Seismic hazard site assessment in Kitimat, British Columbia, via Bernstein-polynomial-based inversion of surface-wave dispersion

by

Jeremy M. Gosselin
B.Sc., University of Victoria, 2014

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of

MASTER OF SCIENCE

in the School of Earth and Ocean Sciences

Seismic hazard site assessment in Kitimat, British Columbia, via Bernstein-polynomial-based inversion of surface-wave dispersion

by

Jeremy M. Gosselin
B.Sc., University of Victoria, 2014

Supervisory Committee

Dr. Stan Dosso, Co-Supervisor
(School of Earth and Ocean Sciences)

Dr. John Cassidy, Co-Supervisor
(School of Earth and Ocean Sciences and Natural Resources Canada)

Dr. Jan Dettmer, Departmental Member
(School of Earth and Ocean Sciences)

Supervisory Committee

Dr. Stan Dosso, Co-Supervisor
(School of Earth and Ocean Sciences)

Dr. John Cassidy, Co-Supervisor
(School of Earth and Ocean Sciences and Natural Resources Canada)

Dr. Jan Dettmer, Departmental Member
(School of Earth and Ocean Sciences)

ABSTRACT

This thesis applies a fully nonlinear Bayesian inversion methodology to estimate shear-wave velocity (V_S) profiles and uncertainties from surface-wave dispersion data extracted from ambient seismic noise. In the inversion, the V_S profile is parameterized using a Bernstein polynomial basis, which efficiently characterizes general depth-dependent gradients in the soil/sediment column. Bernstein polynomials provide a stable parameterization in that small perturbations to the model parameters (basis-function coefficients) result in only small perturbations to the V_S profile. The inversion solution is defined in terms of the marginal posterior probability for V_S as a function of depth, estimated using Metropolis-Hastings sampling with parallel tempering.

This methodology is validated via inversion of synthetic dispersion data as well as previously-considered data inverted using different parameterizations. The approach considered here is better suited than layered modelling approaches in applications where smooth gradients in geophysical parameters are expected, and/or the observed data are diffuse and not sensitive to fine-scale discrete layering (such as surface-wave dispersion). The Bernstein polynomial representation is much more general than other gradient-based models such that the form of the gradients are determined by the data, rather than by subjective parameterization choice.

The Bernstein inversion methodology is also applied to dispersion data processed from passive array recordings collected in the coastal community of Kitimat, British Columbia. The region is the proposed site of several large-scale industrial development projects and has great economic and environmental significance for Canada. The inversion results are consistent with findings from other geophysical studies in the region and are used in a site-specific seismic hazard analysis. The level of ground-motion amplification expected to occur during an earthquake due to near-surface V_S structure is probabilistically quantified, and predicted to be significant compared to reference (hard ground) sites.

Contents

List of Tables

List of Figures

ACKNOWLEDGEMENTS

This work would not have been possible without the contributions of many individuals. In particular I would like to thank:

Stan Dosso and John Cassidy, for guiding and mentoring me through two degrees and fostering my scientific career. I am grateful for your insight, intellect, and humour as well as your willingness to always make time for me. Lastly, I appreciate all of the opportunities you have afforded me during my time at UVic, including the chance to work on such interesting and relevant projects.

Jan Dettmer, for supporting my research and for always asking the right questions.

Camille Brillon, for your help in planning and executing the field program in Kitimat.

Sheri Molnar, Heather Crow, André Pugin, and Jorge Quijano for useful discussions, advice, and motivation.

Dan Perera, for building the portable seismometer array used in this work.

My fellow graduate students and researchers at UVic and NRCan, for your camaraderie and for making these last two years truly fulfilling.

This work was funded through scholarships from the Natural Sciences and Engineering Research Council, the University of Victoria Faculty of Graduate Studies and School of Earth and Ocean Sciences, and the Canadian Society of Exploration Geophysics.

"Somewhere, something incredible is waiting to be known."

-Carl Sagan

DEDICATION

This thesis is dedicated to my parents Robert and Nathalie.
Without your constant encouragement, support, and pride in my achievements
I would not be where I am today.

Chapter 1

Introduction

1.1 Overview and motivation

Earthquakes are one of the most powerful forces of nature on earth and the ground shaking they cause can result in significant loss of life and destruction of property and infrastructure. These ground motions depend on the earthquake magnitude, depth, distance, and focal mechanism as well as the specific structure the seismic waves travel through from the earthquake source to the ground surface. Ground motions at a particular site are also strongly influenced by local geology. The amplitude and duration of shaking can be increased when seismic waves are trapped, focused, and resonate within sedimentary basins (Bard & Bouchon, 1980; Graves et al., 1998). Furthermore, ground motions can be amplified if seismic waves travel through material of decreasing impedance such as soft, unconsolidated, near-surface soils/sediments (Anderson et al., 1986, 1996). An example of this effect is illustrated in Figure 1.1, which shows ground accelerations recorded at four sites in Ottawa, Ontario (overlaying different geology) caused by the magnitude 5.4 Charlevoix-region earthquake near Rivière-du-Loup, Quebec in 2005 (Kakaand & Atkinson, 2005; Abbott & Samson, 2012). Despite being approximately the same distance from the earthquake focus, the four sites experienced significantly different amplitudes and durations of ground motions due to the effects of near-surface materials. Specifically, the presence of a thick soil layer overlying bedrock increases shaking amplitude by up to a factor of 20 compared to recordings on bedrock.

Investigation of site-specific seismic hazards requires knowledge of the near-surface soil/rock rigidity conditions, which are typically quantified by the shear-wave velocity

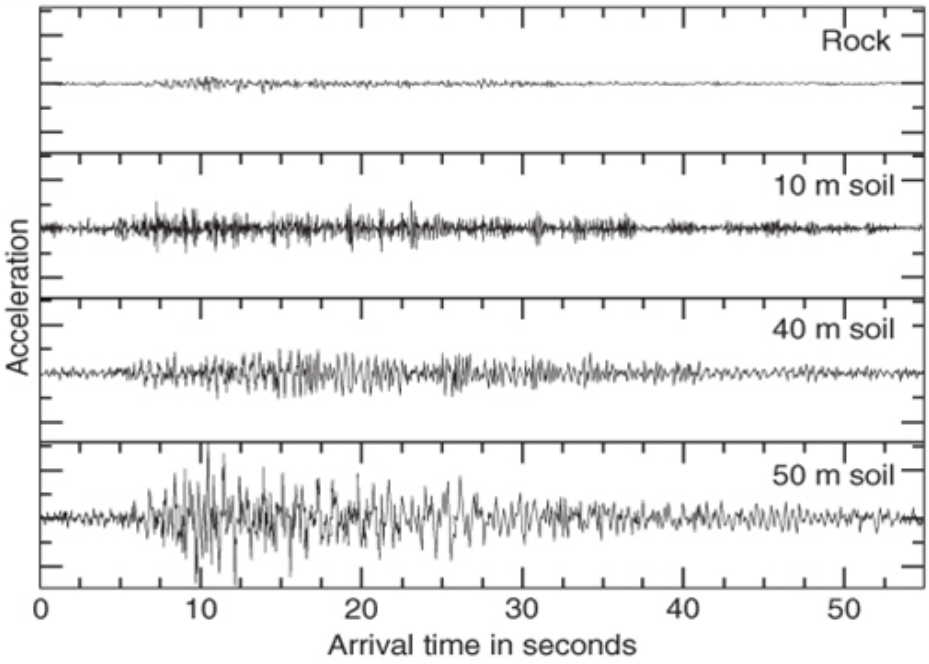

Figure 1.1: Ground acceleration at four sites in Ottawa, Ontario, during the 2005 Charlevoix earthquake (from Abbott & Samson, 2012).

(V_S) structure over the upper tens of metres. Site-specific hazards also depend on topography (basin structures) and earthquake-specific source properties. The building codes in many countries, including Canada, rely on this knowledge in developing engineering and design regulations which protect buildings and infrastructure (Humar, 2015). The federal government of Canada also relies on this knowledge in developing the most accurate national seismic hazard models possible (Adams et al., 2015). Specifically, these codes and models are based on a representative parameter called V_{S30}, which is the travel-time-average of V_S over the upper thirty metres.

Several methods exist for estimating near-surface V_S structure (and V_{S30}) at a site. This thesis considers the estimation of V_S profiles from the inversion of surface-wave dispersion data. The dispersive nature of surface waves is the result of varying penetration depth with frequency. Lower-frequency surface waves are more sensitive to structure at greater depth than higher-frequency waves, and therefore propagate at a different velocity. The V_S profiles, and (to a lesser extent) the compressional-wave velocity (V_P) and density (ρ) profiles, at a site can be inferred from the dispersion data (velocity-frequency relationship). An increasingly popular method for estimating surface-wave dispersion data involves processing recordings of ambient seismic noise

collected on small, portable, two-dimensional arrays of seismometers (Aki, 1957; Lacoss et al., 1969; Wathelet et al., 2008; Molnar et al., 2010). The data-collection process is logistically simple, non-invasive, non-disruptive, and costs less than most other methods (e.g., active-source seismic surveys and boreholes). A commonly-used method for processing dispersion data is the frequency-wavenumber (f-k) approach, which identifies the dominant coherent signal present in the array recordings (Lacoss et al., 1969). This signal is assumed to be the fundamental-mode Rayleigh wave when processing vertical-component recordings of ambient seismic noise. Predictions of fundamental-mode Rayleigh wave dispersion can then be compared to observed dispersion data to infer near-surface soil/rock structure.

The inference of local one-dimensional (1-D) V_S structure from surface-wave dispersion data is an inverse problem. This thesis considers Bayesian inversion of surface-wave dispersion, in which the parameters which describe the earth structure model (V_S profile) are treated as random variables constrained by prior information and data (e.g., Molnar et al., 2010; Dettmer et al., 2012). Bayesian inversion provides a rigorous approach to quantifying uncertainties for V_S profiles, which is important in accurately estimating uncertainty in seismic hazard analyses (Cornou et al., 2006). The solution to the inverse problem is described by the posterior probability density (PPD) of the model parameters, which must be sampled numerically using Markov-chain Monte Carlo (MCMC) methods (Mosegaard & Tarantola, 1995; Brooks et al., 2011). The V_S profile and estimates of the profile uncertainty are determined from the PPD samples.

An appropriate model must be selected to describe the V_S profile within the inversion. This thesis builds upon previous studies which considered Bayesian inversion of surface-wave disperion by modelling the V_S profile using a Bernstein polynomial basis, which efficiently characterizes general depth-dependent V_S gradients in the soil/sediment column (Quijano et al., 2016). The ability to effectively model gradient V_S structure is useful for seismic hazard applications (which investigate soil/sediment structure) since V_S in sediments continuously increases with depth due to the weight of overlying material (Budhu, 2007). This representation of earth structure is much more general than other gradient-based models such that the form of the profile is determined by the data, rather than by a subjective model choice.

This thesis considers the application of these methods in the coastal community of Kitimat, British Columbia (BC), for the purpose of quantifying the level (and variability) of ground-motion amplification due to near-surface sediments which is expected to occur during an earthquake. The Kitimat region is the proposed site of several

large-scale industrial development projects (e.g., possible pipeline terminus and export facilities for heavy oil and natural gas). Consequently, the region has great economic and environmental significance for Canada. Earthquake recordings (Figure 1.2) from recently installed seismometer stations indicate significant amplification occurs in the region (Brillon, 2016b). Site-specific seismic hazard analysis is important in this developing community.

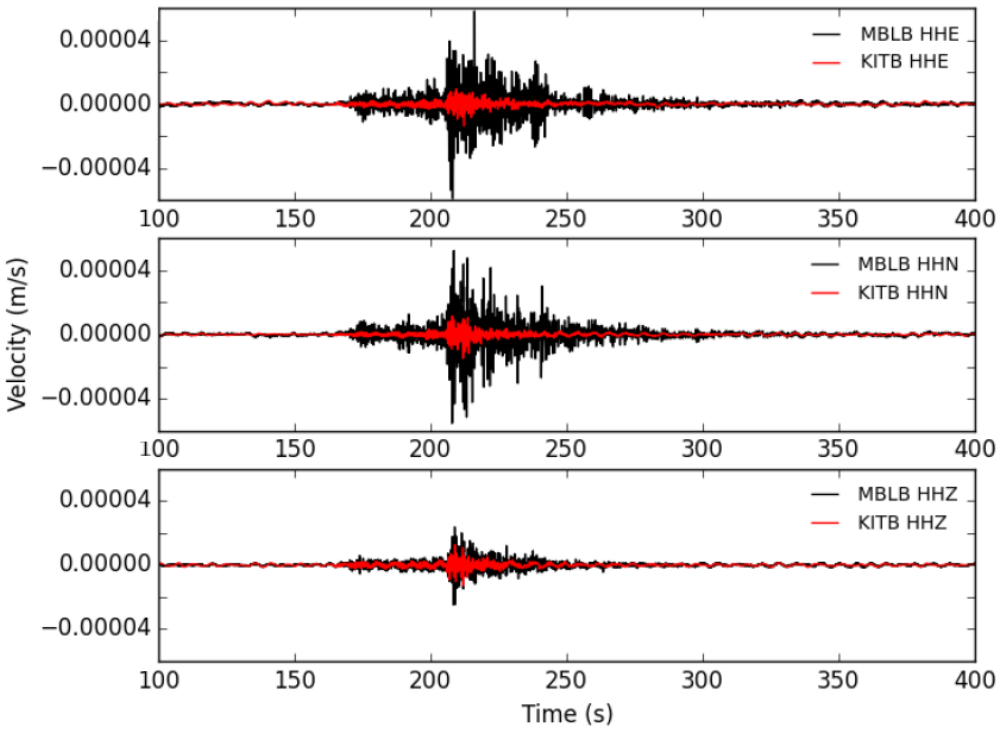

Figure 1.2: Ground velocity at KITB (bedrock) and MBLB (soil) seismometer stations (in the Kitimat region) during the magnitude 3.6 Haida Gwaii earthquake on December 12, 2015 (from Brillon, 2016b). HHE, HHN, and HHZ are the east-, north-, and vertical-component recordings, respectively.

1.2 Thesis work summary

The main body of this thesis consists of two chapters which are intended to be submitted as individual articles in scientific journals. As such, these chapters are meant to be (relatively) self-contained studies and therefore there is some repetition in terms of introductory material, theory and data. The following outline summarizes this work.

Chapter 2 describes the development and application of a fully nonlinear Bayesian inversion methodology to estimate the 1-D V_S structure and uncertainties from surface-wave dispersion data. In the inversion, the V_S profile is parameterized using a Bernstein polynomial basis for efficient characterization of general V_S gradients. The results of the inversion are interpreted from marginal posterior probability profiles for V_S, estimated using MCMC sampling. This work includes the development and testing of FORTRAN software which implements this inversion methodology, in combination with a forward computation (data prediction) routine from Wathelet (2005). This methodology is applied to synthetic dispersion data as well as data processed from ambient seismic noise collected on the Fraser River Delta in BC. The measured data have been previously studied by Molnar et al. (2010), Dettmer et al. (2012), and Molnar et al. (2013) and are addressed in this chapter to validate the Bernstein-polynomial method via comparison with previous inversion methodologies as well as nearby invasive V_S measurements. The results of this work will benefit the earthquake engineering community by demonstrating a general and efficient method for modelling depth-dependent V_S gradients in sediments, for the purpose of seismic hazard analysis. This methodology will be widely applicable in geophysical inverse problems concerned with modelling gradient structure.

Chapter 3 describes the collection, processing, and inversion of ambient seismic noise recordings collected at three sites in Kitimat, BC. The recordings are processed using the f-k method to estimate dispersion curves of fundamental-mode Rayleigh waves. This work included the organization and execution of a field program to collect passive seismic data in the study region, as well as the development and testing of FORTRAN software to apply f-k processing. The dispersion curves are inverted using the methodology described in Chapter 2 to estimate posterior probability profiles for V_S. The rigorous, quantitative determination of V_S structure and uncertainties is then translated to site-specific seismic haz-

ard analyses (Molnar et al., 2013). Specifically, the inversion results are used to quantitatively assess seismic site classification via V_{S30}, within the context of the National Building Code of Canada. This parameter is then used to estimate the linear amplification factor of peak ground velocity, peak ground acceleration, and spectral acceleration during an earthquake due to local 1-D V_S structure. Lastly, the inversion results are used as input for ground-motion simulation software which calculates the amplification and resonance of vertically-incident horizontally-polarized shear waves (Boore, 2005). The results of this work will be useful for planners and regulatory agencies in the development of seismic risk models as well as seismic hazard mitigation plans.

Chapter 4 summarizes the main results of this thesis and discusses recommendations for future work.

Chapter 2

A gradient-based model parameterization using Bernstein polynomials in Bayesian inversion of surface-wave dispersion

2.1 Background

The level of shaking experienced during an earthquake is strongly dependent upon the local, near-surface geophysical properties. In particular, as seismic waves propagate through material of decreasing impedance, the resistance to motion decreases and wave amplitude increases (Anderson et al., 1986, 1996). Amplification can also occur at particular frequencies due to resonance within near-surface low-velocity layers. Knowledge of the local, near-surface geophysical properties is critical for understanding seismic amplification and resonance hazards. The shear-wave velocity (V_S) profile over the upper tens of metres is representative of the soil/sediment rigidity at a site, and can be used to predict the expected ground response to earthquake shaking. Thus, estimating V_S is important for seismic site classification and studies of public safety related to earthquake hazards. The V_S profile at a site can be determined by invasive methods such as boreholes or seismic cone penetration tests (SCPT), or can be estimated from non-invasive, active or passive seismic observations made at the surface. Passive techniques involving small, portable, two-dimensional (2-D) seismic arrays are increasingly popular because they often have lower demands in cost and logistics com-

pared to other methods (Aki, 1957; Lacoss et al., 1969; Wathelet et al., 2008; Molnar et al., 2010). This technique involves processing the array recordings of ambient seismic noise to estimate the phase velocity of fundamental-mode Rayleigh waves as a function of frequency (the dispersion curve), which can be inverted to estimate the V_S profile and, with lower resolution, the compressional-wave velocity (V_P) and density (ρ) profiles.

Understanding and quantifying the uncertainties in estimated V_S profiles, and their effects on seismic site classification uncertainties, is of significant interest to developers, planners, and public officials internationally and has been identified as a critical issue for passive seismic array methods (Cornou et al., 2006). Bayesian inference provides a rigorous approach to uncertainty quantification for nonlinear inverse problems (like dispersion inversion). Bayesian inversion is a probabilistic approach in which the parameters that describe the model (e.g. the V_S profile) are constrained by the observed data and prior information. The solution is described by properties of the posterior probability density (PPD) of the model parameters, which is generally estimated numerically using Markov-chain Monte Carlo (MCMC) methods (Mosegaard & Tarantola, 1995; Brooks et al., 2011). Nonlinear Bayesian inversion provides a rigorous approach to quantifying uncertainties in nonlinear inversion which avoids linearization errors and subjective regularizations that can preclude meaningful uncertainty analysis in linearized approaches. Bayesian inversion has been applied previously to dispersion data processed from passive array recordings (Molnar et al., 2010; Dettmer et al., 2012; Molnar et al., 2013). Within a Bayesian framework, the most probable V_S profile as well as quantitative measures of the profile uncertainty can be determined from the PPD.

In any inversion, an appropriate parameterization is required to describe the model that represents the physical system of interest (in this case, the V_S profile). In many approaches to nonlinear geophysical inversion, the earth-structure profiles are parameterized as a stack of homogeneous layers. However, this may not be an appropriate parameterization in cases where smooth gradients are expected, and can introduce non-physical discontinuities in the inversion result. This could potentially have an impact on the estimated ground response to earthquake shaking, particularly if these non-physical discontinuities produce resonance due to trapped modes at particular frequencies (that are not the true resonant frequencies at a site). Furthermore, Wathelet et al. (2008) suggest the penetration of surface wave dispersion data may be limited by significant impedance contrasts. Consequently, parameterizations based on multiple

impedance contrasts are not believed to be consistent with data resolution. Molnar et al. (2010) considered Bayesian inversion of array dispersion data using parameterizations described by linear and power-law functional forms to represent V_S gradient structures in near-surface soils/sediments. Results in good agreement with borehole and SCPT measurements were obtained, but in this approach the model form (power-law or linear gradient) is imposed on the solution rather than determined from the data. Further, while often reasonably applicable, these gradient forms have limited flexibility to capture general earth structure. Dettmer et al. (2012) considered Bayesian inversion of the same data using a trans-dimensional (trans-D) model parameterization based on a stack of homogeneous layers, where the dimension of the model (number of layers) is treated as unknown and sampled within the inversion (Green, 1995). This approach is designed to objectively treat unknown parameterizations and includes parameterization uncertainty within the profile uncertainty estimates. However, as it is based on discontinuous uniform layers, it may not represent smooth gradients well. Furthermore, surface-wave dispersion data are typically not sensitive to fine-scale discontinuous structure due to the diffusive nature of surface-waves.

This chapter considers a general and efficient approach to parameterizing smooth gradient-based profiles in terms of a Bernstein polynomial representation (Quijano et al 2016). A Bernstein polynomial is defined as a weighted sum of Bernstein basis functions, which are themselves polynomials with several desirable features for representing general smooth functions (Farouki & Rajan, 1989; Farouki, 2012). The Bernstein basis functions are pre-defined for a given polynomial order, and the inversion estimates the coefficients which weight the individual basis functions. In this way, general V_S structure is represented using fewer parameters than a classical layered model parameterization. The distinctive advantageous property of the Bernstein polynomial, over other polynomial forms (e.g., splines), is its stability. Specifically, small perturbations to the model parameters (basis-function coefficients) result in only small, localized perturbations to the V_S profile. In fact, the Bernstein polynomial has optimal stability for a given polynomial order (Farouki & Rajan, 1989; Farouki, 2012; Quijano et al., 2016).

The Bernstein polynomial was originally formulated for applications in theoretical mathematics with later applications in computer-aided design, differential equation solutions, and control theory (Gordon & Riesenfeld, 1974; Bhattia & Bracken, 2007; Farouki, 2012; Basirat & Shahdadi, 2013). Quijano et al. (2016) parameterized a seabed geoacoustic model using a Bernstein polynomial in Bayesian inversion of ocean-acoustic bottom-loss data. This chapter introduces the Bernstein polynomial parameterization

to geophysical inversion by estimating V_S profiles from fundamental-mode Rayleigh wave dispersion data within a Bayesian framework. This methodology is applied to synthetic dispersion data as well as passive array data recorded on the Fraser River Delta in British Columbia (BC), Canada. The measured data have been previously studied by Molnar et al. (2010) and Dettmer et al. (2012) and are considered in this chapter for comparison with previous inversion methodologies, as well as co-located borehole and SCPT measurements.

2.2 Bayesian inversion methodology

This section provides an overview of the Bayesian inversion theory and implementation used in this study, as well as a description of the treatment of prior knowledge and data errors. A more complete review of Bayesian inversion can be found in Mosegaard & Tarantola (1995), Brooks et al. (2011), and Dosso et al. (2014).

Let $\mathbf{d}$ be a vector of N data (in this case, the Rayleigh-wave dispersion) and $\mathbf{m}$ be a vector of M model parameters (geophysical parameter profiles), with both assumed to be random variables. Bayes' theorem can be written as

$$P(\mathbf{m}|\mathbf{d},H) = \frac{P(\mathbf{m}|H)\,P(\mathbf{d}|\mathbf{m},H)}{P(\mathbf{d}|H)} \tag{2.1}$$

$$= \frac{P(\mathbf{m}|H)\,P(\mathbf{d}|\mathbf{m},H)}{\int_{\mathcal{M}} P(\mathbf{m}'|H)\,P(\mathbf{d}|\mathbf{m}',H)\,d\mathbf{m}'}, \tag{2.2}$$

where H is the choice of model. The term $P(\mathbf{m}|H)$ is the probability of a set of model parameters $\mathbf{m}$, given model H, independent of the data, and represents the prior density. $P(\mathbf{d}|\mathbf{m},H)$ is the conditional probability of $\mathbf{d}$ given $\mathbf{m}$. In practice, once the data are measured $\mathbf{d}$ represents a fixed realization of the random variable, and this term is interpreted as the likelihood of $\mathbf{m}$, written $\mathcal{L}(\mathbf{m})$ (for fixed model H). $P(\mathbf{m}|\mathbf{d},H)$ is the PPD, representing the probability density over the model parameters given the data, prior information, and choice of H. $P(\mathbf{d}|H)$ is the probability of the data, for a given model H, independent of $\mathbf{m}$. This term is often referred to as the Bayesian evidence and provides normalization over the parameter space. Bayesian evidence can be interpreted as the likelihood of model H for the measured data. In this case, H is considered to be a Bernstein polynomial (of unknown order) representation of geophysical parameter profiles. The Bernstein polynomial order used in the inversion is determined by estimating an approximation to Bayesian evidence for various polyno-

mial orders, and selecting the one that maximizes evidence. This is discussed further in the next section.

Estimating properties of the PPD such as the maximum *a posteriori* (MAP) model $\hat{\mathbf{m}}$ and marginal probability densities provides the model parameters and associated uncertainties. However, for non-linear inverse problems, analytic solutions generally do not exist and numerical methods are required. Specifically, the Metropolis-Hastings algorithm is a MCMC method used here to draw a series of dependent, asymptotically-unbiased samples from the PPD (Metropolis et al., 1953; Hastings, 1970; Brooks et al., 2011). Metropolis-Hastings sampling is applied by drawing a random set of model parameters $\mathbf{m}'$ from a proposal distribution $Q(\mathbf{m}'|\mathbf{m})$ dependent on the current set of model parameters $\mathbf{m}$, and accepting the new parameters with probability

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

In this study, the prior probability density consists of a bounded uniform distribution (such that $P(\mathbf{m}) = P(\mathbf{m}')$) constraining parameters to geophysically-realistic values. Wide bounds are applied to allow the data information, as opposed to the prior information, to primarily determine the solution. Proposal densities are chosen to be Gaussian distributed and centred on the current model such that they are symmetric ($Q(\mathbf{m}'|\mathbf{m}) = Q(\mathbf{m}|\mathbf{m}')$). The acceptance criterion, eq. (2.3), simplifies to a ratio of likelihoods with these prior and proposal densities. A sufficiently large set of MCMC samples can be used to approximate the PPD.

The likelihood function is formulated by specifying the statistical distribution of the data errors. The data error distribution is often not known independently, as it must include both theory and measurements errors (which typically cannot be distinguished). In many cases, a multi-variate Gaussian distribution is a reasonable assumption (supported by the Central Limit Theorem), and is applied here and subsequently validated (Dosso et al., 2006). In particular, the likelihood function is given by

$$\mathcal{L}(\mathbf{m}, \mathbf{C}_d) = \frac{1}{(2\pi)^{N/2}|\mathbf{C}_d|^{1/2}}\,\exp\left[-\mathbf{r}(\mathbf{m})^\top \mathbf{C}_d^{-1}\mathbf{r}(\mathbf{m})/2\right], \tag{2.4}$$

where

$$\mathbf{r}(\mathbf{m}) = \mathbf{d} - \mathbf{d}(\mathbf{m}) - \tilde{\mathbf{d}}(\mathbf{m}, a) \tag{2.5}$$

are the total data residuals, $\mathbf{d}(\mathbf{m})$ are the data predicted for model $\mathbf{m}$, and $\mathbf{C}_d$ is

the data error covariance matrix. The total residuals account for a first-order autoregressive (AR) process to compensate for error correlations according to (Shumway & Stoffer, 2000)

$$\tilde{d}_i(\mathbf{m}, a) = a(d_{i-1} - d_{i-1}(\mathbf{m})), \tag{2.6}$$

where a is the AR parameter, which is considered unknown in the inversion (Dettmer et al., 2012; Steininger, 2013). With this representation of the data residuals, the data error covariance matrix $\mathbf{C}_d$ is taken to be diagonal with unknown variances. In this study, the N data are divided into N_D subsets with N_i data in the i^{th} subset. Data errors for a given subset are assumed to have the same variance, with the variances σ_i^2 treated as unknowns. Instead of treating the variances as explicit unknown parameters, a likelihood function is applied which samples implicitly over maximum-likelihood estimates of variance. A complete derivation of this likelihood function (from eq. (2.4)) is given in Dosso & Wilmut (2005) and leads to

$$\mathcal{L}(\mathbf{m}) \propto \exp\left[-\frac{1}{2}\sum_{i=1}^{N_D} N_i \log_e |\mathbf{r}_i(\mathbf{m})|^2\right], \tag{2.7}$$

where $\mathbf{r}_i$ are the total residuals for the i^{th} data subset. This procedure assumes that the total residuals are uncorrelated and Gaussian distributed (Dettmer et al., 2012). However, care must be taken with the AR approach. For AR parameters a approaching unity, the error model can compensate for poor-fitting models which are then accepted into the Markov chain. These AR models are therefore considered to provide unreasonable covariance estimates. Similar to Dettmer et al. (2012), this issue is addressed by rejecting AR models predicting a standard deviation of $\tilde{\mathbf{d}}(\mathbf{m}, a)$ greater than a threshold set to 3 times the residual standard deviation (without AR) for the previous model in the Markov Chain. *A posteriori* statistical tests can be applied to the standardized residuals to check the assumptions of the data error model.

In this work, the data are treated in terms of slowness (the reciprocal of phase velocity) as passive array methods typically estimate the wavenumber of propagating surface-waves, and errors in slowness scale linearly with errors in the estimated wavenumber. For data errors with the same variance within a given subset (frequency band), this has the effect of increasing the phase-velocity errors at low frequencies, and is believed to be a more physically-reasonable representation of the frequency dependence of the data errors than previous work which treated phase velocities as data.

PPDs can potentially be multimodal in nonlinear inverse problems. This can lead to

inefficient sampling as the Markov chain can become isolated in high-probability regions of the parameter space. This issue is addressed here by applying parallel tempering within MCMC sampling (Geyer, 1991; Dosso et al., 2012; Sambridge, 2014). Parallel tempering is implemented by simultaneously running additional Markov chains, with successively relaxed proposal and acceptance probabilities, such that these chains explore the parameter space more widely. The Markov chains interact through probabilistic interchanges, allowing them to swap locations within the parameter space. This allows the original (unrelaxed) Markov chain to jump to different regions of the model space more efficiently. Only the samples from the original Markov chain provide an unbiased sampling of the PPD; hence, only these samples are kept.

2.3 Model parameterization

The ability to represent general earth structure using a simple functional form is a useful characteristic of a model parameterization in Bayesian inversion. This section describes the formulation of a gradient-based model parameterization, using Bernstein polynomials, which is applied within the Bayesian inversion methodology described in the previous section. Specifically, a polynomial representation of specific parameter depth profiles are estimated within the inversion. The Bernstein polynomial representation of the depth-dependant model parameter over the interval $0 < z < z_0$ is given by

$$u(\tilde{z}) = \sum_{j=0}^{J^u} g_j^u b_j(\tilde{z}, J^u), \tag{2.8}$$

where u is the parameter being represented in polynomial form (e.g, u can represent V_S, V_P/V_S and ρ), $\tilde{z} = z/z_0$ is the normalized depth (z_0 is the maximum depth of the polynomial representation), J^u is the order of the Bernstein polynomial for parameter u, g_j^u are coefficients, and

$$b_j(\tilde{z}, J^u) = \binom{J^u}{j} (1-\tilde{z})^{(J^u-j)} \tilde{z}^j \tag{2.9}$$

are the Bernstein basis functions. Because the basis functions are fixed, only the coefficients and maximum depth z_0 are treated as unknowns within the inversion. Figure 2.1 shows the Bernstein basis functions for polynomials of orders 1 to 5 over the normalized depth. From the figure it is apparent that, within a given polynomial order,

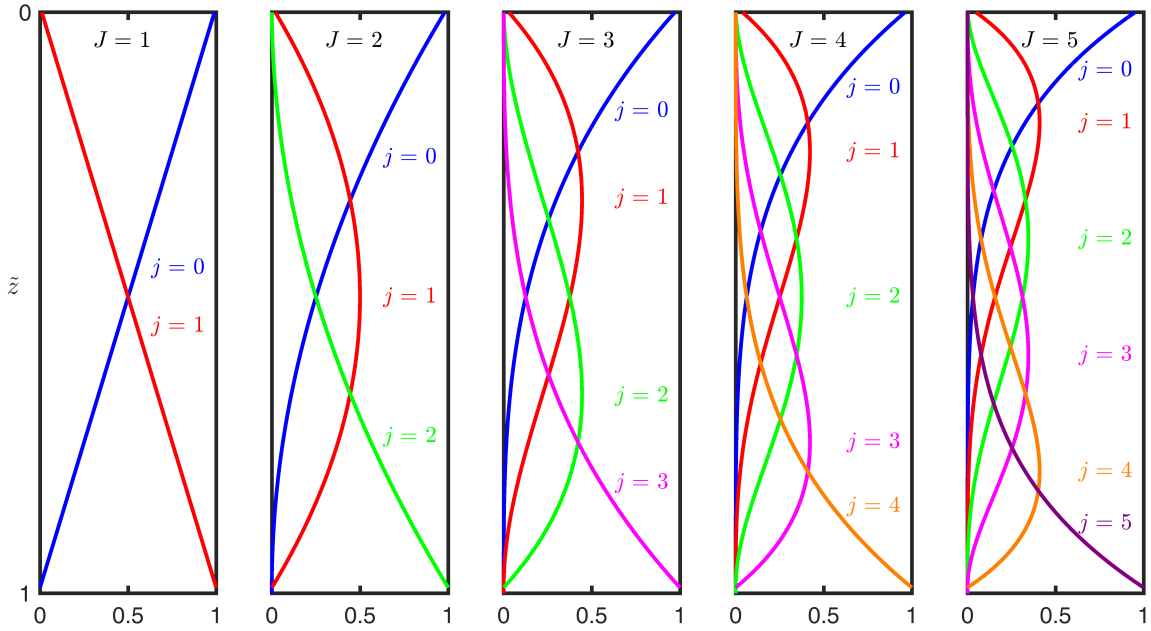

Figure 2.1: Basis functions for Bernstein polynomials of orders 1 to 5.

the basis functions are unimodal with localized peaks at successively greater depths. This property leads the Bernstein polynomial to be an effective and stable model parameterization in that a perturbation to a model parameter (basis function coefficient) will only significantly alter the model over a localized depth range, a property that is highly desirable for inversions based on MCMC sampling as it allows for detailed and efficient exploration of the model parameter space. In fact, the Bernstein polynomial has optimal stability for a given polynomial order (Farouki & Rajan, 1989; Farouki, 2012; Quijano et al., 2016). The Bernstein basis functions are designed such that they sum to unity at all depths (without coefficients). As a result, the width of the uniform prior bound on the coefficients is equivalent to the width of the prior on the V_S profile. However, uniform priors on the coefficients do not lead to a uniform prior for the V_S profile. The V_S prior that results from uniform bound priors on the Bernstein coefficients can be mapped out by applying MCMC sampling based only on the prior (i.e., the data are assumed to have infinite uncertainties). Examples of these V_S priors are given in Figure 2.2 for polynomial orders of 2 and 5. The V_S priors are uniform at (normalized) depths of 0 and 1, but in between favour V_S values in the middle of the bounded range. However, the priors are not strongly peaked and are not expected to strongly effect the PPD compared to the likelihood (data information).

In this study, the forward computation (calculation of predicted dispersion data)

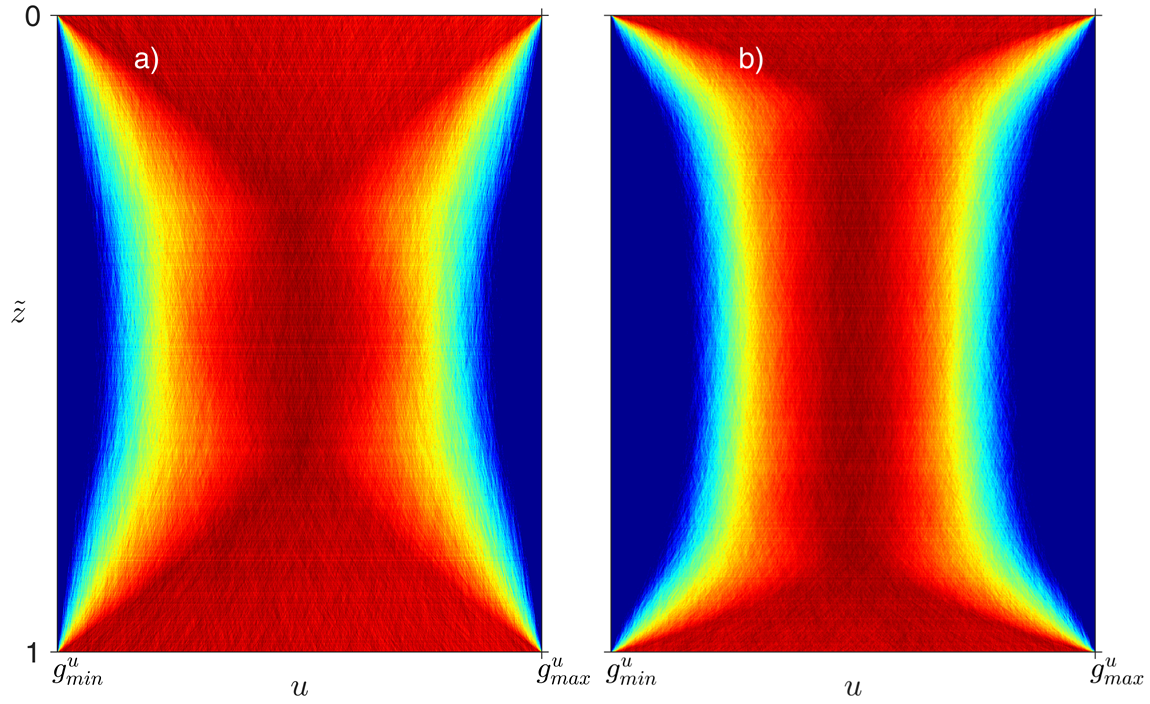

Figure 2.2: Prior density profiles for parameter u using Bernstein polynomials of order (a) $J^u = 2$ and (b) $J^u = 5$. Warm colours (red) represent high probability and cool colours (blue) represent low probability. Probabilities are normalized independently at every depth for display purposes.

is performed using the computer routine developed by Wathelet (2005). This routine assumes the earth-structure model is composed of a stack of homogeneous layers of differing geophysical parameters (layer thickness, V_S, V_P, and ρ) and applies the Thomson-Haskell propagator-matrix method to efficiently calculate the dispersion (Thomson, 1950; Haskell, 1953; Knopoff, 1964; Gilbert & Backus, 1966). It is well known that dispersion data are less sensitive to V_P and ρ than to V_S. In this study, both V_S and V_P/V_S are modelled as Bernstein polynomials with coefficients treated as unknowns in the inversion. Both polynomial profiles span the same depth and overlie a half-space. As such, the V_S and V_P/V_S values in the half-space are also treated as unknowns in the inversion. To improve efficiency and constrain density values, ρ is computed directly from an empirical relationship with V_P (Gardner et al., 1974). A conversion from a Bernstein polynomial to a representative stack of homogeneous layers is required for the forward computation (data prediction). Similar to Molnar et al. (2010), a logarithmically-increasing depth partition is used here since the resolution

of surface-wave dispersion data decays exponentially with depth. The thickness of the earth model which is represented in polynomial form is z_0. The thickness of the i^{th} layer of L layers in the partition is given by $lb^{(i-1)}$, where l is the (user-specified) thickness of the first layer, and b is the solution to

$$z_0 = \sum_{i=1}^{L} lb^{(i-1)} = \frac{l(1-b^L)}{1-b}. \tag{2.10}$$

This equation is solved efficiently for b using the bisection method as z_0 increases monotonically with b.

The choice of polynomial order is an important component of the inversion. Using too small an order can result in an under-parameterized model, where the data may be under-fit and model structure unresolved. Conversely, too large of a polynomial order can result in overfitting the data and including spurious (unconstrained) structure in the model. A quantitative approach to finding the optimal Bernstein polynomial order for both the V_S and V_P/V_S profiles is considered here (Quijano et al., 2016). Specifically, the Bayesian information criterion (BIC) is computed for the maximum-likelihood model in the Markov chain as an approximation to Bayesian evidence (Schwartz, 1978; Kass & Raftery, 1995). Given the uniform prior densities, the maximum-likelihood model corresponds to the MAP model $\hat{\mathbf{m}}$. The BIC is given by

$$\mathrm{BIC} = -2\log_e(\mathcal{L}(\hat{\mathbf{m}})) + M\log_e(N), \tag{2.11}$$

where $\mathcal{L}(\hat{\mathbf{m}})$ is the likelihood of the MAP model, M is the number of model parameters, and N is the number of data. The parameterization, or polynomial order, with the lowest BIC represents the simplest model consistent with the resolving power of the data.

2.4 Synthetic test case

This section illustrates the Bernstein polynomial parameterization using simulated Rayleigh-wave dispersion data computed for an earth structure model with V_S increasing with depth according to a power-law and, V_P/V_S increasing near the surface according to a power-law (the ratio is expected to increase near the surface as V_S decreases more quickly in unconsolidated materials), overlaying a half-space. A high-quality dataset of 40 logarithmically-spaced slowness values were computed between 1

Bernstein parameter	min	max
g^{V_S} (m/s)	50	1000
g^{V_P/V_S}	1.4	3
z_0 (m)	20	150
Half-space V_S (m/s)	500	1000
Half-space V_P/V_S	1.4	3
a	0	0.9

Trans-D parameter	min	max
Layer V_S (m/s)	50	1000
Layer V_P/V_S	1.4	3
Layer depth (m)	0	150
Number of layers	2	20
a	0	0.9

Table 2.1: Prior bounds for model parameters in Bernstein-polynomial and Trans-D inversions of synthetic dispersion data.

and 12 Hz. Correlated errors were simulated using Gaussian-distributed errors with standard deviation 10^{-4} s/m (in slowness) and AR parameter 0.6. Bayesian inversions were carried out using the methodology described above as well as using the trans-D layered approach of Dettmer et al. (2012) (with the same error model as the Bernstein inversion) for comparison of the effects of the choice of model parameterization on the inversion results. Figure 2.3 shows the results of the BIC analysis performed for the Bernstein polynomial approach. The BIC is a minimum for $J^{V_S} = 3$ and $J^{V_P/V_S} = 1$ and so the final inversion results are shown for this parameterization.

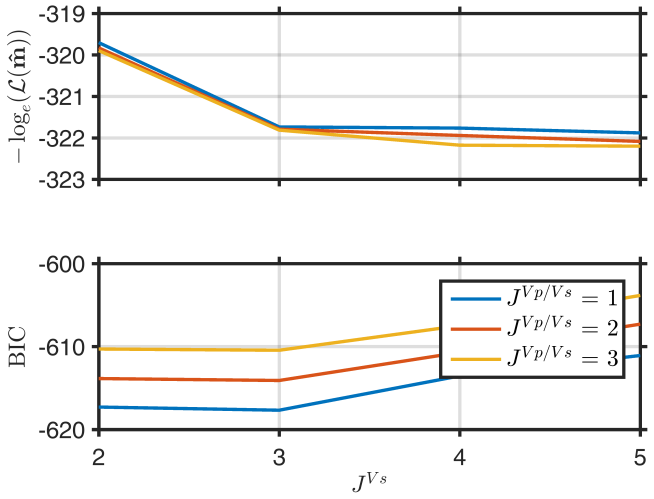

Figure 2.3: BIC analysis for synthetic dispersion data. The likelihood of the MAP model $\mathcal{L}(\hat{\mathbf{m}})$ is also shown as a function of Bernstein-polynomial order.

Figure 2.4 shows the fit to the data from the inversions (plotted as phase velocities, although slownesses were inverted). The red stars are the synthetic dispersion data

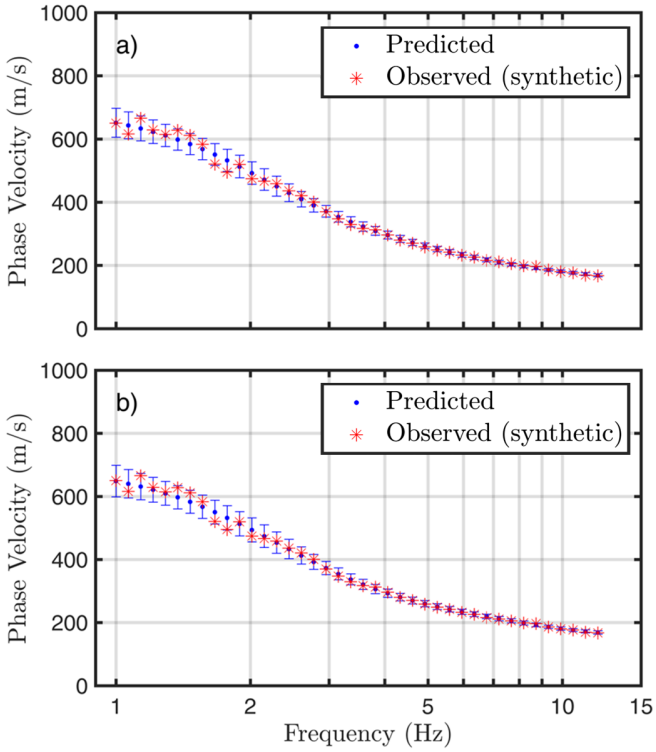

Figure 2.4: Fit to synthetic dispersion data using (a) Bernstein and (b) trans-D model parameterizations. The red stars are the synthetic data (with added errors), and the blue dots are the mean predicted data (with two standard deviation error bars) from the MCMC samples.

(with added errors) and the blue dots with error bars are the mean predicted dispersion data (± 2 standard deviations) from the MCMC samples. Both the Bernstein polynomial and trans-D models provide a good fit to the dispersion data, with similar variance in dispersion data produced for models in the Markov chains. The main results for the inversion are considered in terms of marginal probability profiles for V_S (the marginal probability profiles for V_P/V_S are not shown due to the lack of sensitivity of V_P/V_S to the data) using the Bernstein polynomial model parameterization and the trans-D parameterization (Figure 2.5). The solid lines represent the profile of V_S used to generate the synthetic dispersion data (i.e., the true model). The inversion results for V_S using the Bernstein-polynomial model are in good agreement with the true model, and the form of the profile (smooth, continuous function of depth above the half-space) is well represented. The trans-D inversion, shown in Figure 2.5(b), produces V_S values which are generally close to the true values; however, the recovered profile indicates discrete, uniform layering over the upper 30 m which differs significantly from the true

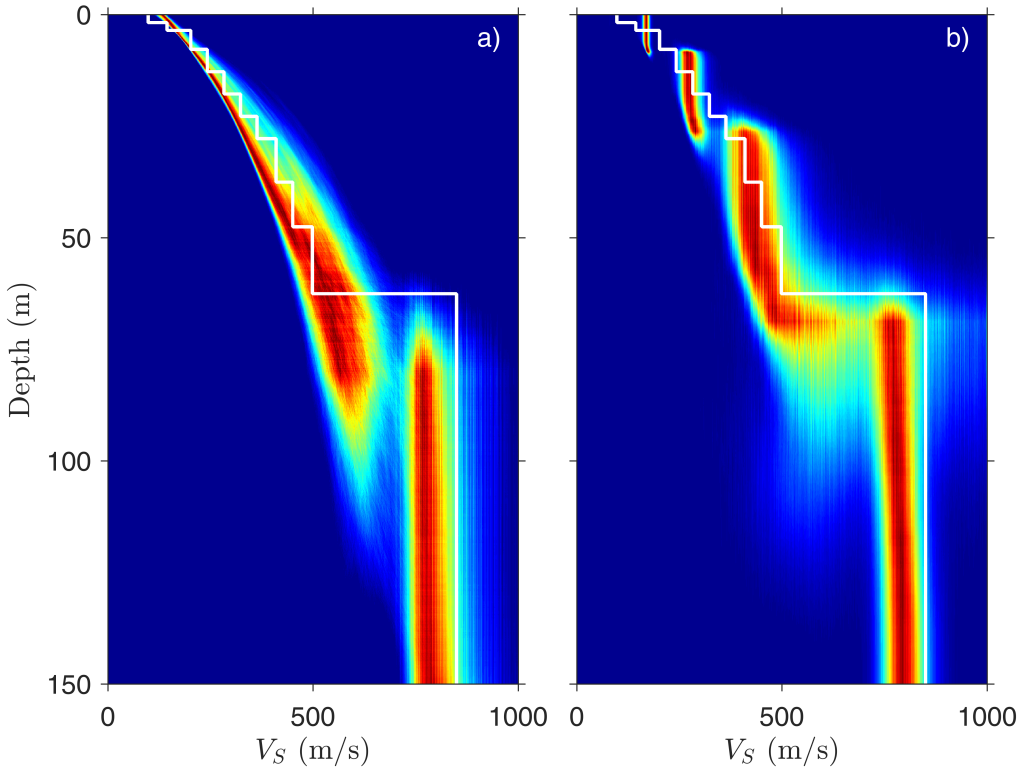

Figure 2.5: Marginal probability profiles for V_S from inversions of synthetic data using the Bernstein-polynomial model in (a), and trans-D inversion in (b).

smooth gradients. Such layering is not required to fit the data, but is a consequence of the choice of model parameterization. The trans-D approach has the added benefit of including the model complexity (number of layers) as an unknown in the inversion. As a result, the uncertainty of the parameterization is included within the estimated PPD. This property, combined with the fact that the trans-D inversion requires more parameters (thickness, V_S, and V_P/V_S values in up to 20 layers) than the Bernstein polynomial, are likely why the uncertainties appear larger in the trans-D inversion results. This synthetic case illustrates that a layered model parameterization may not be the ideal choice for seismic site response studies where accurate knowledge of the near-surface properties are of particular importance. However, the Bernstein-polynomial representation has limitations in that it suggests lower uncertainties due to a fixed parameterization, which can produce less-conservative uncertainty estimates for seismic hazard assessments. Furthermore, this parametrization can potentially miss sharp transitions at sites where structure is unknown and V_S discontinuities exist.

2.5 Fraser River Delta dispersion data

This section considers the inversion of Rayleigh-wave dispersion data extracted from passive array recordings from the Fraser River Delta near Vancouver, BC (Figure 2.6). The delta is composed of up to 300 m of unconsolidated Holocene deltaic sediments from the Fraser River overlying harder Pleistocene glacial material. The combination of the proximity to the northern extent of the Cascadia subduction zone, and the thick sequence of deltaic sediments, results in significant amplification and liquefaction hazards in this region (Hunter et al., 1998). The data were recorded on five portable broad-band three-component seismographs which were arranged in six array configurations with station separation distances ranging from 5 m to 180 m (inset of Figure 2.6(d)). The passive array recordings were processed according to the procedure described in Molnar et al. (2010) to produce 52 approximately logarithmically-spaced dispersion data between 1.2 and 6.5 Hz. Similar to Molnar et al. (2010), the dispersion data are segmented according to assumed changes in error processes related to different array configurations. Specifically, three frequency bands are specified with boundaries at 4.2 and 5.3 Hz. A different standard deviation is sampled implicitly, and a different AR parameter is sampled explicitly, within each frequency band as described is section 2.2.

This section applies Bayesian inversion with the Bernstein polynomial parameterization to these data, which have been previously considered by Molnar et al. (2010) and Dettmer et al. (2012). Model parameter prior bounds are listed in table 2.2. These data are examined in this chapter as they represent a high-quality observed dataset suitable for comparison with previous inversion methodologies, as well as for comparison with co-located borehole and SCPT measurements. Molnar et al. (2010) considered these data by applying Bayesian inversion with linear and power-law model parameterizations to represent gradient structures in the soils/sediments. The BIC was used to identify the power-law model as the preferred parameterization. However, the power-law parameterization has limited flexibility in representing gradient structures, and this approach forces the geophysical profiles to be power laws even if the dispersion data can resolve deviations from such structure. The Bernstein-polynomial parameterization is much more general such that the form of the gradients are determined by the data, rather than by subjective parameterization choice. Dettmer et al. (2012) also considered these data by applying a trans-D Bayesian inversion such that, in every dimension, the model is represented by a discontinuous stack of uniform layers. How-

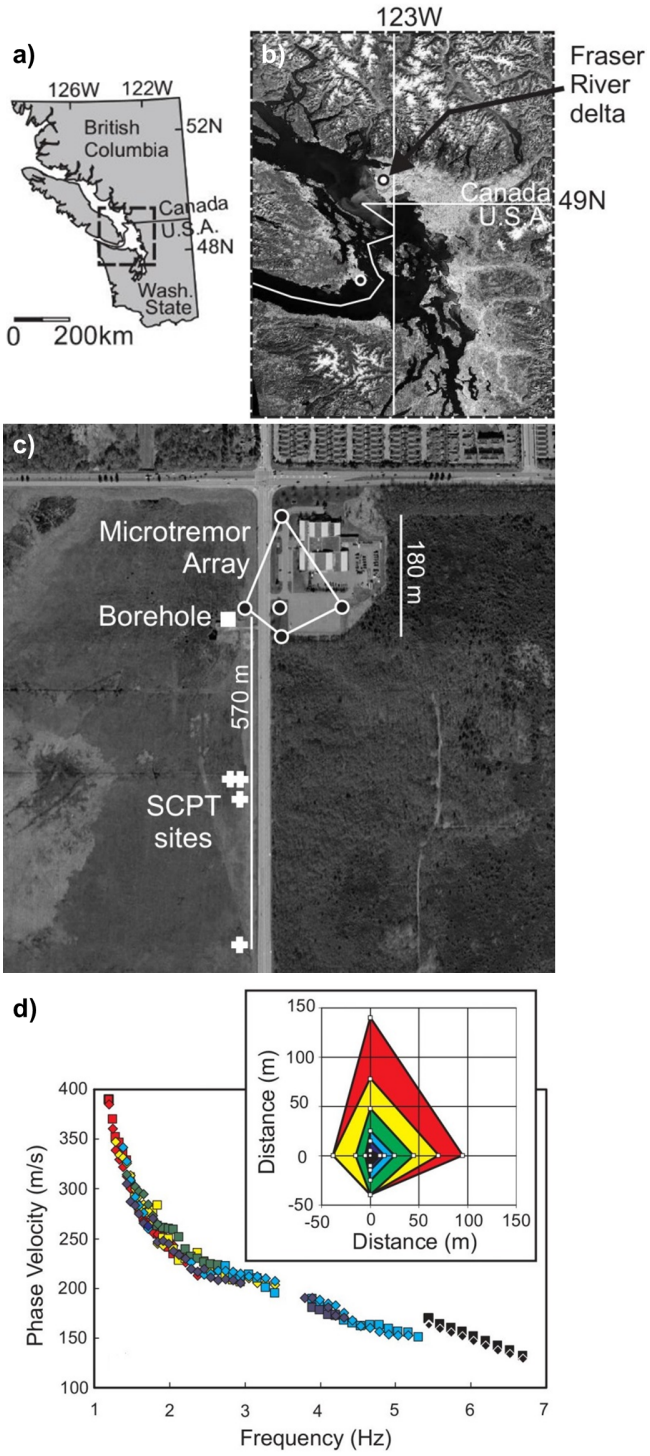

Figure 2.6: (a) and (b) Location of Fraser River Delta passive array site. (c) Array configurations and (d) processed dispersion data, with corresponding colour scales (from Molnar et al., 2010).

Parameter	min	max
g^{V_S} (m/s)	50	800
g^{V_P/V_S}	1.4	3
z_0 (m)	20	200
Half-space V_S (m/s)	500	1500
Half-space V_P/V_S	1.4	3
a	0	0.9

Table 2.2: Prior bounds for model parameters in inversion of Fraser River Delta dispersion data.

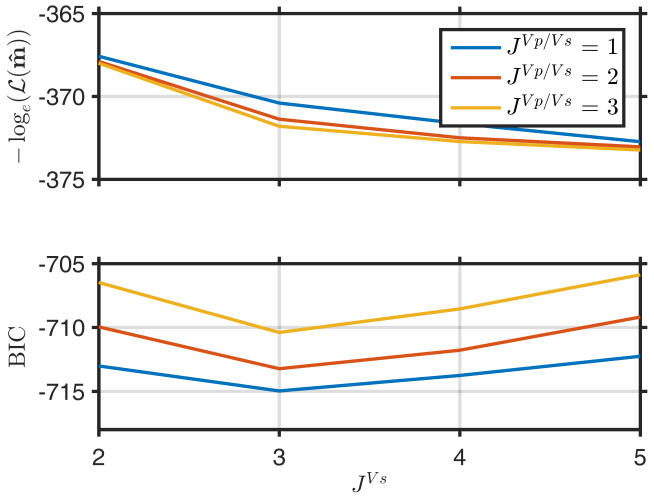

Figure 2.7: BIC analysis for Fraser River Delta dispersion data. The likelihood of the MAP model $\mathcal{L}(\hat{\mathbf{m}})$ is also shown as a function of Bernstein-polynomial order.

ever, discrete uniform layering structure is not expected at this site as V_S in sediments increases with total vertical stress, even for constant lithology (Budhu, 2007).

Figure 2.7 shows the results of the BIC analysis performed for the Bernstein polynomial parameterization. The BIC is a minimum for $J^{V_S} = 3$ and $J^{V_P/V_S} = 1$ and so the final inversion results are shown for this parameterization. Figure 2.8(a) shows the marginal probability profile for V_S, which represent the main results of the inversion (the marginal probability profile for V_P/V_S is not shown due to its lack of sensitivity to the data). Figure 2.8(b) and 2.8(c) show the marginal probability profiles for inversion results from Dettmer et al. (2012) and Molnar et al. (2010), respectively. The V_S marginal profile from the Bernstein-polynomial inversion shows clear depth-dependent

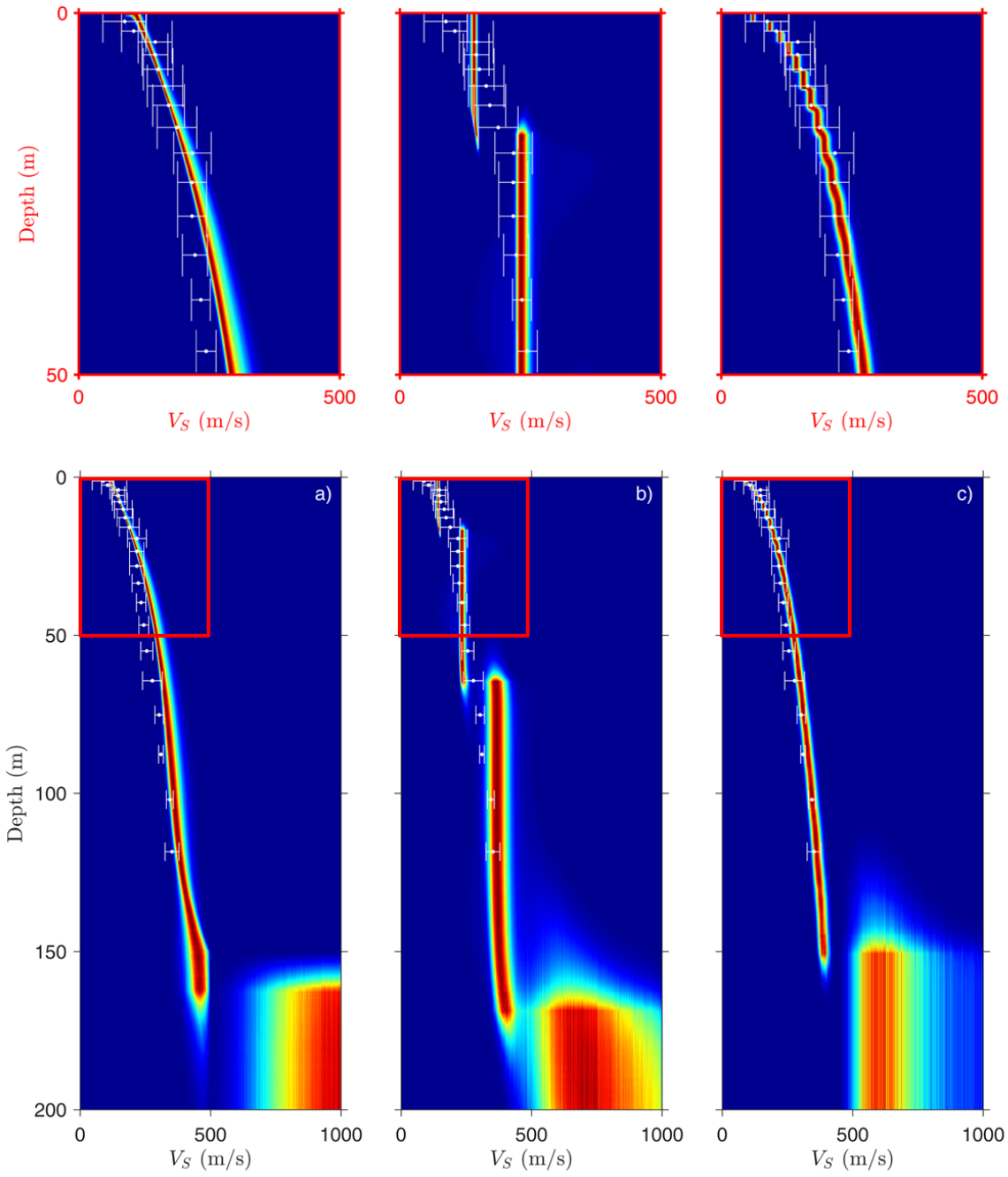

Figure 2.8: Marginal probability profiles for V_S from Fraser River Delta dispersion data using Bernstein model parameterizations (a), as well as trans-D inversion results from Dettmer et al. (2012) (b), and inversion results using a power-law model from Molnar et al. (2010) (c). The top panels show the near-surface V_S marginal profiles for the range delineated in red in the main panels. The white dots are the mean (with one standard deviation error bars) of the available SCPT and borehole measurements (Dallimore et al., 1995; Hunter et al., 1998; Molnar et al., 2010).

structure, with velocities increasing with depth, particularly in the near surface. The Bernstein-polynomial inversion produces a smooth, continuous gradient that is similar to a power-law (with some deviation from this fixed functional form). In particular, the Bernstein-polynomial inversion allows for V_S to increase at depths greater than 120 m, where a power-law does not. The transition to the half-space is estimated to occur between 150 m and 180 m. Uncertainty in the V_S structure generally increases with depth, as is expected from the decaying resolution of surface-wave dispersion data with depth. The dispersion inversion results are compared to invasive V_S measurements from a borehole and SCPT data collected near the passive array (Figure 2.6(c)). The white dots with error bars are the mean of the available SCPT and borehole measurements (± 1 standard deviation) (Dallimore et al., 1995; Hunter et al., 1998; Molnar et al., 2010). The inversion results are in good agreement with the invasive V_S estimates with detailed structure generally captured within the uncertainties in the inversion results. It should be noted that the inversion results represent an average solution over the spatial extent of the passive array. The V_S values estimated from invasive measurements represent an average of point estimates that are spatially distributed (with the greatest separation between measurements being 570 m) and not perfectly co-located with the passive array. V_S values estimated from invasive measurements also have their own associated uncertainty, which are not included in the error bars in Figure 2.8.

The inversion results from Molnar et al. (2010) indicate near-surface V_S of 70-90 m/s increasing to 325-425 m/s at the base of the modelled power-law layer, which is estimated to be 110-160 m thick overlaying a half-space. The inversion results from Dettmer et al. (2012) indicate the V_S profile is composed of three layers with discontinuities at approximately 17 m and 65 m depth, overlaying a half-space which begins at 140-170 m depth. Velocities in the near-surface layers are approximately constant with depth and are consistent with average values over the corresponding depth ranges from results from Molnar et al. (2010) as well as results from this study. There is no geological explanation for the sharp velocity discontinuities at 17 m and 65 m depth, other than to model the increase in V_S with depth. With all parameterization approaches, the depth and velocities in the half-space are poorly resolved with uncertainties of tens of metres and hundreds of metres per second, respectively. Depending on the geological conditions, the depth resolution limit for dispersion data processed from passive array recordings is typically taken to be on the order of the maximum station separation distance which in this case is 180 m (Park et al., 1999; Wathelet et al., 2008). It is possible (likely) that the transition to the half-space does not represent a physical

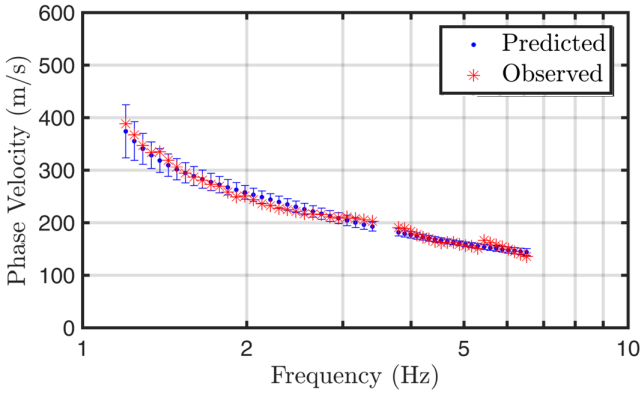

Figure 2.9: Data fit to Fraser River Delta dispersion data. The red stars are the observed dispersion data, and the blue dots are the mean predicted data (with two standard deviation error bars) from the MCMC samples.

discontinuity to a homogeneous layer, but instead marks the resolution limit of the dispersion data (Molnar et al., 2010). Results from this work are generally in good agreement with previous studies. However, the results shown here allow the gradient form of the results to be determined by the data, rather than by subjective parameterization choice. In particular, the Trans-D inversion results do not agree with the invasive measurements as well as the Bernstein-polynomial, or Power-law, inversion results and is not the best type of model for this application. The invasive measurements indicate that a power-law model is a good parameterization choice in this case as the power-law inversion results appear to fit the invasive measurements relatively better than Bernstein-polynomial inversion results, due to sharper curvature in the V_S profile near the surface. This indicates that the prior information built into the functional form of the power-law model is likely correct. However, the Bernstein-polynomial inversion results have less near-surface curvature because the data are not able to resolve the sharper near-surface gradient. This is most likely due to the lack of high-frequency dispersion data (> 6.5 Hz), which constrain shallow V_S structure.

Figure 2.9 shows a good fit to the dispersion data from the Bernstein-polynomial inversion results (in fact, all inversions fit the data in a similar manner), with reasonable variance for the data predictions (plotted as phase velocities, although slownesses were inverted). To validate the assumptions for the error process, statistical tests are applied to the standardized residuals of the MAP model for the Bernstein-polynomial inversion. Figure 2.10 shows the histogram and autocorrelation function of the standardized residuals of the MAP model. If the error process is Gaussian, then the histogram of

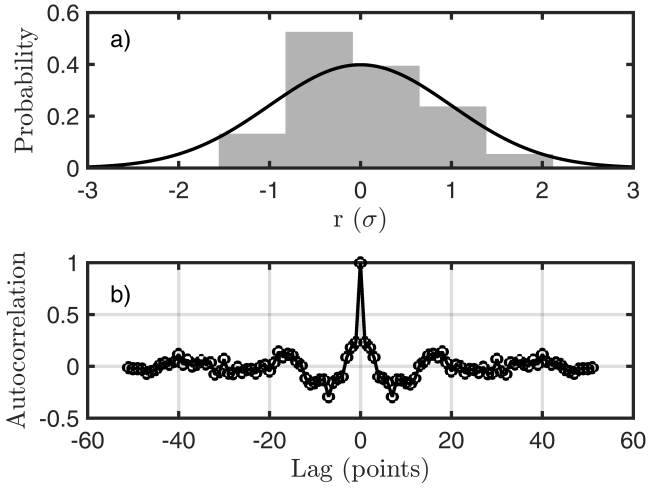

Figure 2.10: (a) Histogram and (b) autocorrelation of standardized residuals from the MAP model. The solid line in (a) is the standard normal distribution.

the standardized residuals will approximate the standard normal Gaussian (the solid line in Figure 2.10(a)). The Kolmogorov-Smirnov (KS) test was applied to quantitatively assess the validity of this assumption and provided a p-value of 0.26 for the null hypothesis of Gaussian-distributed errors (Freund, 1967; Dosso et al., 2006). The width of the central peak of the autocorrelation of the standardized residuals (Figure 2.10(b)) gives an indication of the correlation in data errors. Ideally, the AR model will account for correlation and the resulting standardized residuals will be uncorrelated. The runs test was applied to quantitatively assess the validity of this assumption and provided a p-value of 0.12 for the null hypothesis of uncorrelated errors (Freund, 1967; Dosso et al., 2006). In this case, neither null hypothesis is rejected (using the typical threshold of $p < 0.05$ for rejection) and the data error assumptions in the inversion appear justified.

2.6 Summary

This chapter considered Bayesian inversion of surface-wave dispersion data using a general and efficient approach to parameterizing smooth gradient-based profiles in terms of a Bernstein-polynomial representation. The Bernstein-polynomial parameterization may be better suited than parameterizations based on discrete layers (including trans-dimensional parameterizations) for nonlinear inversions when general smooth gradients

in geophysical parameters are expected, such as surface-wave dispersion. Further, the Bernstein approach is more general than parameterizations based on a prior specification of the gradient type (e.g., a power law) in that the form of the gradient is determined by the data, rather than by a subjective parameterization choice. The Bernstein polynomial parameterization is defined as a weighted sum of Bernstein basis functions, where the basis-function coefficients are estimated to describe the earth structure. This chapter shows that this form of model parameterization is effective at resolving smooth, depth-dependent gradient structures and is a stable functional form, as small perturbations to basis-function coefficients result in only small, localized perturbations to the overall model. The Bernstein polynomial parameterization was implemented within a Bayesian approach which treats the model parameters as random variables constrained by prior information and observation. The solution to the inversion is described by properties of the PPD. Specifically, the marginal distribution profile of V_S, which was estimated numerically using MCMC sampling. This provides an estimate of the most probable V_S structure as well as a quantitative measure of its uncertainty. The BIC was applied to quantitatively determine the appropriate Bernstein polynomial order consistent with the resolving power of the data.

This methodology was applied to synthetic fundamental-mode Rayleigh-wave dispersion data. The data were also considered with a trans-D approach to compare the effects of the choice of model parameterization on the inversion results. This synthetic test illustrates how the use of a Bernstein polynomial parameterization effectively represents smooth gradients in near-surface soil/sediment properties while the trans-D approach introduces discontinuous layered structure which is undesirable in this application (both inversions fit the data to a similar level). The Bernstein-polynomial approach also better resolves near-surface low-V_S structure compared to trans-D inversion, which is the most critical in earthquake site response studies. The Bernstein polynomial parameterization was also applied to previously-considered dispersion data processed from passive array recordings collected on the Fraser River Delta in BC. Molnar et al. (2010) considered these data by applying Bayesian inversion with a power-law model that effectively models gradient structure. However, this model has limited flexibility and forces the geophysical profiles to be power laws even if the dispersion data are better fit by a different gradient structure. Dettmer et al. (2012) considered these data by applying a trans-D Bayesian inversion based on stacked homogeneous layers that includes non-physical discrete layering in the inversion results. Results from this work are in good agreement with co-located SCPT and borehole measurements, and ef-

fectively represent smooth, continuous V_S structure while allowing the gradient form of the V_S profile to be determined by the data, rather than by subjective parameterization choice.

Chapter 3

Seismic hazard site assessment in Kitimat, British Columbia, from Bayesian inversion of surface-wave dispersion

3.1 Background

Estimating site-specific seismic amplification and resonance hazards is important for engineers and planners in designing plans that can mitigate damage caused by earthquakes, protect critical infrastructure, and potentially save lives. These site-specific hazards depend predominantly on near-surface geophysical properties such as shear-wave velocity (V_S), compressional-wave velocity (V_P), and density (ρ). It is well known that seismic-wave amplitude increases as waves propagate through material of decreasing impedance, such as unconsolidated near-surface sediments (Anderson et al., 1986, 1996). Amplification can also occur at particular frequencies due to resonance within near-surface low-velocity layers. Seismic body-waves can also become trapped, focused, and converted into long-period surface waves due to three-dimensional (3-D) basin structures which can significantly amplify and extend the duration of earthquake shaking (Bard & Bouchon, 1980; Graves et al., 1998). For these reasons, estimating the near-surface structure of geophysical properties (particularly V_S) is critical for understanding and predicting site-specific seismic amplification and resonance hazards.

Currently, many site-response classifications and ground-motion prediction equa-

Site classification	Description	V_{S30} range (m/s)
A	Hard rock	> 1500
B	Rock	$760 - 1500$
C	Very dense soil and soft rock	$360 - 760$
D	Stiff soil	$180 - 360$
E	Soil with soft clay	< 180
F	Site-specific analysis required	

Table 3.1: V_{S30} site classification used in 2015 National Building Code of Canada (Humar, 2015).

tions (GMPE) rely on the travel-time average of V_S over the upper 30 m, known as V_{S30} (e.g., Boore & Atkinson, 2008; Seyhan & Stewart, 2014). By definition, V_{S30} is weighted more heavily toward low-velocity material, which generally cause seismic-wave amplification. The National Earthquake and Hazards Reduction Program (NEHRP) uses a classification scheme for site-specific seismic hazard according to six different site classes (A-F) based primarily on ranges of estimated or measured V_{S30} values (FEMA Provision 750, 2009). The National Building Code of Canada adopts a similar classification scheme (Table 3.1) and assigns site amplification factors of earthquake ground acceleration based on this classification (Finn & Wightman, 2003; NBCC, 2010; Humar, 2015). These amplification factors are used for design procedures and are incorporated in Canada-wide seismic hazard models (Adams et al., 2015). The V_{S30} classification scheme reduces the dependence of site amplification on earthquake properties and propagation effects by averaging over empirical amplification measurements from many earthquake recordings (Borcherdt, 1994). However, there is not universal agreement that V_{S30} is a valid proxy for seismic amplification, particularly for complex V_S profiles with impedance discontinuities within the upper 30 m (Castellaro et al., 2008; Molnar et al., 2013).

V_S profiles at a site can be determined by invasive methods such as boreholes or seismic cone penetration tests (SCPT). V_S profiles can also be inferred from non-invasive, active or passive seismic methods (as discussed in Chapter 2). Passive seismic methods have become increasingly popular due to their reduced demands in terms of cost and logistics. An attractive approach is to estimate the dispersion of fundamental-mode Rayleigh waves using array recordings of ambient seismic noise (Aki, 1957; Lacoss et al., 1969; Wathelet et al., 2008; Molnar et al., 2010). Dispersion occurs due to the

nature of propagation of Rayleigh waves, where lower-frequency waves are more sensitive to geophysical properties at greater depth than higher-frequency waves. Hence, the velocity-frequency relationship (dispersion curve) of Rayleigh waves can be inverted to estimate profiles of geophysical properties.

More costly, active-source seismic methods for estimating two-dimensional (2-D) V_S models include high-resolution, multi-component seismic reflection surveys (Pugin et al., 2009, 2013). Such surveys use a vibratory source and a dense linear array of three-component geophones. In near-surface studies (e.g., investigating groundwater or natural hazards) a 'Minivibe' truck capable of producing strong shear waves is typically used as a vibratory source. V_S values are estimated from these surveys using a modified Dix inversion which assumes homogeneous isotropic layers and flat horizontal reflectors in the subsurface, which may not be applicable in all environments (Dix, 1955; Pugin et al., 2009).

This chapter applies passive seismic array processing to estimate dispersion data at three locations in Kitimat, British Columbia (BC), Canada (Figure 3.1). The dispersion data are inverted using the methodology discussed in Chapter 2. Specifically, Bayesian inference is applied to probabilistically estimate parameters of a geophysical model using observed dispersion data and prior information. In the inversion, the geophysical model is parameterized as smooth, gradient-based profiles in terms of a Bernstein-polynomial representation (Farouki & Rajan, 1989; Farouki, 2012; Quijano et al., 2016). The most-probable V_S profile as well as quantitative measures of the profile uncertainty are determined within the Bayesian framework. The inversion results are considered here in terms of marginal probability profiles for V_S (and V_P/V_S) which are estimated numerically using Markov-chain Monte Carlo (MCMC) sampling of the posterior probability density (PPD) of the model parameters (Mosegaard & Tarantola, 1995; Brooks et al., 2011). Inversion results for V_S are compared to estimates from an active-source seismic reflection survey conducted by the Geological Survey of Canada (GSC) (Pugin et al., 2016). From the V_S marginal probability profiles, probability densities for the travel-time average of V_S to an arbitrary depth z (called V_{SZ} here) is computed as a representation of the average soil/sediment conditions at each site (Molnar et al., 2013). In addition, probability densities of V_{S30}, and site amplification factors dependent on V_{S30}, are computed to provide estimates (with quantitative uncertainties) of the seismic-hazard site classification and predicted site amplification at the three Kitimat sites. Lastly, the V_S marginal density profiles are used to compute probabilistic amplification and resonance spectra for horizontally-polarized shear (S_H)

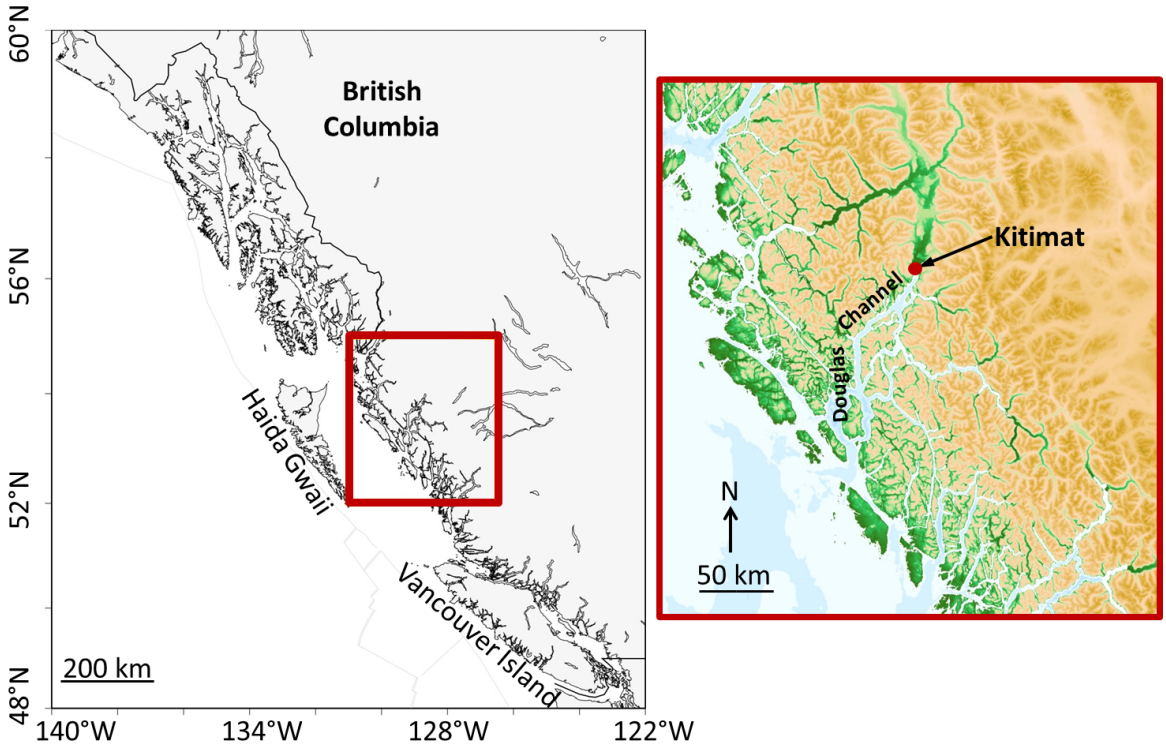

Figure 3.1: Location of Kitimat, BC, Canada.

waves (Molnar et al., 2013).

The town of Kitimat is located on the farthest inland reach of one of the largest fjords on the BC coast, the Douglas Channel (Figure 3.1). The Douglas Channel and its inland expression, which extends over 90 km northward, form a single continuous glacial valley (Dolmage, 1956). Bedrock geology in the region consists mainly of granitic rocks of the Canadian Cordillera Coast Mountain Batholith, consistent with bedrock found along the BC coast from Vancouver to Alaska (Dolmage, 1956). Thick deposits of glacial till overlie bedrock and were left behind during deglaciation of the Cordilleran ice sheet in the late Pleistocene (Bornhold, 1983; Clague, 1984). As glaciers retreated, marine and deltaic sediments were deposited on the isostatically-depressed coastal region before sea level dropped to present day levels. The region is also overlain with Holocene deltaic and surficial floodplain sediments from the Kitimat river (Dolmage, 1956; Clague, 1977). Past drillings southwest of the Municipal District of Kitimat indicate sediment thicknesses can exceed 100 m (Dolmage, 1956), while studies within the Douglas Channel indicate Holocene sediments are up to 60 m thick and overlie several hundred metres of older glacial material (Bornhold, 1983).

There exists a relative gap in the knowledge of the seismicity of the region, compared to coastal areas to the north and south, due to a lack of seismic instrumentation. Studies since the installation of new instruments in the region in 2014 do not indicate any previously unknown areas of seismicity (Brillon, 2016b). Recent studies have identified a fault-like structure in the Douglas Channel which might have triggered submarine landslides during the early Holocene, suggesting the fault is active (Conway et al., 2012). Of more obvious concern, Kitimat is located within 300 km of Haida Gwaii and the Queen Charlotte Fault zone, which is one of the most seismically-active regions in Canada and site of the two largest instrumentally-recorded earthquakes in Canadian history with magnitudes of 8.1 and 7.8 in 1949 and 2012, respectively (Bostwick, 1984; Cassidy et al., 2014). The 2012 event was felt as far away as Alberta and the Yukon (over 1000 km distance).

The GSC undertakes studies of natural geological hazards such as earthquakes, slope failures, and tsunamis within Canada. The Kitimat region and surrounding shoreline are the sites for several proposed large-scale industrial development projects including export facilities for heavy oil and natural gas. As such, the region has great environmental and economic significance for Canada, and the GSC is currently investigating geological hazards in the region (e.g., Conway et al., 2012; Crow et al., 2015; Brillon et al., 2015; Brillon, 2016a,b; Pugin et al., 2016). The findings from this research will be useful for the GSC as well as for planners and regulatory agencies to mitigate the hazards caused by amplification and resonance for the environment, the community, and critical infrastructure.

3.2 Field work and data processing

This section describes the field work and data processing procedures used to estimate dispersion data from passive seismic array recordings at three sites in Kitimat, BC (Figure 3.2). The locations for the array recordings were chosen based on a combination of factors including sufficient space to carry out the survey, relatively flat topography, proximity to potential noise sources, and accessibility. At each site, a cross-shaped array of five seismic stations was deployed in four different array sizes with inter-station distances (distance to the centre station) ranging from 5 m to 45 m. Figure 3.3 shows a photograph of a single instrument installation. Each station consists of a three-component broadband seismometer connected to a digitizer and recorder, a battery power supply, and a GPS antenna for synchronous timing. To deploy the

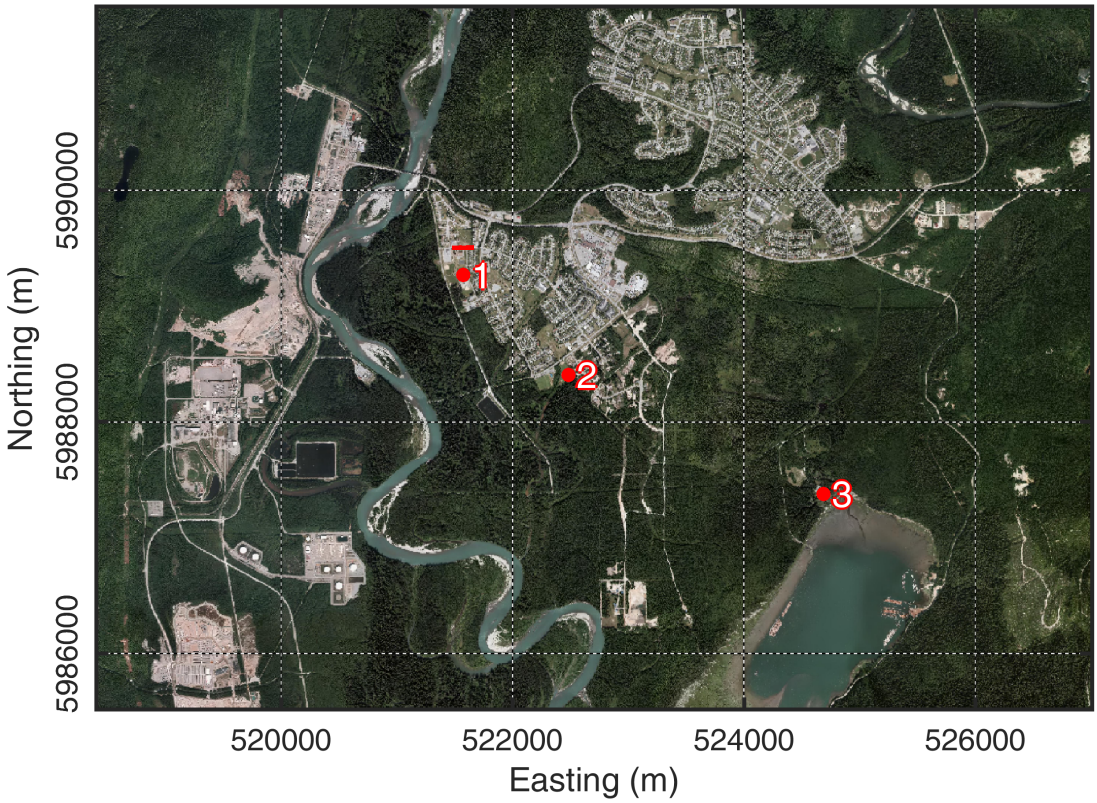

Figure 3.2: Locations of the three array sites in Kitimat, BC (air photo downloaded from KITI-MAP, 2016). The red line north of site 1 shows the location of the seismic-reflection survey from Pugin et al. (2016) used for comparison to inversion results.

seismometers, the surficial layer of vegetation (grass, moss, etc.) was removed and a ceramic tile laid down to achieve better coupling to the ground surface. Approximately one hour of ambient seismic noise was recorded for each array configuration. The multiple array configurations (with different station separations) allow for a relatively wide dispersion frequency band to be estimated, as described below.

This study applies frequency-wavenumber (f-k) processing to estimate the slowness (reciprocal of phase velocity) of the fundamental-mode Rayleigh wave as a function of frequency from passive seismic array recordings (Lacoss et al., 1969). Only vertical-component recordings are used here, such that ambient-noise recordings are assumed to be dominated by Rayleigh waves. In f-k processing, the coherence of the signal across the array is estimated for various phase velocities and azimuths, where individual station recordings are shifted in time according to a particular wavenumber vector $\mathbf{k} = (k_x, k_y)$. The coherence of the signal is estimated over a specific frequnecy band

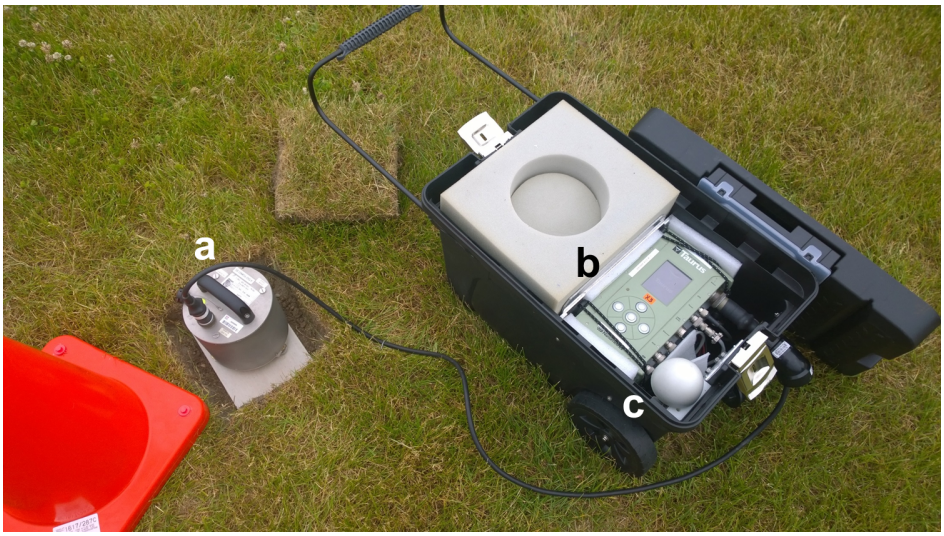

Figure 3.3: Seismic station installation including: three-component broadband seismometer (a), digitizer and recorder (b), and GPS antenna (c).

$[f_l, f_h]$ from the semblance coefficient by

$$S(f, \mathbf{k}) = \frac{\sum_{f=f_l}^{f=f_h} \left[\sum_{i=1}^{i=N} X_i(f) \exp(j\mathbf{k} \cdot \mathbf{r}_i) \right]^2}{N \sum_{f=f_l}^{f=f_h} \sum_{i=1}^{i=N} \left[X_i(f) \exp(j\mathbf{k} \cdot \mathbf{r}_i) \right]^2}, \tag{3.1}$$

where N is the number of stations in the array, $X_i(f)$ is the recording at station i at frequency f, $\mathbf{r}_i$ is the vector of coordinates for station i, and j is the imaginary unit (Neiddell & Taner, 1971; Douze & Laster, 1979). In this way, $\exp(j\mathbf{k} \cdot \mathbf{r}_i)$ is the Fourier-domain equivalent to the time delay at station i for a harmonic plane Rayleigh wave propagating with phase velocity c and azimuth θ given by

$$c = \frac{2\pi f}{|\mathbf{k}|}, \tag{3.2}$$

$$\theta = \arctan\left(\frac{k_x}{k_y}\right). \tag{3.3}$$

Each frequency of interest is considered individually in f-k processing. The ambient noise recordings are time windowed with window lengths of 40 times the period of interest (the reciprocal of the frequency of interest), with 10 percent overlap between windows. For a given time window, the semblance coefficient is computed over the frequency band $[f_l, f_h] = [0.9f_c, 1.1f_c]$, where f_c is the centre frequency (frequency of interest). A brute-force method is used to search for the maximum coherent signal on a

2-D grid of wavenumber vectors. An estimate of phase velocity is then computed from the wavenumber vector of the maximum coherent signal (according to eq. 3.2). The median value of the set of estimated phase velocities (from the set of time windows) is taken as the observed dispersion datum (the median is used as the distributions of estimated phase velocities are often asymmetrical). The process is repeated for all frequencies of interest.

Dispersion data were computed here for 50 logarithmically-spaced frequencies between 1 and 12 Hz for all array configurations, at every site. However, only the data within a specific wavenumber band $[|\mathbf{k_{min}}|, |\mathbf{k_{max}}|]$ (defined by the resolution and aliasing limits of the array) are kept. These limits depend on the inter-station distances and spatial distribution of the stations and are estimated here using the methodology described by Wathelet et al. (2008). Specifically, the resolution and aliasing limits of the array are estimated by calculating the theoretical array response function according to Woods & Lintz (1973) and Asten & Henstridge (1984)

$$R(\mathbf{k}) = \frac{1}{N^2} \left| \sum_{i=1}^{N} \exp(-j\mathbf{k} \cdot \mathbf{r}_i) \right|^2 . \tag{3.4}$$

Figure 3.4 illustrates the process for estimating the wavenumber limits for an array configuration with 5 m station separation. The minimum wavenumber $|\mathbf{k_{min}}|$ is estimated from the half-peak of the centre lobe of the array response function. When the centre lobe is narrow, $|\mathbf{k_{min}}|$ is small and the array is capable of distinguishing waves travelling at similar wavenumbers. The maximum wavenumber $|\mathbf{k_{max}}|$ is estimated where side lobes in the array response function are equivalent to the half-peak of the centre lobe. Aliasing occurs for wavenumbers greater than $|\mathbf{k_{max}}|$ due to the summation of these side lobes. The use of multiple array configurations, with varying station separations, produces dispersion data over a greater wavenumber band (and, consequently, a greater frequency band). The wavenumber limits for each array configuration are listed in Table 3.2. Figures 3.5, 3.6, and 3.7 show the array configurations and processed dispersion data for Kitimat sites 1, 2, and 3, respectively. The dispersion data of a particular colour were processed from the array configuration of the corresponding colour (the centre station being common to all array configurations). The red line in Figure 3.5 represents a section of a multi-component seismic-reflection survey used for comparison of inversion results at site 1.

Inter-station distance (m)	$\lvert\mathbf{k_{min}}\rvert$ (rad/m)	$\lvert\mathbf{k_{max}}\rvert$ (rad/m)
45	0.0325	0.1103
20	0.0689	0.2493
10	0.1343	0.4987
5	0.2645	0.9973

Table 3.2: Wavenumber limits for all array configurations.

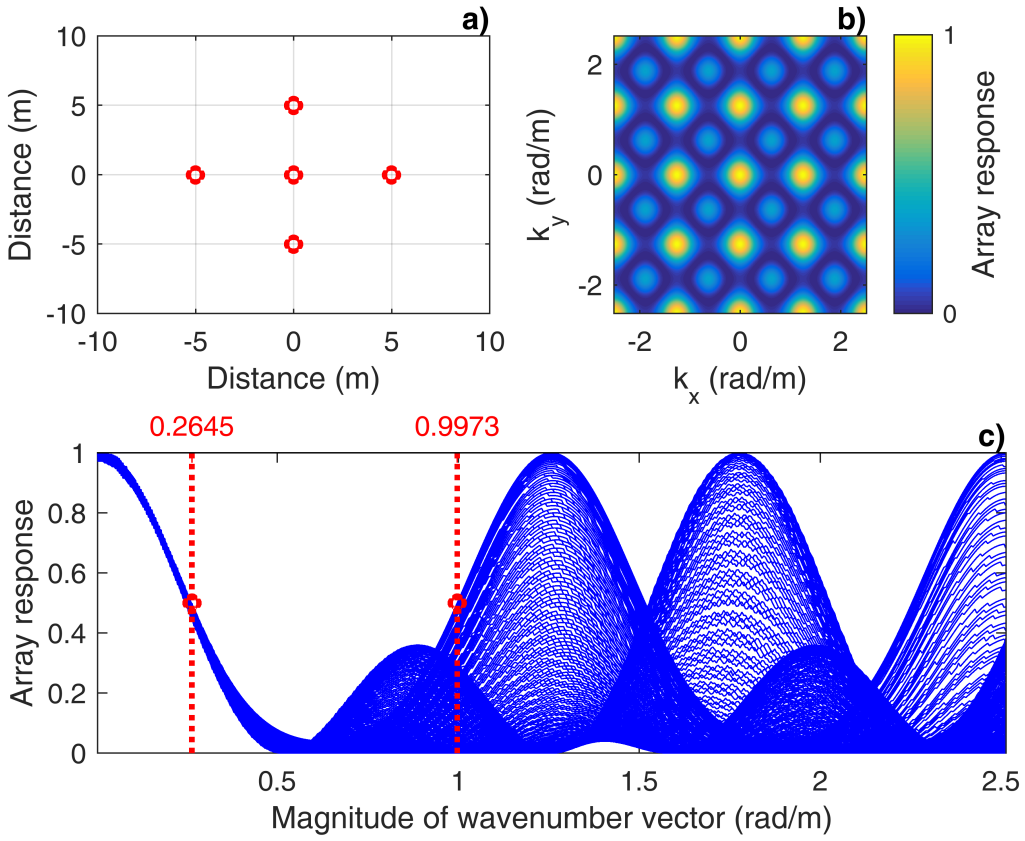

Figure 3.4: Array resolution and aliasing limits. (a) Station configuration with 5 m station separation, (b) theoretical array response function calculated on a 2-D wavenumber grid, and (c) theoretical array response as a function of magnitude of wavenumber vector. Estimated wavenumber limits are shown by the red dotted lines in (c).

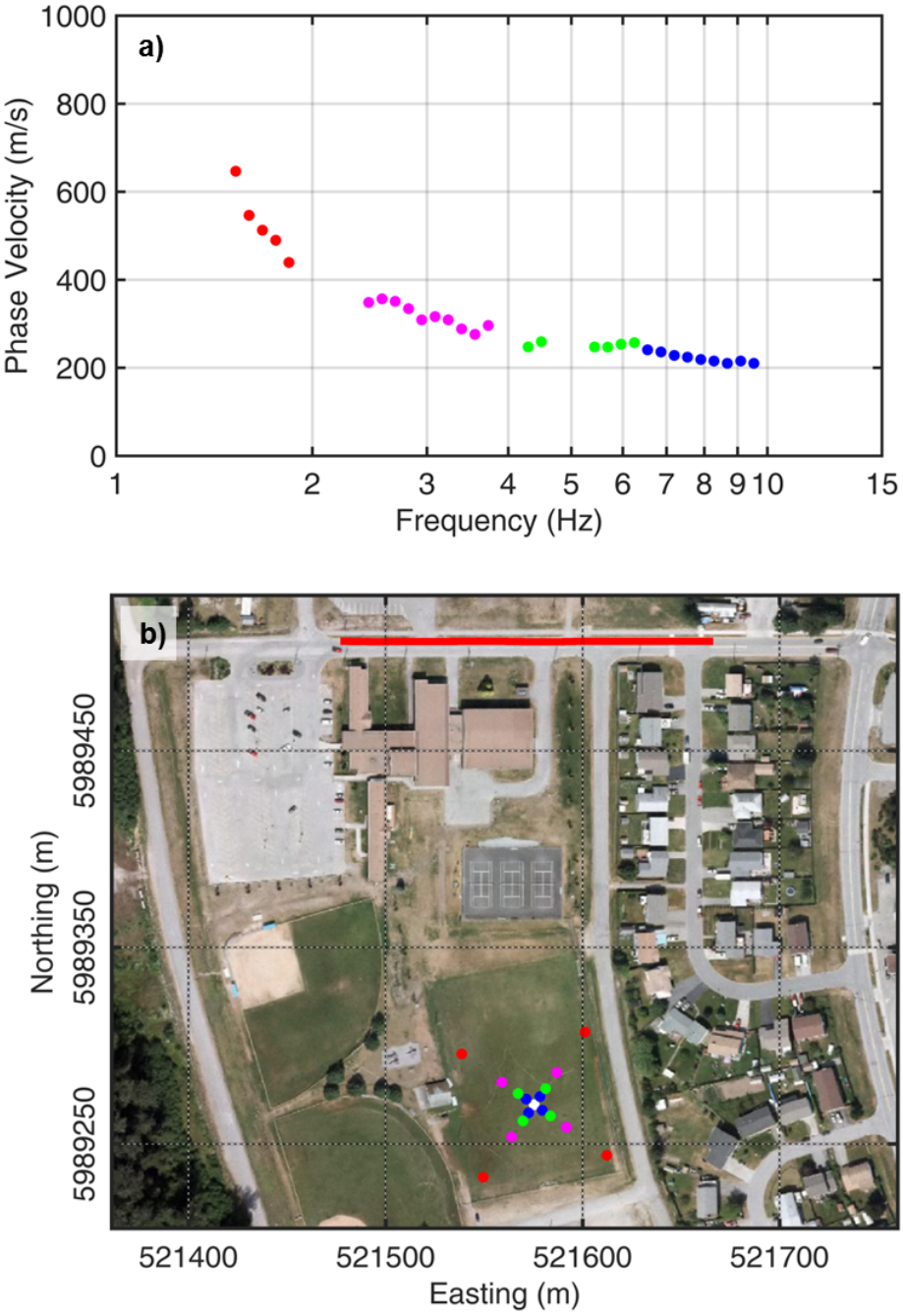

Figure 3.5: (a) Processed dispersion data, and (b) array configurations with corresponding colours, at Kitimat site 1 (air photo downloaded from KITI-MAP, 2016). The red line shows the location of the active-source seismic-reflection survey from Pugin et al. (2016) used for comparison to site 1 results.

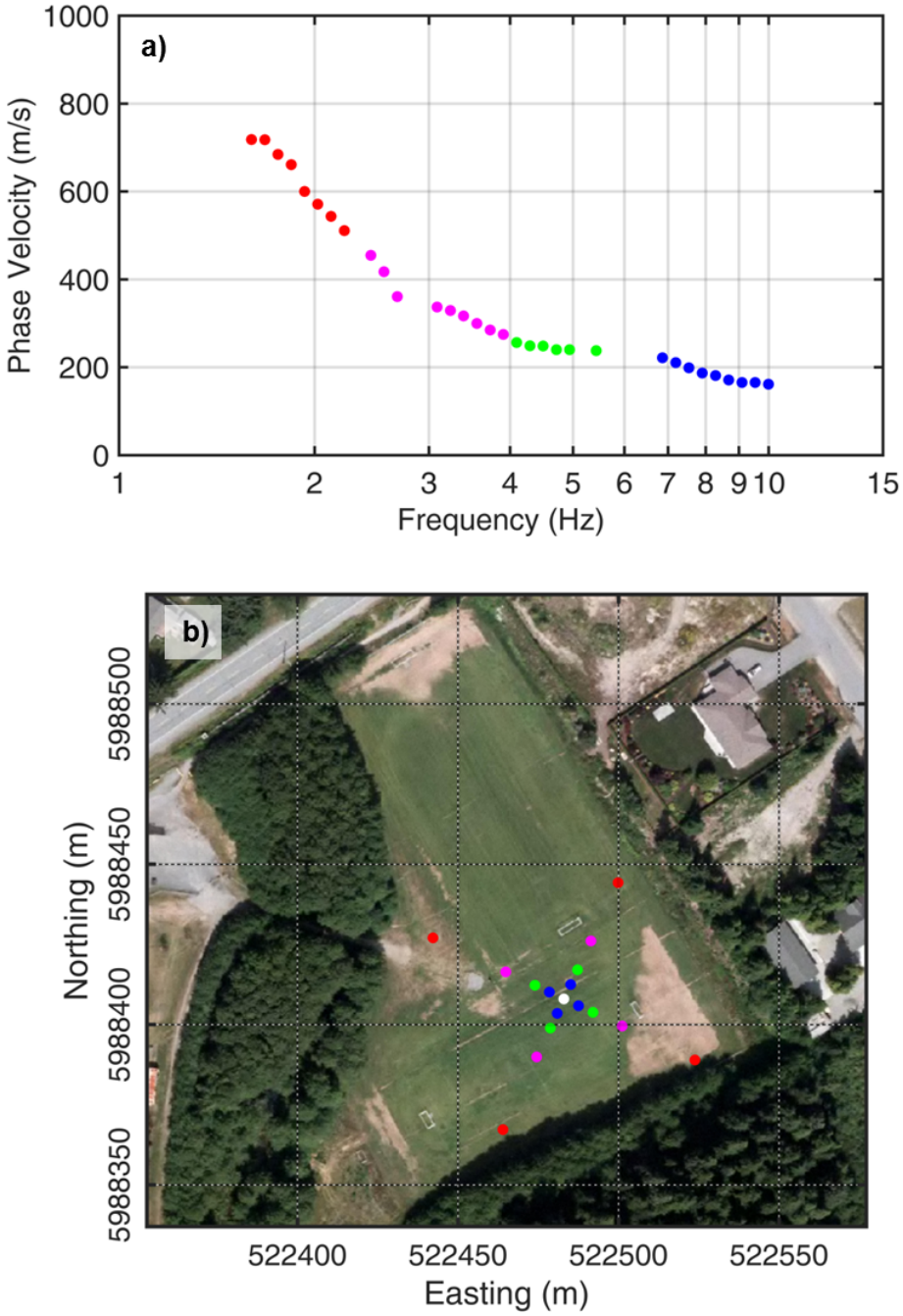

Figure 3.6: (a) Processed dispersion data, and (b) array configurations with corresponding colours, at Kitimat site 2 (air photo downloaded from KITI-MAP, 2016).

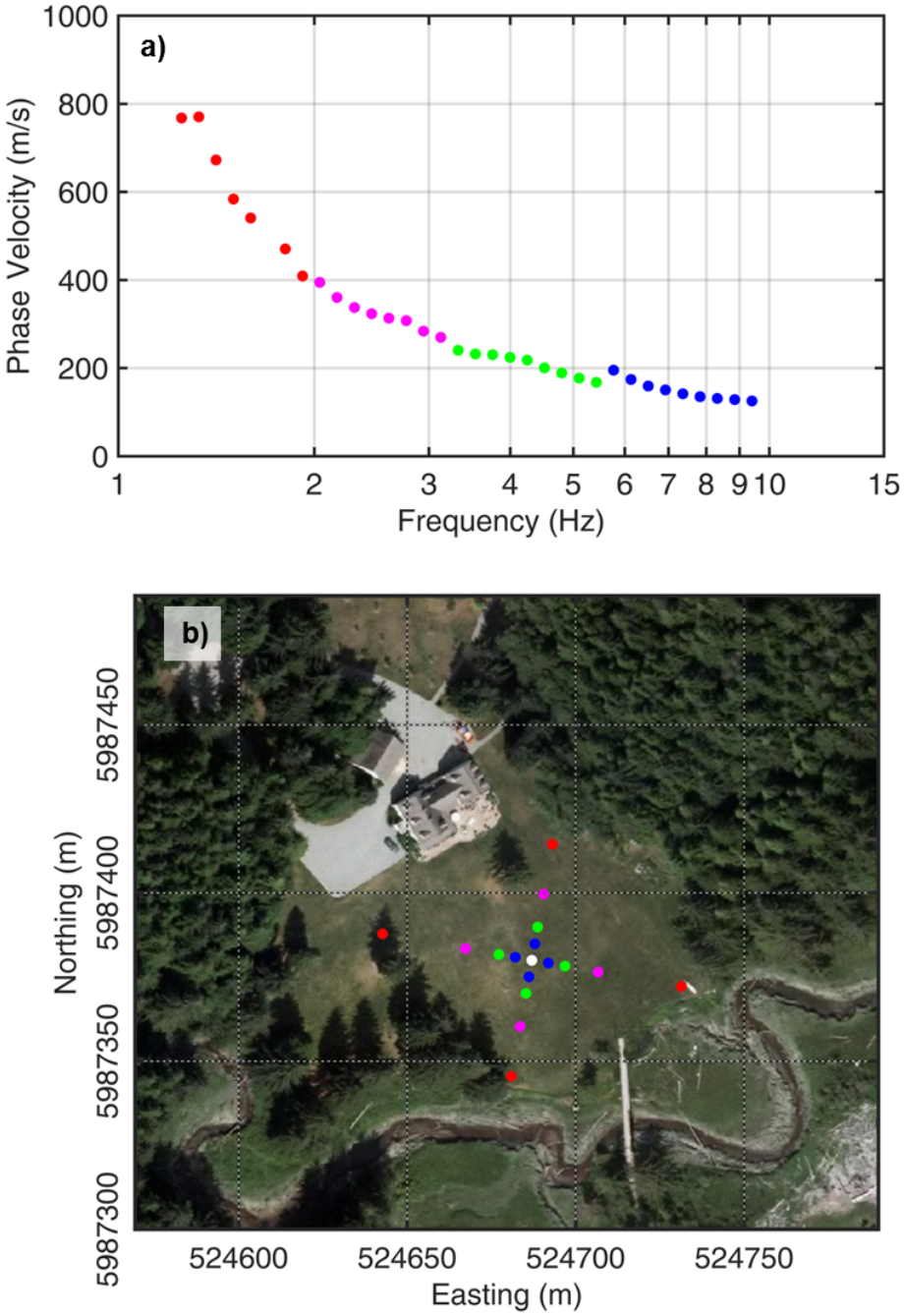

Figure 3.7: (a) Processed dispersion data, and (b) array configurations with corresponding colours, at Kitimat site 3 (air photo downloaded from KITI-MAP, 2016).

3.3 Inversion of surface-wave dispersion data

This section considers the inversion of Rayleigh-wave dispersion data extracted from passive array recordings from three sites in Kitimat, BC (discussed in section 3.2). The dispersion data are inverted using the methodology described in the previous chapter (sections 2.2 and 2.3). Specifically, a nonlinear Bayesian inversion method is used to estimate the maximum *a posteriori* (MAP) V_S and V_P/V_S profiles at each site, as well as marginal probability densities. These are estimated from samples of the PPD of the parameters that describe the earth structure model. The V_S and V_P/V_S profiles are parameterized in terms of a Bernstein-polynomial representation (overlaying a half-space). The Bernstein polynomial parameterization is defined as a weighted sum of Bernstein basis functions, where the basis-function coefficients are estimated in order to describe the V_S and V_P/V_S profiles (Quijano et al., 2016). This representation is a general and efficient approach to parameterizing smooth gradient-based profiles, and is considered appropriate here as gradient-based profiles are expected in the sediment sequence because V_S increases with total vertical stress, even for constant lithology (Budhu, 2007). This is also believed to be consistent with the resolution of surface-wave dispersion data, which are not sensitive to fine-scale structure.

The PPD of the earth structure model (basis-function coefficients) is sampled numerically using the MCMC method of Metropolis-Hastings sampling (Metropolis et al., 1953; Hastings, 1970; Brooks et al., 2011). Parallel tempering is applied to provide wide and efficient sampling within the inversion (Geyer, 1991; Dosso et al., 2012; Sambridge, 2014). The prior probability density of the basis-function coefficients are bounded uniform distributions over geophysically-realistic values. Wide prior bounds are applied to allow the data information to primarily determine the solution. Prior bounds for model parameters are summarized in Table 3.3. The Bernstein basis functions are designed such that they sum to unity at all depths. As a result, the width of the prior bounds on the coefficients spans the equivalent width of the prior for the V_S profile (with a slighty altered density, as discussed in section 2.3). Data errors are assumed to be Gaussian-distributed with variance sampled implicitly within the inversion (Dosso & Wilmut, 2005). Data error correlations are considered by applying a first-order auto-regressive (AR) model, where the AR coefficient is sampled explicitly within the inversion (Dettmer et al., 2012; Steininger, 2013). The data are treated in terms of slowness (the reciprocal of phase velocity) because the f-k processing discussed in section 3.2 estimates the wavenumber of propagating Rayleigh waves, and errors in

Parameter	min	max
V_S Bernstein-polynomial coefficients (m/s)	50	800
V_P/V_S Bernstein-polynomial coefficients	1.4	3
Half-space transition depth z_0 (m)	20	150
Half-space V_S (m/s)	500	1500
Half-space V_P/V_S	1.4	3
AR data-error coefficient	0	0.9

Table 3.3: Prior bounds for model parameters in inversion of Kitimat dispersion data.

slowness scale linearly with errors in the estimated wavenumber. For data errors with the same variance, this has the effect of increasing the phase-velocity errors at low frequencies, and is believed to be a physically-reasonable representation of the frequency dependence of the data errors. Due to the limited number of dispersion data, a single variance and AR process is used for the entire slowness dispersion curve (unlike the inversion discussed in section 2.5). The calculation of predicted dispersion data for a candidate earth model is performed using the software developed by Wathelet (2005). This software requires density profiles as an input, which are computed directly from an empirical relationship with V_P (Gardner et al., 1974).

The inversion is performed for several Bernstein-polynomial orders given by J^{V_S} and J^{V_P/V_S} for the V_S and V_P/V_S profiles, respectively. The Bayesian information criterion (BIC) is computed for the maximum-likelihood model in the Markov chains for each inversion to determine the simplest model consistent with the resolving power of the data (Schwartz, 1978; Kass & Raftery, 1995). Given the uniform prior densities, the maximum-likelihood model corresponds to the MAP model $\hat{\mathbf{m}}$ (most probable set of model parameters in the Markov chain). The optimal polynomial order is the one that minimizes the BIC, and so the final inversion results are shown for this parameterization. Figure 3.8 shows the results of the BIC analysis performed for the dispersion data from the Kitimat sites. The BIC is a minimum for $J^{V_S} = 2$ and $J^{V_P/V_S} = 1$ for sites 1 and 3, and is a minimum for $J^{V_S} = 3$ and $J^{V_P/V_S} = 1$ for site 2.

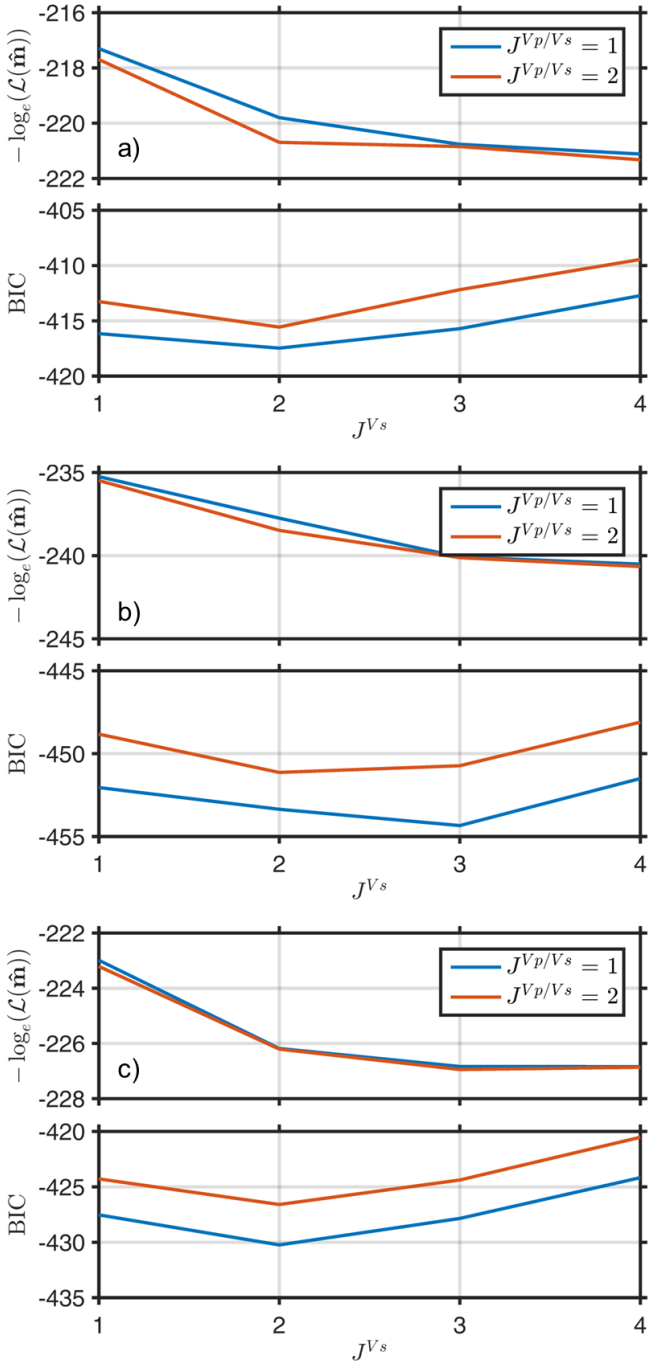

Figure 3.8: BIC analysis for Kitmat dispersion data at sites 1, 2, 3 shown in (a), (b), (c), respectively. The likelihood of the MAP model $\mathcal{L}(\hat{\mathbf{m}})$ is also shown as a function of Bernstein-polynomial order.

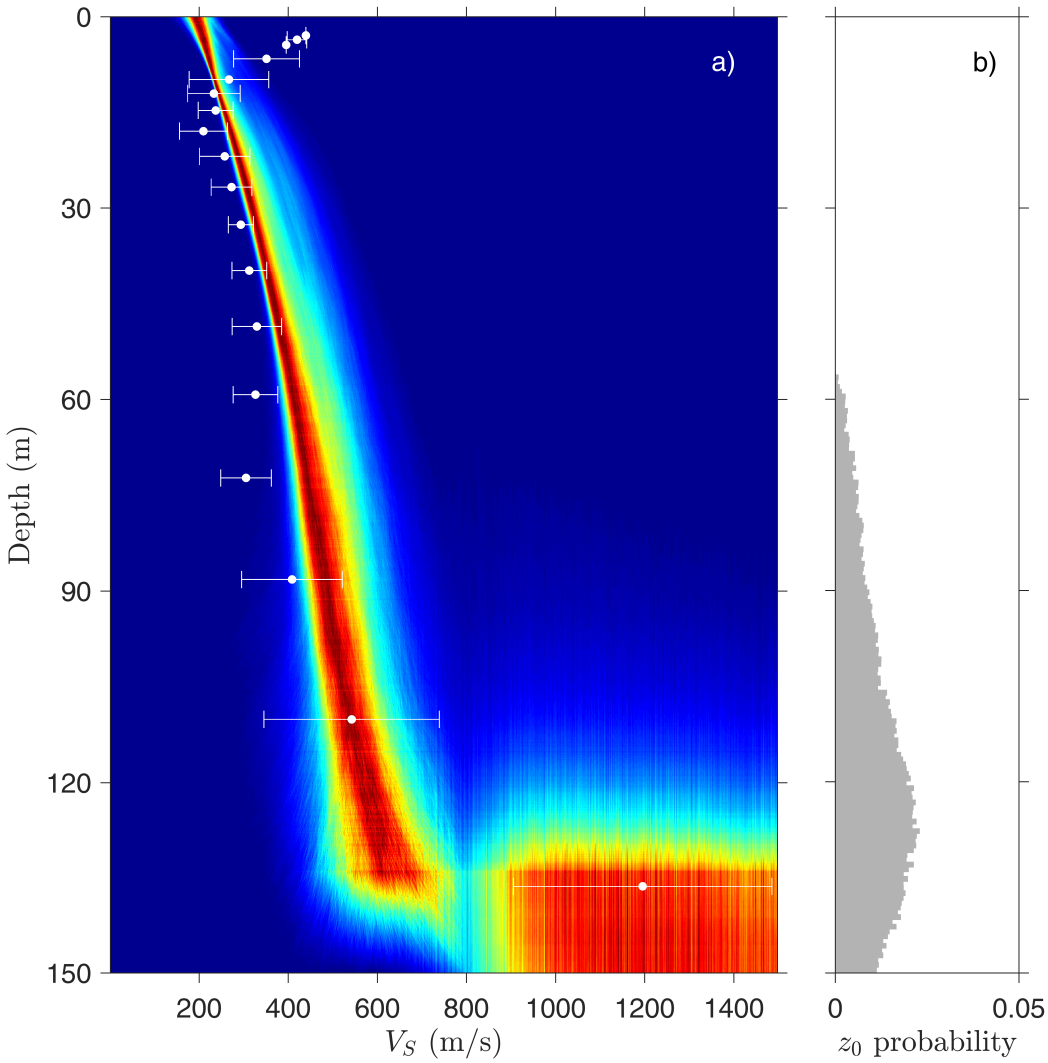

Figure 3.9: (a) Marginal probability profile for V_S, and (b) marginal probability density of half-space transition depth z_0 at Kitmat site 1. In (a), probabilities are normalized independently at every depth for display purposes. Warm colours (red) in (a) represent high probability and cool colours (blue) represent low probability. The white dots in (a) indicate mean values (with 1 standard deviation error bars) estimated from a nearby multi-component seismic reflection survey from Pugin et al. (2016).

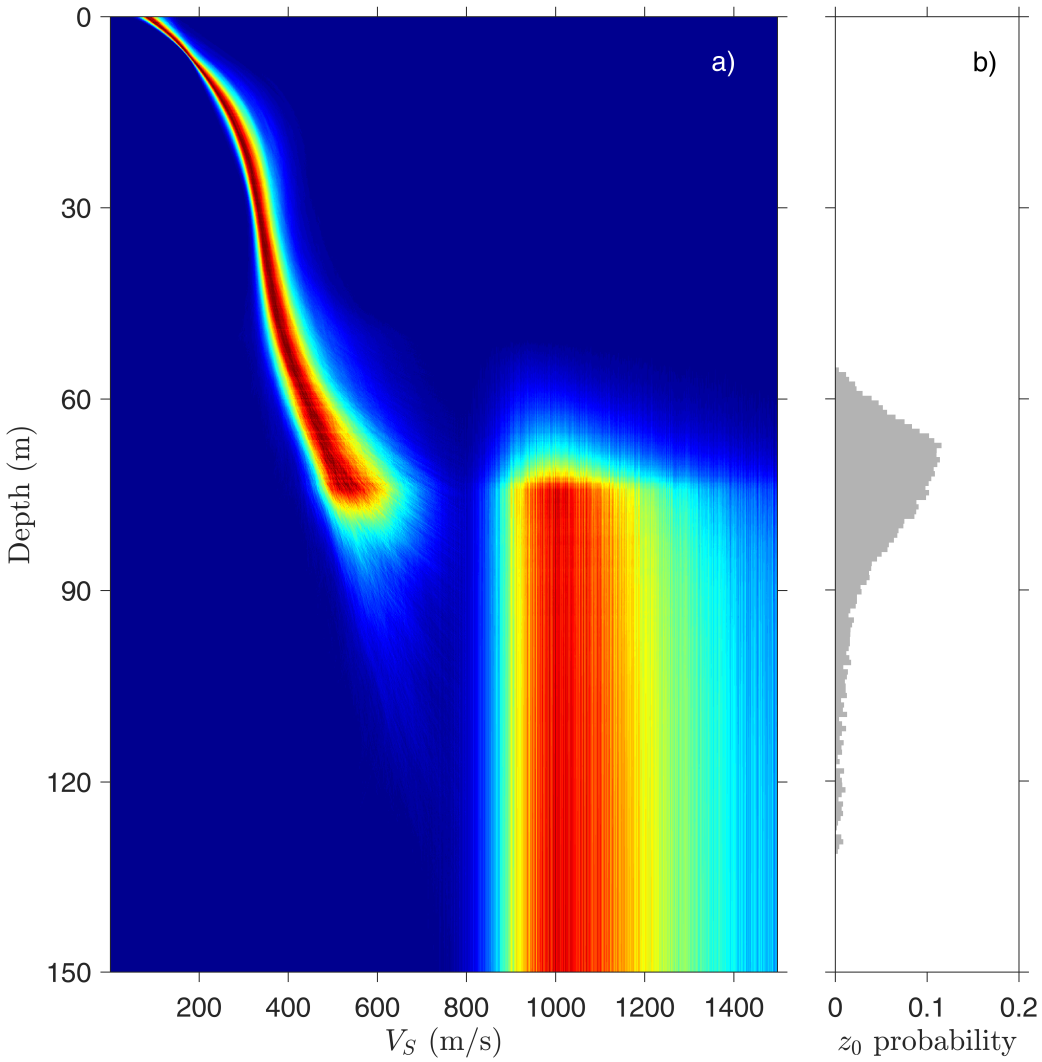

Figure 3.10: (a) Marginal probability profile for V_S, and (b) marginal probability density of half-space transition depth z_0 at Kitmat site 2.

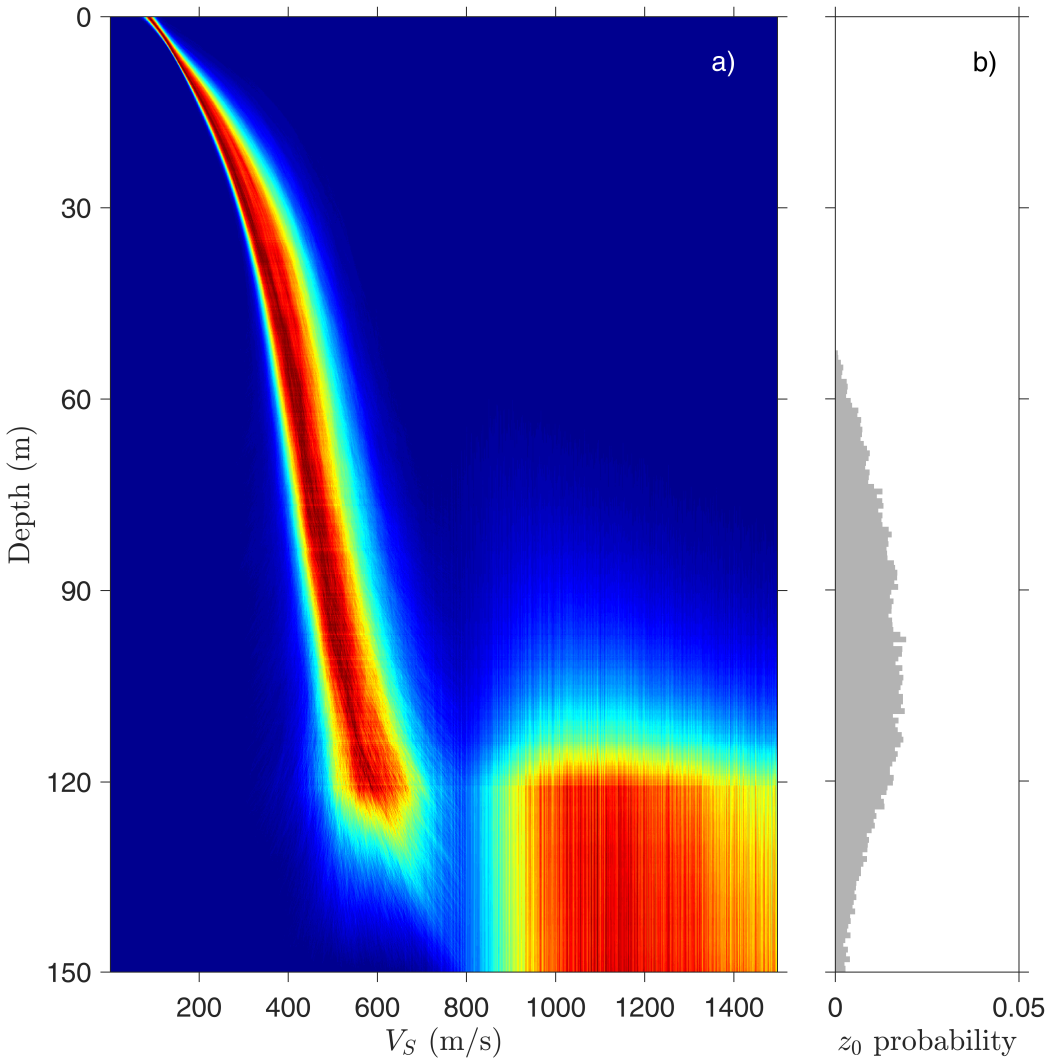

Figure 3.11: (a) Marginal probability profile for V_S, and (b) marginal probability density of half-space transition depth z_0 at Kitmat site 3.

Figures 3.9, 3.10, and 3.11 show the results of the inversion in the form of marginal probability profiles for V_S at Kitimat sites 1, 2, and 3, respectively (the marginal probability profiles for V_P/V_S are not shown due to the lack of sensitivity of V_P/V_S to the dispersion data). The V_S marginal profiles at each site shows clear depth-dependent structure with velocities increasing with depth. Results for sites 2 and 3 show particularly low near-surface V_S values (<100 m/s) with sharper shallow gradients. This is expected from the high-frequency content of the dispersion data for sites 2 and 3 (Figures 3.6 and 3.7), which show lower Rayleigh-wave phase velocities than at site 1. Geological mapping of the region suggests the three sites should be overlain by similar surficial material (mostly sandy-soil deposits) (Clague, 1977, 1984).

Uncertainty in the V_S profiles generally increases with depth at each site, as expected from the decaying resolution of dispersion data with depth. The estimated depth of the transition to the half-space, and the velocity in the half-space, are not well-constrained. The MAP values ($\pm$ 1 standard deviation) for the depth of the transition to the half-space z_0 are 127 m $\pm$ 26 m, 67 m $\pm$ 11 m, and 97 m $\pm$ 24 m at sites 1, 2, and 3, respectively (Figures 3.9(b), 3.10(b), and 3.11(b)). The shallow z_0 depth at site 2 is expected from the dispersion data, which show a transition to higher phase velocities occurring at a higher frequency at site 2 (Figure 3.6), compared to sites 1 and 3 (Figures 3.5 and 3.7). The MAP values for the velocity in the half-space are 1159 m/s $\pm$ 197 m/s, 1049 m/s $\pm$ 168 m/s, and 1047 m/s $\pm$ 186 m/s at sites 1, 2, and 3, respectively. These half-space velocities are consistent (within uncertainties) between sites. Similar uncertainties have been observed in other studies that considered Bayesian inversion of surface-wave dispersion data (Molnar et al., 2010; Dettmer et al., 2012).

Depending on the geological conditions, the depth resolution limit for dispersion data from passive array recordings is typically taken to be on the order of the maximum station separation distance (Park et al., 1999; Wathelet et al., 2008), which in this case is 90 m. This is a possible explanation for the large uncertainties for z_0 estimated at sites 1 and 3, compared to site 2. The dispersion data at site 2 appear to resolve more detailed structure as indicated by the higher Bernstein polynomial order of 3 from the BIC analysis. The result is that gradients in the V_S profile are greater before z_0 at site 2 (between 40 m to 80 m depth in Figure 3.10), than at sites 1 and 3. Similar to other studies that considered Bayesian inversion of surface-wave dispersion data (Molnar et al., 2010; Dettmer et al., 2012), it is possible that the transition to the half-space does not represent a physical discontinuity to a homogeneous layer, but

instead marks the resolution limit of the dispersion data. However, past drilling in the region suggest sediments of a similar thickness (Dolmage, 1956), and the agreement in half-space velocities between sites suggests the inversion resolves the physical base of the surficial sediments (albeit with high uncertainties).

The white dots in Figure 3.9 are V_S values extracted from a nearby multi-component seismic-reflection survey (Pugin et al., 2016). The survey was conducted with 72 three-component geophones (216 channels) spaced 1.5 m apart beginning 3 m from the source (for a total array length of 109.5 m), and the source and geophone array were moved 4.5 m between source shots. Shear-wave velocity values were estimated at various discrete depths and at many offsets along the survey line using a modified Dix inversion, and indicate lateral variability in V_S structure (Dix, 1955; Pugin et al., 2009, 2016). The values shown here are the mean velocity ($\pm$ 1 standard deviation) within 25 logarithmically-spaced depth intervals, across the entire seismic-reflection line (similar to the procedure used by Molnar et al. (2010) for averaging SCPT and borehole data). The error bars provide an estimate of the variability in the V_S values estimated from the active-source survey, as well as an approximate measure of the uncertainty due to the assumption of laterally homogeneous structure. This assumption is made in both the processing of active-source reflection data as well as within the inversion of passive-array dispersion data. Results from the seismic-reflection survey indicate limited lateral variability in the depth to basement along this track. However, results from the seismic-reflection survey along different tracks in Kitimat reveal buried bedrock valleys. The passive-array results represent an average solution over the spatial extent of the array, and lateral inhomogeneity may contribute to the estimated uncertainties in V_S structure.

Results from passive-array dispersion data at site 1 are in relatively good agreement with the active-source results, particularly the trend in the V_S profile between 20 m and 100 m depth. The Vs estimates from noise inversion are slightly higher than those from reflection inversion over this depth range, but differences are within uncertainties of the two approaches. The significant deviation between the two results at depths $<$20 m is most likely due to the fact that the reflection survey was conducted on a road surface as opposed to the open grass field where the passive-array was installed. The road asphalt overlies hard, compacted gravels and road-construction materials that are expected to produce a high-V_S signature to some depth. The depth at which the active-source results indicate the base of the sediments (approximately 110 m to 130 m) is consistent with the estimated depth of z_0 from the passive-array results. However, as

Kitimat site	KS test (p-value)	runs test (p-value)
1	0.85	0.36
2	0.76	0.12
3	0.13	0.17

Table 3.4: Summary of *a posteriori* statistical tests for validation of data error assumptions.

previously stated, the uncertainty in the estimated depth from the passive-array results is significant. The results from the active-source survey indicate z_0 likely corresponds with a physical boundary with another medium (bedrock or hard consolidated sediments) providing further evidence that inversion results at all Kitimat sites resolve the base of the surficial sediments. It should be noted that the V_S values estimated from the reflection survey are along a single transect that is not co-located with the passive array (approximately 200 m north). It is also well known that velocities estimated from seismic reflection surveys have their own associated uncertainties (not quantified by Pugin et al., 2016) due to measurement errors and processing assumptions (e.g., homogeneous isotropic layers, flat horizontal reflectors, etc.), and are biased based on values estimated in the near-surface (Dix, 1955; Pugin et al., 2009).

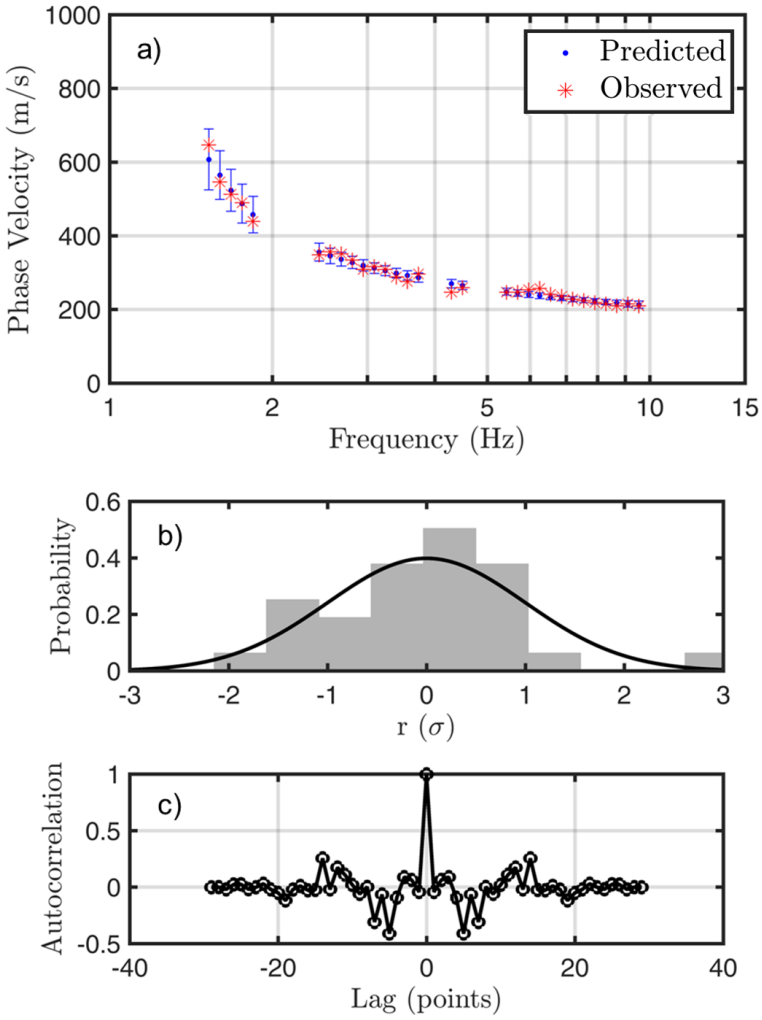

Figure 3.12: (a) Data fit to dispersion data from Kitimat site 1. The red stars are the observed dispersion data, and the blue dots with error bars are the mean predicted data (with 2 standard deviation error bars) from the MCMC samples. (b) Histogram and (c) autocorrelation of standardized residuals from the MAP model. The solid line in (b) is the standard normal distribution.

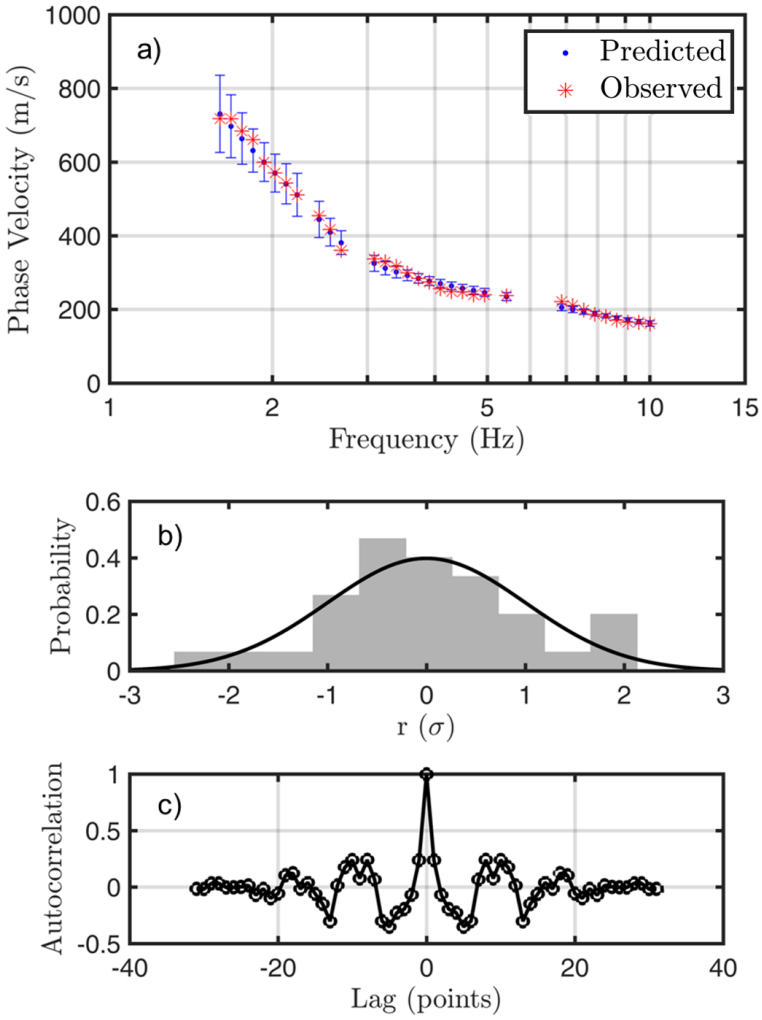

Figure 3.13: (a) Data fit to dispersion data from Kitimat site 2. The red stars are the observed dispersion data, and the blue dots with error bars are the mean predicted data (with 2 standard deviation error bars) from the MCMC samples. (b) Histogram and (c) autocorrelation of standardized residuals from the MAP model. The solid line in (b) is the standard normal distribution.

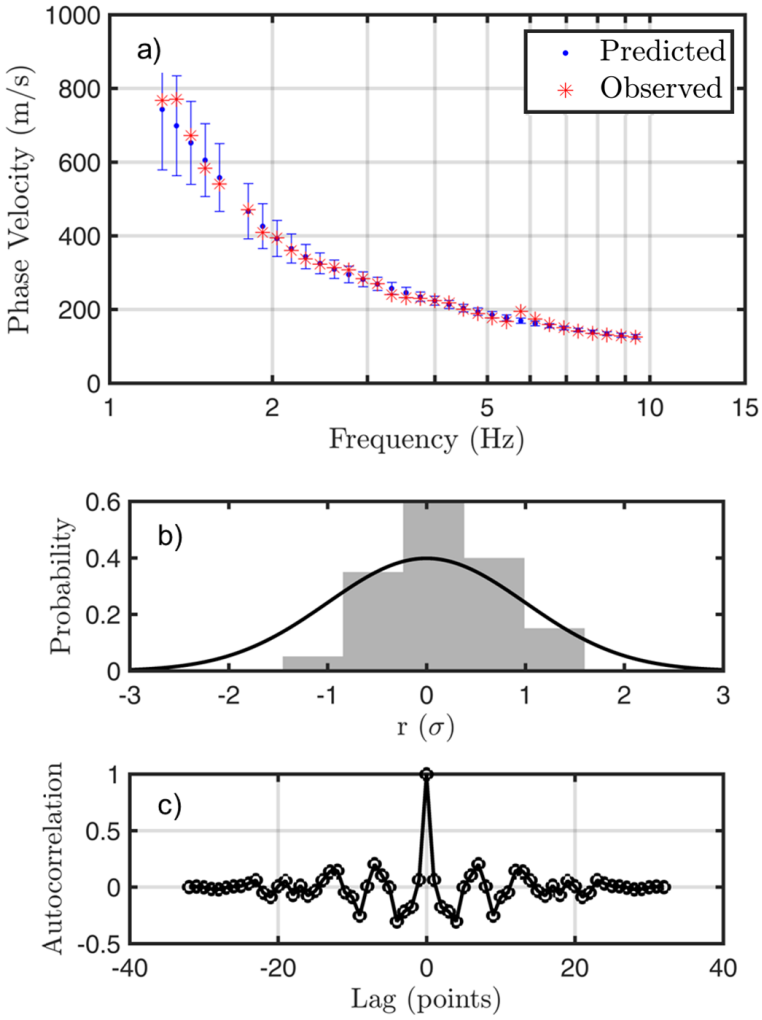

Figure 3.14: (a) Data fit to dispersion data from Kitimat site 3. The red stars are the observed dispersion data, and the blue dots with error bars are the mean predicted data (with 2 standard deviation error bars) from the MCMC samples. (b) Histogram and (c) autocorrelation of standardized residuals from the MAP model. The solid line in (b) is the standard normal distribution.

Panel (a) in Figures 3.12, 3.13, and 3.14 show the fit to the observed dispersion data at Kitimat sites 1, 2, and 3, respectively (plotted as phase velocities, although slownesses were inverted). The red stars are the observed dispersion data and the blue dots with error bars are the mean predicted dispersion data (± 2 standard deviations) from the MCMC samples. Inversion results at all sites show a good fit to the observed data with reasonable variance for the data predictions. The assumptions of the data error model within the inversion are validated by applying *a posteriori* statistical tests to the standardized residuals (differences between the observed and predicted data, divided by estimated standard deviations and accounting for the AR process, as described in section 2.2). Panels (b) and (c) in Figures 3.12, 3.13, and 3.14 show histograms and autocorrelation functions of the standardized residuals of the MAP models at sites 1, 2, and 3, respectively. If the error process is Gaussian, then the histograms will approximate the standard normal Gaussian (appears to be consistent at all sites). The Kolmogorov-Smirnov (KS) test was applied to quantitatively assess the validity of the assumption (null hypothesis) of Gaussian-distributed data errors (Freund, 1967; Dosso et al., 2006). The results of KS tests (p-values) are summarized in Table 3.4. Using the typical threshold of $p < 0.05$ (for rejection of the null hypothesis), this assumptions appears to be justified in the inversion. The width of the central peak of the autocorrelation of the standardized residuals gives an indication of the correlation in data errors. Ideally, the AR model will account for correlation and the resulting standardized residuals will be uncorrelated. Qualitatively, this appears to be valid as the central peak of the autocorrelation function is only 1-3 points wide at every site. The runs test was applied to quantitatively assess the validity of the assumption (null hypothesis) of random data errors (Freund, 1967; Dosso et al., 2006), and results (p-values) are summarized in Table 3.4.

3.4 Seismic hazard site assessment

Earthquake ground motions are influenced by the seismic source, the propagation path through the earth, and local sediment/soil conditions. This section considers site-specific seismic hazards in terms amplification due to V_S profiles estimated from surface-wave dispersion data (Molnar et al., 2013). Specifically, V_S posterior probability profiles estimated at three sites in Kitimat (discussed in section 3.3) are considered here. Probability density profiles of V_{SZ} and distributions of V_{S30} are calculated from the V_S profiles. V_{S30} distributions are assessed in terms of the National Building Code

of Canada seismic site classification, and V_{S30}-dependent amplification factors are computed from a semi-empirical model. Lastly, the V_S marginal density profiles are used to estimate the amplification and resonance spectra of horizontally-polarized shear S_H waves (which typically cause the most damaging shaking). This section does not consider amplification due to 3-D basin structure, or earthquake-specific rupture and propagation effects which are beyond the scope of this research. However, all of these factors are needed in order to predict the true ground motion during an earthquake.

The travel-time-weighted average of V_S from the surface to a depth z, V_{SZ}, is calculated by

$$V_{SZ} = \frac{z}{\int_0^z \frac{dz}{V_S(z)}}. \tag{3.5}$$

Probability profiles for V_{SZ} are calculated for the V_S profile samples drawn from the PPD for each site (Figures 3.15(a), 3.16(a), and 3.17(a)). Complicated structure in the V_{SZ} profile can suggest average-V_S-based classification schemes and amplification predictions are not suitable. Specifically, if large near-surface V_S discontinuities exist (that would significantly affect seismic amplification and resonance), the arbitrary choice of 30 m for seismic hazard classification would not be justified. This is specifically why the prior bound on the depth of the base of the sediments z_0 was set shallower than thirty metres. Surface-wave dispersion data do not generally resolve detailed structure beyond a significant impedance contrast such as the sediment-basement boundary (Wathelet et al., 2008). If such a boundary existed in the upper thirty metres, the data would have required z_0 to be shallow. However, the V_{SZ} marginal profiles at the three Kitimat sites show smooth near-surface structure, and V_{S30} is believed to be a reasonable proxy for site amplification here. As such, V_{S30}-dependent site classification and amplification predictions should provide meaningful information. Distributions of V_{S30} indicate that site classification D (stiff soil), adopted by the National Building Code of Canada, is representative of each site (Figures 3.15(b), 3.16(b), and 3.17(b)). The probabilities for site classification C, D, and E at every site are shown as percentages in Figures 3.15(b), 3.16(b), and 3.17(b)).

Recent GMPEs use V_{S30} as a convenient parameter (representative of soil/sediment rigidity conditions) to calculate site amplification. Seyhan & Stewart (2014) developed a model for the site amplification term in the Boore et al. (2014) GMPE using a combination of empirical ground-motion observations and ground-motion simulations. The model describes ground-motion scaling with V_{S30} as well as nonlinear soil effects. The nonlinear response of soil/sediment is due to the reduction in shear strength (and

increased damping) as shear strain increases, and modifies the frequency-dependent response of site amplification (Hardin & Drnevich, 1973). Including nonlinear effects requires knowledge of earthquake and propagation-path properties. Nonlinear effects generally decreases site amplification predictions and are not included here to provide conservative estimates of site amplification factors, and eliminate the requirement of predictions for earthquake-specific scenarios. The linear site amplification term F_{lin} is calculated according to

$$\log_e(F_{lin}) = \begin{cases} b \log_e\left(\frac{V_{S30}}{V_{ref}}\right) & V_{S30} \leq V_b \\ b \log_e\left(\frac{V_b}{V_{ref}}\right) & V_{S30} > V_b \end{cases}, \tag{3.6}$$

where b is a coefficient describing the linear V_{S30} scaling, V_b is the velocity at which site amplification no longer scales with V_{S30}, and V_{ref} is the V_{S30} value for a reference (hard ground) site where amplification is unity (taken to be 760 m/s as prescribed by Seyhan & Stewart (2014) and Boore et al. (2014)). The use of 760 m/s for V_{ref} also provides conservative amplification estimates within Canada as the National Building Code of Canada uses site class C as a reference site (Table 3.1), and 760 m/s is the upper bound of the V_{S30} range for this site class (Adams et al., 2015; Humar, 2015). The coefficients V_b and b are included within a software routine by Seyhan & Stewart (2014).

Site amplification factors are computed for peak ground velocity (PGV), peak ground acceleration (PGA), and spectral acceleration (SA) response at periods of engineering interest using the V_{S30} probability densities for Kitimat sites 1, 2, and 3 (Figures 3.15, 3.16, and 3.17, respectively). The SA response describes the maximum acceleration of a damped harmonic oscillator (with a specific natural frequency) with a single degree of freedom. This SA response generally provides a more useful prediction of the motion of a structure during an earthquake than PGA or PGV. Mean amplification factors and uncertainties for PGV, PGA and SA (at 2 s period) are summarized in Table 3.5. As is expected, the higher near-surface V_S values at site 2 result in higher V_{S30} estimates and, consequently, smaller amplification predictions than at the other two sites.

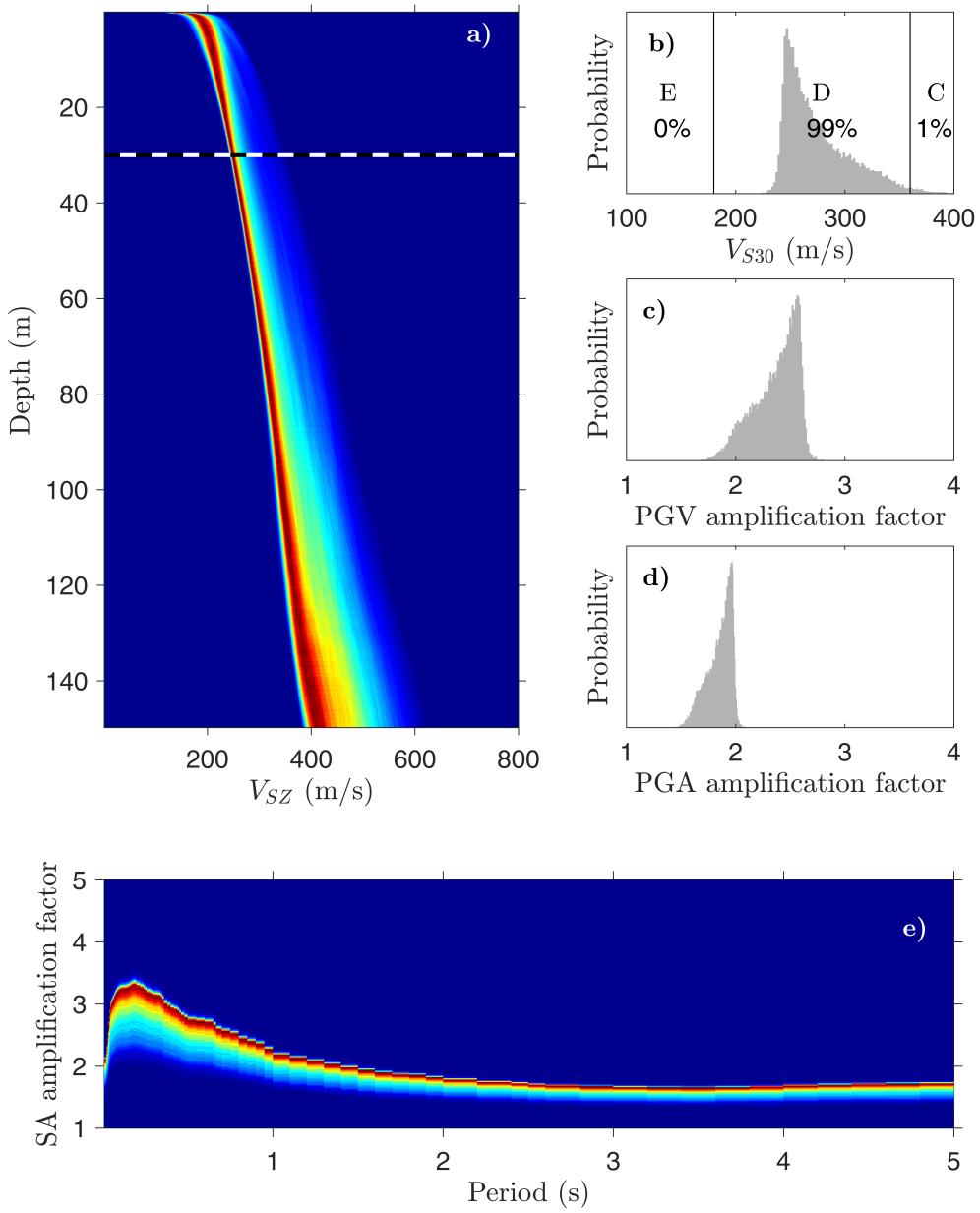

Figure 3.15: (a) V_{SZ} probability density profiles, (b) V_{S30} distribution, and V_{S30}-dependent amplification factors for (c) PGV, (d) PGA, and (e) SA at Kitimat site 1. V_{S30} site classification boundaries (and probabilities) are shown in (b). The dashed line in (a) marks thirty metres depth.

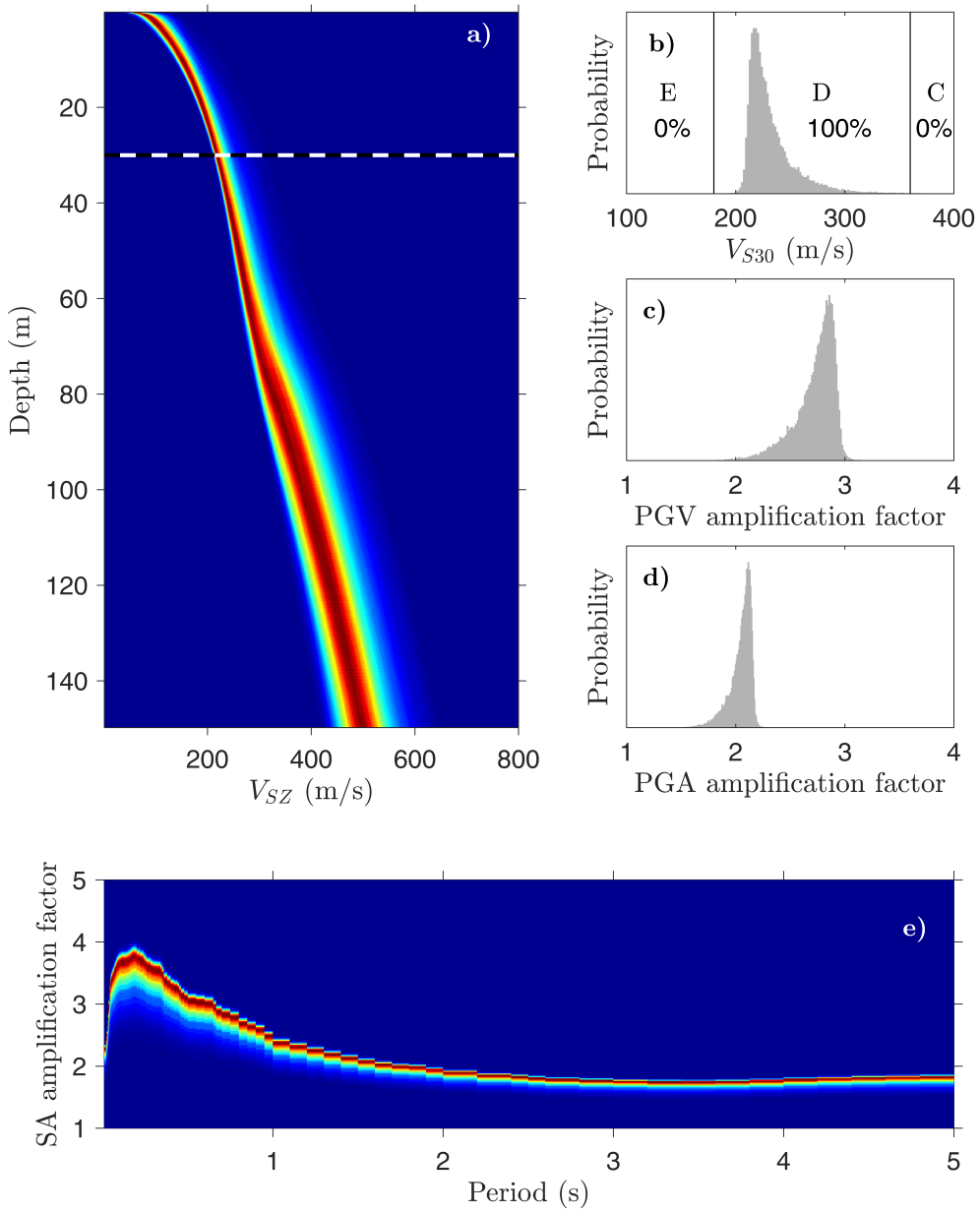

Figure 3.16: (a) V_{SZ} probability density profiles, (b) V_{S30} distribution, and V_{S30}-dependent amplification factors for (c) PGV, (d) PGA, and (e) SA at Kitimat site 2. V_{S30} site classification boundaries (and probabilities) are shown in (b). The dashed line in (a) marks thirty metres depth.

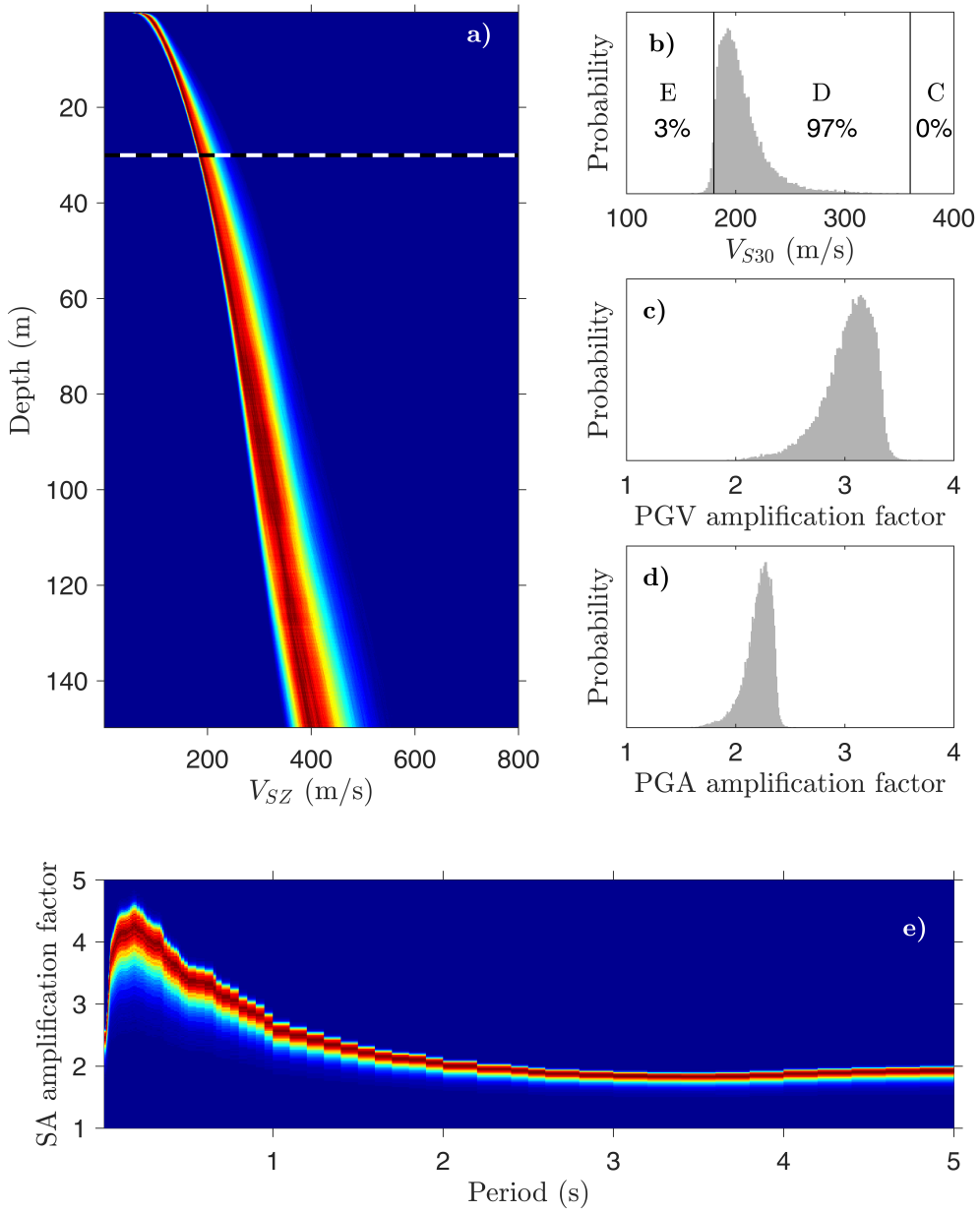

Figure 3.17: (a) V_{SZ} probability density profiles, (b) V_{S30} distribution, and V_{S30}-dependent amplification factors for (c) PGV, (d) PGA, and (e) SA at Kitimat site 3. V_{S30} site classification boundaries (and probabilities) are shown in (b). The dashed line in (a) marks thirty metres depth.

Kitimat site	PGV	PGA	SA	f_{SH} (Hz)	S_H Amplification
1	2.4 ± 0.20	1.8 ± 0.11	1.8 ± 0.13	1.1 ± 0.26	2.9 ± 0.62
2	2.7 ± 0.18	2.0 ± 0.10	1.9 ± 0.08	1.5 ± 0.23	3.6 ± 0.72
3	3.0 ± 0.23	2.2 ± 0.12	2.1 ± 0.10	1.2 ± 0.26	3.5 ± 0.75

Table 3.5: Mean (± 1 standard deviation) linear amplification factors for PGV, PGA and SA (at 2 s period), as well as fundamental S_H wave resonance frequency f_{SH} and amplification (at f_{SH}) at all Kitimat sites.

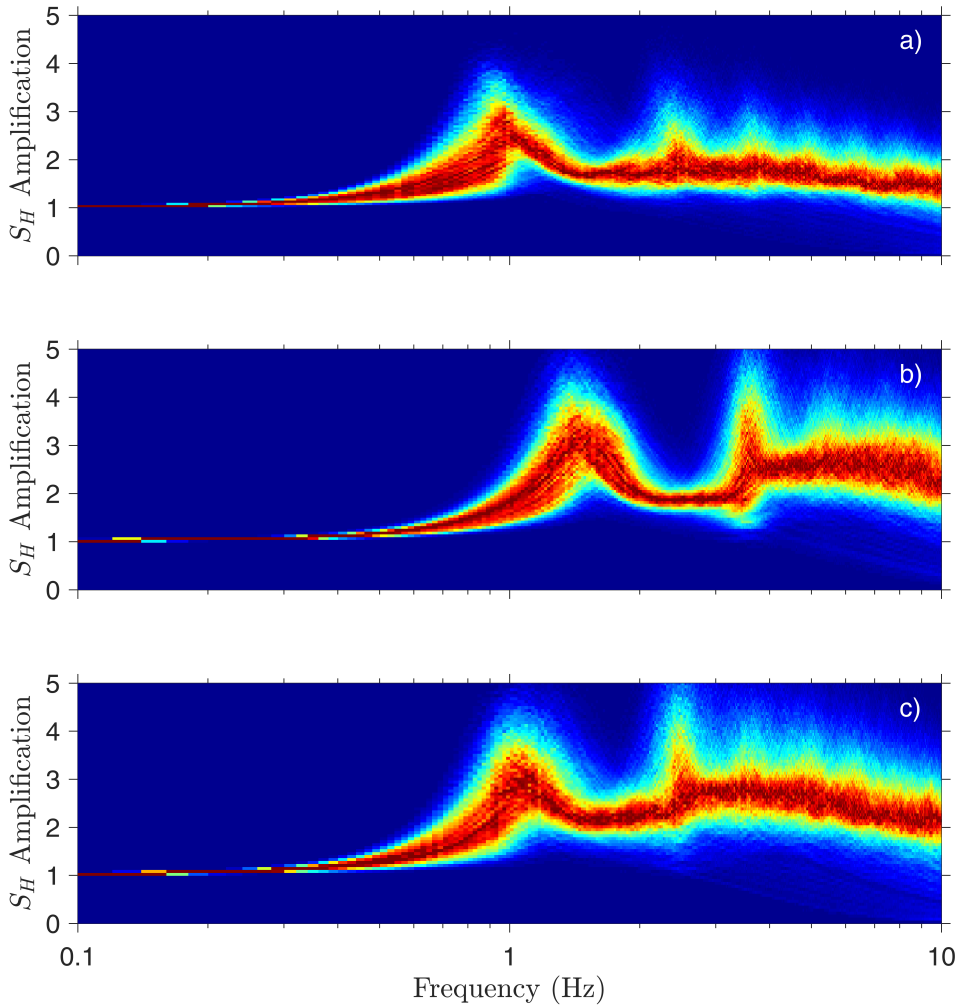

Figure 3.18: Probability densities of predicted amplification of S_H waves at Kitimat sites 1, 2, 3 in (a), (b), (c), respectively.

Amplification and resonance of S_H waves is considered using ground-motion simulation software developed by Boore (2005). This routine assumes the earth-structure model is composed of a stack of homogeneous layers of differing geophysical parameters (layer thickness, V_S, density, and quality factor Q) and uses the Thomson-Haskell method to calculate the transfer function of a vertically-incident plane S_H wave. Density and V_P/V_S values were not well constrained within the inversion (discussed in section 3.3). As a result, density values used to compute S_H wave amplification and resonance are calculated directly from V_S using an empirical relationship (Brocher, 2005). The quality factor Q (which accounts for attenuation) is drawn from a Gaussian distribution with mean 20 and standard deviation 10 (with an imposed lower bound of 2). This is believed to be representative of the sediments in Kitimat and consistent with previous studies that simulated amplification and resonance of S_H waves in Holocene deltaic sediments (Brocher, 2008; Molnar et al., 2013). Figure 3.18 shows probability densities for predicted S_H-wave amplification at the three sites in Kitimat. The mean and uncertainties for fundamental resonance frequency f_{SH}, and amplification (at f_{SH}) for each site are listed in Table 3.5. These frequencies are directly related to the estimated depth of the base of the sediments. At site 2, where this transition occurs at the shallower depth, resonance occurs at a higher frequency. Higher-order resonant frequencies are not well determined due to the large uncertainties in the estimated depth and velocity of underlying material. The level of amplification at high frequencies is lower at site 1 than at sites 2 and 3 due to the higher near-surface V_S values estimated at this site. The highest amplification is estimated at site 2. This is likely due to the combined effects of low near-surface V_S values and relatively shallow sediment base, compared to sites 1 and 3. The relatively large uncertainties in estimated amplification (at f_{SH}) are likely due to the large uncertainty in the estimated velocity of the material underlying surficial sediments, since it is this main impedance contrast which produces fundamental resonance and amplification. Knowledge of the resonant frequencies at these sites has practical uses in engineering and design.

3.5 Summary

This chapter considered the collection, processing and inversion of passive seismic array recordings collected at three sites in Kitimat, BC. The data were inverted within a Bayesian (probabilistic) framework using a general and efficient approach to parameterizing smooth gradient-based V_S profiles in terms of a Bernstein-polynomial represen-

tation. A MCMC method was used to sample the PPD of the model parameters, and the results are represented by Vs marginal probability profiles, at each site. Inversion results are in relatively good agreement with V_S estimates from a nearby active-source seismic-reflection survey (Pugin et al., 2016). In particular, the active-source survey indicates a transition to high-velocity material beneath the surficial sediments at a similar depth to that estimated from the passive-array method.

The inversion results were used to calculate marginal probability profiles of V_{SZ} (V_S averaged to depth z) at every site, as well as probability distributions of V_{S30} (a parameter commonly used for seismic site classification and amplification prediction). The V_{S30} values indicate that site classification D (stiff soil), according to the National Building Code of Canada, is representative of all sites (with a high degree of certainty). Linear site amplification terms (that are V_{S30}-dependent) were computed for PGV, PGA, and SA using a semi-empirical model and indicate site amplification factors of 2.4-3.0 for PGV, and 1.8-2.2 for PGA. The amplification and resonance of S_H waves was estimated using the full V_S marginal density profiles and provided estimates of fundamental resonance frequencies of 1.1-1.5 Hz and amplification factors at the fundamental frequency of 2.9-3.6.

This chapter applied rigorous, quantitative methods to estimate V_S profiles (and profile uncertainties) using passive seismic data, for seismic hazard analysis. It should be noted that the inversion methodology applied here has limitations and is likely not suitable in regions where discrete layering structure is expected (that the data are sensitive to). However, the methods and findings discussed here are useful for planners and regulatory agencies to mitigate the hazard of seismic amplification and resonance in Kitimat, BC.

Chapter 4

Conclusions

4.1 Summary of results

This thesis develops and applies inversion methodologies for estimating shear-wave velocity (V_S) profiles from surface-wave dispersion data, for use in site-specific seismic hazard analysis. The inversion is developed within a Bayesian framework, in which the parameters that describe the earth structure model (V_S profile) are treated as random variables constrained by prior information and data. The solution to the inverse problem is described by the posterior probability density (PPD) of the model parameters. A Markov-chain Monte Carlo (MCMC) method is used to draw dependent, asymptotically-unbiased samples from the PPD in order to estimate marginal probability profiles for V_S.

The inversion implements a general and efficient approach to parameterizing smooth V_S gradients using a Bernstein-polynomial representation. This representation is defined by a weighted sum of Bernstein-basis functions, where the basis-function coefficients are treated as unknowns in the inversion. The Bernstein-polynomial is more general than other gradient-based models (e.g., power-law models) while requiring fewer parameters to be estimated than many other gradient-based models (e.g., spline models). This representation also has the distinctive property that small perturbations to the model parameters (basis-function coefficients) result in only small, localized perturbations to the V_S profile. The stability of this representation is desirable in MCMC sampling as it allows for efficent, detailed exploration of the model parameter space.

Chapter 2 develops the inversion methodology and validates its use via inversion of synthetic dispersion data. The synthetic test compares inversion results using the

Bernstein-polynomial inversion approach to the trans-dimensional (trans-D) approach based on uniform layered structure. The test illustrates how the approach considered in this thesis is better suited than layered modelling approaches in applications where smooth gradients in geophysical parameters are expected, and/or the observed data are diffuse and not sensitive to fine-scale discrete layering, such as surface-wave dispersion. The Bernstein-polynomial inversion approach is also applied to previously considered data processed from ambient seismic noise collected on the Fraser River Delta in British Columbia (BC). The Bernstein-inversion results are compared to results from Bayesian inversion with a power-law model which effectively models gradient structure but is restrictive and forces the V_S profile to be a power law even if the dispersion data are better fit by a different gradient structure. The Bernstein-polynomial approach is also compared to results of trans-D Bayesian inversion, which is flexible but produces an undesirable discrete layering structure in the V_S marginal density profile. The Bernstein-polynomial inversion approach produces results that are in good agreement with co-located invasive V_S measurements. These inversion results illustrate how the Bernstein representation allows the data information to determine general gradient structures, as opposed to imposing a subjective parameterization choice. The Bernstein-polynomial inversion approach has widespread applicability, particularly in investigations of soil/sediment structure (such as earthquake engineering and seismic hazard analysis) where smooth gradients in geophysical properties are expected.

Chapter 3 presents surface-wave dispersion data collected in Kitimat, BC, for use in site-specific seismic hazard analysis. Passive seismic array surveys were conducted at three sites in Kitimat, and dispersion data were estimated from ambient-noise recordings using frequency-wavenumber processing. The Bernstein-polynomial inversion is used to estimate V_S structure from these dispersion data and provides results that are in good agreement with V_S values estimated from a nearby active-source seismic-reflection survey. The agreement between the results of the two types of surveys provides confidence in the use of the Bernstein-polynomial inversion results for site-specific seismic hazard analysis. Specifically, the V_S marginal probability profiles are used to calculate probability distributions of V_{S30} (a commonly-used parameter for seismic site classification and seismic amplification prediction). Within the context of the National Building Code of Canada, the inversion results indicate with high confidence that site classification D (stiff soil) is representative of all sites investigated in Kitimat. Linear amplification factors for peak ground velocity and peak ground acceleration are estimated to be 2.4-3.0 and 1.8-2.2, respectively. The inversion results are used to

predict the amplification and resonance of horizontally-polarized shear (S_H) waves and indicate fundamental resonance frequencies of 1.1-1.5 Hz and amplification factors at the fundamental frequency of 2.9-3.6. This thesis does not consider amplification of seismic waves due to three-dimensional basin structure, or earthquake-specific rupture and propagation effects which are beyond the scope of this research. However, the seismic hazard analysis results presented will be useful in the development of seismic hazard mitigation plans as well as for predicting the true ground motion for specific earthquake scenarios.

4.2 Recommendations for future work

The main contributions presented in this thesis provide a natural direction for future research. This work implemented a rigorous and objective approach to modelling V_S structure using a weighted sum of Bernstein-basis functions. The procedure requires carrying out inversions for several polynomial orders (number of basis functions) to determine the most-probable order. As such, inversion results presented here do not include the uncertainty in the polynomial order. A potential path for future research would be to investigate a trans-dimensional Bernstein parameterization, where the polynomial order is included as an unknown in the inversion. However, the implementation of this concept appears to be very challenging. Specifically, all of the Bernstein-basis functions change when the polynomial order is changed (see Figure 2.1 in Chapter 2). As a result, dimension jumps will likely never be accepted since the V_S profile changes completely when the order changes (unlike layered trans-D inversion where only one layer is changed during dimension jumps). Furthermore, the Bernstein polynomial is likely not suitable in regions where discrete layering structure is expected (that the data are sensitive to). A fully general parameterization that can incorporate gradients and/or discrete layering (and quantify the uncertainty between them) would be ideal.

The diffusivity of surface-wave dispersion data results in large uncertainties in the estimated depth of the base of the sediments (as noted in previous studies as well as all results in this thesis). Recordings of ambient seismic noise contain additional useful information for constraining V_S structure beyond fundamental-mode Rayleigh-wave dispersion. These include horizontal-to-vertical spectral ratios (HVSR) as well as higher-order Rayleigh-wave modes and Love-wave dispersion. HVSR has been used to estimate the depth of the base of sediments (e.g., Tokimatsu et al., 1991), and has been

applied in joint analysis with surface-wave dispersion to improve V_S profile estimates (e.g., Parolai et al., 2005). However, there is not universal agreement on the physical processes which produce the features observed in HVSR (Rayleigh-wave ellipticity, S_H-wave resonance, etc.). It is not clear how the processing methods used to estimate dispersion from ambient noise translate data uncertainties from the recordings to the phase-velocity estimates. A potential path for future research would be to investigate new methods in data processing which can objectively quantify and translate data uncertainties from seismometer recordings to the marginal probability density of V_S.

Lastly, within the context of seismic hazards in Kitimat, an obvious direction for future research would be to perform additional surveys throughout the Kitimat valley to study variability in soils/sediments in the region. The recent installation of a soil seismometer in the region also presents an opportunity to study empirical amplification observations by comparing recordings from this new instrument to recordings from the existing seismometer in Kitimat (installed on bedrock).

Bibliography

Abbott, P. L. & Samson, C., 2012. *Natural Disasters*, 2nd Canadian Edition, McGraw-Hill.

Adams, J., Halchuk, S., Allen, T., & Rogers, G. C., 2015. Canada 5th generation seismic hazard model, as prepared for the 2015 National Building Code of Canada, *Proceedings of The 11th Canadian Conference on Earthquake Engineering*, Victoria, Canada.

Aki, K., 1957. Space and time spectra of stationary stochastic waves, with special reference to microtremors, *Bull. Earthq. Res. Inst.*, **35**, 415–456.

Anderson, J. G., Bodin, P., Brune, J. N., Prince, J., Singh, S. K., Quass, R., & Onate, M., 1986. Strong motion from the Michoacan, Mexico, earthquake, *Science*, **233**, 1043–1049.

Anderson, J. G., Lee, Y., Zeng, Y., & Day, S., 1996. Control of strong motion by the upper 30 meters, *Bull. Seismol. Soc. Am.*, **86**, 1749–1759.

Asten, M. W. & Henstridge, J. D., 1984. Array estimators and use of microseisms for reconnaissance of sedimentary basins, *Geophysics*, **49**, 1828–1837.

Bard, P. Y. & Bouchon, M., 1980. The two-dimensional resonance of sediment-filled valleys, *Bull. Seismol. Soc. Am.*, **75**, 519–541.

Basirat, B. & Shahdadi, M. A., 2013. Numerical solution of nonlinear integro-differential equations with initial conditions by Bernstein operational matrix of derivative, *Int. J. Mod. Nonlin. Theory App.*, **2**, 141–149.

Bhattia, M. I. & Bracken, P., 2007. Solutions of differential equations in a Bernstein polynomial basis, *J. Comput. Appl. Math.*, **205**, 272–280.

Boore, D. M., 2005. SMSIM FORTRAN programs for simulating ground motions from earthquakes: Version 2.3, USGS open file report, 96-80-A, Online Manual, www.daveboore.com.

Boore, D. M. & Atkinson, G. M., 2008. Ground-motion prediction equations for the average horizontal component of PGA, PGV, and 5 percent damped PSA at spectral periods between 0.01 s and 10.0 s, *Earthq. Spectra*, **24**, 99–138.

Boore, D. M., Stewart, J. P., Seyhan, E., & Atkinson, G. M., 2014. NGA-West2 equations for predicting PGA, PGV, and 5 percent damped PSA for shallow crustal earthquakes., *Earthq. Spectra*, **30**, 1057–1085.

Borcherdt, R. D., 1994. Estimates of site-dependent response spectra for design (methodology and justification), *Earthq. Spectra*, **10**, 617–653.

Bornhold, B. D., 1983. Sedimentation in Douglas Channel and Kitimat Arm, *Can. Tech. Rep. Hydrography Ocean Sci.*, **18**, 88–114.

Bostwick, T. K., 1984. *A re-examination of the August 22, 1949 Queen Charlotte earthquake*, Master's thesis, University of British Columbia, British Columbia, Canada.

Brillon, C., 2016a. Baseline assessment of seismic hazard in British Columbia's north coast, Tech. Rep. Open File 7994, Geological Survey of Canada.

Brillon, C., 2016b. North Coast geohazards - 2016 seismology update, Tech. Rep. Open File 8052, Geological Survey of Canada.

Brillon, C., Cote, M. M., & Hunter, J. A., 2015. HVSR analysis of preliminary Kitimat ambient noise survey, Tech. Rep. Open File 7793, Geological Survey of Canada.

Brocher, T. M., 2005. Empirical relations between elastic wavespeeds and density in the earth's crust, *Bull. Seismol. Soc. Am.*, **95**, 2081–2092.

Brocher, T. M., 2008. Compressional and shear-wave velocity versus depth relations for common rock types in northern California, *Bull. Seismol. Soc. Am.*, **98**, 950–968.

Brooks, S., Gelman, A., Jones, G., & Meng, X. L., 2011. *Handbook of Markov Chain Monte Carlo*, CRC press.

Budhu, M., 2007. *Soil Mechanics and Foundations*, 2nd edition, Wiley, Hoboken, NJ.

Cassidy, J. F., Rogers, G. C., & Hyndman, R. D., 2014. An overview of the 28 October 2012 M 7.7 earthquake in Haida Gwaii, Canada: A tsunamigenic thrust event along a predominantly strike-slip margin, *Pure Appl. Geophys.*, **171**, 3457–3457.

Castellaro, S., Mulargia, F., & Rossi, P. L., 2008. VS30: Proxy for seismic amplification, *Seismol. Res. Lett.*, **79**, 540–543.

Clague, J., 1984. Deglaciation of Prince Rupert - Kitimat area, British Columbia, *Can. J. Earth Sci.*, **22**, 256–265.

Clague, J. J., 1977. Surficial geology, Kitimat, British Columbia, Tech. Rep. Open File 470, Geological Survey of Canada.

Conway, K. W., Barrie, J. V., & Thomson, R. E., 2012. Submarine slope failures and tsunami hazard in coastal British Columbia: Douglas Channel and Kitimat Arm, Tech. Rep. Current Research 2012-10, Geological Survey of Canada.

Cornou, C., Ohrnberger, M., Boore, D. M., Kudo, K., & Bard, P. Y., 2006. Derivation of structural models from ambient vibration array recordings: results from an international blind test, *Proceedings 3rd International Symposium on the Effects of Surface Geology on Seismic Motion*, Grenoble, France.

Crow, H. L., Hunter, J., Brewer, K., Brillon, C., Cote, M., Allen, T., Cassidy, J. F., & Hayek, S., 2015. Soft soil response investigations in Kitimat, BC: some preliminary results, *Proceedings of The 11th Canadian Conference on Earthquake Engineering*, Victoria, Canada.

Dallimore, S. R., Edwardson, K. A., Hunter, J. A., Clague, J. J., Meldrum, J. L., & Luternauer, J. L., 1995. Composite geotechnical logs for two deep boreholes in the Fraser River delta, British Columbia, Tech. Rep. Open File 3018, Geological Survey of Canada.

Dettmer, J., Molnar, S., Steininger, G., Dosso, S. E., & Cassidy, J. F., 2012. Trans-dimensional inversion of microtremor array dispersion data with hierarchical autoregressive error models, *Geophys. J. Int.*, **188**, 719–734.

Dix, C. H., 1955. Seismic velocities from surface measurements, *Geophysics*, **20**, 68–86.

Dolmage, V., 1956. Geology of Kitimat area in British Columbia, report to the Township of Kitimat, 1956, 29p.

Dosso, S. E. & Wilmut, M. J., 2005. Data uncertainty estimation in matched-field geoacoustic inversion, *J. Ocean. Eng.*, **31**, 470–479.

Dosso, S. E., Nielsen, P. L., & Wilmut, M. W., 2006. Data error covariance in matched-field inversion, *J. Acoust. Soc. Am.*, **119**, 208–219.

Dosso, S. E., Holland, C. W., & Sambridge, M., 2012. Parallel tempering for strongly nonlinear geoacoustic inversion, *J. Acoust. Soc. Am.*, **132**, 3030–3040.

Dosso, S. E., Dettmer, J., Steininger, G., & Holland, C. W., 2014. Efficient trans-dimensional Bayesian inversion for geoacoustic profile estimation, *Inverse Probl.*, **30**, 114018.

Douze, E. J. & Laster, S. J., 1979. Statistics of semblance, *Geophysics*, **44**, 1999–2003.

Farouki, R. T., 2012. The Bernstein polynomial basis: A centennial retrospective, *Comput. Aided Geom. D.*, **29**, 379–419.

Farouki, R. T. & Rajan, V. T., 1989. On the numerical condition of polynomials in Bernstein form, *Comput. Aided Geom. D.*, **4**, 191–216.

FEMA Provision 750, 2009. NEHRP (national earthquake hazards reduction program) recommended seismic provisions for new buildings and other structures, Building Seismic Safety Council, Washington, DC, 406.

Finn, W. D. L. & Wightman, A., 2003. Ground motion amplification factors for the proposed 2005 edition of the National Building Code of Canada, *Can. J. Civil Eng.*, **30**, 272–278.

Freund, J. E., 1967. *Modern Elementary Statistics*, Prentice-Hall, Englewood Cliffs, NJ.

Gardner, G. H. F., Gardner, L. W., & Gregory, A. R., 1974. Formation velocity and density: The diagnostic basics for stratigraphic traps, *Geophysics*, **39**, 770–780.

Geyer, C. J., 1991. Markov Chain Monte Carlo maximum likelihood, *Proceedings of the 23rd Symposium on the Interface, American Statistical Association*, New York, pp. 156163.

Gilbert, F. & Backus, G., 1966. Propagator matrices in elastic wave and vibration problems, *Geophysics*, **31**, 326–332.

Gordon, W. J. & Riesenfeld, R. F., 1974. Bernstein-Bezier methods for the computer-aided design of free form curves and surfaces, *J. Assoc. Comp. Mach.*, **21**, 293–310.

Graves, R. W., Pitarka, A., & Sommerville, P., 1998. Ground-motion amplification in the Santa Monica area: Effects of shallow basin-edge structure, *Bull. Seismol. Soc. Am.*, **88**, 1224–1242.

Green, P. J., 1995. Reversible jump Markov chain Monte Carlo computation and Bayesian model determination, *Biometrika*, **82**, 711–732.

Hardin, B. O. & Drnevich, V. P., 1973. Shear modulus and damping in soils, *J. Soil Mech. Found. Div.*, **99**, 849–862.

Haskell, N. A., 1953. The dispersion of surface waves on multilayered media, *Bull. Seism. Soc. Am.*, **43**, 17–34.

Hastings, W. K., 1970. Monte Carlo sampling methods using Markov chains and their applications, *Biometrika*, **57**, 97–109.

Humar, J., 2015. Background to some of the seismic design provisions of the 2015 National Building Code of Canada, *Can. J. Civil Eng.*, **42**, 940–952.

Hunter, J. A., Burns, R. A., Good, R. L., & Pelletier, C. F., 1998. A compilation of shear-wave velocities and borehole geophysical logs in unconsolidated sediments of the Fraser River delta, British Columbia, Tech. Rep. Open File 3622, Geological Survey of Canada.

Kakaand, S. I. & Atkinson, G. M., 2005. ShakeMap for the M 5.4, 6 March 2005 Rivire-du-Loup, Qubec earthquake, *Seismol. Res. Lett.*, **76**, 743–751.

Kass, R. E. & Raftery, A. E., 1995. Bayes factors, *J. Am. Statist. Assoc.*, **90**, 773–795.

KITI-MAP, 2016. KITI-MAP GIS Webmapping, Municipal District of Kitimat, www.kitimat.ca.

Knopoff, L., 1964. A matrix method for elastic wave problems, *Bull. Seism. Soc. Am.*, **54**, 431–438.

Lacoss, R. T., Kelly, E. J., & Toksoz, M. N., 1969. Estimation of seismic noise structure using arrays, *Geophysics*, **34**, 21–38.

Metropolis, N., Rosenbluth, A., Rosenbluth, M., & Teller, A. T. A. E., 1953. Equations of state calculations by fast computing machines, *J. Chem. Phys.*, **21**, 1087–1092.

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

Molnar, S., Cassidy, J. F., & Dosso, S. E., 2013. Uncertainty of linear earthquake site amplification via Bayesian inversion of surface seismic data, *Geophysics*, **78**, WB37–WB48.

Mosegaard, K. & Tarantola, A., 1995. Monte Carlo sampling of solutions to inverse problems, *J. geophys. Res.*, **100**, 12431–12447.

NBCC, 2010. National Building Code of Canada 2010, Institute for Researchin Construction, National Research Council of Canada.

Neiddell, N. & Taner, M. T., 1971. Semblance and other coherency measures for multichannel data, *Geophysics*, **36**, 482–497.

Park, C. B., Miller, R. D., & Xia, J., 1999. Multichannel analysis of surface waves, *Geophysics*, **64**, 800–808.

Parolai, S., Picozzi, M., Richwalski, S. M., & Milkereit, C., 2005. Joint inversion of phase velocity dispersion and H/V ratio curves from seismic noise recordings using a genetic algorithm, considering higher modes, *Geophys. Res. Lett.*, **32**, L01303.

Pugin, A. J. M., Pullan, S. E., & Hunter, J. A., 2009. Multicomponent high-resolution seismic reflection profiling, *The Leading Edge*, **28**, 1248–1261.

Pugin, A. J. M., Crow, H., Dietiker, B., Brewer, K., Cartwright, T., Brillon, C., Hunter, J. A., Gosselin, J., & Allen, T., 2016. Soft soil response investigations in Kitimat, BC: Geophysical data sets, Tech. Rep. Open File, in prep., Geological Survey of Canada.

Pugin, A. M., Brewer, K., Cartwright, T., Pullan, S. E., Perret, D., Crow, H., & Hunter, J. A., 2013. Near surface s-wave seismic reflection profiling: New approaches and insights, *First Break*, **31**, 49–60.

Quijano, J. E., Dosso, S. E., Dettmer, J., & Holland, C. W., 2016. Geoacoustic inversion for the seabed transition layer using a Bernstein polynomial model, *J. Acoust. Soc. Am.*, in review.

Sambridge, M., 2014. A parallel tempering algorithm for probabilistic sampling and multimodal optimization, *Geophys. J. Int.*, **196**, 357–374.

Schwartz, G., 1978. Estimating the dimensions of a model, *Ann. Stat.*, **6**, 461–464.

Seyhan, E. & Stewart, J. P., 2014. Semi-empirical nonlinear site amplification from NGA-West2 data and simulations, *Earthq. Spectra*, **30**, 1241–1256.

Shumway, R. & Stoffer, D., 2000. *Time Series Analysis and Its Applications*, Springer, New York, NY.

Steininger, G., 2013. *Determination of seabed acoustic scattering properties by trans-dimensional bayesian inversion*, Ph.D. thesis, University of Victoria, British Columbia, Canada.

Thomson, W. T., 1950. Transmission of elastic waves through a stratified solid medium, *J. Appl. Phys.*, **21**, 89–93.

Tokimatsu, K., Kuwayama, S., Tamura, S., & Miyadera, Y., 1991. Vs determination from steady state Rayleigh wave method, *Soils Found.*, **31**, 153–163.

Wathelet, M., 2005. *Array recordings of ambient vibrations: surface-wave inversion*, Ph.D. thesis, University of Liege, Wallonia, Belgium.

Wathelet, M., Jongmans, D., Ohrnberger, M., & Bonnefoy-Claudet, S., 2008. Array performances for ambient vibrations on a shallow structure and consequences over VS inversion, *J. Seism.*, **12**, 1–19.

Woods, J. W. & Lintz, P. L., 1973. Plane waves at small arrays, *Geophysics*, **34**, 1023–1041.