseismic earthquakes recorded between 1993-2019 at five long-running seismograph stations in the forearc of the Cascadia subduction zone (fig. 5.1) and calculate vertically (S_V) and horizontally (S_H) polarized RFs. The raypaths of the scattered seismic waves that produce the RFs are nearly vertical. As a result, RFs are sensitive to the velocity structure directly beneath the recording station and sample the medium repeatedly. We use these repeated observations to constrain time-varying velocity structure

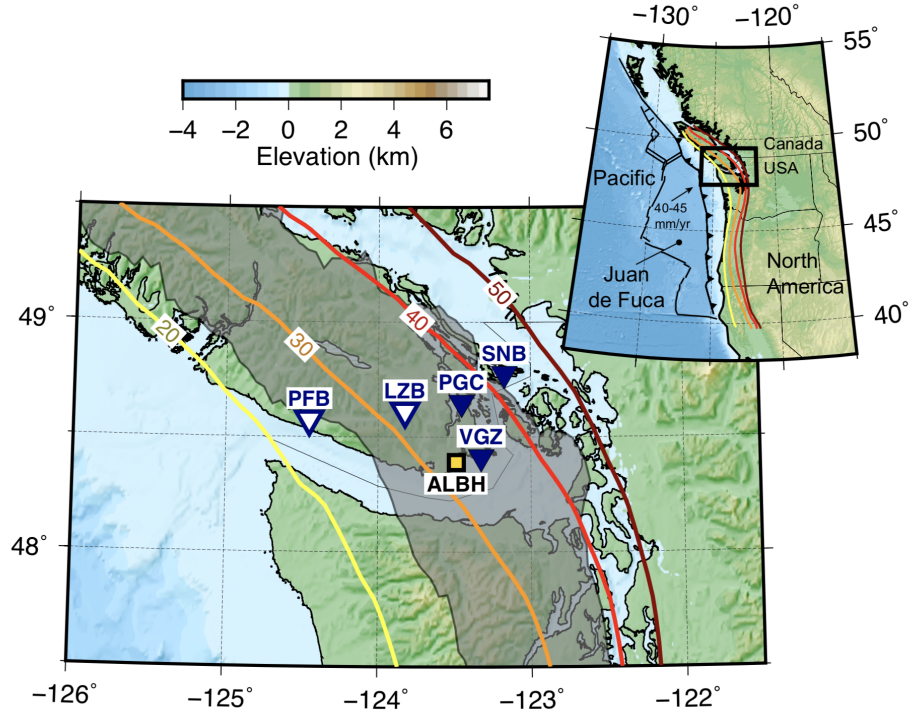


Figure 5.1: The study region in the forearc of the Cascadia subduction zone. Station locations considered for RF analysis are shown by triangles (solid triangles mark stations where changes in velocity structure are observed). The yellow square marks the location of the GPS station used to estimate ETS timing. Contours mark the depth of the plate interface inferred from the deformation front, up and downdip limits of tremor, and volcanic arc (Bostock et al., 2019). The gray shaded region delineates the ETS source region interpreted from tremor density (Bostock et al., 2019). The inset shows the tectonic setting of the Cascadia subduction zone.

in relation to ETS. The time resolution of RF data is low as they depend on the random and infrequent occurrence of large, distant earthquakes. However, combining RF attributes before and after 21 observed ETS events between 1994-2019 allows for observation of changes in material properties beneath the stations.

5.2 Results

The complexity of the structural environment can make it difficult to identify temporal changes in RF data. Dipping layers and/or anisotropy cause variations in the vertical incidence and back-azimuth angles of incoming waves, which results in periodicity of observed RF data with respect to the back-azimuth of the recorded earthquakes. We account for this by performing a harmonic decomposition of all S_V and S_H data, and produce a data-predicting model for a given input back-azimuth angle. We calculate the cross-correlation coefficients (CC) between observed and predicted RFs, remove outliers, and sort them chronologically. Using 15-day time bins before and after the midpoints of the ETS episodes determined using GPS data, we observe a statistically-significant (95% confidence level) reduction of 5-10% in the mean CC values of S_V data following ETS at three of the five stations (fig. 5.2; PGC, SNB, VGZ). This implies a systematic error in the harmonic decomposition model to accurately predict RFs following ETS, and suggests that the velocity structure following ETS is different from the velocity structure averaged over the entire recording period. However, this analysis does not indicate where in the depth profile the change in velocity structure is occurring. No significant changes in CC values are observed for the S_H data, suggesting changes occur predominantly in the background isotropic structure (see Materials and methods).

To investigate the origin of the changes in CC values for S_V data, we convert the amplitudes of the direct Ps phases for each individual RF into estimates of shear-wave velocity contrasts (δ_{V_S}) across the top and bottom of the LVL using a geometric model of the subducting slab (Bostock et al., 2019). This accounts for effects of back-azimuth angle and slowness on slab-normal incidence (corrects for geometrical variations in the amplitude

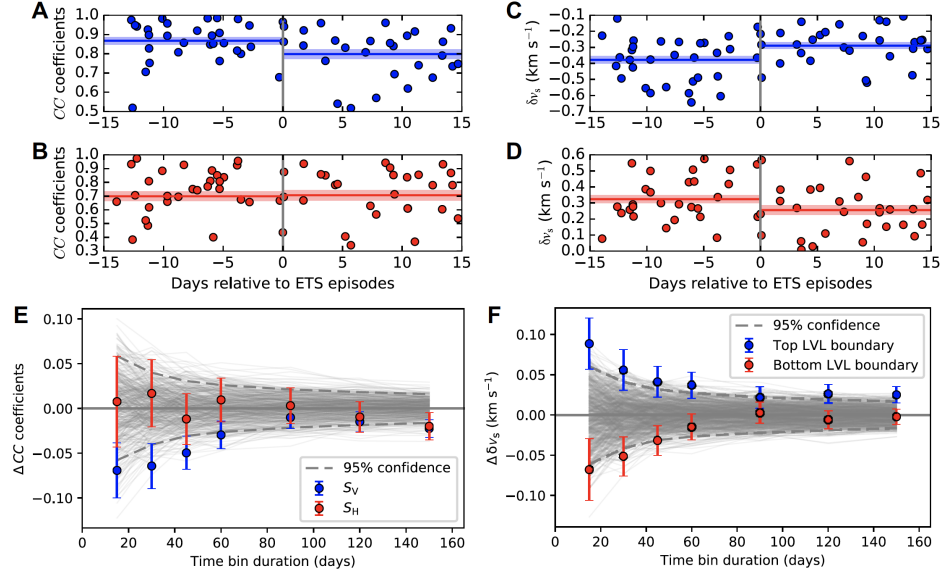


Figure 5.2: ETS-related changes in velocity structure beneath station PGC. Changes in mean CC values for 15-day time bins before and after ETS are shown for S_V (A) and S_H (B) data. Changes in mean δV_S before and after ETS are shown for the top (C) and bottom (D) boundaries of the LVL. Mean values and estimated standard errors in A-D are shown by horizontal lines and shaded regions, respectively. Changes in RF data as a function of time-bin duration are shown for CC analysis (E) and δV_S analysis (F). Bootstrap resampling of receiver function times relative to ETS are shown by grey lines (500 of 20000 samples), and 95% confidence levels are shown by dashed lines in E and F.

of RF data), and reduces the data information content to model parameters that reflect bulk structural properties of the LVL. After removing long-term amplitude variations in δ_{V_S} estimates and outliers, we observe a statistically significant (95% confidence) increase in the mean δ_{V_S} of 0.05-0.1 km/s across the top boundary of the LVL following ETS for the same stations that revealed a change in CC values (fig. 5.2). A reduction in the average δ_{V_S} across the bottom boundary of the LVL is observed following ETS for these stations, though our confidence in rejecting the null hypothesis is lower in this case. We note that the two stations that do not reveal changes in velocity structure are those where the megathrust is at shallower depths. The lack of observation may be structural in nature, though it may also be a consequence of lower-quality (and fewer) data.

Using longer duration time bins shows that the changes remain observable at statistically significant levels but at decaying magnitudes, suggesting that any change in the velocity structure is short-lived (fig. 5.2). However, we are unable to determine the absolute time scale for which the change persists with these data. Bootstrap analyses (by randomly resampling the dates of the events used to compute the RF data) of the CC and δ_{V_S} estimates as a function of time-bin width provide further evidence that the observed decaying change in these quantities is robust. The lack of scatter (relative to bootstrap samples) in the trend for estimated changes in δ_{V_S} across the bottom boundary of the LVL suggests this observation may also be real. Considering changes in δ_{V_S} across both the top and bottom boundaries would imply the change is predominantly within the LVL. Finally, we observe statistically significant changes in δ_{V_S} estimates using higher frequency corners of RF data (up to 0.7 Hz), suggesting that they occur over a scale length of 3.5 km.

5.3 Discussion

Recent studies that observe temporal changes in seismic velocities coincidentally with large earthquakes (e.g., Husen & Kissling, 2001; Peng & Ben-Zion, 2006; Pei et al., 2019) interpret their results in terms of rock damage from strong ground motion, static stress redistribution, and/or fluid processes. Our observations of temporal changes in seismic velocities are

associated with aseismic slow slip during ETS, which has a static stress drop several orders of magnitude smaller than regular earthquakes (Gao et al., 2012). ETS does not produce strong ground shaking, and seismically-induced rock damage would only be expected in the uppermost part of the crust (Peng & Ben-Zion, 2006), far above the LVL. Furthermore, neither rock damage or stress redistribution can explain the rapid recovery of our observed velocity changes. Alternatively, the ‘fault-valve’ model (Sibson, 1990, 2013) has been used to explain observations of changes in seismic velocities associated with large earthquakes along subduction megathrusts (Husen & Kissling, 2001). This model describes breaching of low-permeability materials within the megathrust shear zone during slip, which may be trapping and pressurizing fluids. Breaching of these seals opens permeable pathways and allows for fluid migration within and across the megathrust. Low-permeability is eventually re-established and fluid pressurization re-accumulates. It has been suggested that aseismic slow slip (during ETS) may be linked to a similar model (Nakajima & Uchida, 2018; Warren-Smith et al., 2019), which would necessitate periodic fluctuations in near lithostatic P_f .

We suggest that the observed V_S changes within the LVL are the result of temporary reductions in P_f following ETS (fig. 5.3), consistent with the fault-valve model. This is supported by observations of temporal changes in bulk seismic attenuation coincident with inferred slow slip in Japan, which suggest episodic fluid drainage of the plate interface (Nakajima & Uchida, 2018). However, these observations are made in a colder subduction zone that does not display typical ETS behavior. Time-dependent measurements of the stress ratio within the subducting oceanic crust have recently been found to correlate with shallow slow slip events, and suggest reductions in P_f on the order of several MPa (Warren-Smith et al., 2019). Assuming that all of our observed velocity change occurs within the LVL at high P_f conditions, we estimate that the increase in V_S of 0.1 km/s corresponds to a P_f drop of 1-10 MPa (Christensen, 1984), consistent with estimates from stress ratio fluctuations (Warren-Smith et al., 2019). In addition, models of fluid compartmentalization within the fracture network of the subducting crust or shear zone may produce changes in P_f of several MPa when low-permeability barriers between compartments are ruptured (Byerlee, 1993; Shapiro et al., 2018). The fluid flux from the subducting slab at ETS depths

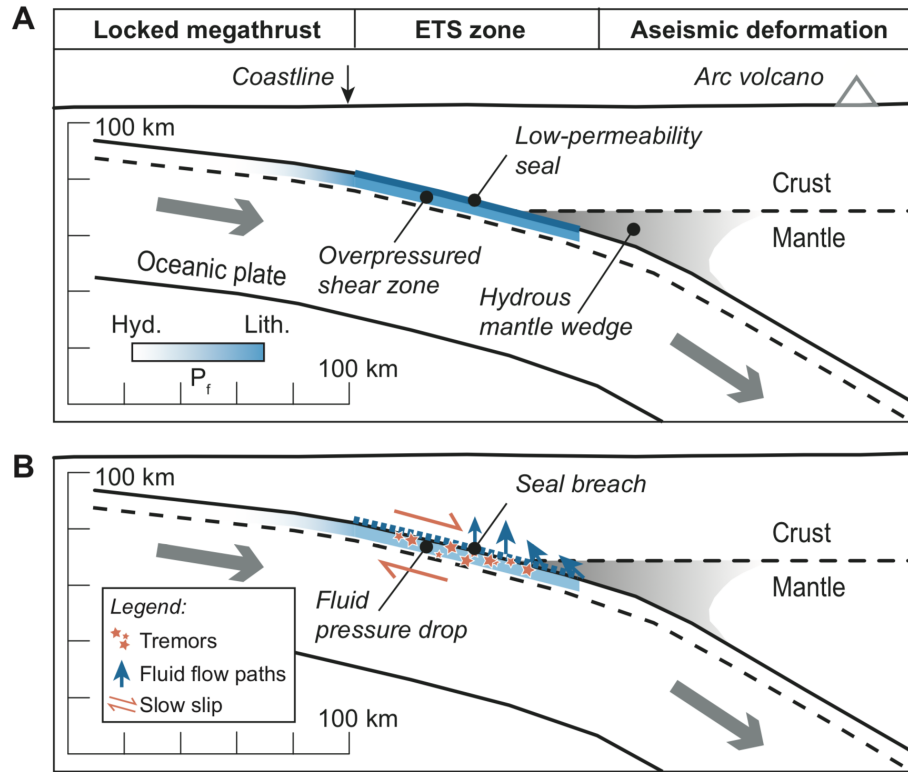


Figure 5.3: Conceptual model of ETS-related fluctuations in P_f in Cascadia. During inter-slip periods (A), the deep extension of megathrust faults where ETS occurs is characterized by high P_f . During ETS (B), low-permeability seals within the plate interface shear zone are breached, which opens permeable pathways for fluid migration within and across the megathrust causing a drop in P_f . Low-permeability barriers are re-established and the cycle repeats. This figure is adapted from Peacock et al. (2011).

is low given the short recurrence time of ETS (Hyndman & Peacock, 2003). However, a fracture network model does not necessitate a high flux of fluids from the subducting plate, as high P_f can also be achieved through localized sealing of fluid compartments (Byerlee, 1993), which may be a result of fault healing or hydrothermal mineral precipitation (Hyndman et al., 2015; Audet & Bürgmann, 2014; Savard et al., 2018).

Finally, the observation that only stations located in the down dip extension of the slow slip source region experience a change in V_S correlates with greater fluid availability near the mantle wedge corner (Hyndman & Peacock, 2003; Hyndman et al., 2015). Fluids may ultimately be mobilized into the forearc crust during this process (fig. 5.3), consistent with Helium isotopic signatures from crustal fluids in Cascadia (McCrory et al., 2016). Enhanced vertical fluid mobility following slow megathrust rupture would result in greater silica precipitation and explain the concentration of seismicity and reduced electrical resistivity in the forearc crust above the mantle wedge corner (Wannamaker et al., 2014; Hyndman et al., 2015; Nakajima & Uchida, 2018; Bostock et al., 2019). These results lend support to models that favor hydrologic control of slow slip at the deep extension of megathrust faults.

5.4 Materials and methods

5.4.1 Receiver function analysis

Data used in this study come from permanent broadband seismic stations of the Canadian National Seismograph Network located in the forearc of the Cascadia subduction zone (fig. 5.1). We use 120 s-long, three-component recordings of the P-wave coda of teleseismic earthquakes with magnitudes above 5.8 that are located at epicentral distances between 30° and 90° . At each station we compile all available data between 1993 and 2019 with high signal-to-noise ratio (>5 dB) on the vertical component. Seismograms are decimated to 5 Hz, low-pass filtered using a corner frequency of 1 Hz, rotated to the vertical-radial-transverse orientations and then decomposed into up-going P, S_V and S_H wave modes

(Bostock, 1998). The P mode is deconvolved from the S modes using Wiener spectral deconvolution (Audet, 2010) to obtain S_V and S_H component RFs. This deconvolution procedure has been shown to reduce the seasonal amplitude variations in the power-spectral density function of RF data due to increased noise levels in the winter and is important in this context to avoid mapping seasonal variations into the ETS-related signal (Audet, 2010). RFs are then filtered at corner frequencies of 0.05 Hz to 0.5 Hz. Figure C.1 shows all S_V data for station PGC sorted by back-azimuth, showing the signature of the low-velocity layer.

5.4.2 Harmonic decomposition

In the presence of horizontal discontinuities in seismic velocity structure separating two isotropic layers, the timing and amplitude of the various phases depend only on the vertical incidence angle (characterized by the slowness vector) of incoming P waves. In this case, there is no energy on the S_H component of RFs; the signal is entirely contained on the S_V component and displays no variation with back-azimuth. In the presence of either dipping discontinuities or anisotropy in seismic velocities, RF data will show variations in both timing and amplitude of the converted phases as a function of slowness and back-azimuth (Bianchi et al., 2010; Schulte-Pelkum & Mahan, 2014; Audet, 2015). These patterns are generally characterized by periodic variations in amplitude with back-azimuth observed in both S_V and S_H components, with a characteristic 90° phase shift between them. The harmonic decomposition method exploits this periodicity by decomposing RF data into a set of periodic functions with back-azimuth that are described by $\sin(k\phi)$ and $\cos(k\phi)$, where k is the harmonic degree and ϕ is back-azimuth (Bianchi et al., 2010; Audet, 2015).

Prior to harmonic decomposition, we convert all RF data from time to depth using an average seismic velocity profile under Vancouver Island (Ramachandran et al., 2006). This step reduces the influence of small phase delays produced by variations in horizontal slowness (Audet, 2015). We specify a maximum depth of 60 km to avoid contamination by the free-surface reverberations and include only the direct Ps conversions. Here we limit the decomposition to $k = 0, 1, 2$, which captures most of the signal for simple models of

dipping layers and low order anisotropy (Bianchi et al., 2010; Schulte-Pelkum & Mahan, 2014; Audet, 2015). The system of equations is written as:

$$\begin{bmatrix} S_{V1}(z) \\ \vdots \\ S_{VN}(z) \\ S_{H1}(z) \\ \vdots \\ S_{HN}(z) \end{bmatrix} = \begin{bmatrix} 1 & \cos(\phi_1) & \sin(\phi_1) & \cos(2\phi_1) & \sin(2\phi_1) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & \cos(\phi_N) & \sin(\phi_N) & \cos(2\phi_N) & \sin(2\phi_N) \\ 0 & \cos(\phi_1 + \frac{\pi}{2}) & \sin(\phi_1 + \frac{\pi}{2}) & \cos(2\phi_1 + \frac{\pi}{4}) & \sin(2\phi_1 + \frac{\pi}{4}) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cos(\phi_N + \frac{\pi}{2}) & \sin(\phi_N + \frac{\pi}{2}) & \cos(2\phi_N + \frac{\pi}{4}) & \sin(2\phi_N + \frac{\pi}{4}) \end{bmatrix} \begin{bmatrix} A(z) \\ B_1(z) \\ B_2(z) \\ C_1(z) \\ C_2(z) \end{bmatrix}. \quad (5.1)$$

The term A captures the signal that is independent of back-azimuth and corresponds to a simple mean of all S_V component data. The terms B_1 and B_2 describe RF signals that vary by one cycle over all back-azimuths, and mainly reflect either a dipping interface or hexagonal anisotropy with a plunging axis of symmetry. The terms C_1 and C_2 describe variations of two cycles over back-azimuths and are normally interpreted as hexagonal anisotropy with a sub-horizontal axis of symmetry.

We perform the harmonic decomposition using all available RF data by inverting equation (5.1) using singular value decomposition (fig. C.2). The harmonic components combine to provide a data-predicting representation of the structure beneath each station, averaged over the station deployment time (> 25 years). Once these components are obtained, we reverse the process and calculate predicted S_V and S_H data for each teleseismic source in the RF data set. This allows us to quantitatively compare observed and predicted RFs by calculating CC coefficients as a function of teleseismic source ordered chronologically for both the S_V and S_H data sets separately (fig. C.3). We then remove outliers within each data set using a median absolute deviation threshold of 2.5.

5.4.3 Interface velocity contrast

At an interface separating two materials with different seismic velocities, RF amplitudes vary according to the interface-normal incidence angle, or interface-parallel slowness. For a

dipping interface, corrections are therefore required to convert vertical incidence and back-azimuth angles into interface-normal incidence. We use a smooth slab interface model (Bostock et al., 2019) and extract the strike and dip angles of the slab surface to carry out the angle corrections. This procedure is performed by gridding the slab contours using a spline interpolation under tension with minimum curvature and measuring the directional derivative and slab dip at the locations of the seismic stations considered herein (fig. C.4 and table C.1). We then extract the amplitudes of the main P-to-S converted phases (Ps) on the S_V component from the top and bottom of the low-velocity layer to estimate the S-wave velocity contrast at those interfaces. We convert the Ps RF amplitudes to estimates of S-wave velocity contrasts for plane waves at near-normal incidence (Aki & Richards, 2002):

$$Ps = \frac{p}{2q_\alpha} \left[(1 - 2\beta^2 p^2 - 2\beta^2 q_\alpha q_\beta) \frac{\Delta\rho}{\rho} - 4\beta^2 (p^2 + q_\alpha q_\beta) \frac{\Delta\beta}{\beta} \right], \quad (5.2)$$

where α and β are the P- and S-wave velocities, respectively, p is the slab-parallel slowness, q_α and q_β are the slab-normal slowness values for P and S waves, respectively, and ρ density. Although the amplitudes are sensitive to both S-wave velocity and density contrasts, the equation is an order of magnitude more sensitive to the contrast in S-wave velocity; the second term on right-hand side of equation (5.2). We therefore ignore variations in density contrast and simply invert the amplitudes for $\frac{\Delta\beta}{\beta}$, which we scale to an absolute velocity contrast $\delta_{V_S} = \frac{\Delta\beta}{\beta} \beta$ using an average upper mantle S-wave velocity of $\beta = 4.6$ km/s. We estimate δ_{V_S} for each RF and sort these values chronologically. We then remove a third-order polynomial fit to the estimated δ_{V_S} values to remove long-term variations as well as outliers using a median absolute deviation threshold of 2.5 (fig. C.5).

5.4.4 ETS times from GPS data

To estimate changes in the seismic velocity structure relative to ETS, we require precise estimates of ETS times from 1993 to 2019 obtained in a consistent way. In this case, however, we do not require the formal inversion of GPS data since we are only interested in the

timing of slow slip in the vicinity of the seismic network. We estimate ETS start, middle, and end times using GPS data from station ALBH, located nearest station VGZ (fig. 5.1). We use the NAM08 reference frame solution from UNAVCO, but similar results are derived using alternative reference models (e.g., IGS08). We then detrended and smoothed the raw east-component GPS times series using a rolling window average of 14 days and extract shorter time series for dates surrounding known ETS events (Dragert et al., 2004) and estimated ETS events from visual inspection (fig. C.6). For each shorter ETS time series, we fit a 5th order polynomial curve and estimate its first and second derivatives. We accept dates where the first derivative is approximately 0 to mark the beginning and end of the GPS reversal (i.e., the start and end of the ETS event). Since these dates are obtained from smoothed and processed data, they are likely over-estimates (i.e., errors of 5-10 days), but with no consequence to our analysis. The date where the second derivative is approximately 0 is taken as the mid-point of the given event, and represents the time of maximum slip rate during ETS that we use as reference in our seismic data analysis (fig. C.6). These dates are summarized in table C.2 and were also verified to correspond with times when tremors during slow slip are located within 50 km of station PGC (Wech & Creager, 2008; Idehara et al., 2014). ETS start and end dates for the events in 1994 and 1995 could not be determined from the available GPS data and the mid-point dates were determined independently (H. Dragert, personal communication, 2018).

5.4.5 Temporal stacking

RF data sample a medium repeatedly beneath a seismic station; however, temporal sampling is highly irregular and the results can be biased by the teleseismic source distribution (Audet, 2010; Porritt & Yoshioka, 2017). Furthermore, RFs may contain significant noise that is difficult to quantify. Observing time-varying changes in seismic velocity structure using these data therefore requires stacking over multiple ETS events and evaluating the resulting stacks using statistical inference. Using the dates estimated from the GPS data (see above), we bin RF times into pre- and post-ETS events for all events in the catalogue. We use the mid-point of ETS events as reference, and define time bins of 15, 30, 45, 60,

90, 120, and 150 days. We then stack the RF-derived quantities (i.e., either the CC or δ_{V_S} values) to estimate the sample mean values and standard errors. We perform a two-sample t -test for the difference in sample mean values and calculate p -values to evaluate the statistical significance (p -values for station PGC are provided in table C.3). Taking the mid-point of individual slow slip events also implies that any potential bias due to tremor activity on the individual waveforms will similarly affect the pre- and post-ETS time bins. We verified that randomly perturbing the ETS times using a normal distribution with standard deviation of 5 days does not change the statistical significance of our results.

5.4.6 Bootstrap analysis

In addition to the t -test, we further test the null hypothesis by randomly shuffling the dates of the RF data (without replacement) 20,000 times and re-processing the randomized data sets for both CC and δ_{V_S} estimates. This is performed for all time bins considered and shown in figure 5.2 for station PGC. We calculate the 95% confidence envelope from the randomized data sets and find that the difference in mean values is highly scattered when considering single randomized data sets as a function of time-bin duration. Such scattering is not observed with the real data sets, further confirming that our results are not due to chance. Figure C.7 shows the bootstrap analysis for stations LZB, PFB, SNB and VGZ for the change in δ_{V_S} values. Among these, station SNB and VGZ show trends that are consistent with those observed at station PGC (i.e., low scatter; comparable but lower change in δ_{V_S}). These stations are located along similar (or slightly greater) plate interface contours. Results for station LZB and PFB are not statistically different than a purely random process.

Chapter 6

Seismic constraint on heterogeneous deformation and stress state in the forearc of the Hikurangi subduction zone, New Zealand

6.1 Introduction

The Hikurangi subduction zone (HSZ) spans nearly 500 km along the eastern coast of the North Island of New Zealand, and marks the collisional boundary between the Pacific and Australian tectonic plates (fig. 6.1). The Pacific plate subducts transpressively westward beneath the North Island, at the southern end of the Tonga-Kermadec-Hikurangi subduction system. The section of the Pacific plate subducting along the HSZ consists of a Cretaceous large igneous province (the Hikurangi Plateau) with a progressively southward thickening (10-15 km thick) buoyant oceanic crust (Mochizuki et al., 2019). This ultimately resists subduction and causes a complex southward transition into the strike-slip Alpine Fault, across the South Island of New Zealand.

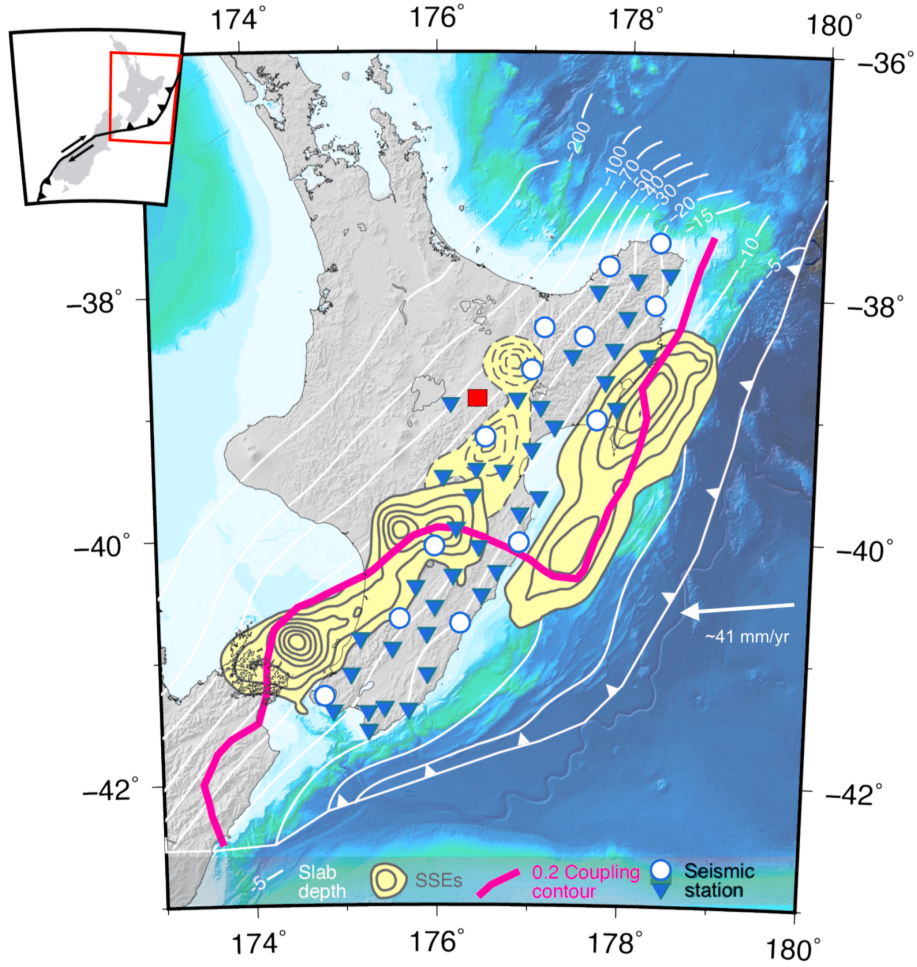


Figure 6.1: Tectonic setting and seismic station distribution. Inverted blue triangles and white circles show the locations of short-period and broadband seismic stations used in this work, respectively. The red square shows the location of station MRHZ. White contour lines show the depth to the plate interface in kilometres (Williams et al., 2013). The white arrow shows the motion of the Pacific plate relative to the Australian plate (Kreemer et al., 2003). The pink contour line delineates the boundary between the locked (southwest) and creeping (northeast) parts of the subduction zone (the 0.2 coupling contour from Wallace, 2020). The yellow contour patches show the locations of major slow slip events (SSEs).

The interface (megathrust fault) between the Pacific and Australian plates at the HSZ is believed to be capable of producing magnitude 9.0 earthquakes and associated tsunamis, with geological evidence of four megathrust ruptures over the past 2000 years (Pizer et al., 2021). However, Global Navigation Satellite Systems (GNSS) data over the North Island suggest that only the southern part of the megathrust is currently locked and accumulating elastic strain that can be released in a large earthquake (Wallace et al., 2004). In contrast, the Australian and Pacific plates are believed to be uncoupled, and the megathrust creeping, in the northern part of the HSZ (fig. 6.1). Furthermore, slow-slip events (SSEs) on the Hikurangi megathrust fault have been observed (Wallace, 2020) following the pattern of plate locking along the margin but generally downdip of the locked part of the fault (fig. 6.1). These observations are consistent with modelling that suggests SSEs occur in a transitional frictional state (Liu, 2013). Changes in the depth/location of SSEs as well as the inferred state of plate coupling along the margin indicate that the physical/mechanical conditions within the subduction zone change significantly from north to south (Wallace, 2020).

Knowledge of the structural and mechanical conditions of currently locked and creeping segments of megathrust faults is important for understanding the seismic hazards posed by these tectonic systems. This research uses passive seismic data to constrain the tectonic grain over the forearc of the HSZ in order to address how present-day and inherited deformation are distributed within the context of the spectrum of megathrust behaviour along the margin. Specifically, this work examines the directional dependence of seismic scattering structure, and explores the spatial variability of these data in relation to forearc deformation, plate coupling, and subduction geometry and morphology.

6.2 Data and Methods

Distant earthquakes produce teleseismic P-waves that travel through the Earth’s mantle and arrive almost vertically beneath a seismograph station. Some of the P-wave energy will be scattered by structure beneath the seismograph station, and will be converted into S-

wave phases that appear on the horizontal-component channels. Deconvolving the vertical component from the two horizontal components of motion isolates these converted S-waves (by approximating the vertical-component as a source wavelet). Hence, a 3-component teleseismic P-waveform contains a record of the structure beneath a seismic station in the form of a P-S receiver function (RF) (Burdick & Langston, 1977; Langston, 1979). A single RF is a waveform that approximates the impulse response of the structure beneath a seismic station. The timing of a particular pulse in the RF waveform constrains the depth to the structure that generated the pulse. The amplitude of the pulse constrains the magnitude of the velocity contrast across the structure.

In this work, we compute RFs for 53 long-running seismic stations (fig. 6.1) part of the national seismograph network of New Zealand (GEONET) using recordings of earthquakes with magnitudes >5.5 that occurred between January 2000 and March 2020, within epicentral distances of $30^\circ - 90^\circ$ for P arrivals, and $90^\circ - 150^\circ$ for PP arrivals, relative to each station. For each earthquake, we use the iasp91 Earth model (Kennett & Engdahl, 1991) to extract 120 s time windows centered on the predicted arrival (P or PP). Note that many of the GEONET stations used in this work are short-period instruments, which are typically inferior to broadband instruments in RF applications. However, many recent studies have successfully applied RF analysis using similar instruments (e.g., Niu et al., 2005). All time windows are initially corrected for their instrument response. Furthermore, given the long recording times available for the instruments considered in this work, we only consider events with signal-to-noise ratios greater than 5 dB (relative to pre-event noise) over the initial 30 s after the predicted arrival on the vertical component (measured over the frequency band of 0.05-1 Hz). Horizontal components are then rotated into radial and transverse orientations, relative to the event location. In order to isolate P-S converted phases (i.e., compute RFs), we use a modified Wiener spectral deconvolution that is regularized by the pre-event covariant noise spectrum between vertical and horizontal components. This has the advantage of minimizing potential contamination of the RFs due to seasonal noise effects (Audet, 2010). RFs are then bandpassed filtered using a second-order Butterworth filter between 0.05 and 0.75 Hz. Examples of radial and transverse RFs computed at station MRHZ are shown in figure 6.2, sorted by event back-azimuth (BAZ).

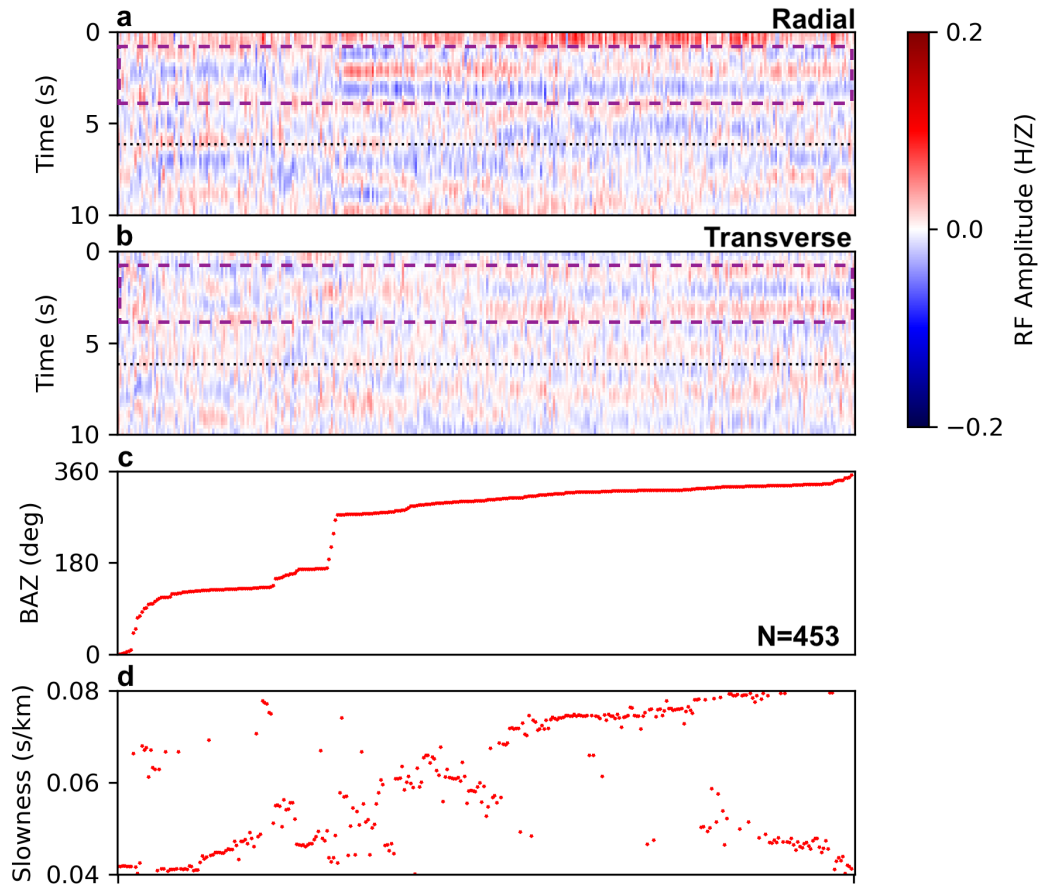


Figure 6.2: Receiver function data at station MRHZ. Individual radial (a) and transverse (b) receiver functions are shown with back-azimuths (BAZ) (c) and slowness (d) for the corresponding events. Data are sorted by BAZ. The number of events considered for this station (after quality control) is listed in (c). The black dotted line shows the expected arrival time of a P-to-S converted phase generated at the depth of the plate interface beneath the station, based on the plate interface model by Williams et al. (2013) and the seismic velocity model by Eberhart-Phillips et al. (2020). Structure at earlier times (i.e., shallower depths) are within the overriding forearc of the Australian plate. The purple dashed box highlights shallow crustal phases that change amplitude/polarity with respect to BAZ.

For isotropic media with horizontal velocity discontinuities, all of the teleseismic P-waveform energy is contained within the radial plane (i.e., no energy appears on transverse RFs). However, in the presence of material anisotropy or dipping structure, some of the teleseismic P-waveform energy can be refracted out of the radial plane and onto the transverse plane. Consequently, the observation of coherent pulses on transverse-component RFs is indicative of anisotropy and/or dipping structure beneath a seismic station. For all stations considered in this work, we observe coherent pulses on transverse-component RFs at P-S delay times (i.e., depths) associated with the crust of the overriding Australian plate (fig. 6.2).

In the presence of anisotropy or dipping structure, the amplitude and timing of pulses on both the radial and transverse RFs depends on the orientation/geometry of the structure (i.e., strike of dipping structure or fast-axis direction of anisotropy) relative to the BAZ of the event used to compute the RF. Forward modelling of RFs suggests that dipping structure or plunging axis of anisotropy characterized by hexagonal symmetry produces pulses that vary in amplitude and timing with a 360-degree periodicity with respect to BAZ (Porter et al., 2011; Schulte-Pelkum & Mahan, 2014). Such pulses are observed in our RF dataset (fig. 6.2). A heuristic approach to constrain the orientation of the structure beneath a seismic station, without requiring complex waveform modelling, is to decompose the RF data at a given station into a set of harmonics with respect to BAZ (Bianchi et al., 2010; Schulte-Pelkum & Mahan, 2014; Audet, 2015; Schulte-Pelkum et al., 2020). A set of N radial and transverse RFs ($R_i(t)$ and $T_i(t)$ for $i = 1, \dots, N$), that are functions of time t , can be approximated by 5 harmonic components according to

$$\begin{bmatrix} R_1(t) \\ \vdots \\ R_N(t) \\ T_1(t) \\ \vdots \\ T_N(t) \end{bmatrix} = \begin{bmatrix} 1 & \cos(\phi_1 - \alpha) & \sin(\phi_1 - \alpha) & \cos(2(\phi_1 - \alpha)) & \sin(2(\phi_1 - \alpha)) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & \cos(\phi_N - \alpha) & \sin(\phi_N - \alpha) & \cos(2(\phi_N - \alpha)) & \sin(2(\phi_N - \alpha)) \\ 0 & \cos(\phi_1 - \alpha + \frac{\pi}{2}) & \sin(\phi_1 - \alpha + \frac{\pi}{2}) & \cos(2(\phi_1 - \alpha) + \frac{\pi}{4}) & \sin(2(\phi_1 - \alpha) + \frac{\pi}{4}) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cos(\phi_N - \alpha + \frac{\pi}{2}) & \sin(\phi_N - \alpha + \frac{\pi}{2}) & \cos(2(\phi_N - \alpha) + \frac{\pi}{4}) & \sin(2(\phi_N - \alpha) + \frac{\pi}{4}) \end{bmatrix} \begin{bmatrix} A(t) \\ B_1(t) \\ B_2(t) \\ C_1(t) \\ C_2(t) \end{bmatrix}. \quad (6.1)$$

In the equation above, ϕ_i is the BAZ of the event that produced the i^{th} radial and transverse RFs. From the transformation matrix in equation (6.1), we can see that the $A(t)$ harmonic component represents the mean amplitude of the radial RFs over all BAZ, and therefore approximates the isotropic structure beneath the station. The $B_1(t)$ and $B_2(t)$ harmonic components represent structure that displays a periodicity of 360 degrees with respect to BAZ (i.e., dipping structure or plunging axis of anisotropy). The last two harmonic components, $C_1(t)$ and $C_2(t)$, represent structure that displays a periodicity of 180 degrees with respect to BAZ, which is often attributed to anisotropy with a horizontal axis of hexagonal symmetry (i.e., azimuthal anisotropy). In practice, it is difficult to constrain these components as RFs are required from earthquakes over a complete range of BAZ directions in order to adequately sample the 180-degree periodicity.

In a geographic reference frame, the coordinate system of the BAZ harmonics is defined by sine and cosine functions that are aligned with the north-south and east-west directions. This implies that dipping structure (or plunging anisotropy) that is aligned in a north-south orientation will be described by only one of the $B(t)$ components (i.e., the complementary $B(t)$ component contains no signal). Structure that is not aligned in either a north-south or east-west direction will produce a signal on both $B(t)$ components. By introducing the phase shift α in equation (6.1), the reference frame can be rotated to determine the orientation of the coordinate system that produces maximum signal amplitude on a particular $B(t)$ component (Audet, 2015), and therefore the orientation of geologic structure. In this study, we search for the angle α that maximizes signal variance on the $B_2(t)$ component in order to estimate the strike orientations of dipping structure, or the trends of the axes of plunging anisotropy. We specifically analyze RF data over the time window from 0.5 seconds (to avoid the direct-P arrival) to the P-S delay time associated with the depth of the plate interface (based on the model by Williams et al., 2013), which is unique beneath each station. The P-S delay times are estimated using a unique velocity profile for each station, based on a 3D seismic tomography model of New Zealand (Eberhart-Phillips et al., 2020). Consequently, our results represent the dominant orientation of structure within the crust of the overriding Australian plate.

6.3 Results

We calculate harmonic decompositions of RF data for all stations over the forearc of the HSZ. For each station, we estimate the orientation of the strike of dipping structure, or trend of plunging anisotropy, within the overriding Australian plate (fig. 6.3a). In order to determine the uncertainty in the estimated orientations, we perform a bootstrap analysis by randomly re-sampling the RF data for each station (with replacement) and recomputing the harmonic decompositions 100 times. The black bow ties in figure 6.3a represent the 80 percent confidence interval on the estimated orientation for each station. Overall, our results show good agreement between proximal stations throughout the study region.

Throughout the southern HSZ, our estimated orientations are generally aligned with the margin as well as the strike of the plate interface (fig. 6.1). However, the orientations of structure beneath stations in the northern HSZ are predominantly aligned in a WSW-ENE direction. We are not able to distinguish between plunging anisotropy and dipping structure from this analysis alone. Furthermore, the estimated orientations represent the alignment of the dominant structure (or structures) beneath each station within the overriding plate. Hence, our analysis is not able to distinguish multiple dipping structures or multiple anisotropic layers (or combinations thereof).

Figure 6.3b shows the magnitudes of the $B_2(t)$ harmonic components as a percentage of the total signal in the RF data for each station (over the specified time window corresponding to depths of the overriding Australian plate). These magnitude values represent the amount of signal variance in the data that is described by a 360-degree periodicity with respect to BAZ. Hence, this describes the relative strength of the signal associated with dipping structure or plunging anisotropy. For stations in the northern HSZ, approximately 25-30% of the RF data variance can be described by dipping structure or plunging anisotropy. In the northern part of the southern HSZ, over 40% of the RF dataset can be described by dipping structure or plunging anisotropy, decreasing to approximately 20% towards the southernmost point of the North Island.

The estimated orientations, and relative magnitudes, from the harmonic decompositions of RF data suggest significant margin-parallel heterogeneity in tectonic fabric within

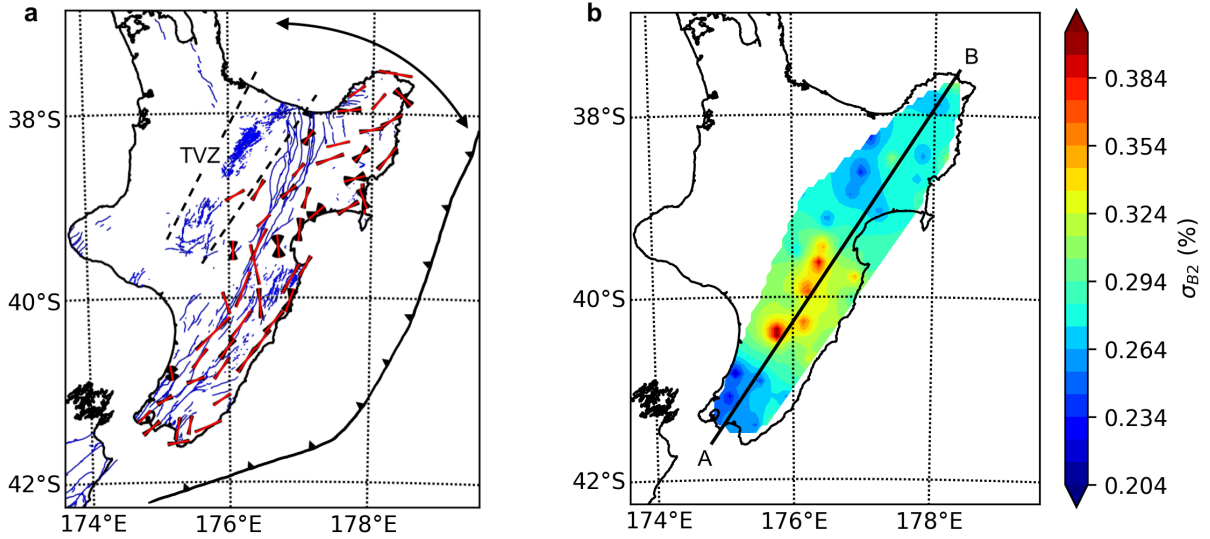


Figure 6.3: Receiver function harmonic decomposition orientations (a) and magnitudes (b). The red bars in (a) show the estimated strike orientations of dipping structure, or trend orientations of plunging anisotropy, beneath each station. The lengths of the red bars are scaled according to the magnitudes of the harmonic signals shown in (b). The black bow ties in (a) show the 80 percent confidence intervals around the estimated orientations. The surface traces of known active faults are shown in blue in (a). The offshore deformation front is shown as a thick black line in (a). The Taupo Volcanic Zone (TVZ) is delineated by dashed lines and is characterized by rifting and extensive normal faulting. The black arrow in (a) illustrates extension and rotation of the forearc in the northern Hikurangi subduction zone. The percentage of the total receiver function amplitudes (at each station) that is characterized by dipping structure or plunging anisotropy (over time windows corresponding to the overriding Australian plate) is shown in (b). The transect A-B is shown in figure 6.4.

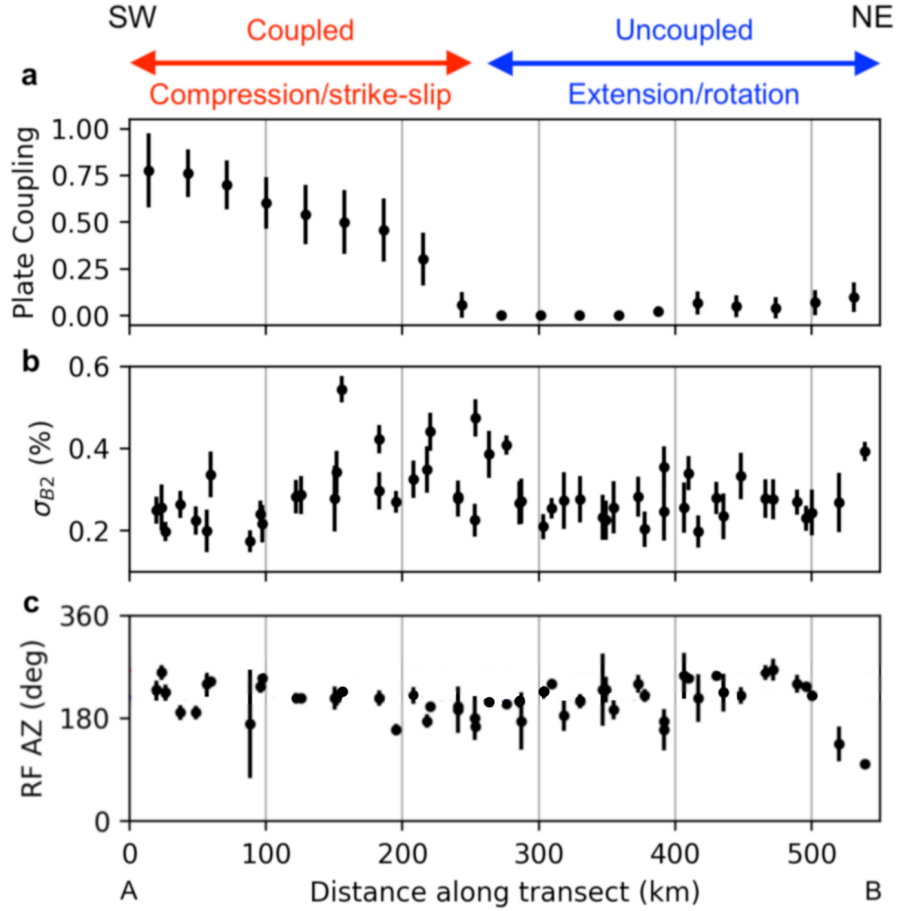


Figure 6.4: Along-strike heterogeneity in forearc deformation. The average value of plate coupling (Wallace, 2020) over the forearc along the transect A-B (shown in figure 6.3), in 25km intervals, is shown in (a). The error bars in (a) are the standard deviation of plate coupling values over the forearc, perpendicular to the transect (within 25km). The relative magnitude of the $B_2(t)$ harmonic component for each station, projected onto the transect A-B, is shown in (b). Similarly, receiver function harmonic decomposition orientations are shown in (c). Error bars in (b) and (c) show the 80 percent confidence intervals from bootstrap analysis.

the forearc. From forward modelling of RF data (Frederiksen & Bostock, 2000), we identify several possible causes for increases in the relative magnitude of the $B_2(t)$ harmonic component including (1) an increase in the magnitude of a seismic velocity contrast across a dipping structure, (2) an increase in the strength of anisotropy, (3) the superposition of multiple structures with the same alignment, (4) an increase in the dip/plunge angle of structure, and lastly (5) a decrease in the relative contribution of other harmonic components. It is unlikely that the variability in our results over the entire forearc can be explained by a single factor. Several recent studies observe along-margin variability in seismic velocities (e.g., Reyners & Eberhart-Phillips, 2009; Eberhart-Phillips et al., 2010), seismic attenuation (e.g., Eberhart-Phillips et al., 2017), and electrical resistivity (e.g., Heise et al., 2013), suggesting our estimated changes in the relative magnitude of the $B_2(t)$ harmonic components may be partially attributed to heterogeneity in bulk geophysical properties along the margin. However, based on our estimated structural orientations, we suggest that the dominant source of the $B_2(t)$ harmonic signal is different between the southern HSZ and the northern HSZ, and that observed changes across the southern HSZ (where the plate interface is coupled) likely reflect changes in subduction morphology.

Motion of the Pacific plate relative to the Australian plate is oblique to the margin (Kreemer et al., 2003), with the obliquity increasing southward (fig. 6.1). Margin-parallel relative plate motion is accommodated predominantly by a general clockwise rotation of the entire forearc (Wallace et al., 2004), as well as by a belt of major dextral strike-slip faults spanning the entire forearc (Litchfield et al., 2014), as shown in figure 6.1. Motion on these faults is believed to accommodate a significant portion of margin-parallel relative plate motion in the southern HSZ, decreasing to almost no motion in the northern HSZ (Litchfield et al., 2014). Active-source seismic imaging suggests that these faults are splayed, and likely intersect the plate interface at depth (e.g., Henrys et al., 2013). Hence, these faults are expected to be dipping structures, and shearing along these faults could impart a foliation within the crust of the overriding Australian plate, generally aligned with their strike orientation (which is similar to the strike of the plate interface in this region). We suggest that the observed margin-parallel variations in $B_2(t)$ harmonic orientations and magnitudes in the southern HSZ, where the plate interface is coupled, are due to

changes in the morphology of the plate interface as well as a progressive increase in strike-slip deformation southwards (fig. 6.4). We also note that the southward decrease in $B_2(t)$ harmonic magnitudes in the southern HSZ (figs. 6.3b and 6.4b) correlates with a significant change in alignment of the deformation front offshore and corresponding changes in plate interface strike (fig. 6.1). Consequently, the geometry and morphology of deformation within the overriding plate is expected to vary accordingly across this region, altering the geometry, and potentially the strength, of crustal foliations. However, long time periods may be required in order to impart foliations within the forearc crust, and the orientations of structures relative to the strike of the plate interface (and offshore deformation front) are not constant throughout the geologic history of the HSZ.

In the northern HSZ, where the plate interface is not coupled, we suggest that the orientations estimated from the harmonic decomposition of RF data are not related to subduction directly. Instead, we propose that the general clockwise rotation of orientations, towards a WSW-ENE direction, may be related to extension and rifting of the Taupo Volcanic Zone (TVZ, fig. 6.3a) further inland. We note that our results are consistent with estimates of fast-axis directions of anisotropy derived from measurements of shear-wave splitting from local earthquakes by Illsley-Kemp et al. (2019), who also observe a rotation in structural orientations along the margin towards a similar WSW-ENE direction in the northern HSZ. The orientations may represent structural fabrics imparted by past deformation, which have subsequently been rotated clockwise with the overall rotation of the forearc (Wallace et al., 2004). Alternatively, extension in a NNW-SSE direction suggests that the direction of maximum horizontal stress is oriented in an orthogonal (i.e., WSW-ENE) direction, consistent with our observations. In this case, our results may represent the preferential opening of fluid-filled microfractures that are oriented in a WSW-ENE direction, within the extensional forearc crust. This would result in a fast-axis of anisotropy consistent with our observations. Furthermore, geochemical analysis of onshore spring waters throughout the forearc suggests greater permeability of the upper plate in the northern HSZ compared to the southern HSZ, which may be related to changes in tectonic stress state (i.e., compression vs. extension) along the margin (Reyes et al., 2010; Barnes et al., 2019). Lastly, changes in stress state offer a possible explanation for

the aforementioned observations of along-strike variability in bulk geophysical properties (e.g., Reyners & Eberhart-Phillips, 2009; Heise et al., 2013; Eberhart-Phillips et al., 2017), which may reflect changes in the abundance and inter-connectivity of crustal fluids.

6.4 Conclusions

The HSZ displays enigmatic margin-parallel changes in plate coupling and SSE locations, suggesting heterogeneity in the physical and frictional state of the megathrust. In this work, we compute RFs from teleseismic data recorded at 53 long-running seismic stations over the forearc of the HSZ in order to constrain seismic scattering structure. We observe a strong BAZ-dependence in the timing and amplitude/polarity of RF pulses at P-S delay times associated with depths of the overriding forearc crust. We decompose the RF data into harmonic components in order to constrain the orientations of associated structures that cause these signals. We observe significant along-strike variability in the magnitudes and orientations of $B_2(t)$ harmonic signals, suggesting heterogeneous deformation throughout the forearc. In the southern HSZ, we suggest that our observations are related to splayed dextral faults cutting through the forearc crust, which change in morphology along strike. In the northern HSZ, where the overriding forearc crust is uncoupled from the subducting Pacific slab, we suggest that our observations are related to past deformation, clockwise rotation, and/or extension of the forearc crust. Consequently, our results provide further evidence of variability in tectonic stress state of the forearc crust along the margin, which is likely related to the state of interseismic coupling. This work yields further evidence of the heterogeneous structure and dynamics of the HSZ, and provides useful constraints on the relationship between megathrust coupling, forearc deformation, and seismic hazards.

Thesis conclusions

This thesis explores several questions in seismology that vary in scope from the development of inverse methods to the tectonics and dynamics of subduction zones in multiple geographic regions. The central theme and motivation of this thesis is to improve our understanding of active tectonics in western Canada, as well as along the North Island of New Zealand. This thesis advocates for uncertainty estimation in passive seismic imaging in order to inform reliable interpretations of lithosphere structure. Furthermore, this thesis produces multiple data products that are valuable for future investigation of the tectonics of western Canada. This thesis also provides valuable seismic constraints that improve our understanding of subduction zone processes. As this work represents a collection of individual studies, the main conclusions and contributions are already summarized within each chapter. Instead, here I focus on limitations and potential opportunities for future research that further develop the ideas presented in this thesis.

· Part I :

Bayesian (probabilistic) approaches to inverse problems provide rigorous uncertainty quantification, and are developed for applications in passive seismic imaging in part I of this thesis. Both chapter 2 and chapter 3 produce intermediate results towards the ultimate goal of resolving a complete 3D seismic velocity model of western Canada in future work. This will further improve our understanding of the transition between northern Cascadia, the southern Canadian Cordillera, and the Haida Gwaii region. This is particularly useful for providing a broad structural context as future seismic

station deployments are planned for these regions. Unfortunately, I was not able to perform this investigation in the time allotted for my doctoral program.

This work predominantly considers bounded uniform prior distributions on model parameters within Bayesian inversion. However, one of the great advantages of Bayesian methods is the ability to incorporate meaningful prior information into the problem. Estève et al. (2021) considered Bayesian inversion of surface-wave dispersion. They applied prior constraint on discontinuous structure based on independent receiver function (RF) observations, and demonstrate improved models (in terms of uncertainty and recovered structure). Future research involving the intermediate products produced in part I of this thesis should incorporate additional model constraints through the use of informative prior information. Alternatively, independent data may be incorporated via joint inversion.

- **Chapter 2** develops a Bayesian approach to surface-wave tomography that estimates azimuthal anisotropy and associated uncertainty of surface-wave group velocities. A model parameterization based on linear interpolation within the Delaunay triangulation of a distributed set of nodes was implemented in this work. This improved inversion algorithm performance and produced more realistic individual model samples when compared with previous approaches to Bayesian tomography. Although this is a physically intuitive model that represents an improvement over previous studies, more efficient (though less intuitive) representations exist such as sparse models based on discrete wavelet transforms (e.g., Fang et al., 2015; Hawkins & Sambridge, 2015). These could be implemented within this Bayesian tomographic framework, and may help reduce the computational demand of this approach. Furthermore, future work could consider decoupling the parameterizations of isotropic and anisotropic structure (e.g., Bodin et al., 2016).

The tomographic method solved for parameters of a data error model that is linearly dependent on propagation path length. The implementation of this model produced interesting results (in terms of dependence on surface-wave pe-

riod), and illustrates the ability of Bayesian inversion to gain insight on data error sources and processes. The method could be improved by including parameters that account for data error correlations for common surface-wave paths (i.e., source terms and station terms), especially since errors in offshore earthquake locations are believed to be significant. However, the inclusion of additional model parameters typically increases the computational cost of the inversion procedure, which is already relatively impractical (and potentially unfeasible) for most problems.

Ultimately, this chapter demonstrates the value of this novel inversion technique for robust determination of anisotropy from surface wave dispersion data. However, this work also highlights that surface-wave tomography is only weakly nonlinear for regional-scale models based on dispersion data at periods greater than ~ 15 s. Consequently, such sophisticated inverse methods may not be appropriate for all problems, particularly given the computational expense of this approach.

- **Chapter 3** develops a Bayesian approach to estimate two-station surface-wave dispersion from measurements of phase spectra, and implements a data error model for Bayesian inversion of circular data. Fortunately, the computational demand for 1D Bayesian inversion is relatively minor. With this approach, surface-wave dispersion can be estimated for thousands of station pairs significantly faster than by manual selection. The greatest limitation of this work is the potential bias due to the unknown propagation distance/path and arrival azimuth of teleseismic surface waves, which is discussed in the chapter.

It is my impression that this work represents a valuable proof-of-concept of Bayesian inversion of circular data. It is challenging to accurately propagate uncertainty through multiple analysis steps (dispersion estimation, tomography for phase-velocity maps, then inversion for pseudo-3D Earth structure). Distortion and bias can be introduced at every step in ways that are difficult to predict and account for. The future of such methods likely involves the direct estimation of 3D structure from phase measurements. Which becomes similar

to full-waveform inversion in the Fourier domain (without amplitude consideration). Such an approach may be feasible for regional surface-wave studies within a Bayesian framework as computation capabilities continue to improve. Lastly, the ideas presented in this chapter can be applied to other problems that consider circular measurements.

· Part II :

The study of subduction zones is valuable from scientific, economic, and societal (i.e., public safety) perspectives. Part II of this thesis analyzed thousands of recordings of distant earthquakes (spanning several decades) from stations over the Cascadia subduction zone (Canada) and the Hikurangi subduction zone (New Zealand). This work studied mechanical conditions in relation to deformation, plate locking, and slow slip events. This work provides useful constraints for advancing our knowledge of these tectonic features, but also merely scratches the surface of this topic.

Methodologically, RF analysis is a valuable tool but possesses significant limitations. The assumptions of planar wavefronts, as well as planar structural discontinuities (over 10s of km), are significant. The method is also severely limited in resolution due to the frequency band of teleseismic body-wave arrivals. Lastly, although we applied a robust Weiner deconvolution to estimate RFs in this work, spectral division is inherently an unstable process. Recent studies have proposed methods to bypass the deconvolution procedure altogether (e.g. Bodin et al., 2014; Dettmer et al., 2015; Frederiksen & Delaney, 2015). Specifically, these strategies have been shown to overcome challenges with deconvolution in settings where the seismic station is situated on top of thick, low-velocity sediments (e.g. Frederiksen, 2020) as well as in marine environments (e.g. Audet, 2016). My interest in seismic methods will likely lead me to pursue research in improving the analysis of teleseismic body-wave recordings.

Studies on subduction zones are considered from a wide range of disciplines (e.g., geochemistry, petrology, seismology, structural geology, geodesy, other geophysical fields, etc.). For example, the results from both chapter 5 and chapter 6 are supported

by geochemical analysis of well waters from the overriding crust (e.g., Reyes et al., 2010; McCrory et al., 2016; Barnes et al., 2019). Multidisciplinary science allows for comprehensive understanding of subduction zone structure and processes, and should be a focus of future research.

- **Chapter 5** suggests that mechanical properties near the plate interface in subduction zones change during slow slip events, based on RF data. The temporal and spatial sampling of this observation is sparse. There are not enough seismic stations with 10-20 years of recorded data that are located directly above the slow slip source region. Furthermore, large earthquakes suitable for RF analysis are relatively infrequent. The conclusions from this work are based on hypothesis testing as well as from combining observations over multiple slow slip events. Warren-Smith et al. (2019) observed a similar phenomenon based on stress state analysis from microseismicity, and were able to resolve changes over individual slow slip events. Such approaches are likely more promising for future study of temporal changes in subduction zone structure. However, evenly distributed microseismicity is not ubiquitous to all subduction zones.
- **Chapter 6** suggests deformation and stress state throughout the forearc of the Hikurangi subduction zone are linked to plate locking. This is based on RF analysis at over 50 seismic stations throughout the region. The Hikurangi subduction zone is complex, and represents a snapshot of the transitional deformation that eventually results in the transform Alpine fault system over the South Island. The overriding plate is heavily deformed, and the subducting oceanic crust is abnormally thick (Mochizuki et al., 2019). Consequently, RFs from stations over the North Island do not clearly reveal the tell-tale low-velocity zone as observed in other subduction zones. RFs in the Hikurangi subduction zone reveal abundant coherent scattering structure. The work presented in this chapter represents a heuristic approach to analyzing these data in order to make broad tectonic inference over the entire region. There remains much to resolve from these data, which may require station-specific interpretation.

Final thoughts

The pace of seismological discovery, and seismic imaging, is tied to technological advancement. Improved instrumentation leads to larger volumes of higher-quality seismic data, sometimes in new locations. Improved computation provides the ability to process these data, as well as model complex seismic wave phenomenology more accurately. New methods are required to examine seismic data from different perspectives to fully exploit the benefits and information available from next-generation measurements. There are many exciting frontiers in seismology.

The British geologist Herbert Harold Read once stated that “*the best geologist is [the one] who has seen the most rocks*”. As academics become increasingly specialized, it can become challenging to retain a broad perspective and an open mind. Specialization advances a field, yet there exists significant value in the knowledge from adjacent fields. Perhaps, today, H. H. Read would say that the best *scientist* is the one who has explored the most *problems*. I view this (to some extent) as a goal for my career. The process of writing and assembling this thesis has helped me reflect on what I have worked on during my doctoral program, which I am proud of in hindsight. This process has also reminded me of my failed ideas (which are numerous and do not appear in this work), as well as (more optimistically) ideas that I have yet to explore.

References

- Aki, K. & Richards, P. G., 2002. *Quantitative seismology*, Freeman, San Francisco.
- Aster, R. C., Borchers, B., & Thurber, C. H., 2018. *Parameter estimation and inverse problems*, Elsevier.
- Astiz, L., Earle, P., & Shearer, P., 1996. Global stacking of broadband seismograms, *Seismological Research Letters*, **67**(4), 8–18.
- Audet, P., 2010. Temporal variations in crustal scattering structure near Parkfield, California, using receiver functions, *Bulletin of the Seismological Society of America*, **100**(3), 1356–1362.
- Audet, P., 2015. Layered crustal anisotropy around the San Andreas fault near Parkfield, California, *Journal of Geophysical Research: Solid Earth*, **120**(5), 3527–3543.
- Audet, P., 2016. Receiver functions using OBS data: Promises and limitations from numerical modelling and examples from the Cascadia Initiative, *Geophysical Journal International*, **205**(3), 1740–1755.
- Audet, P. & Bürgmann, R., 2014. Possible control of subduction zone slow-earthquake periodicity by silica enrichment, *Nature*, **510**(7505), 389–392.
- Audet, P. & Kim, Y., 2016. Teleseismic constraints on the geological environment of deep episodic slow earthquakes in subduction zone forearcs: A review, *Tectonophysics*, **670**, 1–15.

- Audet, P. & Schaeffer, A. J., 2018. Fluid pressure and shear zone development over the locked to slow slip region in Cascadia, *Science advances*, **4**(3), eaar2982.
- Audet, P., Bostock, M., Mercier, J.-P., & Cassidy, J., 2008. Morphology of the Explorer–Juan de Fuca slab edge in northern Cascadia: Imaging plate capture at a ridge-trench-transform triple junction, *Geology*, **36**(11), 895–898.
- Audet, P., Bostock, M. G., Christensen, N. I., & Peacock, S. M., 2009. Seismic evidence for overpressured subducted oceanic crust and megathrust fault sealing, *Nature*, **457**(7225), 76–78.
- Audet, P., Bostock, M. G., Boyarko, D. C., Brudzinski, M. R., & Allen, R. M., 2010. Slab morphology in the Cascadia fore arc and its relation to episodic tremor and slip, *Journal of Geophysical Research: Solid Earth*, **115**(B4).
- Babuska, V. & Cara, M., 1991. *Seismic anisotropy in the Earth*, vol. 10, Springer Science & Business Media.
- Balfour, N., Cassidy, J., Dosso, S., & Mazzotti, S., 2011. Mapping crustal stress and strain in southwest British Columbia, *Journal of Geophysical Research: Solid Earth*, **116**, B03314.
- Balfour, N., Cassidy, J., & Dosso, S., 2012. Crustal anisotropy in the forearc of the northern Cascadia subduction zone, British Columbia, *Geophysical Journal International*, **188**(1), 165–176.
- Bannister, S., Reyners, M., Stuart, G., & Savage, M., 2007. Imaging the Hikurangi subduction zone, New Zealand, using teleseismic receiver functions: crustal fluids above the forearc mantle wedge, *Geophysical Journal International*, **169**(2), 602–616.
- Bao, X., Eaton, D. W., & Gu, Y. J., 2016. Rayleigh wave azimuthally anisotropic phase velocity maps beneath western Canada, *Journal of Geophysical Research: Solid Earth*, **121**(3), 1821–1834.

- Barnes, J. D., Cullen, J., Barker, S., Agostini, S., Penniston-Dorland, S., Lassiter, J. C., Klügel, A., & Wallace, L., 2019. The role of the upper plate in controlling fluid-mobile element (Cl, Li, B) cycling through subduction zones: Hikurangi forearc, New Zealand, *Geosphere*, **15**(3), 642–658.
- Behr, W. M. & Bürgmann, R., 2021. What’s down there? the structures, materials and environment of deep-seated slow slip and tremor, *Philosophical Transactions of the Royal Society A*, **379**(2193), 20200218.
- Bensen, G., Ritzwoller, M., Barmin, M., Levshin, A. L., Lin, F., Moschetti, M., Shapiro, N., & Yang, Y., 2007. Processing seismic ambient noise data to obtain reliable broadband surface wave dispersion measurements, *Geophysical Journal International*, **169**(3), 1239–1260.
- Bensen, G., Ritzwoller, M., & Shapiro, N. M., 2008. Broadband ambient noise surface wave tomography across the United States, *Journal of Geophysical Research: Solid Earth*, **113**(B5).
- Beroza, G. C. & Ide, S., 2011. Slow earthquakes and nonvolcanic tremor, *Annual review of Earth and planetary sciences*, **39**, 271–296.
- Bérubé, J., Rogers, G. C., Ellis, R. M., & Hasselgren, E. O., 1989. A microseismicity study of the Queen Charlotte Islands region, *Canadian Journal of Earth Sciences*, **26**(12), 2556–2566.
- Bianchi, I., Park, J., Piana Agostinetti, N., & Levin, V., 2010. Mapping seismic anisotropy using harmonic decomposition of receiver functions: An application to Northern Apennines, Italy, *Journal of Geophysical Research: Solid Earth*, **115**(B12).
- Bird, A., Rogers, G., & Spence, G., 1997. Earthquakes in the Queen Charlotte Islands region: 1984–1996, *Lithoprobe*, **10**, 39–44.
- Bodin, T. & Sambridge, M., 2009. Seismic tomography with the reversible jump algorithm, *Geophysical Journal International*, **178**(3), 1411–1436.

- Bodin, T., Sambridge, M., Rawlinson, N., & Arroucau, P., 2012. Transdimensional tomography with unknown data noise, *Geophysical Journal International*, **189**(3), 1536–1556.
- Bodin, T., Yuan, H., & Romanowicz, B., 2014. Inversion of receiver functions without deconvolution—Application to the Indian craton, *Geophysical Journal International*, **196**(2), 1025–1033.
- Bodin, T., Leiva, J., Romanowicz, B., Maupin, V., & Yuan, H., 2016. Imaging anisotropic layering with Bayesian inversion of multiple data types, *Geophysical Journal International*, **206**(1), 605–629.
- Bostock, M., 1998. Mantle stratigraphy and evolution of the Slave province, *Journal of Geophysical Research: Solid Earth*, **103**(B9), 21183–21200.
- Bostock, M., 2013. The Moho in subduction zones, *Tectonophysics*, **609**, 547–557.
- Bostock, M. & Christensen, N., 2012. Split from slip and schist: Crustal anisotropy beneath northern Cascadia from non-volcanic tremor, *Journal of Geophysical Research: Solid Earth*, **117**(B8).
- Bostock, M. G., Christensen, N. I., & Peacock, S. M., 2019. Seismicity in Cascadia, *Lithos*, **332**, 55–66.
- Bostwick, T. K., 1984. *A re-examination of the August 22, 1949 Queen Charlotte earthquake*, Ph.D. thesis, University of British Columbia.
- Brandon, M. T., Roden-Tice, M. K., & Garver, J. I., 1998. Late Cenozoic exhumation of the Cascadia accretionary wedge in the Olympic Mountains, northwest Washington State, *Geological Society of America Bulletin*, **110**(8), 985–1009.
- Braunmiller, J. & Nábelek, J., 2002. Seismotectonics of the Explorer region, *Journal of Geophysical Research: Solid Earth*, **107**(B10), ETG–1.
- Brooks, S., Gelman, A., Jones, G., & Meng, X. L., 2011. *Handbook of Markov Chain Monte Carlo*, Chapman & Hall/CRC, Boca Raton, FL.

- Brothers, D. S., Miller, N. C., Barrie, J. V., Haeussler, P. J., Greene, H. G., Andrews, B. D., Zielke, O., Watt, J., & Dartnell, P., 2020. Plate boundary localization, slip-rates and rupture segmentation of the Queen Charlotte Fault based on submarine tectonic geomorphology, *Earth and Planetary Science Letters*, **530**, 115882.
- Burdick, L. J. & Langston, C. A., 1977. Modeling crustal structure through the use of converted phases in teleseismic body-wave forms, *Bulletin of the Seismological Society of America*, **67**(3), 677–691.
- Bürgmann, R., 2018. The geophysics, geology and mechanics of slow fault slip, *Earth and Planetary Science Letters*, **495**, 112–134.
- Byerlee, J. D., 1993. Model for episodic flow of high-pressure water in fault zones before earthquakes, *Geology*, **21**(4), 303–306.
- Calvert, A. J., 2004. Seismic reflection imaging of two megathrust shear zones in the northern Cascadia subduction zone, *Nature*, **428**(6979), 163–167.
- Cao, L., Kao, H., & Wang, K., 2017. Contrasting upper-mantle shear wave anisotropy across the transpressive Queen Charlotte margin, *Tectonophysics*, **717**, 311–320.
- Cassidy, J., 1992. Numerical experiments in broadband receiver function analysis, *Bulletin of the Seismological Society of America*, **82**(3), 1453–1474.
- Cassidy, J. F. & Ellis, R. M., 1993. S wave velocity structure of the northern Cascadia subduction zone, *Journal of Geophysical Research: Solid Earth*, **98**(B3), 4407–4421.
- Cassidy, J. F., Rogers, G. C., & Hyndman, R. D., 2014. An overview of the 28 October 2012 M w 7.7 earthquake in Haida Gwaii, Canada: a tsunamigenic thrust event along a predominantly strike-slip margin, *Pure and Applied Geophysics*, **171**(12), 3457–3465.
- Chib, S., 1995. Marginal likelihood from the Gibbs output, *Journal of the American Statistical Association*, **90**, 1313–132.

- Christensen, N. I., 1984. Pore pressure and oceanic crustal seismic structure, *Geophysical Journal International*, **79**(2), 411–423.
- Clayton, R. W. & Wiggins, R. A., 1976. Source shape estimation and deconvolution of teleseismic bodywaves, *Geophysical Journal International*, **47**(1), 151–177.
- Clowes, R. M., Zelt, C. A., Amor, J. R., & Ellis, R. M., 1995. Lithospheric structure in the southern Canadian Cordillera from a network of seismic refraction lines, *Canadian Journal of Earth Sciences*, **32**(10), 1485–1513.
- Constable, S. C., Parker, R. L., & Constable, C. G., 1987. Occam’s inversion: A practical algorithm for generating smooth models from electromagnetic sounding data, *Geophysics*, **52**(3), 289–300.
- Crampin, S., 1994. The fracture criticality of crustal rocks, *Geophysical Journal International*, **118**(2), 428–438.
- Currie, C. A., Cassidy, J. F., & Hyndman, R. D., 2001. A regional study of shear wave splitting above the Cascadia subduction zone: Margin-parallel crustal stress, *Geophysical Research Letters*, **28**(4), 659–662.
- Dalton, C. A., Gaherty, J. B., & Courtier, A. M., 2011. Crustal VS structure in northwestern Canada: Imaging the Cordillera-craton transition with ambient noise tomography, *Journal of Geophysical Research: Solid Earth*, **116**(B12).
- Darbyshire, F. A. & Lebedev, S., 2009. Rayleigh wave phase-velocity heterogeneity and multilayered azimuthal anisotropy of the Superior Craton, Ontario, *Geophysical Journal International*, **176**(1), 215–234.
- Darbyshire, F. A., Dahl-Jensen, T., Larsen, T. B., Voss, P. H., & Joyal, G., 2018. Crust and uppermost-mantle structure of Greenland and the Northwest Atlantic from Rayleigh wave group velocity tomography, *Geophysical Journal International*, **212**(3), 1546–1569.
- Darot, M. & Bouchez, J., 1976. Study of directional data distributions from principal preferred orientation axes, *The Journal of Geology*, **84**(2), 239–247.

- Davies, G. F., 2001. *Dynamic Earth: Plates, plumes and mantle convection*, American Association of Physics Teachers.
- Debayle, E. & Sambridge, M., 2004. Inversion of massive surface wave data sets: Model construction and resolution assessment, *Journal of Geophysical Research: Solid Earth*, **109**(B2).
- Dettmer, J., Dosso, S. E., & Holland, C., 2010. Trans-dimensional geoacoustic inversion, *Journal of the Acoustical Society of America*, **128**(6), 3393–3405.
- Dettmer, J., Molnar, S., Steininger, G., Dosso, S. E., & Cassidy, J. F., 2012. Trans-dimensional inversion of micro-tremor array dispersion data with hierarchical autoregressive error models, *Geophysical Journal International*, **118**, 719–734.
- Dettmer, J., Dosso, S. E., Bodin, T., Stipčević, J., & Cummins, P. R., 2015. Direct-seismogram inversion for receiver-side structure with uncertain source–time functions, *Geophysical journal international*, **203**(2), 1373–1387.
- Dosso, S. E. & Wilmut, M. J., 2006. Data uncertainty estimation in matched-field geoacoustic inversion, *IEEE Journal of Oceanic Engineering*, **31**(2), 470–479.
- Dosso, S. E. & Wilmut, M. J., 2008. Uncertainty estimation in simultaneous Bayesian tracking and environmental inversion, *Journal of the Acoustical Society of America*, **124**, 82–97.
- Dosso, S. E. & Wilmut, M. J., 2012. Maximum-likelihood and other processors for incoherent and coherent matched-field localization, *Journal of the Acoustical Society of America*, **132**(4), 2273–2285.
- Dosso, S. E., Nielsen, P. L., & Wilmut, M. W., 2006. Data error covariance in matched-field inversion, *Journal of the Acoustical Society of America*, **119**, 208–219.
- Dosso, S. E., Holland, C. W., & Sambridge, M., 2012. Parallel tempering for strongly nonlinear geoacoustic inversion, *Journal of the Acoustical Society of America*, **132**, 3030–3040.

- Dosso, S. E., Dettmer, J., Steininger, G., & Holland, C. W., 2014. Efficient trans-dimensional Bayesian inversion for geoacoustic profile estimation, *Inverse Problems*, **30**, 114018.
- Dragert, H., Wang, K., & James, T. S., 2001. A silent slip event on the deeper Cascadia subduction interface, *Science*, **292**(5521), 1525–1528.
- Dragert, H., Wang, K., & Rogers, G., 2004. Geodetic and seismic signatures of episodic tremor and slip in the northern Cascadia subduction zone, *Earth, planets and space*, **56**(12), 1143–1150.
- Dunkin, J. W., 1965. Computation of modal solutions in layered, elastic media at high frequencies, *Bulletin of the Seismological Society of America*, **55**(2), 335–358.
- Earl, D. J. & Deem, M. W., 2005. Parallel tempering: Theory, applications, and new perspectives, *Physical Chemistry Chemical Physics*, **7**, 3910–3916.
- Eberhart-Phillips, D. & Reyners, M., 2012. Imaging the Hikurangi Plate interface region, with improved local-earthquake tomography, *Geophysical Journal International*, **190**(2), 1221–1242.
- Eberhart-Phillips, D., Reyners, M., Bannister, S., Chadwick, M., & Ellis, S., 2010. Establishing a versatile 3-D seismic velocity model for New Zealand, *Seismological Research Letters*, **81**(6), 992–1000.
- Eberhart-Phillips, D., Bannister, S., & Reyners, M., 2017. Deciphering the 3-D distribution of fluid along the shallow Hikurangi subduction zone using P-and S-wave attenuation, *Geophysical Journal International*, **211**(2), 1032–1045.
- Eberhart-Phillips, D., Bannister, S., Reyners, M., & Henrys, S., 2020. New Zealand Wide model 2.2 seismic velocity and Qs and Qp models for New Zealand. DOI:10.5281/zenodo.3779523.

- Eilon, Z. C. & Forsyth, D. W., 2020. Depth-dependent azimuthal anisotropy beneath the Juan de Fuca plate system, *Journal of Geophysical Research: Solid Earth*, **125**(8), e2020JB019477.
- Estève, C., Gosselin, J., Audet, P., Schaeffer, A., Schutt, D., & Aster, R., 2021. Surface wave tomography of the northern Canadian Cordillera using earthquake Rayleigh wave group velocities, *Journal of Geophysical Research: Solid Earth*, p. e2021JB021960.
- Fang, H., Yao, H., Zhang, H., Huang, Y.-C., & van der Hilst, R. D., 2015. Direct inversion of surface wave dispersion for three-dimensional shallow crustal structure based on ray tracing: methodology and application, *Geophysical Journal International*, **201**(3), 1251–1263.
- Fisher, N. I., 1995. *Statistical analysis of circular data*, cambridge university press.
- Fogel, L. J., Owens, A. J., & Walsh, M. J., 1966. *Artificial intelligence through simulated evolution*, Wiley.
- Frederiksen, A., 2020. Transfer functions between teleseismic data components, *Geophysical Journal International*, **221**(2), 1248–1263.
- Frederiksen, A. & Bostock, M., 2000. Modelling teleseismic waves in dipping anisotropic structures, *Geophysical Journal International*, **141**(2), 401–412.
- Frederiksen, A. & Delaney, C., 2015. Deriving crustal properties from the P Coda without deconvolution: the southwestern Superior Province, North America, *Geophysical Journal International*, **201**(3), 1491–1506.
- Galetti, E., Curtis, A., Baptie, B., Jenkins, D., & Nicolson, H., 2017. Transdimensional Love-wave tomography of the British Isles and shear-velocity structure of the East Irish Sea Basin from ambient-noise interferometry, *Geophysical Journal International*, **208**(1), 36–58.

- Gao, D., Wang, K., Insua, T. L., Sypus, M., Riedel, M., & Sun, T., 2018. Defining megathrust tsunami source scenarios for northernmost Cascadia, *Natural Hazards*, **94**(1), 445–469.
- Gao, H., Schmidt, D. A., & Weldon, R. J., 2012. Scaling relationships of source parameters for slow slip events, *Bulletin of the Seismological Society of America*, **102**(1), 352–360.
- García-Jerez, A., Seivane, H., Navarro, M., Martínez-Segura, M., & Piña-Flores, J., 2019. Joint analysis of Rayleigh-wave dispersion curves and diffuse-field HVSR for site characterization: the case of El Ejido town (SE Spain), *Soil Dynamics and Earthquake Engineering*, **121**, 102–120.
- Gilbert, F. & Backus, G. E., 1966. Propagator matrices in elastic wave and vibration problems, *Geophysics*, **31**(2), 326–332.
- Gilks, W. R., Richardson, S., & Spiegelhalter, D. J., 1996. *Markov Chain Monte Carlo in Practice*, Chapman and Hall, London, United Kingdom.
- Gosselin, J. M., Cassidy, J. F., & Dosso, S. E., 2015. Shear-wave velocity structure in the vicinity of the 2012 M w 7.8 Haida Gwaii earthquake from receiver function inversion, *Bulletin of the Seismological Society of America*, **105**(2B), 1106–1113.
- Gosselin, J. M., Dosso, S. E., Cassidy, J. F., Quijano, J. E., & Molnar, S., 2017. A gradient-based model parameterization using Bernstein polynomials in Bayesian inversion of surface-wave dispersion, *Geophysical Journal International*, **211**, 528–540.
- Gosselin, J. M., Cassidy, J. F., Dosso, S. E., & Brillon, C., 2018. Probabilistic seismic-hazard site assessment in Kitimat, British Columbia, from Bayesian inversion of surface-wave dispersion, *Canadian Geotechnical Journal*, **55**(7), 928–940.
- Gosselin, J. M., Audet, P., Schaeffer, A. J., Darbyshire, F. A., & Estève, C., 2021. Azimuthal anisotropy in Bayesian surface wave tomography: application to northern Cascadia and Haida Gwaii, British Columbia, *Geophysical Journal International*, **224**(3), 1724–1741.

- Green, P. J., 1995. Reversible jump Markov chain Monte Carlo computation and Bayesian model determination, *Biometrika*, **82**(4), 711–732.
- Green, P. J., 2003. Trans-dimensional Markov chain Monte Carlo, *Oxford Statistical Science Series*, pp. 179–198.
- Hannemann, K., Papazachos, C., Ohrnberger, M., Savvaidis, A., Anthymidis, M., & Lontsi, A. M., 2014. Three-dimensional shallow structure from high-frequency ambient noise tomography: New results for the Mygdonia basin-Euroseistest area, northern Greece, *Journal of Geophysical Research: Solid Earth*, **119**(6), 4979–4999.
- Hansen, P. C., 1992. Analysis of discrete ill-posed problems by means of the L-curve, *Society for Industrial and Applied Mathematics Review*, **34**(4), 561–580.
- Hansen, R. T., Bostock, M. G., & Christensen, N. I., 2012. Nature of the low velocity zone in Cascadia from receiver function waveform inversion, *Earth and Planetary Science Letters*, **337**, 25–38.
- Harris, R. N. & Wang, K., 2002. Thermal models of the middle America trench at the Nicoya Peninsula, Costa Rica, *Geophysical Research Letters*, **29**(21), 6–1.
- Haskell, N. A., 1953. The dispersion of surface waves on multilayered media, *Bulletin of the Seismological Society of America*, **43**(1), 17–34.
- Hastings, W. K., 1970. Monte Carlo sampling methods using Markov chains and their applications, *Biometrika*, **57**, 97–109.
- Hawkins, R. & Sambridge, M., 2015. Geophysical imaging using trans-dimensional trees, *Geophysical Journal International*, **203**(2), 972–1000.
- Hawkins, R. & Sambridge, M., 2019. An adjoint technique for estimation of interstation phase and group dispersion from ambient noise cross correlations, *Bulletin of the Seismological Society of America*, **109**(5), 1716–1728.

- Hawkins, R., Bodin, T., Sambridge, M., Choblet, G., & Husson, L., 2019. Trans-dimensional surface reconstruction with different classes of parameterization, *Geochemistry, Geophysics, Geosystems*, **20**(1), 505–529.
- Hedjazian, N., Bodin, T., & Métivier, L., 2019. An optimal transport approach to linearized inversion of receiver functions, *Geophysical Journal International*, **216**(1), 130–147.
- Heise, W., Caldwell, T. G., Bertrand, E. A., Hill, G. J., Bennie, S. L., & Ogawa, Y., 2013. Changes in electrical resistivity track changes in tectonic plate coupling, *Geophysical Research Letters*, **40**(19), 5029–5033.
- Henrys, S., Wech, A., Sutherland, R., Stern, T., Savage, M., Sato, H., Mochizuki, K., Iwasaki, T., Okaya, D., Seward, A., et al., 2013. SAHKE geophysical transect reveals crustal and subduction zone structure at the southern Hikurangi margin, New Zealand, *Geochemistry, Geophysics, Geosystems*, **14**(7), 2063–2083.
- Herrmann, R. B., 2013. Computer programs in seismology: An evolving tool for instruction and research, *Seismological Research Letters*, **84**(6), 1081–1088.
- Holland, J. H. et al., 1992. *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence*, MIT press.
- Husen, S. & Kissling, E., 2001. Postseismic fluid flow after the large subduction earthquake of Antofagasta, Chile, *Geology*, **29**(9), 847–850.
- Hutchinson, J., Kao, H., Spence, G., Obana, K., Wang, K., & Kodaira, S., 2019. Seismic Characteristics of the Nootka Fault Zone: Results from the Seafloor Earthquake Array Japan–Canada Cascadia Experiment (SeaJade), *Bulletin of the Seismological Society of America*, **109**(6), 2252–2276.
- Hyndman, R., 1995. The Lithoprobe corridor across the Vancouver Island continental margin: the structural and tectonic consequences of subduction, *Canadian Journal of Earth Sciences*, **32**(10), 1777–1802.

- Hyndman, R., 2015. Tectonics and structure of the Queen Charlotte fault zone, Haida Gwaii, and large thrust earthquakes, *Bulletin of the Seismological Society of America*, **105**(2B), 1058–1075.
- Hyndman, R., 2017. Lower-crustal flow and detachment in the North American Cordillera: a consequence of Cordillera-wide high temperatures, *Geophysical Journal International*, **209**(3), 1779–1799.
- Hyndman, R. D. & Peacock, S. M., 2003. Serpentinization of the forearc mantle, *Earth and Planetary Science Letters*, **212**(3-4), 417–432.
- Hyndman, R. D. & Rogers, G. C., 2010. Great earthquakes on Canada’s west coast: a review, *Canadian Journal of Earth Sciences*, **47**(5), 801–820.
- Hyndman, R. D. & Wang, K., 1993. Thermal constraints on the zone of major thrust earthquake failure: The Cascadia subduction zone, *Journal of Geophysical Research: Solid Earth*, **98**(B2), 2039–2060.
- Hyndman, R. D., McCrory, P. A., Wech, A., Kao, H., & Ague, J., 2015. Cascadia subducting plate fluids channelled to fore-arc mantle corner: ETS and silica deposition, *Journal of Geophysical Research: Solid Earth*, **120**(6), 4344–4358.
- Idehara, K., Yabe, S., & Ide, S., 2014. Regional and global variations in the temporal clustering of tectonic tremor activity, *Earth, Planets and Space*, **66**(1), 1–10.
- Illsley-Kemp, F., Savage, M. K., Wilson, C. J., & Bannister, S., 2019. Mapping stress and structure from subducting slab to magmatic rift: crustal seismic anisotropy of the North Island, New Zealand, *Geochemistry, Geophysics, Geosystems*, **20**(11), 5038–5056.
- Inzunza, D. A., Montalva, G. A., Leyton, F., Prieto, G., & Ruiz, S., 2019. Shallow ambient-noise 3d tomography in the Concepción Basin, Chile: Implications for low-frequency ground motions, *Bulletin of the Seismological Society of America*, **109**(1), 75–86.
- Jackson, D. D., 1972. Interpretation of inaccurate, insufficient and inconsistent data, *Geophysical Journal International*, **28**(2), 97–109.

- Janiszewski, H. A., Gaherty, J. B., Abers, G. A., Gao, H., & Eilon, Z. C., 2019. Amphibious surface-wave phase-velocity measurements of the Cascadia subduction zone, *Geophysical Journal International*, **217**(3), 1929–1948.
- Kao, H., Behr, Y., Currie, C. A., Hyndman, R., Townend, J., Lin, F.-C., Ritzwoller, M. H., Shan, S.-J., & He, J., 2013. Ambient seismic noise tomography of Canada and adjacent regions: Part I. Crustal structures, *Journal of Geophysical Research: Solid Earth*, **118**(11), 5865–5887.
- Karato, S., Jung, H., Katayama, I., & Skemer, P., 2008. Geodynamic significance of seismic anisotropy of the upper mantle: new insights from laboratory studies, *Annual Review of Earth Planetary Sciences*, **36**, 59–95.
- Kennett, B. & Engdahl, E., 1991. Traveltimes for global earthquake location and phase identification, *Geophysical Journal International*, **105**(2), 429–465.
- Kennett, B. L., Engdahl, E., & Buland, R., 1995. Constraints on seismic velocities in the Earth from traveltimes, *Geophysical Journal International*, **122**(1), 108–124.
- Kikuchi, M. & Kanamori, H., 1982. Inversion of complex body waves, *Bulletin of the Seismological Society of America*, **72**(2), 491–506.
- Knopoff, L., 1964. A matrix method for elastic wave problems, *Bulletin of the Seismological Society of America*, **54**(1), 431–438.
- Knopoff, L., 1972. Observation and inversion of surface-wave dispersion, *Tectonophysics*, **13**(1-4), 497–519.
- Kolb, J. & Lekić, V., 2014. Receiver function deconvolution using transdimensional hierarchical Bayesian inference, *Geophysical Journal International*, **197**(3), 1719–1735.
- Kolínský, P., Bokelmann, G., & Group, A. W., 2021. On the wobbles of phase-velocity dispersion curves, *Geophysical Journal International*, **224**(3), 1477–1504.

- Kovach, R. L., 1978. Seismic surface waves and crustal and upper mantle structure, *Reviews of Geophysics*, **16**(1), 1–13.
- Kreemer, C., Holt, W. E., & Haines, A. J., 2003. An integrated global model of present-day plate motions and plate boundary deformation, *Geophysical Journal International*, **154**(1), 8–34.
- Kurtz, R. D., DeLaurier, J. M., & Gupta, J. C., 1990. The electrical conductivity distribution beneath Vancouver Island: A region of active plate subduction, *Journal of Geophysical Research: Solid Earth*, **95**(B7), 10929–10946.
- Langston, C. A., 1979. Structure under Mount Rainier, Washington, inferred from teleseismic body waves, *Journal of Geophysical Research: Solid Earth*, **84**(B9), 4749–4762.
- Lark, R., Clifford, D., & Waters, C., 2014. Modelling complex geological circular data with the projected normal distribution and mixtures of von Mises distributions, *Solid Earth*, **5**(2), 631–639.
- Lay, T., Kanamori, H., Ammon, C. J., Koper, K. D., Hutko, A. R., Ye, L., Yue, H., & Rushing, T. M., 2012. Depth-varying rupture properties of subduction zone megathrust faults, *Journal of Geophysical Research: Solid Earth*, **117**(B4).
- Lebedev, S. & Van Der Hilst, R. D., 2008. Global upper-mantle tomography with the automated multimode inversion of surface and S-wave forms, *Geophysical Journal International*, **173**(2), 505–518.
- Lebedev, S., Adam, J. M.-C., & Meier, T., 2013. Mapping the Moho with seismic surface waves: a review, resolution analysis, and recommended inversion strategies, *Tectonophysics*, **609**, 377–394.
- Lee, A., 2010. Circular data, *Wiley Interdisciplinary Reviews: Computational Statistics*, **2**(4), 477–486.
- Leonard, L. J. & Bednarski, J. M., 2014. Field survey following the 28 October 2012 Haida Gwaii tsunami, *Pure and Applied Geophysics*, **171**(12), 3467–3482.

- Levenberg, K., 1944. A method for the solution of certain non-linear problems in least squares, *Quarterly of Applied Mathematics*, **2**(2), 164–168.
- Levshin, A. L., Ritzwoller, M. H., & Resovsky, J. S., 1999. Source effects on surface wave group travel times and group velocity maps, *Physics of the Earth and Planetary Interiors*, **115**(3-4), 293–312.
- Ligorria, J. P. & Ammon, C. J., 1999. Iterative deconvolution and receiver-function estimation, *Bulletin of the Seismological Society of America*, **89**(5), 1395–1400.
- Litchfield, N., Van Dissen, R., Sutherland, R., Barnes, P., Cox, S., Norris, R., Beavan, R., Langridge, R., Villamor, P., Berryman, K., et al., 2014. A model of active faulting in New Zealand, *New Zealand Journal of Geology and Geophysics*, **57**(1), 32–56.
- Liu, Y., 2013. Numerical simulations on megathrust rupture stabilized under strong dilatancy strengthening in slow slip region, *Geophysical Research Letters*, **40**(7), 1311–1316.
- Liu, Y. & Rice, J. R., 2007. Spontaneous and triggered aseismic deformation transients in a subduction fault model, *Journal of Geophysical Research: Solid Earth*, **112**(B9).
- Luo, Y., Yang, Y., Xu, Y., Xu, H., Zhao, K., & Wang, K., 2015. On the limitations of interstation distances in ambient noise tomography, *Geophysical Journal International*, **201**(2), 652–661.
- Magrini, F., Diaferia, G., Boschi, L., & Cammarano, F., 2020. Arrival-angle effects on two-receiver measurements of phase velocity, *Geophysical Journal International*, **220**(3), 1838–1844.
- Malinverno, A., 2002. Parsimonious Bayesian Markov chain Monte Carlo inversion in nonlinear geophysical problems, *Geophysical Journal International*, **151**, 675–688.
- Mardia, K. V. & Jupp, P. E., 2009. *Directional statistics*, vol. 494, John Wiley & Sons.
- Marquardt, D. W., 1963. An algorithm for least-squares estimation of nonlinear parameters, *Journal of the Society of Industrial and Applied Mathematics*, **11**(2), 431–441.

- Mazzotti, S., Dragert, H., Henton, J., Schmidt, M., Hyndman, R., James, T., Lu, Y., & Craymer, M., 2003. Current tectonics of northern Cascadia from a decade of GPS measurements, *Journal of Geophysical Research: Solid Earth*, **108**(B12).
- McCrory, P. A. & Wilson, D. S., 2013. A kinematic model for the formation of the Siletz-Crescent forearc terrane by capture of coherent fragments of the Farallon and Resurrection plates, *Tectonics*, **32**(3), 718–736.
- McCrory, P. A., Blair, J. L., Waldhauser, F., & Oppenheimer, D. H., 2012. Juan de Fuca slab geometry and its relation to Wadati-Benioff zone seismicity, *Journal of Geophysical Research: Solid Earth*, **117**(B9).
- McCrory, P. A., Constantz, J. E., Hunt, A. G., & Blair, J. L., 2016. Helium as a tracer for fluids released from Juan de Fuca lithosphere beneath the Cascadia forearc, *Geochemistry, Geophysics, Geosystems*, **17**(6), 2434–2449.
- McGillivray, P. R. & Oldenburg, D., 1990. Methods for calculating Fréchet derivatives and sensitivities for the non-linear inverse problem: a comparative study 1, *Geophysical Prospecting*, **38**(5), 499–524.
- McLellan, M., Schaeffer, A. J., & Audet, P., 2018. Structure and fabric of the crust and uppermost mantle in the northern Canadian Cordillera from Rayleigh-wave tomography, *Tectonophysics*, **724**, 28–41.
- Meade, B. J. & Loveless, J. P., 2009. Predicting the geodetic signature of Mw 8 slow slip events, *Geophysical Research Letters*, **36**(1).
- Meier, T., Dietrich, K., Stöckhert, B., & Harjes, H.-P., 2004. One-dimensional models of shear wave velocity for the eastern Mediterranean obtained from the inversion of Rayleigh wave phase velocities and tectonic implications, *Geophysical Journal International*, **156**(1), 45–58.
- Menke, W., 2018. *Geophysical data analysis: Discrete inverse theory*, Academic press.

- Métivier, L., Brossier, R., Merigot, Q., Oudet, E., & Virieux, J., 2016. An optimal transport approach for seismic tomography: Application to 3d full waveform inversion, *Inverse Problems*, **32**(11), 115008.
- Metropolis, N., Rosenbluth, A., Rosenbluth, M., & Teller, A. H., 1953. Equations of state calculations by fast computing machines, *Journal of Chemical Physics*, **21**, 1087–1092.
- Mochizuki, K., Sutherland, R., Henrys, S., Bassett, D., Van Avendonk, H., Arai, R., Kodaira, S., Fujie, G., Yamamoto, Y., Bangs, N., et al., 2019. Recycling of depleted continental mantle by subduction and plumes at the Hikurangi Plateau large igneous province, southwestern Pacific Ocean, *Geology*, **47**(8), 795–798.
- Molnar, S., Dosso, S. E., & Cassidy, J. F., 2010. Bayesian inversion of microtremor array dispersion data in Southwestern British Columbia, *Geophysical Journal International*, **183**, 923–940.
- Molnar, S., Dosso, S. E., & Cassidy, J. F., 2013. Uncertainty of earthquake site amplification via Bayesian inversion of surface seismic data, *Geophysics*, **78**, WB37–WB48.
- Montagner, J.-P. & Anderson, D. L., 1989. Petrological constraints on seismic anisotropy, *Physics of the Earth and planetary interiors*, **54**(1-2), 82–105.
- Montagner, J.-P. & Guillot, L., 2002. Seismic anisotropy and global geodynamics, *Reviews in mineralogy and geochemistry*, **51**(1), 353–385.
- Mori, N., Takahashi, T., Yasuda, T., & Yanagisawa, H., 2011. Survey of 2011 Tohoku earthquake tsunami inundation and run-up, *Geophysical research letters*, **38**(7).
- Moschetti, M., Ritzwoller, M., & Shapiro, N., 2007. Surface wave tomography of the western United States from ambient seismic noise: Rayleigh wave group velocity maps, *Geochemistry, Geophysics, Geosystems*, **8**, Q08010.
- Mosegaard, K. & Tarantola, A., 1995. Monte Carlo sampling of solutions to inverse problems, *Journal of Geophysical Research: Solid Earth*, **100**, 12431–12447.

- Mosher, S. G., Audet, P., & L’Heureux, I., 2014. Seismic evidence for rotating mantle flow around subducting slab edge associated with oceanic microplate capture, *Geophysical Research Letters*, **41**(13), 4548–4553.
- Nakajima, J. & Uchida, N., 2018. Repeated drainage from megathrusts during episodic slow slip, *Nature Geoscience*, **11**(5), 351–356.
- Nedimović, M. R., Hyndman, R. D., Ramachandran, K., & Spence, G. D., 2003. Reflection signature of seismic and aseismic slip on the northern Cascadia subduction interface, *Nature*, **424**(6947), 416–420.
- Nelder, J. A. & Mead, R., 1965. A simplex method for function minimization, *Computer Journal*, **7**(4), 308–313.
- Nelson, A. R., Atwater, B. F., Bobrowsky, P. T., Bradley, L.-A., Clague, J. J., Carver, G. A., Darienzo, M. E., Grant, W. C., Krueger, H. W., Sparks, R., et al., 1995. Radiocarbon evidence for extensive plate-boundary rupture about 300 years ago at the Cascadia subduction zone, *Nature*, **378**(6555), 371–374.
- Nelson, J., Colpron, M., Israel, S., Bissig, T., Rusk, B., & Thompson, J., 2013. The cordillera of British Columbia, Yukon, and Alaska: tectonics and metallogeny, *Tectonics, Metallogeny, and Discovery: The North American Cordillera and Similar Accretionary Settings: Society of Economic Geologists Special Publication*, **17**, 53–109.
- Niu, F., Levander, A., Ham, S., & Obayashi, M., 2005. Mapping the subducting Pacific slab beneath southwest Japan with Hi-net receiver functions, *Earth and Planetary Science Letters*, **239**(1-2), 9–17.
- Nolet, G., 2012. *Seismic tomography: With applications in global seismology and exploration geophysics*, vol. 5, Springer Science & Business Media.
- Obara, K. & Kato, A., 2016. Connecting slow earthquakes to huge earthquakes, *Science*, **353**(6296), 253–257.

- Obara, K., Hirose, H., Yamamizu, F., & Kasahara, K., 2004. Episodic slow slip events accompanied by non-volcanic tremors in southwest Japan subduction zone, *Geophysical Research Letters*, **31**(23).
- Owens, T. J., Zandt, G., & Taylor, S. R., 1984. Seismic evidence for an ancient rift beneath the Cumberland Plateau, Tennessee: A detailed analysis of broadband teleseismic P waveforms, *Journal of Geophysical Research: Solid Earth*, **89**(B9), 7783–7795.
- O’Neill, A. & Matsuoka, T., 2005. Dominant higher surface-wave modes and possible inversion pitfalls, *Journal of Environmental and Engineering Geophysics*, **10**(2), 185–201.
- Parker, R. L., 1977. Understanding inverse theory, *Annual Review of Earth and Planetary Sciences*, **5**(1), 35–64.
- Parker, R. L., 1994. *Geophysical inverse theory*, vol. 1, Princeton university press.
- Parolai, S., Richwalski, S. M., Milkereit, C., & Fäh, D., 2006. S-wave velocity profiles for earthquake engineering purposes for the Cologne area (Germany), *Bulletin of Earthquake Engineering*, **4**(1), 65–94.
- Pawlak, A., Eaton, D. W., Darbyshire, F., Lebedev, S., & Bastow, I. D., 2012. Crustal anisotropy beneath Hudson Bay from ambient noise tomography: Evidence for post-orogenic lower-crustal flow?, *Journal of Geophysical Research: Solid Earth*, **117**, B08301.
- Peacock, S. M., Christensen, N. I., Bostock, M. G., & Audet, P., 2011. High pore pressures and porosity at 35 km depth in the Cascadia subduction zone, *Geology*, **39**(5), 471–474.
- Pei, S., Niu, F., Ben-Zion, Y., Sun, Q., Liu, Y., Xue, X., Su, J., & Shao, Z., 2019. Seismic velocity reduction and accelerated recovery due to earthquakes on the Longmenshan fault, *Nature Geoscience*, **12**(5), 387–392.
- Peng, Z. & Ben-Zion, Y., 2006. Temporal changes of shallow seismic velocity around the Karadere-Düzce branch of the north Anatolian fault and strong ground motion, *Pure and Applied Geophysics*, **163**(2), 567–600.

- Phinney, R. A., 1964. Structure of the Earth’s crust from spectral behavior of long-period body waves, *Journal of Geophysical Research*, **69**(14), 2997–3017.
- Piana Agostinetti, N. & Malinverno, A., 2010. Receiver function inversion by trans-dimensional Monte Carlo sampling, *Geophysical Journal International*, **181**(2), 858–872.
- Piana Agostinetti, N., Giacomuzzi, G., & Malinverno, A., 2015. Local three-dimensional earthquake tomography by trans-dimensional Monte Carlo sampling, *Geophysical Journal International*, **201**(3), 1598–1617.
- Picozzi, M. & Albarello, D., 2007. Combining genetic and linearized algorithms for a two-step joint inversion of Rayleigh wave dispersion and H/V spectral ratio curves, *Geophysical Journal International*, **169**(1), 189–200.
- Pizer, C., Clark, K., Howarth, J., Garrett, E., Wang, X., Rhoades, D., & Woodroffe, S., 2021. Paleotsunamis on the southern Hikurangi subduction zone, New Zealand, show regular recurrence of large subduction earthquakes, *The Seismic Record*, **1**(2), 75–84.
- Porritt, R. W. & Yoshioka, S., 2017. Evidence of dynamic crustal deformation in Tohoku, Japan, from time-varying receiver functions, *Tectonics*, **36**(10), 1934–1946.
- Porter, R., Zandt, G., & McQuarrie, N., 2011. Pervasive lower-crustal seismic anisotropy in Southern California: Evidence for underplated schists and active tectonics, *Lithosphere*, **3**(3), 201–220.
- Ramachandran, K., Hyndman, R., & Brocher, T., 2006. Regional P wave velocity structure of the northern Cascadia subduction zone, *Journal of Geophysical Research: Solid Earth*, **111**(B12).
- Rawlinson, N. & Sambridge, M., 2004. Wave front evolution in strongly heterogeneous layered media using the fast marching method, *Geophysical Journal International*, **156**(3), 631–647.
- Rawlinson, N., Fichtner, A., Sambridge, M., & Young, M. K., 2014. Seismic tomography and the assessment of uncertainty, *Advances in Geophysics*, **55**, 1–76.

- Ray, A., Sekar, A., Hoversten, G. M., & Albertin, U., 2016. Frequency domain full waveform elastic inversion of marine seismic data from the Alba field using a Bayesian trans-dimensional algorithm, *Geophysical Journal International*, **205**(2), 915–937.
- Reyes, A., Christenson, B., & Faure, K., 2010. Sources of solutes and heat in low-enthalpy mineral waters and their relation to tectonic setting, New Zealand, *Journal of Volcanology and Geothermal Research*, **192**(3-4), 117–141.
- Reyners, M. & Eberhart-Phillips, D., 2009. Small earthquakes provide insight into plate coupling and fluid distribution in the Hikurangi subduction zone, New Zealand, *Earth and Planetary Science Letters*, **282**(1-4), 299–305.
- Ristau, J., Rogers, G. C., & Cassidy, J. F., 2007. Stress in western Canada from regional moment tensor analysis, *Canadian Journal of Earth Sciences*, **44**(2), 127–148.
- Ritzwoller, M. H. & Levshin, A. L., 1998. Eurasian surface wave tomography: Group velocities, *Journal of Geophysical Research: Solid Earth*, **103**(B3), 4839–4878.
- Rogers, G. & Dragert, H., 2003. Episodic tremor and slip on the Cascadia subduction zone: The chatter of silent slip, *Science*, **300**(5627), 1942–1943.
- Rubinstein, J. L., Vidale, J. E., Gomberg, J., Bodin, P., Creager, K. C., & Malone, S. D., 2007. Non-volcanic tremor driven by large transient shear stresses, *Nature*, **448**(7153), 579–582.
- Rubinstein, J. L., La Rocca, M., Vidale, J. E., Creager, K. C., & Wech, A. G., 2008. Tidal modulation of nonvolcanic tremor, *Science*, **319**(5860), 186–189.
- Sambridge, M., 2014. A parallel tempering algorithm for probabilistic sampling and multimodal optimization, *Geophysical Journal International*, **196**, 357–374.
- Sambridge, M. & Mosegaard, K., 2002. Monte Carlo methods in geophysical inverse problems, *Reviews of Geophysics*, **40**, 3–29.

- Sambridge, M., Braun, J., & McQueen, H., 1995. Geophysical parametrization and interpolation of irregular data using natural neighbours, *Geophysical Journal International*, **122**(3), 837–857.
- Sambridge, M., Gallagher, K., Jackson, A., & Rickwood, P., 2006. Trans-dimensional inverse problems, model comparison and the evidence, *Geophysical Journal International*, **167**, 528–542.
- Satake, K., Shimazaki, K., Tsuji, Y., & Ueda, K., 1996. Time and size of a giant earthquake in Cascadia inferred from Japanese tsunami records of January 1700, *Nature*, **379**(6562), 246–249.
- Satake, K., Wang, K., & Atwater, B. F., 2003. Fault slip and seismic moment of the 1700 Cascadia earthquake inferred from Japanese tsunami descriptions, *Journal of Geophysical Research: Solid Earth*, **108**(B11), 2535.
- Savage, M. K., 1998. Lower crustal anisotropy or dipping boundaries? Effects on receiver functions and a case study in New Zealand, *Journal of Geophysical Research: Solid Earth*, **103**(B7), 15069–15087.
- Savage, M. K., 1999. Seismic anisotropy and mantle deformation: what have we learned from shear wave splitting?, *Reviews of Geophysics*, **37**(1), 65–106.
- Savage, M. K., Park, J., & Todd, H., 2007. Velocity and anisotropy structure at the Hikurangi subduction margin, New Zealand from receiver functions, *Geophysical Journal International*, **168**(3), 1034–1050.
- Savard, G., Bostock, M. G., & Christensen, N. I., 2018. Seismicity, metamorphism, and fluid evolution across the Northern Cascadia fore arc, *Geochemistry, Geophysics, Geosystems*, **19**(6), 1881–1897.
- Savard, G., Bostock, M., Hutchinson, J., Kao, H., Christensen, N., & Peacock, S., 2020. The northern terminus of Cascadia subduction, *Journal of Geophysical Research: Solid Earth*, **125**(6), e2019JB018453.

- Schaeffer, A. & Lebedev, S., 2014. Imaging the north american continent using waveform inversion of global and usarray data, *Earth and Planetary Science Letters*, **402**, 26–41.
- Schaeffer, A., Lebedev, S., & Becker, T., 2016. Azimuthal seismic anisotropy in the Earth’s upper mantle and the thickness of tectonic plates, *Geophysical Supplements to the Monthly Notices of the Royal Astronomical Society*, **207**(2), 901–933.
- Scholz, C., 1988. The brittle-plastic transition and the depth of seismic faulting, *Geologische Rundschau*, **77**(1), 319–328.
- Scholz, C. H., 1998. Earthquakes and friction laws, *Nature*, **391**(6662), 37–42.
- Scholz, C. H., 2019. *The mechanics of earthquakes and faulting*, Cambridge university press.
- Schulte-Pelkum, V. & Mahan, K. H., 2014. A method for mapping crustal deformation and anisotropy with receiver functions and first results from USArray, *Earth and Planetary Science Letters*, **402**, 221–233.
- Schulte-Pelkum, V., Caine, J. S., Jones III, J. V., & Becker, T. W., 2020. Imaging the tectonic grain of the northern Cordillera orogen using Transportable Array receiver functions, *Seismological Research Letters*, **91**(6), 3086–3105.
- Schwarz, G., 1978. Estimating the dimension of a model, *Annals of Statistics*, **6**, 461–464.
- Shapiro, N. M., Campillo, M., Kaminski, E., Vilotte, J.-P., & Jaupart, C., 2018. Low-frequency earthquakes and pore pressure transients in subduction zones, *Geophysical Research Letters*, **45**(20), 11–083.
- Shearer, P. M., 2019. *Introduction to seismology*, Cambridge university press.
- Sibson, R. H., 1990. Conditions for fault-valve behaviour, *Geological Society, London, Special Publications*, **54**(1), 15–28.
- Sibson, R. H., 2013. Stress switching in subduction forearcs: Implications for overpressure containment and strength cycling on megathrusts, *Tectonophysics*, **600**, 142–152.

- Smith, A., Hyndman, R., Cassidy, J., & Wang, K., 2003. Structure, seismicity, and thermal regime of the Queen Charlotte transform margin, *Journal of Geophysical Research: Solid Earth*, **108**(B11), 2539.
- Smith, M. L. & Dahlen, F., 1973. The azimuthal dependence of Love and Rayleigh wave propagation in a slightly anisotropic medium, *Journal of Geophysical Research*, **78**(17), 3321–3333.
- Soomro, R., Weidle, C., Cristiano, L., Lebedev, S., Meier, T., & Group, P. W., 2016. Phase velocities of Rayleigh and Love waves in central and northern Europe from automated, broad-band, interstation measurements, *Geophysical Journal International*, **204**(1), 517–534.
- Soyer, W. & Unsworth, M., 2006. Deep electrical structure of the northern Cascadia (British Columbia, Canada) subduction zone: Implications for the distribution of fluids, *Geology*, **34**(1), 53–56.
- Tarantola, A., 2005. *Inverse problem theory and methods for model parameter estimation*, Society for Industrial and Applied Mathematics, Philadelphia, PA.
- Thomson, W. T., 1950. Transmission of elastic waves through a stratified solid medium, *Journal of Applied Physics*, **21**(2), 89–93.
- Treitel, S. & Lines, L., 2001. Past, present, and future of geophysical inversion—A new millennium analysis, *Geophysics*, **66**(1), 21–24.
- Um, J. & Thurber, C., 1987. A fast algorithm for two-point seismic ray tracing, *Bulletin of the Seismological Society of America*, **77**(3), 972–986.
- Underwood, M. B., 2007. Sediment Inputs to Subduction Zones, in *The seismogenic zone of subduction thrust faults*, pp. 42–85, Columbia University Press.
- Van Alstine, D. R., 1980. Analysis of the modes of directional data with particular reference to palaeomagnetism, *Geophysical Journal International*, **61**(1), 101–113.

- Van Laarhoven, P. J. & Aarts, E. H., 1987. Simulated annealing, in *Simulated annealing: Theory and applications*, pp. 7–15, Springer.
- Vinnik, L., Makeyeva, L., Milev, A., & Usenko, A. Y., 1992. Global patterns of azimuthal anisotropy and deformations in the continental mantle, *Geophysical Journal International*, **111**(3), 433–447.
- Vinnik, L. P., Farra, V., & Romanowicz, B., 1989. Azimuthal anisotropy in the Earth from observations of SKS at Geoscope and NARS broadband stations, *Bulletin of the Seismological Society of America*, **79**(5), 1542–1558.
- Wallace, L. M., 2020. Slow slip events in New Zealand, *Annual Review of Earth and Planetary Sciences*, **48**, 175–203.
- Wallace, L. M. & Eberhart-Phillips, D., 2013. Newly observed, deep slow slip events at the central Hikurangi margin, New Zealand: Implications for downdip variability of slow slip and tremor, and relationship to seismic structure, *Geophysical Research Letters*, **40**(20), 5393–5398.
- Wallace, L. M., Beavan, J., McCaffrey, R., & Darby, D., 2004. Subduction zone coupling and tectonic block rotations in the North Island, New Zealand, *Journal of Geophysical Research: Solid Earth*, **109**(B12).
- Wang, K., Wells, R., Mazzotti, S., Hyndman, R. D., & Sagiya, T., 2003. A revised dislocation model of interseismic deformation of the Cascadia subduction zone, *Journal of Geophysical Research: Solid Earth*, **108**(B1).
- Wang, K., Dragert, H., Kao, H., & Roeloffs, E., 2008. Characterizing an “uncharacteristic” ETS event in northern Cascadia, *Geophysical Research Letters*, **35**(15).
- Wannamaker, P. E., Evans, R. L., Bedrosian, P. A., Unsworth, M. J., Maris, V., & McGary, R. S., 2014. Segmentation of plate coupling, fate of subduction fluids, and modes of arc magmatism in Cascadia, inferred from magnetotelluric resistivity, *Geochemistry, Geophysics, Geosystems*, **15**(11), 4230–4253.

- Warren-Smith, E., Fry, B., Wallace, L., Chon, E., Henrys, S., Sheehan, A., Mochizuki, K., Schwartz, S., Webb, S., & Lebedev, S., 2019. Episodic stress and fluid pressure cycling in subducting oceanic crust during slow slip, *Nature Geoscience*, **12**(6), 475–481.
- Wech, A. G. & Creager, K. C., 2008. Automated detection and location of Cascadia tremor, *Geophysical Research Letters*, **35**(20).
- Williams, C. A., Eberhart-Phillips, D., Bannister, S., Barker, D. H., Henrys, S., Reyners, M., & Sutherland, R., 2013. Revised interface geometry for the Hikurangi subduction zone, New Zealand, *Seismological Research Letters*, **84**(6), 1066–1073.
- Xia, J., Miller, R. D., & Park, C. B., 1999. Estimation of near-surface shear-wave velocity by inversion of Rayleigh waves, *Geophysics*, **64**(3), 691–700.
- Xia, J., Miller, R. D., Park, C. B., & Tian, G., 2002. Determining Q of near-surface materials from Rayleigh waves, *Journal of Applied Geophysics*, **51**(2-4), 121–129.
- Xia, J., Miller, R. D., Park, C. B., & Tian, G., 2003. Inversion of high frequency surface waves with fundamental and higher modes, *Journal of Applied Geophysics*, **52**(1), 45–57.
- Yamanaka, H. & Ishida, H., 1996. Application of genetic algorithms to an inversion of surface-wave dispersion data, *Bulletin of the Seismological Society of America*, **86**(2), 436–444.
- Yeck, W. L., Sheehan, A. F., Stachnik, J. C., & Lin, F.-C., 2017. Offshore Rayleigh group velocity observations of the South Island, New Zealand, from ambient noise data, *Geophysical Journal International*, **209**(2), 827–841.
- Yoshizawa, K. & Kennett, B., 2002. Determination of the influence zone for surface wave paths, *Geophysical Journal International*, **149**(2), 440–453.
- Zhu, L. & Kanamori, H., 2000. Moho depth variation in southern California from teleseismic receiver functions, *Journal of Geophysical Research: Solid Earth*, **105**(B2), 2969–2980.

Zulfakriza, Z., Saygin, E., Cummins, P., Widiyantoro, S., Nugraha, A. D., Lühr, B.-G., & Bodin, T., 2014. Upper crustal structure of central Java, Indonesia, from transdimensional seismic ambient noise tomography, *Geophysical Journal International*, **197**(1), 630–635.

APPENDICES

Appendix A

Supplementary materials for:
Azimuthal anisotropy in Bayesian
surface-wave tomography

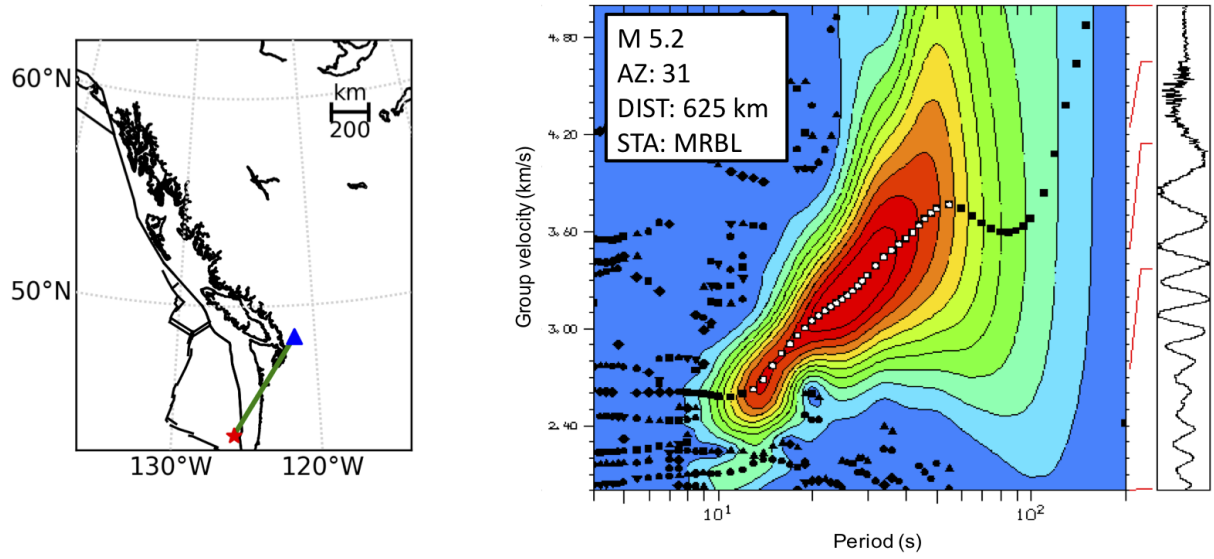


Figure A.1: Example of surface wave dispersion processing from a regional earthquake recording. The location of the earthquake and the recording station are shown by the red star and the blue triangle, respectively. The right panel shows the display of the CPS program (Herrmann, 2013). White markers show the manually-selected period band for this event at this station. M, AZ, DIST, and STA refer to the event magnitude, source-to-station azimuth, distance, and station name, respectively.

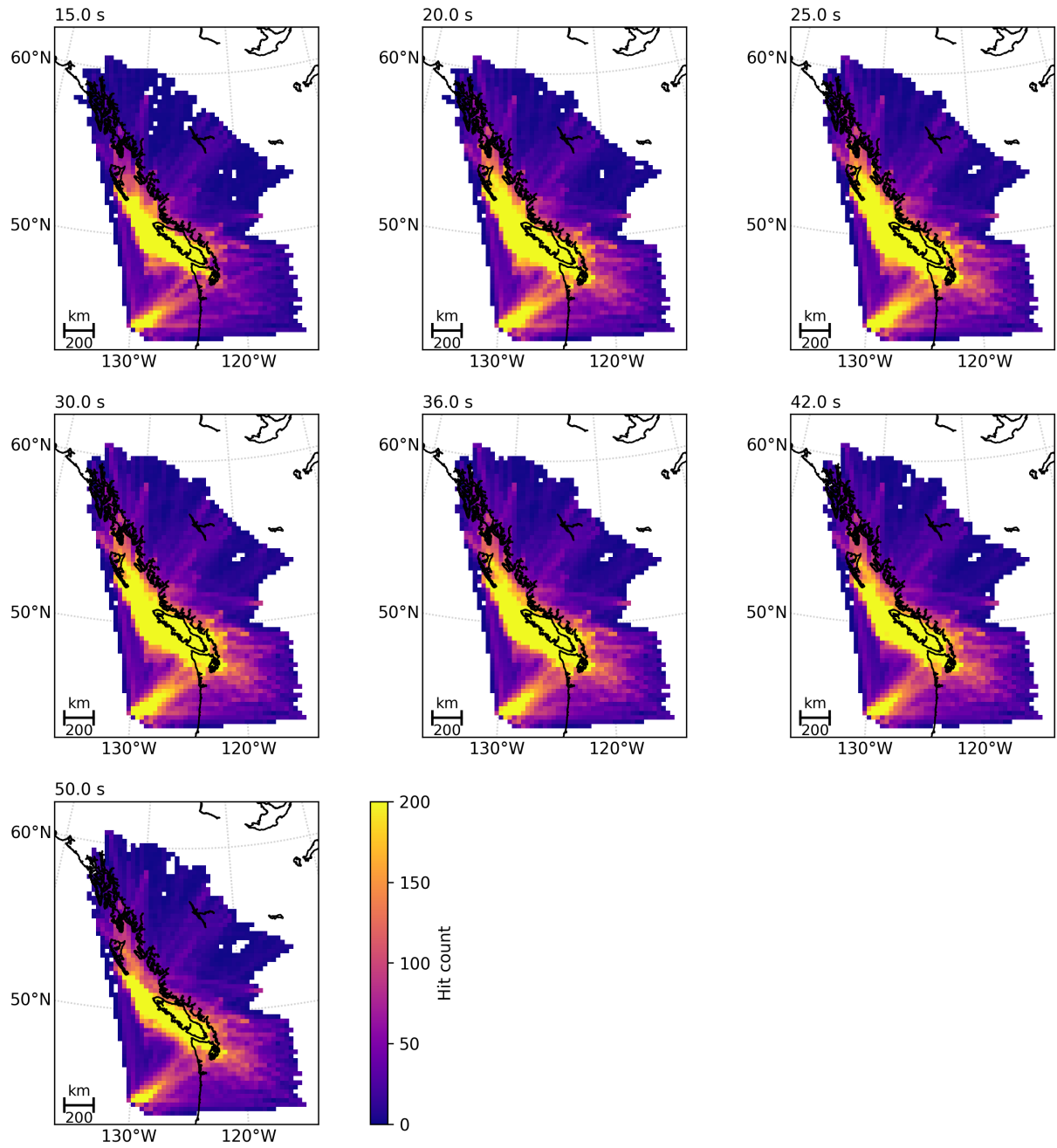


Figure A.2: Surface wave path coverage at periods considered in this work. Hitcounts are shown through the underlying model integration grid (discussed in the text).

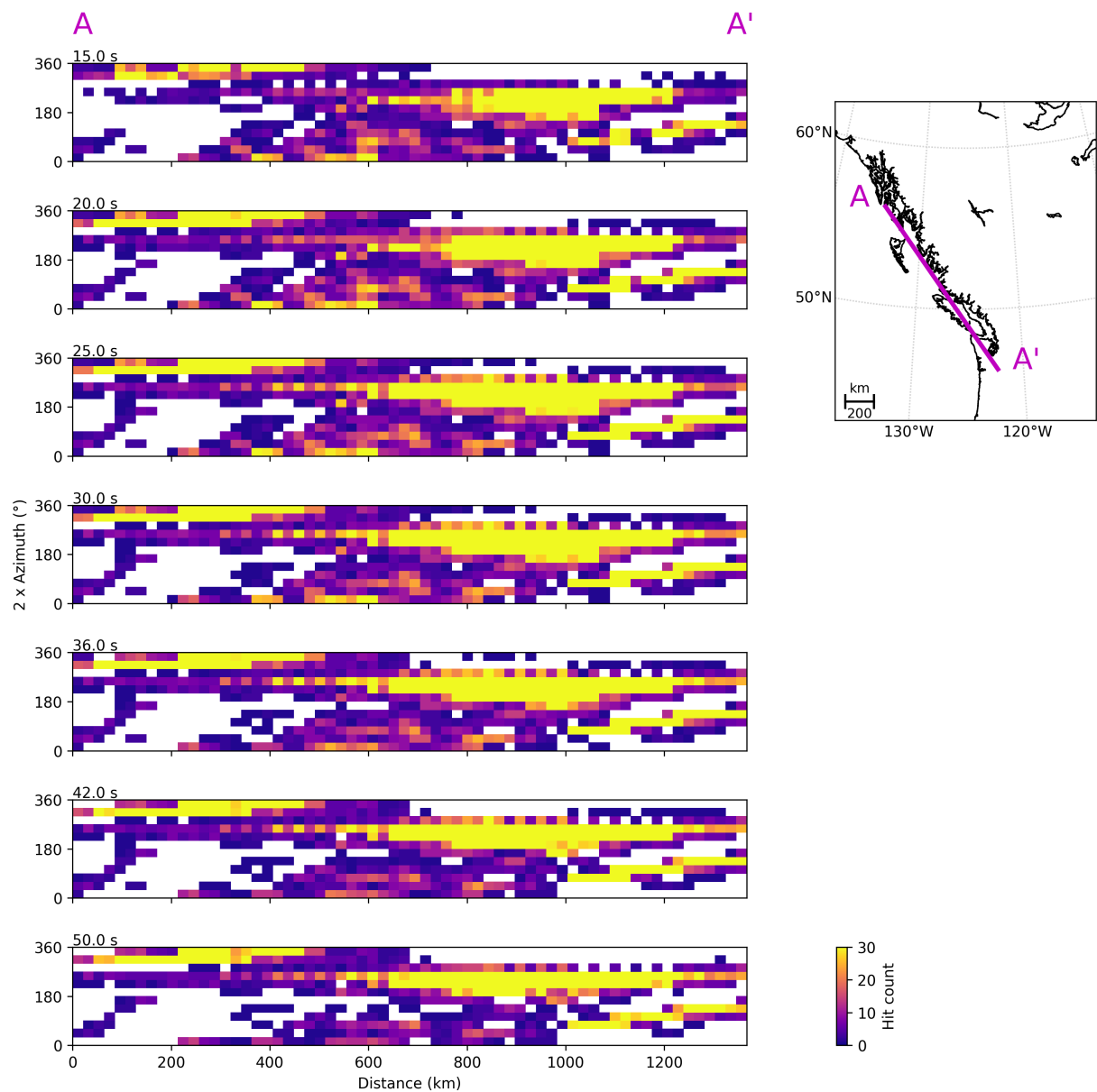


Figure A.3: Surface wave azimuthal path coverage at periods considered in this work over Vancouver Island and Haida Gwaii. Histograms of source-to-station azimuths through the underlying model integration grid along the transect A-A' are shown. Azimuthal path coverage is best between central Vancouver Island and Haida Gwaii.

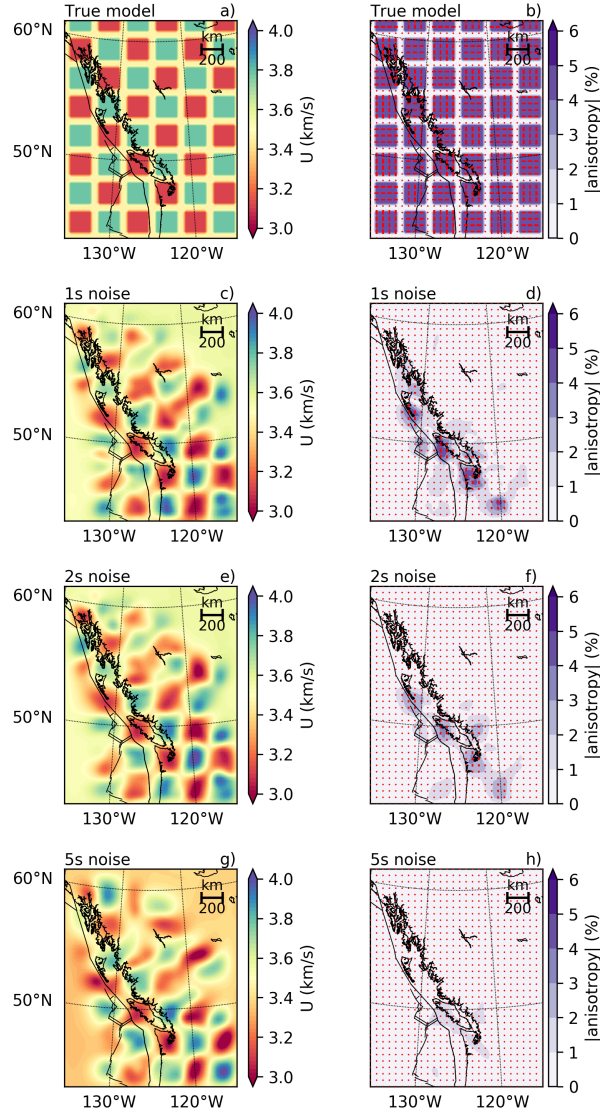


Figure A.4: Results of synthetic Bayesian inversion tests under varying noise conditions. The isotropic and anisotropic components of the true model are shown in a) and b), respectively. Mean isotropic velocities estimated from the ensemble of rjMcMC samples for inversion tests with Gaussian-distributed noise with 1, 2, and 5 second standard deviations are shown in c), e), and g) respectively. Mean azimuthal anisotropy estimated from the ensemble of rjMcMC samples for inversion tests with Gaussian-distributed noise with 1, 2, and 5 second standard deviations are shown in d), f), and h) respectively. Red lines delineate the fast direction of anisotropy. 179

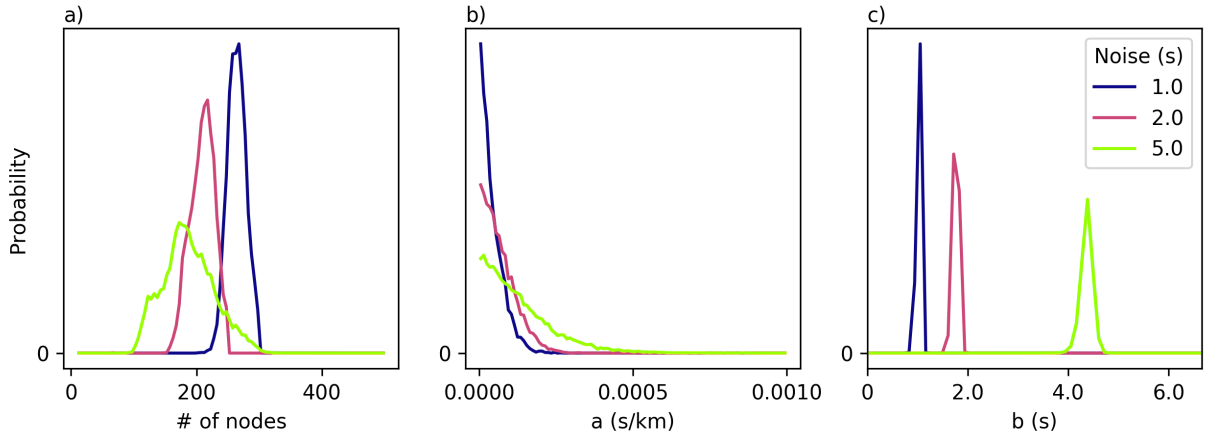


Figure A.5: Hyper-parameter estimates from synthetic Bayesian inversion tests under varying noise conditions. The number of nodes in the Delaunay triangulation with linear interpolation is shown in a). The linear coefficient between unknown data errors and travel path length is shown in b). The constant term for unknown data error is shown in c).

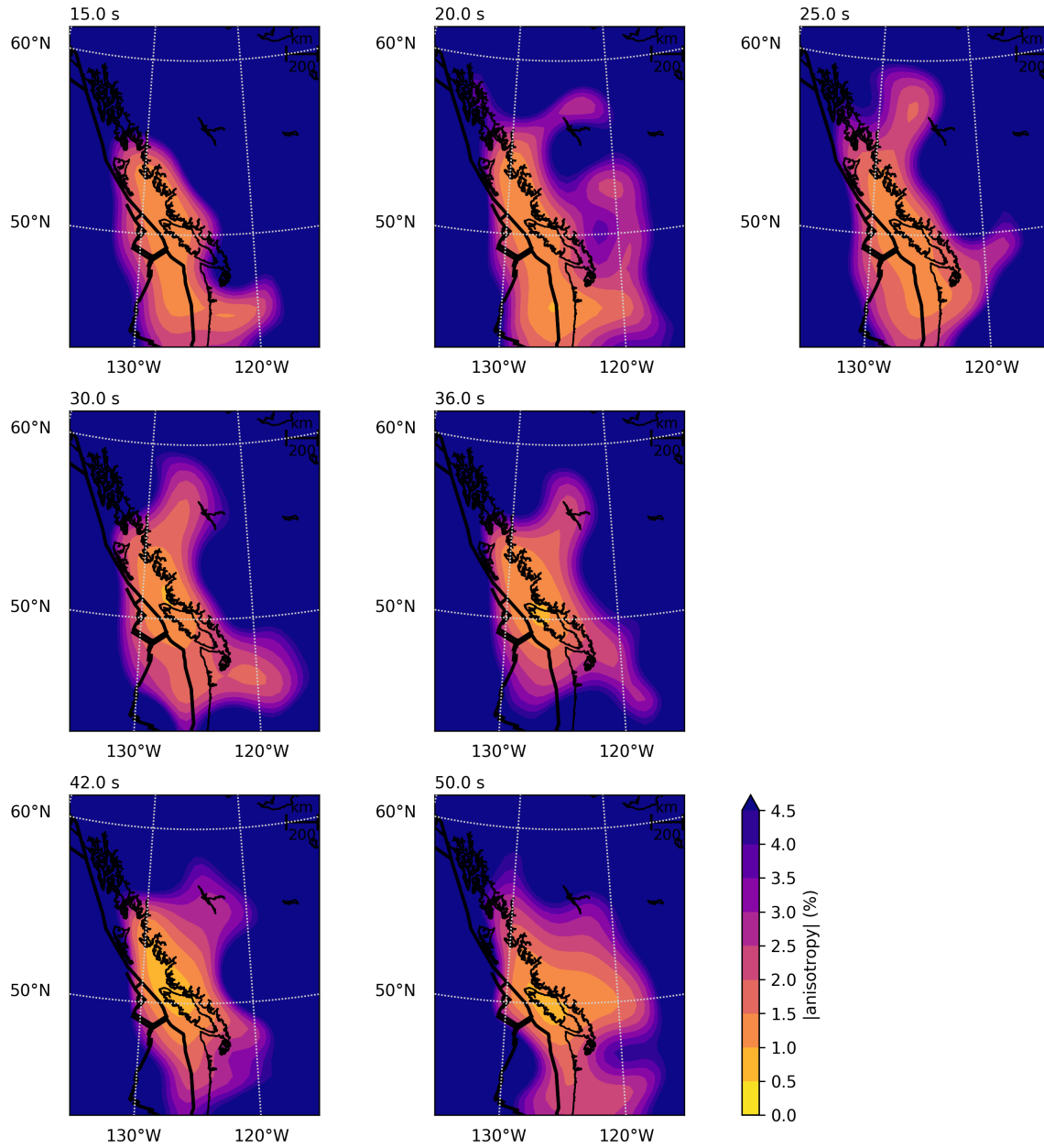


Figure A.6: Uncertainty of azimuthal anisotropy magnitude over northern Cascadia and Haida Gwaii. The estimated uncertainty of anisotropy magnitudes estimated from the ensemble of rjMCMC samples for inversions at periods 15-50 s are shown (as percent anisotropy with respect to isotropic velocity). In general, anisotropy magnitudes are poorly resolved.

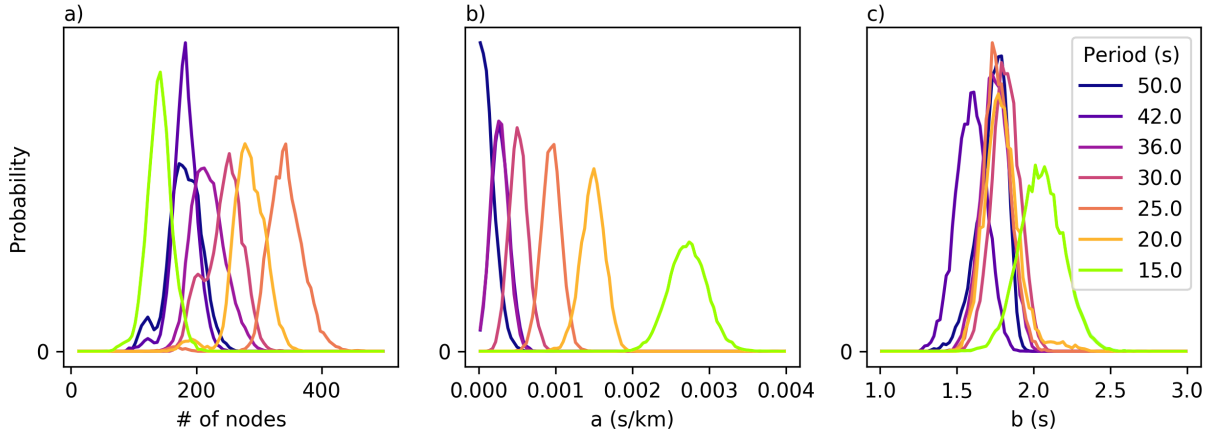


Figure A.7: Hyper-parameter estimates from Bayesian inversions after first inversion iteration (assuming great circle ray paths). The number of nodes in the Delaunay triangulation with linear interpolation is shown in a). The linear coefficient between unknown data errors and travel path length is shown in b). The constant term for unknown data error is shown in c).

Table A.1: Prior bounds for model parameters in Bayesian tomographic inversion. C_0 is the isotropic group velocity. A_1 and B_1 are the anisotropic terms for group velocity with azimuthal periodicity of π .

Parameter	Min	Max
C_0 (km/s)	2.0	4.5
A_1 (km/s)	-0.3	0.3
B_1 (km/s)	-0.3	0.3
Number of model nodes	10	500

Table A.2: Simple crustal model for surface wave sensitivity kernel estimation (see main text for detail).

Depth (km)	Vs (km/s)	Vp (km/s)
0.0	3.14	5.50
1.0	3.14	5.50
2.0	3.14	5.50
3.0	3.14	5.50
4.0	3.43	6.00
5.0	3.46	6.05
6.0	3.49	6.11
7.0	3.52	6.16
8.0	3.55	6.22
9.0	3.58	6.27
10.0	3.62	6.33
11.0	3.65	6.38
12.0	3.67	6.43
13.0	3.71	6.49
14.0	3.74	6.54
15.0	3.77	6.60
16.0	3.80	6.65
17.0	3.83	6.71
18.0	3.86	6.76
19.0	3.89	6.81
20.0	3.93	6.87
21.0	3.95	6.92
22.0	3.99	6.98
23.0	4.02	7.03
24.0	4.05	7.09
25.0	4.08	7.14

26.0	4.11	7.19
27.0	4.14	7.25
28.0	4.17	7.30
29.0	4.21	7.36
30.0	4.23	7.41
31.0	4.27	7.47
32.0	4.30	7.52
33.0	4.33	7.57
34.0	4.36	7.63
35.0	4.39	7.68
36.0	4.42	7.74
37.0	4.45	7.79
38.0	4.49	7.85
39.0	4.51	7.90
40.0	4.63	8.10

Appendix B

Supplementary materials for:
Circular statistics in Bayesian
inversion for two-station
phase-velocity dispersion estimation

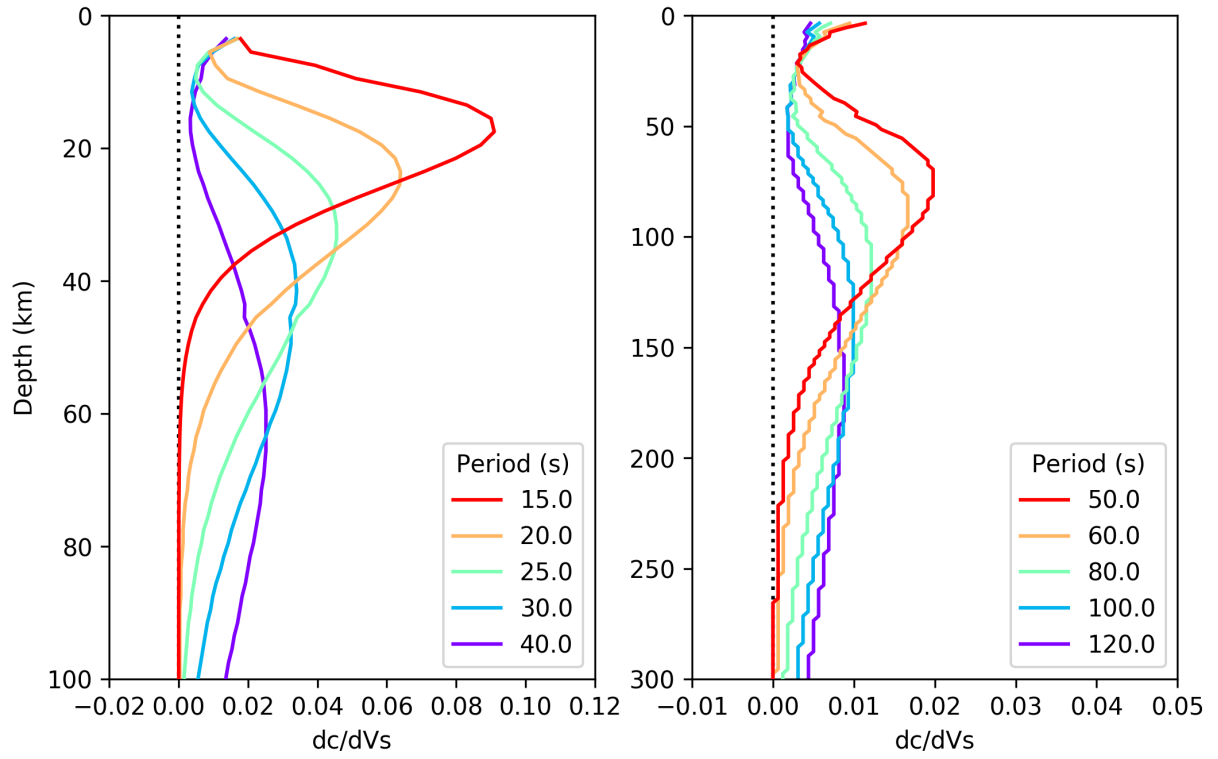


Figure B.1: Fundamental-mode Rayleigh wave phase-velocity depth sensitivity for periods considered in this work, calculated using AK135 Earth model (Kennett et al., 1995). Note the different depth and sensitivity axis scales between the two panels.

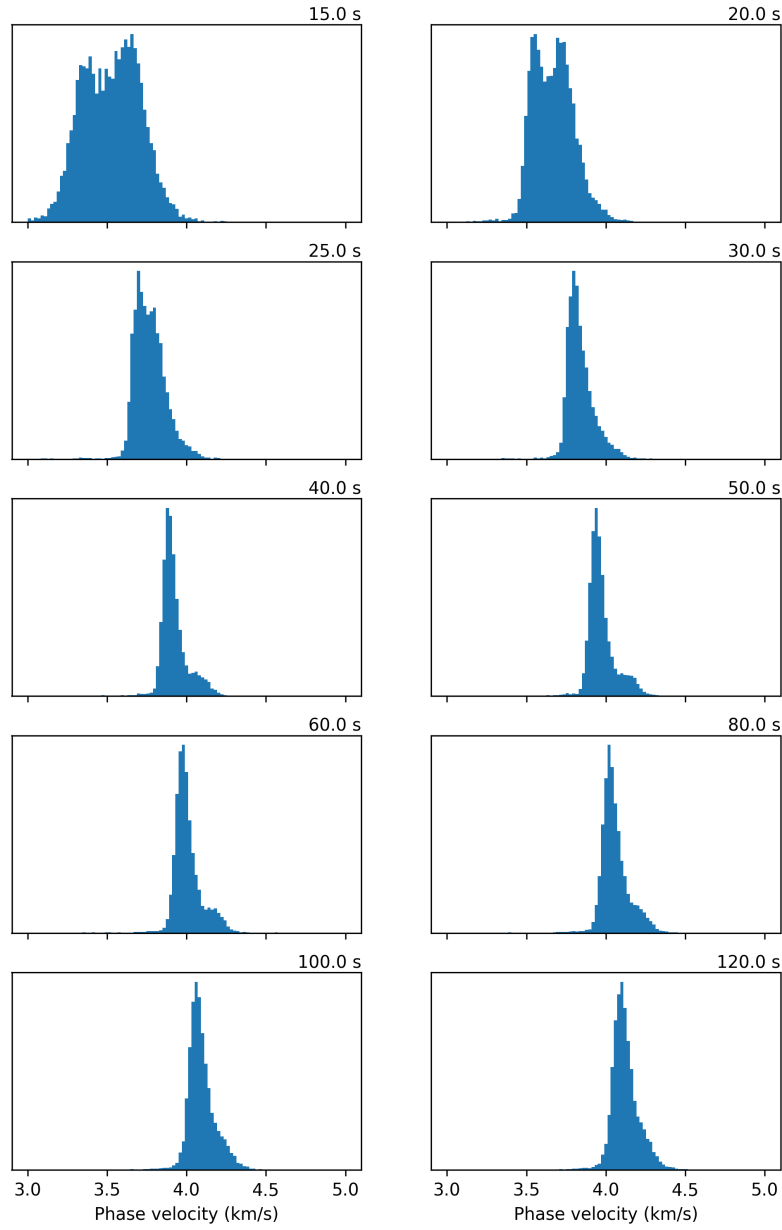


Figure B.2: Mean predicted surface-wave dispersion from rjMcMC sampling for all stations pairs over British Columbia.

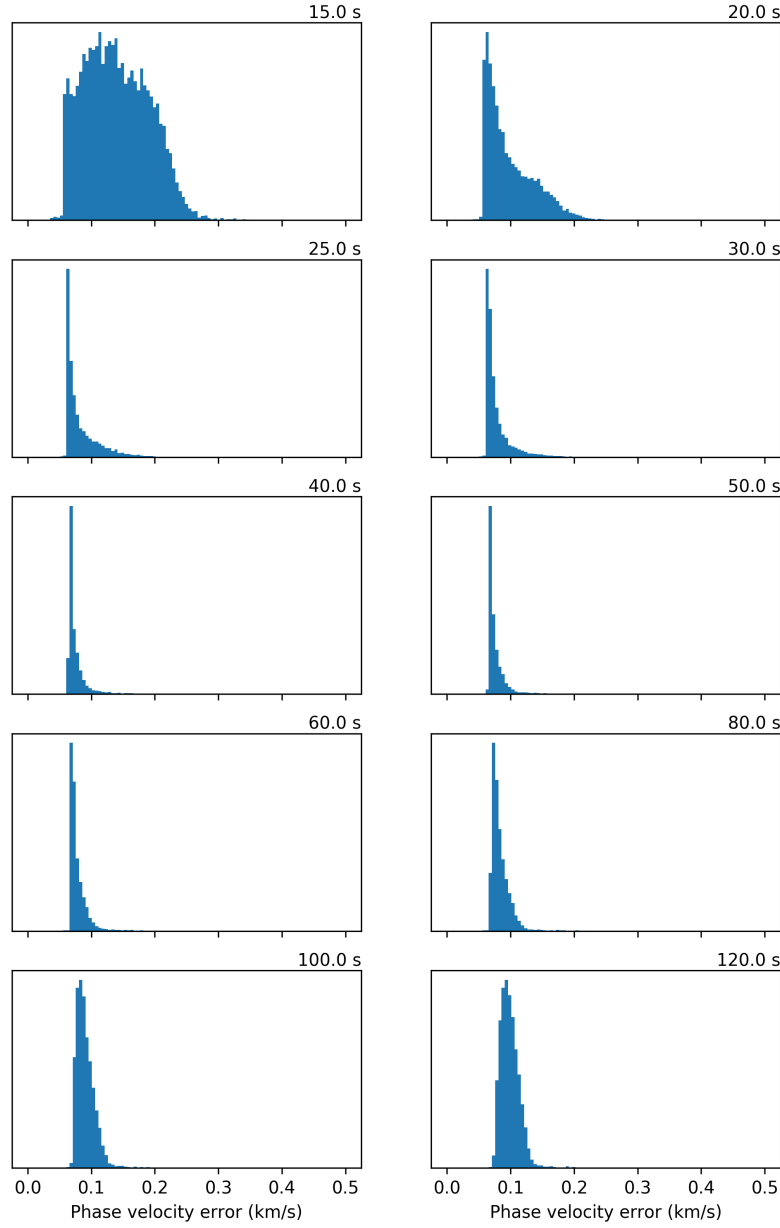


Figure B.3: Standard deviation of predicted surface-wave dispersion from rjMCMC sampling for all stations pairs over British Columbia.

Appendix C

Supplementary materials for: Seismic evidence for megathrust fault-valve behavior during episodic tremor and slip

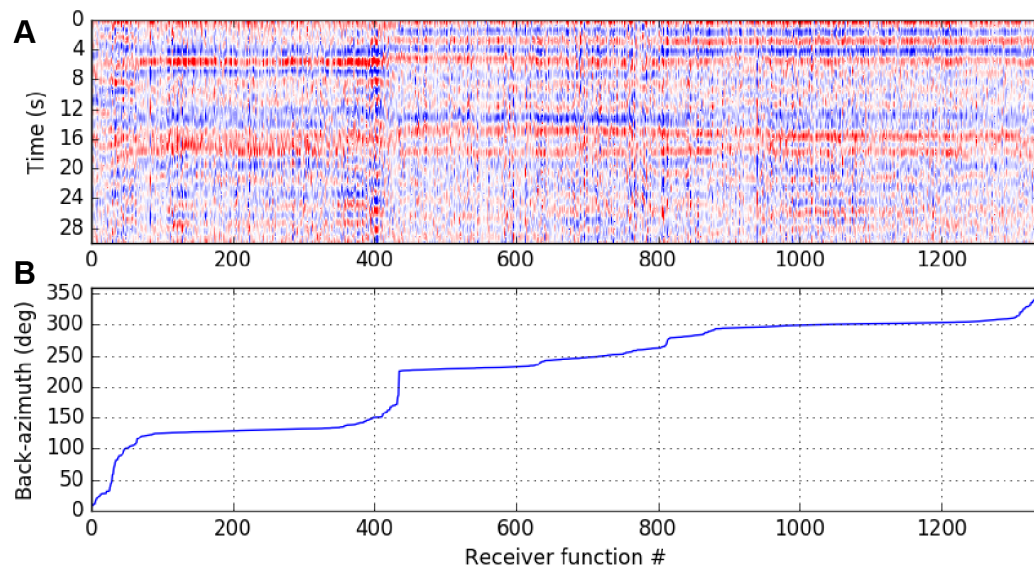


Figure C.1: S_V receiver functions for station PGC. These data are filtered between 0.05 and 0.5 Hz and are sorted by back-azimuth, showing the signature of the low-velocity layer.

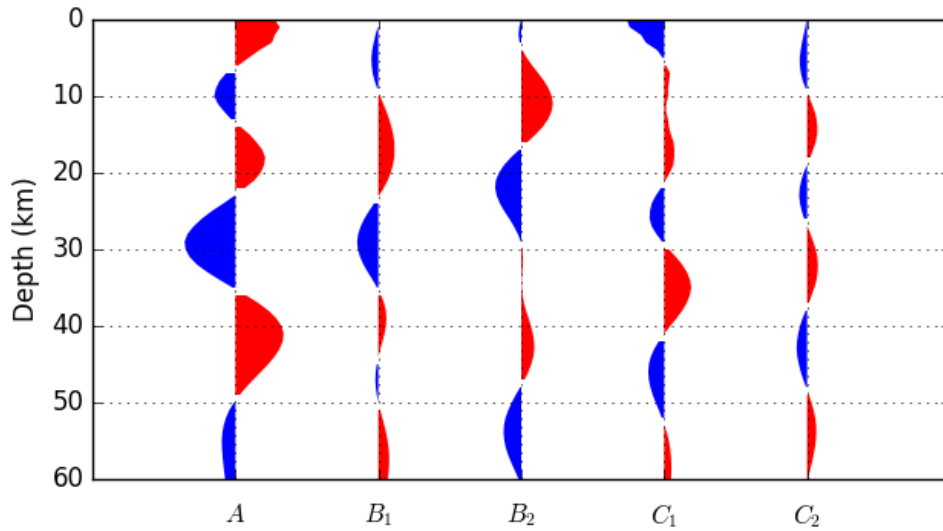


Figure C.2: Harmonic components for station PGC. These components are obtained from the harmonic decomposition of all S_V and S_H receiver functions. The A component displays large amplitude negative-positive pulses at approximate depths of 30-40 km that represent the low-frequency signature of the low-velocity layer from the direct Ps conversion.

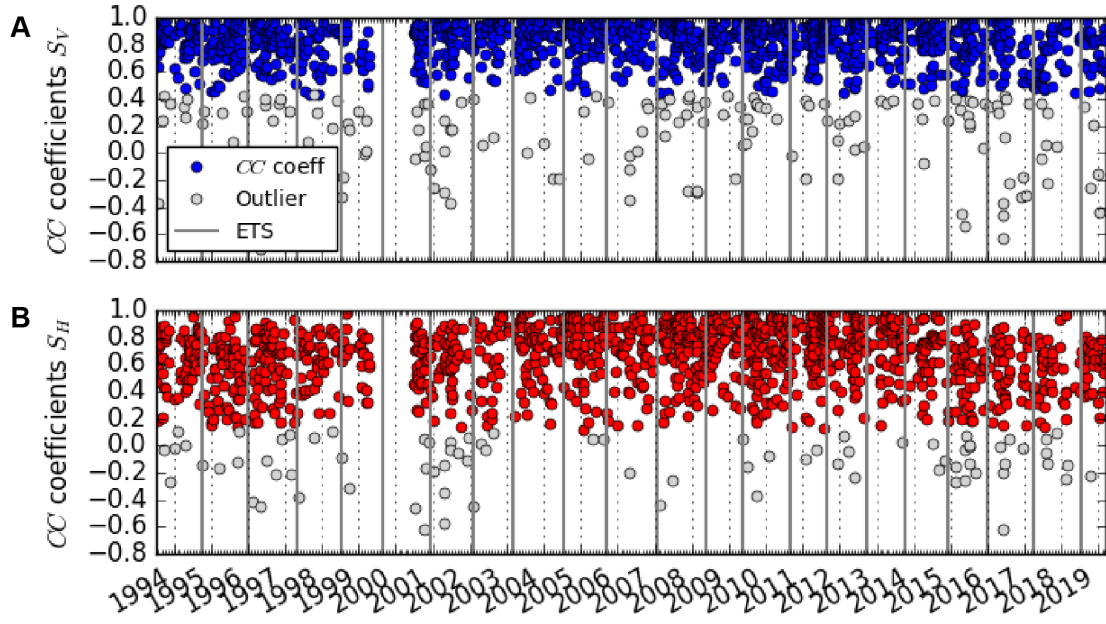


Figure C.3: Cross-correlation coefficients between observed and predicted receiver functions. (A) and (B) show S_V and S_H component receiver functions, respectively, sorted chronologically. Predicted receiver functions are obtained by forward solving equation 5.1 using the back-azimuth of each teleseismic earthquake. The ETS times are shown as vertical grey bars.

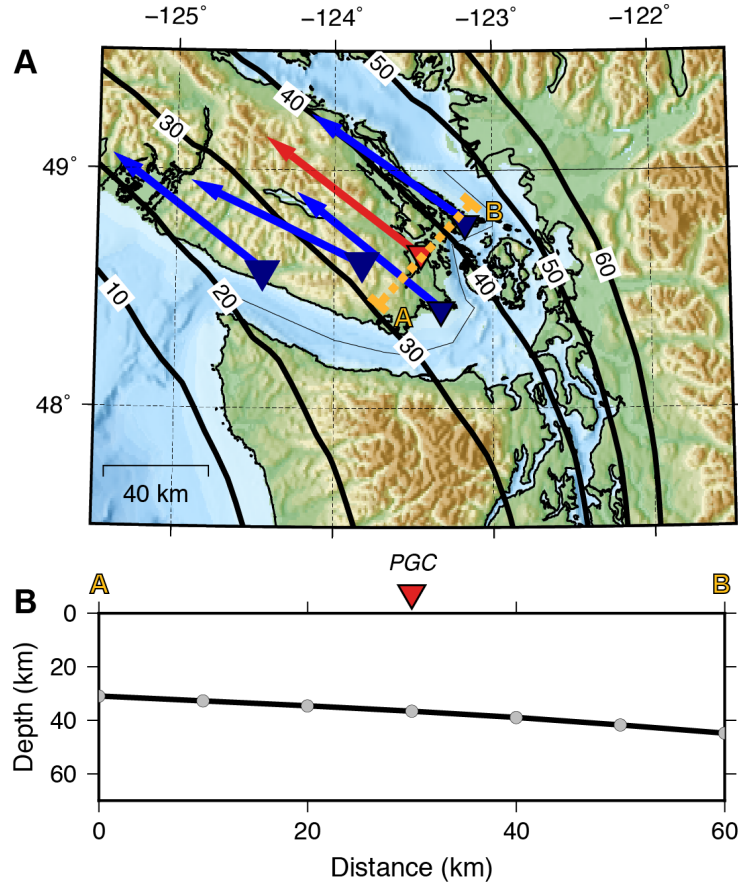


Figure C.4: Example calculation of strike and dip angles at each station location. Map of the northern Cascadia subduction zone showing the slab contours (Bostock et al., 2019) and the strike directions as blue or red vectors (A) obtained by calculating the dip direction from the maximum directional derivative of the slab model and subtracting 90° (right-hand rule). The dip angle is obtained by extracting a 60 km profile of the slab depth along the dip direction (A) and measuring the slope angle (B).

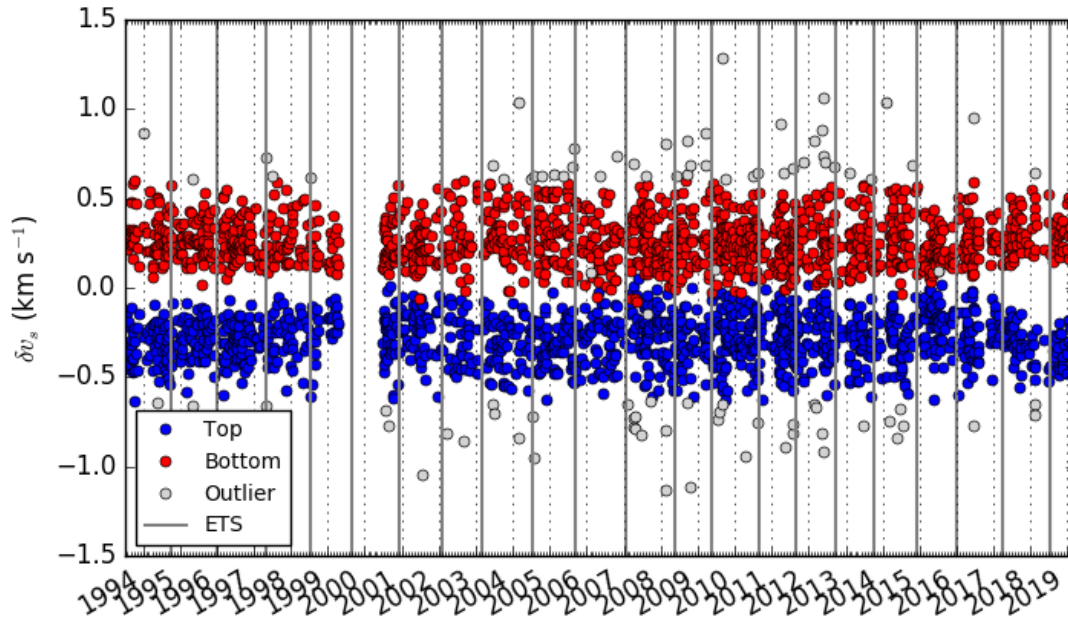


Figure C.5: Estimated S-wave velocity contrast across the top and bottom of the LVL. These estimates are obtained using the amplitude of Ps conversions observed on S_V receiver functions, sorted chronologically. The ETS times are shown as vertical grey bars.

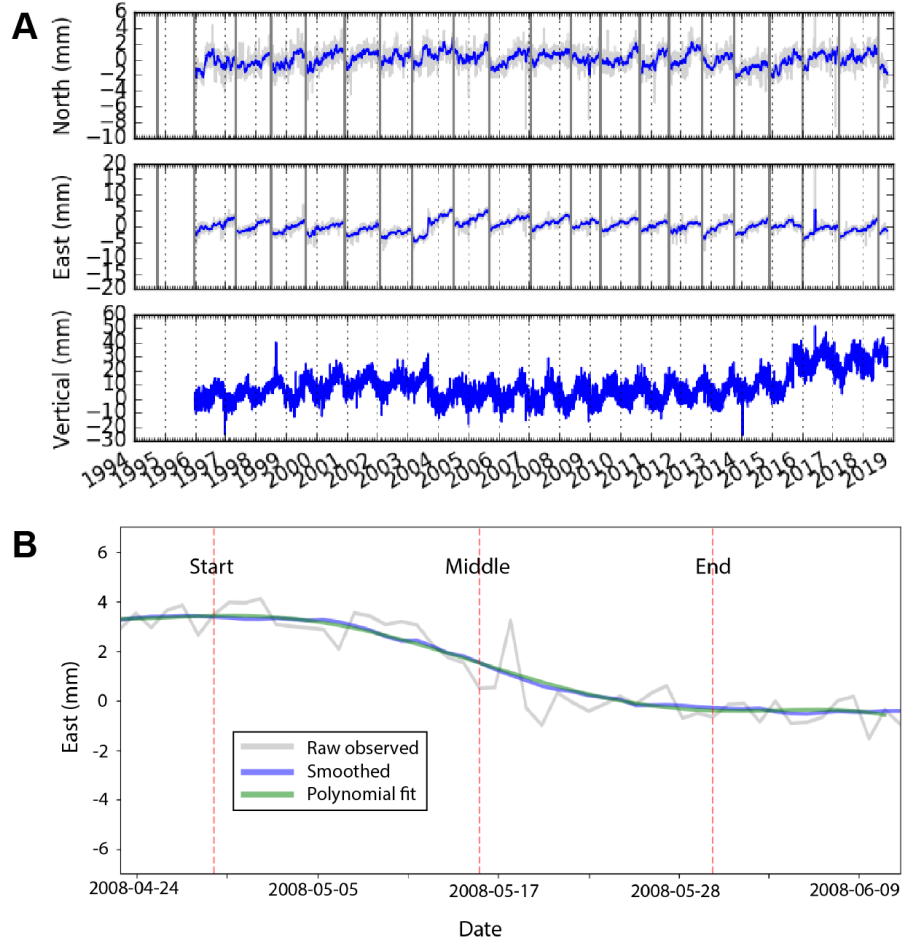


Figure C.6: GPS data used in the determination of ETS event times. (A) North, east and vertical position of GPS station ALBH as a function of time, detrended and smoothed using a 14-day rolling mean. (B) Example of start, end and mid-point times for the ETS event that occurred in 2008. The start and end times are likely overestimates but with no consequence to our temporal seismic analysis.

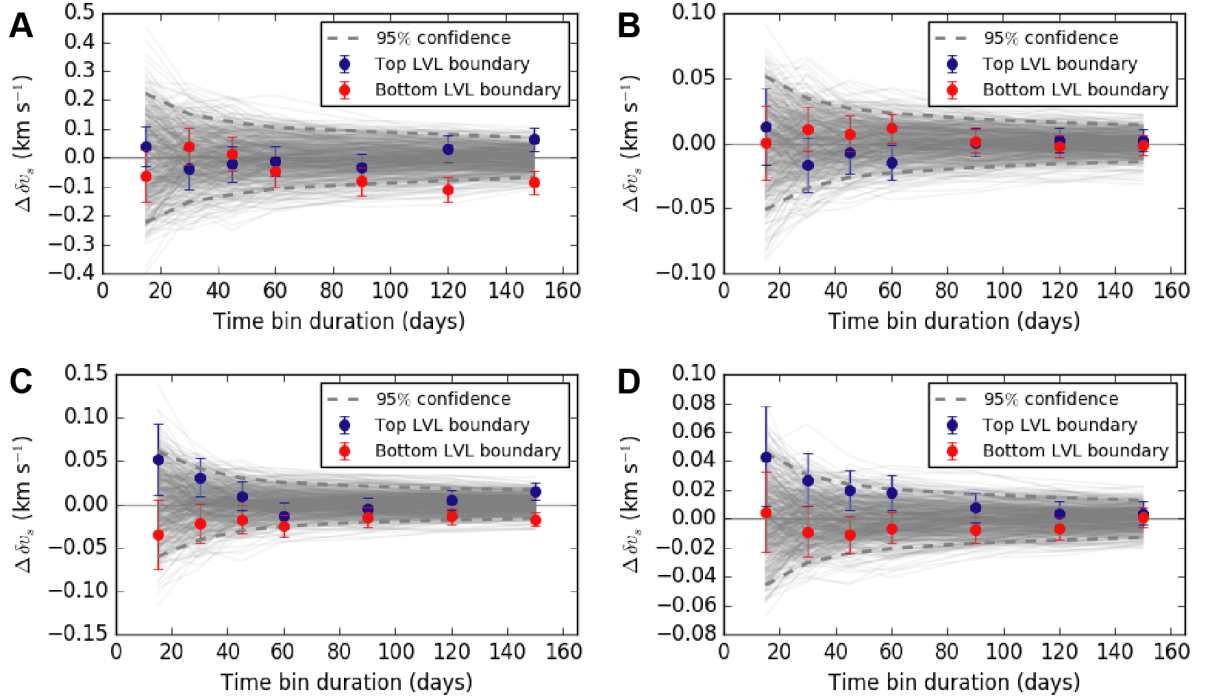


Figure C.7: Changes in δv_s across the top and bottom boundaries of the LVL. These results are shown for stations LVB (A), PFB (B), SNB (C) and VGZ (D) as a function of time bin duration. Bootstrap resampling of receiver function times relative to ETS are shown by grey lines (500 of 20000 samples). 95% confidence estimates from bootstrap analysis are shown by the dashed lines.

Table C.1: Strike and dip angles of the plate interface at each station location. These estimates are based on a smooth slab model (Bostock et al., 2019).

Station	Longitude ($^{\circ}$)	Latitude ($^{\circ}$)	Strike ($^{\circ}$)	Dip ($^{\circ}$)
LZB	236.176	48.6122	302.9	10.7
PFB	235.546	48.5747	305.0	12.4
PGC	236.548	48.6498	304.5	12.4
SNB	236.828	48.7751	307.9	15.5
VGZ	236.675	48.4131	307.2	11.6

Table C.2: Estimated start, end and mid-point dates of ETS events from 1994-2019. These dates are determined using continuous GPS data from station ALBH.

Start date	End date	Mid-point date
N/A	N/A	1994-10-02
N/A	N/A	1995-12-29
1997-04-25	1997-05-19	1997-05-05
1998-06-21	1998-07-24	1998-07-08
1999-08-18	1999-09-08	1999-08-28
2000-11-26	2000-12-17	2000-12-07
2002-01-18	2002-02-18	2002-02-05
2003-02-18	2003-03-13	2003-03-01
2004-06-25	2004-07-25	2004-07-12
2005-08-28	2005-10-01	2005-09-12
2007-01-16	2007-02-12	2007-01-29
2008-04-30	2008-05-29	2008-05-16
2009-05-03	2009-05-29	2009-05-15
2010-08-13	2010-09-09	2010-08-24
2011-08-11	2011-09-02	2011-08-23
2012-09-01	2012-09-28	2012-09-18
2013-09-17	2013-10-17	2013-10-01
2014-11-28	2014-12-23	2014-12-10
2015-12-20	2016-01-12	2016-01-01
2017-03-14	2017-04-08	2017-03-26
2018-06-25	2018-07-19	2018-07-06

Table C.3: p -values from the t -tests for the change in CC coefficients and δ_{V_S} values following ETS events measured at station PGC.

Time-bin width (days)	Δ CC (S_V)	Δ CC (S_H)	$\Delta\delta_{V_S}$ (top)	$\Delta\delta_{V_S}$ (bottom)
15	0.008	0.38	0.004	0.04
30	0.002	0.37	0.016	0.019
45	0.002	0.33	0.017	0.046
60	0.016	0.36	0.011	0.18
90	0.18	0.46	0.054	0.58
120	0.1	0.27	0.012	0.3
150	0.013	0.09	0.007	0.4

Appendix D

Supplementary materials for: Seismic
constraint on heterogeneous
deformation and stress state in the
forearc of the Hikurangi subduction
zone, New Zealand

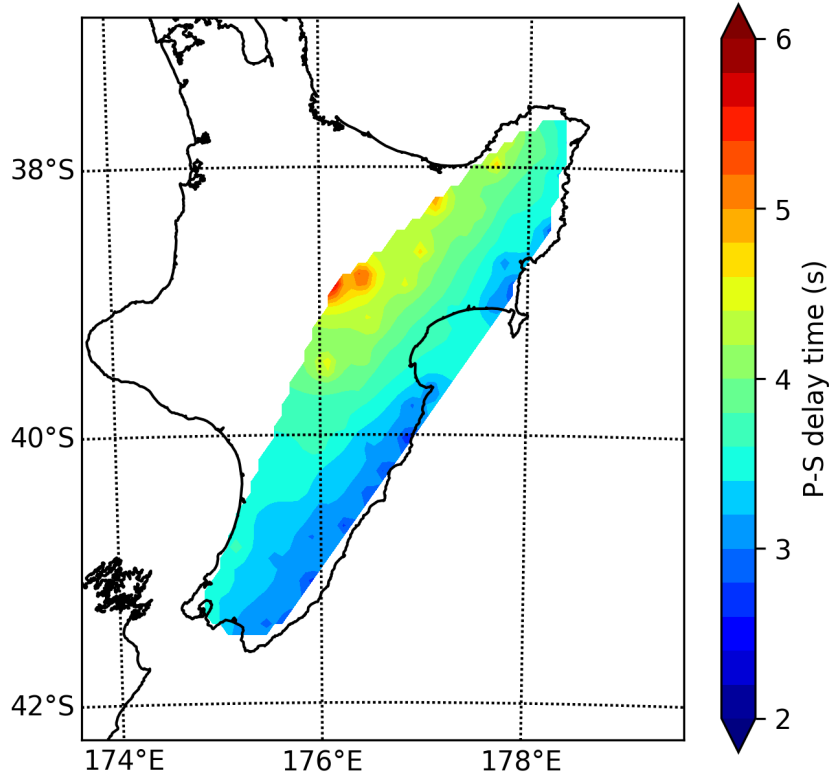


Figure D.1: Predicted P-S delay times of a converted S phase generated at a depth of the plate interface. Delay times are calculated using the plate interface model by Williams et al. (2013) and the seismic velocity model by Eberhart-Phillips et al. (2020). Because the Pacific plate is dipping towards the northwest, the phase from the plate interface is expected to arrive progressively later.

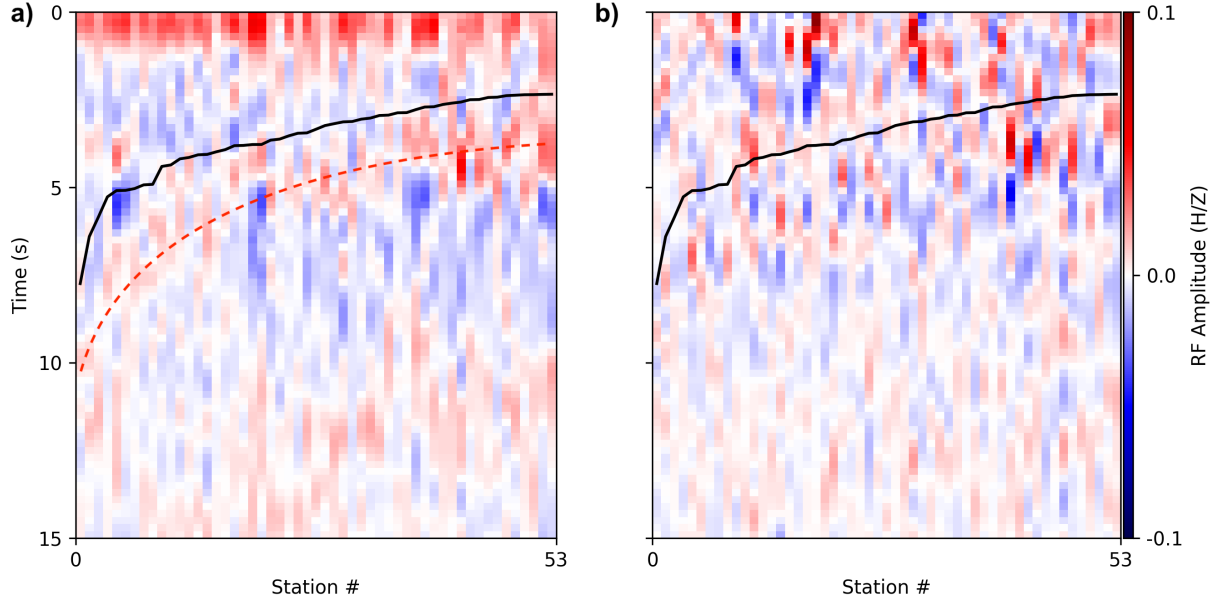


Figure D.2: A (a) and B_2 (b) harmonic components for all stations considered in this work. Stations are sorted by decreasing depth of the plate interface beneath each station. The black line represents the expected arrival of a direct converted phase from the plate interface. Delay times are calculated using the plate interface model by Williams et al. (2013) and the seismic velocity model by Eberhart-Phillips et al. (2020). The red dashed line in (a) highlights a coherent positive phase that appears beneath most stations and consistently arrives 2-3 seconds later than the expected arrival from the plate interface. This phase could be the result of a velocity contrast across the downgoing oceanic Moho. Note that the plate interface in the Hikurangi subduction zone is heterogeneous. Furthermore, the thickness of the subducting Pacific plate is not uniform across the margin, which may explain the variability in delay time of this phase between stations. Lastly, the A component represents the stack of the azimuthally binned radial RFs (see main text for details). Consequently, the signature of structure that varies azimuthally may destructively interfere.

Figure D.2 continued: Harmonic components are shown here for the optimal angle that maximizes signal variance on the B_2 component. Note that the equivalent maximum variance is found for B_2 of opposite polarity (i.e., optimal orientation rotated by 180°). Note also that the B_2 component amplitudes are generally greater at shallow depths (i.e., earlier times) as well as in the vicinity of the (predicted) plate interface.

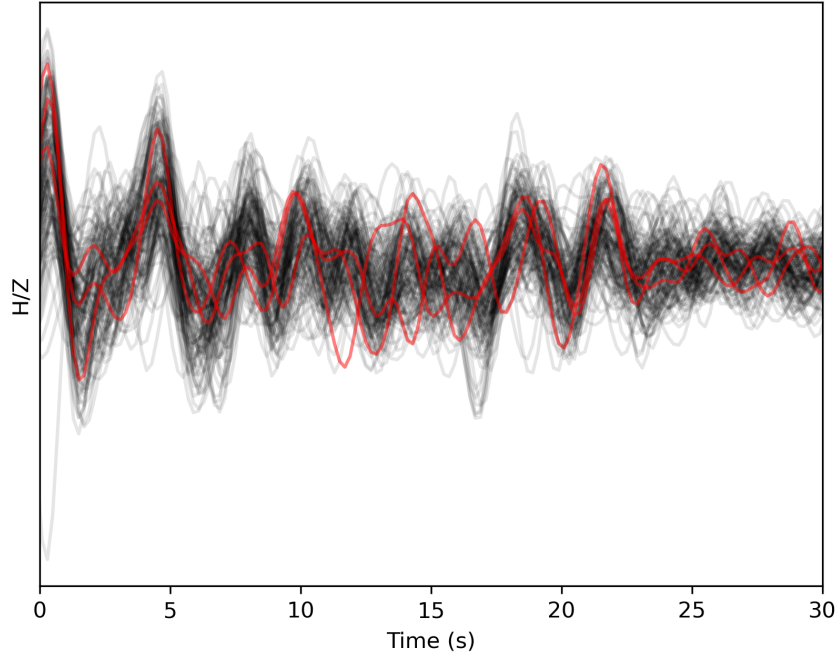


Figure D.3: Comparison of P (black) and PP (red) receiver functions calculated for station BFZ. Data are shown for events with back-azimuth between 270° - 300° . PP receiver functions capture the main phases consistent with the P receiver functions, but are generally of lower quality. Our dataset contains 2-3 orders of magnitude more P receiver functions than PP receiver functions (126 P and 3 PP receiver functions are shown here, respectively). This is primarily due to the quality control threshold based on signal-to-noise ratio used in this work (see text for details).

Table D.1: Results of harmonic decomposition of receiver function data. B_2 values are normalized by the sum of the standard deviations from all harmonic components, and represent the (relative) percent signal in the receiver function data explained by plunging anisotropy or dipping structures.

Station	Lat ($^{\circ}$)	Lon ($^{\circ}$)	α ($^{\circ}$)	α CI_{80} ($^{\circ}$)	B_2 (%)	B_2 CI_{80} (%)
ANWZ	-40.4597	176.4751	214	16.8	0.30	0.053
ARHZ	-39.2631	176.9959	185	25.5	0.27	0.070
BHHZ	-39.4893	176.0629	179	41.5	0.22	0.054
CAW	-41.1072	175.0664	238	22.1	0.20	0.056
CKHZ	-39.6582	177.0777	208	15.5	0.27	0.059
CNGZ	-38.4837	178.2075	218	14.8	0.33	0.070
DVHZ	-40.2989	176.1666	216	6.7	0.42	0.035
HATZ	-38.8908	176.0921	241	10.6	0.25	0.028
HOWZ	-40.8973	175.5151	249	20.0	0.22	0.043
KAHZ	-39.7942	176.8766	211	7.0	0.39	0.056
KRHZ	-39.6432	176.3701	165	6.4	0.47	0.043
KWHZ	-39.424	176.4228	204	5.9	0.41	0.021
MCHZ	-39.4442	176.6967	175	69.0	0.27	0.049
MHGZ	-38.9226	177.8833	176	9.6	0.35	0.040
MRHZ	-38.8337	176.4173	209	11.8	0.28	0.053
MSWZ	-41.4158	175.2495	188	16.4	0.26	0.050
MTHZ	-38.8523	176.8411	229	20.6	0.22	0.050
OGWZ	-40.8167	175.1703	168	95.3	0.17	0.045
PAWZ	-41.3816	175.4269	189	12.1	0.22	0.041
PKGZ	-37.8664	178.0805	220	8.8	0.24	0.064
PLWZ	-41.5692	175.2548	258	14.1	0.26	0.052
PNHZ	-39.9134	176.1998	199	7.5	0.44	0.035
POWZ	-40.3917	175.768	224	7.0	0.54	0.033
PRGZ	-38.9226	177.8833	161	34.8	0.25	0.053

PRHZ	-40.2664	176.6352	220	12.0	0.32	0.047
PRWZ	-40.5547	175.971	214	19.5	0.28	0.075
RAGZ	-38.4956	177.4154	248	8.0	0.34	0.041
RAHZ	-38.9162	177.086	195	14.1	0.26	0.052
RIGZ	-38.7058	177.7622	256	48.2	0.26	0.066
RUGZ	-37.9651	177.6775	264	15.0	0.28	0.041
TIWZ	-40.776	175.8857	214	11.6	0.29	0.037
TKGZ	-38.4392	177.8485	225	49.2	0.23	0.053
TMWZ	-41.1066	175.8905	236	10.8	0.24	0.046
TRWZ	-41.398	175.6879	245	5.5	0.34	0.058
TWGZ	-38.1767	177.9804	260	16.1	0.28	0.049
WHHZ	-39.0768	177.2344	230	65.2	0.23	0.071
WMGZ	-37.8191	178.4149	133	32.6	0.27	0.067
WPHZ	-40.0643	176.4406	176	9.6	0.35	0.050
BFZ	-40.6796	176.2462	215	10.4	0.34	0.036
BHW	-41.4082	174.8711	229	21.2	0.25	0.038
BKZ	-39.1657	176.4925	223	11.5	0.21	0.029
HAZ	-37.7561	177.7826	235	9.6	0.23	0.039
KNZ	-39.0218	177.6737	239	23.2	0.28	0.052
MRZ	-40.6605	175.5785	214	7.7	0.28	0.029
MWZ	-38.334	177.5278	253	5.2	0.28	0.058
MXZ	-37.5623	178.3066	100	5.1	0.39	0.038
PUZ	-38.0715	178.2572	240	13.9	0.27	0.033
PWZ	-40.0297	176.862	196	34.9	0.28	0.050
PXZ	-40.0306	176.8621	200	12.0	0.28	0.028
RTZ	-38.6154	176.9805	218	10.5	0.20	0.032
TSZ	-40.0586	175.9611	160	11.3	0.27	0.034
URZ	-38.2592	177.1109	215	39.3	0.20	0.037
WEL	-41.284	174.7682	224	15.8	0.20	0.020