Determination of Seabed Acoustic Scattering Properties by Trans-Dimensional Bayesian Inversion

by

Gavin Steininger
B.Sc., Simon Fraser University, 2006
M.Sc., University of British Columbia, 2009

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

DOCTOR OF PHILOSOPHY

in the School of Earth and Ocean Sciences

Determination of Seabed Acoustic Scattering Properties by Trans-Dimensional Bayesian Inversion

by

Gavin Steininger
B.Sc., Simon Fraser University, 2006
M.Sc., University of British Columbia, 2009

Supervisory Committee

Dr. S. E. Dosso, Supervisor
(School of Earth and Ocean Sciences)

Dr. A. H. Monahan, Departmental Member
(School of Earth and Ocean Sciences)

Dr. M. J. Wilmut, Departmental Member
(School of Earth and Ocean Sciences)

Dr. J. Zhou, Outside Member
(Department of Mathematics and Statistics)

Dr. J. Dettmer, Additional Member
(School of Earth and Ocean Sciences)

Supervisory Committee

Dr. S. E. Dosso, Supervisor
(School of Earth and Ocean Sciences)

Dr. A. H. Monahan, Departmental Member
(School of Earth and Ocean Sciences)

Dr. M. J. Wilmut, Departmental Member
(School of Earth and Ocean Sciences)

Dr. J. Zhou, Outside Member
(Department of Mathematics and Statistics)

Dr. J. Dettmer, Additional Member
(School of Earth and Ocean Sciences)

ABSTRACT

This thesis develops and applies Bayesian model selection and inversion approaches to acoustic seabed scattering and reflectivity data to estimate scattering and geoacoustic parameters with uncertainties, and to discriminate the relative importance of interface and volume scattering mechanisms. Determining seabed scattering mechanisms and parameters is important for reverberation modelling and sonar performance predictions. This thesis shows that remote acoustic sensing can provide efficient estimates of scattering properties and mechanisms with uncertainties, and is well suited for the development of bottom-scattering databases.

An important issue in quantitative nonlinear inversion is model selection, i.e., specifying the physical theory, appropriate parameterization, and error statistics which

describe the system of interest (acoustic scattering and reflection). The approach developed here uses trans-dimensional (trans-D) Bayesian sampling for both the number of sediment layers and the order (zeroth or first) of auto-regressive parameters in the error model. The scattering and reflection data are inverted simultaneously and the Bayesian sampling is conducted using a population of interacting Markov chains. The data are modelled using homogeneous fluid sediment layers overlying an elastic basement. The scattering model assumes a randomly rough water-sediment interface and random sediment-layer volume heterogeneities with statistically independent von Karman spatial power spectra. A Dirichlet prior distribution that allows the sediment layers and basement to have different numbers of parameters in a trans-D inversion is derived and implemented. The deviance information criterion and trans-D sampling are used to determine the dominant scattering mechanism for a particular data set.

The inversion procedure is developed and validated through several simulated test cases, which demonstrate the following. (i) Including reflection data in joint inversion with scattering data improves the resolution and accuracy of scattering and geoacoustic parameters. (ii) The trans-D auto-regressive model improves scattering parameter resolution and correctly differentiates between strongly and weakly correlated residual errors. (iii) Joint scattering/reflection inversion is able to distinguish between interface and volume scattering as the dominant mechanism.

The inversion procedure is applied to data measured at several survey sites on the Malta Plateau (Mediterranean Sea) to estimate *in-situ* seabed scattering and geoacoustic parameters with uncertainties. Results are considered in terms of marginal posterior probability distributions and profiles, which quantify the effective data-information content to resolve scattering/ geoacoustic structure.

At the first site scattering was assumed (*a priori*) to be dominated by interface roughness. The inversion results indicate well-defined roughness parameters in good agreement with existing measurements, and a multi-layer sediment profile over a high-speed (elastic) basement, consistent with independent knowledge of sand layers over limestone.

At the second site no assumptions were made about the scattering mechanism. The deviance information criterion indicated volume scattering to be the dominant scattering mechanism. The scattering parameters and geoacoustic profile are well resolved. The parameters and preference for volume scattering are consistent with a core extracted at the site which indicated a sediment layer which included large (0.1 m) stones underlying $\sim$1 m of mud at the seafloor.

As a final component of this thesis, a polynomial spline-based parameterization for trans-D geoacoustic inversion is developed for application to sites where sediment gradients (rather than discontinuous layers) dominate. The parameterization is evaluated using data for a third site on the Malta Plateau known to consist of soft mud with smoothly changing geoacoustic properties. The spline parameterization is compared to the standard stack-of-homogeneous-layers parameterization for the inversion of bottom-loss data. Inversion results for both parameterizations are in good agreement with measurements on a sediment core extracted at the site. However, the spline parameterization more accurately resolves the power-law like structure of the core density profile, and represents the preferred model according to the deviance information criterion.

Contents

List of Tables

List of Figures

ACKNOWLEDGEMENTS

I would like to thank:

My supervisor Stan Dosso, for mentoring, support, encouragement, patience and most importantly a willingness to look past irremediably bad spelling.

My coauthors Charles Holland and Jan Dettmer, For providing me with the Fortran codes and data that made this thesis possible.

Office of Naval Research, for funding this work (grant N00014-11-0213).

At the end of a miserable day, instead of grieving my virtual nothing, I can always look at my loaded wastepaper basket and tell myself that if I failed, at least I took a few trees down with me.

David Sedaris

Some humans would do anything to see if it was possible to do it. If you put a large switch in some cave somewhere, with a sign on it saying 'End-of-the-World Switch. PLEASE DO NOT TOUCH', the paint wouldn't even have time to dry.

Terry Pratchett

DEDICATION

This thesis is dedicated to Stephanie Isabelle Hovell nee Wheatley, born 9th May 1914 Sliema, Malta.

Chapter 1

Introduction

1.1 Overview

The overall objective of this thesis is to develop and study a remote acoustic sensing method for determining seabed scattering and geoacoustic parameters with rigorous uncertainty estimates. One of the primary uses of sonar is to detect and classify targets such as submarines, mines, and marine animals. In such cases, acoustic scattering from the seabed is a source of interference which degrades target signals and reduces sonar performance. An accurate model of seabed scattering in a particular environment can help mitigate this performance reduction.[1] In addition, scattering and geoacoustic properties can provide insight into the geological, oceanographic, and biological processes that formed the seabed. Hence, the development of bottom-scattering and geoacoustic databases is of great interest to the naval and oceanographic communities. Unfortunately, direct measurements of seabed scattering and geoacoustic parameters (e.g., sediment cores for geoacoustics and stereoscopic photography or laser imaging for scattering parameters) are expensive and time consuming,[1] and characterize relatively small lateral regions (patch sizes of order 10^0–10^1 m). Alternatively, long-range (10^3–10^4 m) reverberation-based approaches have low data information content and aggregate parameters estimates over their entire range. All of this indicates a significant need to develop remote *in-situ* sensing methods to determine seabed parameters at the meso-scale (10^1–10^2 m).

This thesis develops an acoustic remote sensing approach to quantify seabed scattering and geoacoustic parameters with rigorous uncertainties. This approach does not directly measure the desired parameters but instead estimates them as the solu-

tion to an inverse problem.[2–11] Forward and inverse problems may be broadly defined as data prediction and parameter estimation, respectively. That is, if the function f represents the physical process of interest (e.g., acoustic scattering or reflection) and $\mathbf{m}$ is a set of model parameters representing the system (seabed), then calculating predicted data $\mathbf{d}_{\text{pred}} = f(\mathbf{m})$ represents the forward problem, while solving the same system of equation with an observed data set $\mathbf{d}_{\text{obs}}$ for an estimate of the parameters $\hat{\mathbf{m}}$ is the inverse problem, i.e., $\hat{\mathbf{m}} = f^{-1}(\mathbf{d}_{\text{obs}})$. Note, however, an analytic form for f^{-1} rarely exists.

The acoustic data used to estimate the seabed scattering and geoacoustic properties in this thesis are measures of backscatter[1,12] and reflectivity.[13,14] For both data types an acoustic wave is transmitted though the water column until it encounters the seabed, where it interacts and scatters/reflects; the scattered or reflected component is then quantified. For backscatter data both the acoustic source and receiver are at the same position, and the measured quantity is the acoustic intensity as a function of angle (at the seabed) and frequency.[12,15] For reflection data the source and receiver are at different locations; the measurement is the ratio of reflected to incident wave energy as a function of angle (assuming specular reflection) and frequency.[15,16]

The physics used to model the data (i.e., the forward model) is based on first-order perturbation theory for the scattering data[1,17,18] and either the plane- or spherical-wave reflection coefficient for reflection data.[13,14] In general, acoustic scattering occurs at the water-sediment and sediment-layer interfaces as a result of interface roughness and impedance contrasts, and within the volume of sediment layers due to heterogeneities.[1] Thus, interface and volume scattering depend on parameters defining interface-roughness and volume-heterogeneity spatial spectra as well as geoacoustic profiles. Geoacoustic reflectivity depends only on the geoacoustic profile, but the conjecture here is that the inclusion of reflection-coefficient data in a joint inversion with backscatter data can improve the resolution and accuracy of scattering parameter estimates. Throughout this thesis the seabed is assumed to be range independent (laterally invariant) with relevant parameters varying with only depth.

An important issue is establishing the relative importance between scattering due to rough boundaries at the seafloor and within sediment-layer volumes due to heterogeneities. It has generally been assumed that interface scattering dominates.[1,19] However, recent analysis of long-range reverberation data suggests that volume scattering may be the dominant mechanism in several shallow-water regions.[19] In general, the two mechanisms (interface and volume scattering) produce different relations between

scattering strength and angle at a given frequency. Consequently, mis-identifying the scattering mechanism can result in a biased scattering model.

In this thesis, a Bayesian[20–23] approach is applied to the inverse problem. Bayesian inversion provides a powerful and flexible framework which has been used on a number of geophysical[3,24–34] and geoacoustic[7,35–45] problems. Bayesian methods provide a natural formulation to incorporate available prior information. They also provide rigorous nonlinear uncertainty (probability) distributions for the estimated parameters. The rigorous, quantitative approach to model selection and inversion for parameter estimates and uncertainties carried out here represent the most fundamental improvement over previous scattering inversion work which has generally applied optimization techniques, producing only point-estimates of parameters.[1,46,47]

1.2 A prior preference for Bayesian methods

The details of Bayesian inverse theory are given in the body of this thesis. However, it may be helpful at the outset to provide some basic discussion of Bayesian inference and the philosophy generally associated with it.

The most fundamental component of Bayesian statistics is the conceptualization of the unknown parameters as random variables. This conceptualization is linked to the understanding of probabilities as a degree of certainty or belief[20] instead of as a representation of the long-term frequency or propensity of a physical state.[48] The distinction may be subtle but it is important.

Intuitively, Bayesian inference consists of expressing an initial belief of the unknown model parameters in terms of a prior probability distribution and evolving the distribution through the introduction of data. The posterior (or post data) belief of the unknown parameters is described by a distribution called the posterior probability density (PPD), which contains both the prior and data information. More formally, Bayesian inference can be described in two steps: the first is to formulate the joint probability of the parameters and data, i.e., define $P(\mathbf{m}, \mathbf{d})$; the second is to condition the joint distribution on the data, i.e., calculate the PPD, $P(\mathbf{m}|\mathbf{d})$.[20,49] For Bayesian model selection it is convenient to write the joint distribution of the parameters and data as conditional upon the choice of model denoted $\mathcal{J}$, i.e., $P(\mathbf{m}, \mathbf{d}|\mathcal{J})$. Here the concept of a model is general and represents the choice of the governing physical theory (i.e., the forward model), a set of appropriate model parameters $\mathbf{m}$, and a statistical representation of the error process which together describe the system of

interest (e.g., acoustic scattering or reflection).

Formulating the joint distribution of the parameters and data requires the postulation of a parameter space and the specification of prior and likelihood functions that are well defined over this space. Criticisms of Bayesian statistical methods commonly attack the subjectivity of this stage of Bayesian inference.[50] While defending Bayesian inference in general is outside the scope of this work, this particular criticism can be easily answered in geophysical inversion where the assumptions made in defining the likelihood and priors for Bayesian inversion[33,51–53] are no more informative than those in non-Bayesian methods.[36,54,55]

Conditioning the joint distribution is done using Bayes rule[20]

$$P(\mathbf{m}|\mathcal{J},\mathbf{d}) = \frac{P(\mathbf{m}|\mathcal{J})P(\mathbf{d}|\mathbf{m},\mathcal{J})}{P(\mathbf{d}|\mathcal{J})}. \tag{1.1}$$

When the PPD is conditioned on the observed data, $\mathbf{d}$ is treated as known. Consequently Eq. (1.1) can be interpreted as

$$P(\mathbf{m}|\mathcal{J},\mathbf{d}) = \frac{\pi(\mathbf{m}|\mathcal{J})\mathcal{L}(\mathbf{m}|\mathcal{J})}{\mathcal{Z}(\mathcal{J})}, \tag{1.2}$$

where $\pi(\mathbf{m}|\mathcal{J})$ is the prior distribution, $\mathcal{L}(\mathbf{m}|\mathcal{J})$ is the likelihood function, and $\mathcal{Z}(\mathcal{J})$ is the Bayesian evidence. Commonly in nonlinear geophysical inverse problems the PPD cannot be described analytically and is approximated using numerical sampling techniques.[20,23,56–59] The most common sampling paradigm is Metropolis-Hasting (MH) sampling,[60] which produces a random walk through the parameter space. The walk is conducted such that the long-run distribution of the steps converge to the target distribution, e.g., the PPD. Figure 1.1 shows an example of MH sampling for a one-dimensional target distribution; the distribution of the steps converge to the target as the sampling progresses.

Bayesian evidence is the probability of observing the data marginalized over all possible parameter values, $\mathcal{Z}(\mathcal{J}) = P(\mathbf{d}|\mathcal{J}) = \int_{\mathbf{m}\in\mathcal{M}} P(\mathbf{d},\mathbf{m}|\mathcal{J})d\mathbf{m}$. It may be helpful to think of evidence as the zeroth-dimensional marginal density of the model. Alternatively, if the data are variable (i.e., not consider fixed at observed values) the Bayesian evidence is the marginal predictive density of the data, that is the distribution $P(\mathbf{d}|\mathcal{J})$.[20] As the name suggests, comparisons of Bayesian evidence are a strong indicator of which model has the greatest support by the data and prior information. The model with the highest evidence evaluated at the observed data, i.e., the model

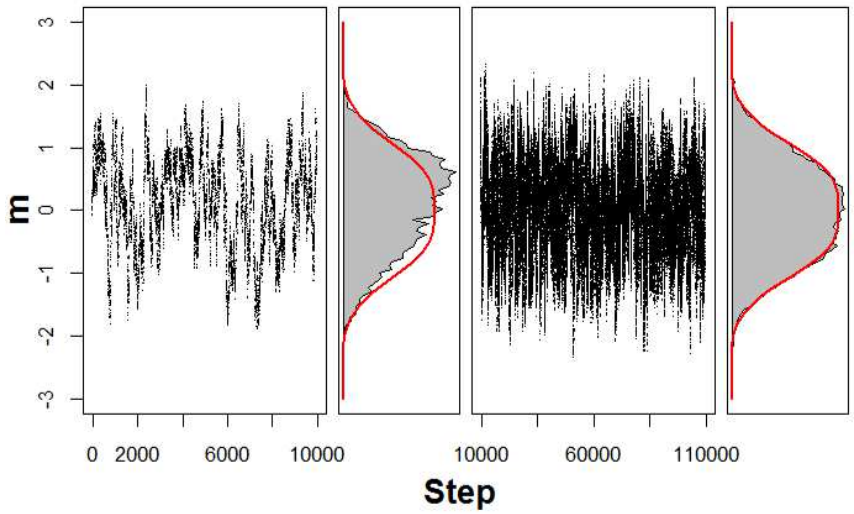

Figure 1.1: Samples (dots) and distribution for Metropolis-Hasting sampling after 10,000 (left) and 110,000 (right) iterations. The red line indicates the target distribution and the shaded area is the sampled distribution.

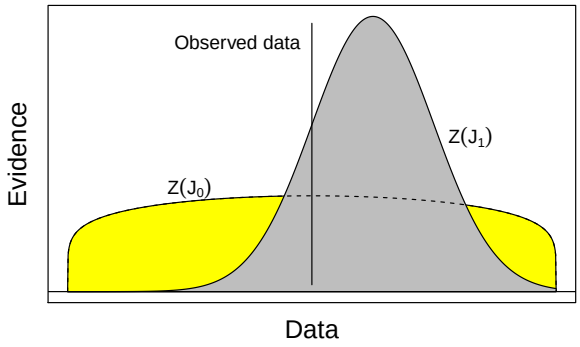

Figure 1.2: Schematic illustration of evidence verses data space for two models J_0 and J_1. The vertical line indicates the observed data (after Ref. [61]).

for which the observations are most probable, should be taken as the preferred model.

Bayesian evidence accounts for both data fit and model parsimony (i.e., a preference for simple models).[49,61] Parsimony is addressed naturally by penalizing needlessly flexible models, i.e., penalizing models that assign evidence over an unnecessarily large region of the data space. This penalty results from the fact that evidence is a normalized distribution over the data space; models that assign probability too widely over the data space will tend to have lower probabilities at the observed data than more focused models. Intuitively, overly-flexible, non-parsimonious models are unde-

sirable because they are not easily disproved. Figure 1.2 illustrates the evidence of two models, J_0 and J_1, as a function of the data space (schematically compressed into one dimension). Model J_0 is less parsimonious as $\mathcal{Z}(J_0)$ is more widely distributed than $\mathcal{Z}(J_1)$. Evidence compares the two models at the observed data (represented by the vertical line in Fig. 1.2). In this example, J_1 is found to have higher evidence (support from the data and prior) than J_0.

Evaluating Bayesian evidence is computationally difficult for many nonlinear problems.[49] Bayesian evidence computation was introduced to ocean acoustic inversion using reverse importance sampling.[35] However, the process there did not indicate a clear choice of preferred model. In addition reverse importance sampling is known to be unstable and inaccurate with large variance for evidence computation.[62] Other methods of computing Bayesian evidence have since been considered in acoustic inversion, including annealed importance sampling[63,64] and nested sampling.[65,66] These methods satisfactorily solved the model selection problem;[64,66] however, they are extremely computationally expensive for high-dimensional inverse problems and are not practical for this work.

To avoid the complexity of direct evaluating the evidence, this thesis uses the deviance information criterion[67] (DIC) and trans-dimensional (trans-D) sampling[33,53,68,69] to address model selection problems. The DIC uses posterior samples to evaluate the fit to data and complexity of a model. Unlike similar criteria, such as the Bayesian information criterion,[70] the DIC accounts for both parameter correlation and nonlinear (non-Gaussian) effects in the PPD when evaluating model complexity. Trans-D sampling does not explicitly evaluate the evidence of models but samples over a collection of models in direct proportion to their support by the evidence. Trans-D sampling is conducted here using the reversible jump Markov chain Monte Carlo (rjMCMC) algorithm,[33,68] an extension to MH sampling. The rjMCMC samples by creating a random walk through the joint parameter model space. At each step of the random walk the parameter values and/or the model (parameterization) may be perturbed. Thus the PPD represents a density function over both the parameterization and the parameter values. Trans-dimensional sampling has the advantage over discrete model selection methods in that the model selection uncertainty is accounted for in the parameter estimate uncertainty. It is this attribute that makes trans-D sampling most appropriate for geoacoustic inversion.

1.3 Outline of Thesis

The body of this thesis is composed of four chapters which correspond to four papers on scattering and/or reflection inversion. Details of the data acquisition, inversion procedures, and results are discussed in each chapter. It should be noted that since the papers were produced as stand-alone works, there is some repetition in the introductory material and theory across the chapters. The following provides a brief overview of this thesis.

Chapter 2 (Published as Ref. [71]) This chapter develops a powerful and general trans-D inversion framework which is applied to a simulation study of joint backscatter and reflection-coefficient inversion. The new developments for this inversion scheme include an empirical prior for geoacoustic parameters, the use of a Dirichlet partition prior which allows for the explicit evaluation of the PPD, and a trans-D sampling procedure for both the number of seabed layers and data residual auto-correlations. The results show that realistic acoustic data can resolve the geoacoustic and interface roughness (scattering) parameters. Joint inversion is found to improve the accuracy of the interface roughness parameters. The use of trans-D auto-regressive sampling is found to reduce posterior uncertainty in the scattering parameters.

Chapter 3 (Published as Ref. [72]) This chapter uses the inversion framework developed in Chapter 2 to estimate seabed interface roughness and geoacoustic parameters using measured data from a survey site on the Malta Plateau in the Mediterranean Sea. The estimated geoacoustic profile is found to be in good agreement with a high-resolution seismic survey at the site. The interface roughness is described in terms of a spatial power-law spectrum (von Karman spectrum) or root-mean-squared roughness and spatial correlation length. The roughness parameters are well resolved and agree with direct measurements of roughness for a variety of seabed locations.

Chapter 4 (To be submitted as Ref. [73]) This chapter develops a rigorous, quantitative, and objective method for determining the dominant seabed scattering mechanism for a particular backscatter data set based on trans-D sampling and the DIC. The volume scattering is modelled with first-order perturbation theory[18] and volume heterogeneities are assumed to be statistically independent of the interface roughness process. The approach is applied to six simulated

test cases where it accurately identifies the dominant scattering mechanism in five cases (the sixth case is ambiguous). The approach is also applied to measured data from the Malta Plateau (different site than used in Chapter 3) where volume scattering is determined as the dominant scattering mechanism.

Chapter 5 (Submitted as Ref. [74]) This chapter develops a novel seabed parameterization for geoacoustic inversion in sediment regimes dominated by gradient-based profiles (rather than discrete layers). The parameterization describes the geoacoustic profiles as discretized polynomial-splines. Each geoacoustic parameter is modelled by its own spline where number of nodes used to define each polynomial-spline is treated as an unknown in the inversion and is estimated using trans-D sampling. The spline parameterization is compared to a stack-of-homogeneous-layers parameterization for simulated data generated with a gradient model and for measured data from a site with a soft mud seabed known to involve gradients from a sediment core. The DIC is used to determine the preferred parameterization. For both simulated and measured data sets the spline parameterization is found to have the greatest support from the data.

Appendices A–E Appendix A describes the derivation of the empirical prior used for the geoacoustic parameters throughout this thesis. Appendix B proves the equivalence of two approaches to represent residual auto-correlation in nonlinear inversion. Appendix C describes novel approaches to chain thinning with Markov chain Monte Carlo sampling which were developed in the course of this work. Appendix D presents an adaptive sampling procedure that uses local gradient information to approximate the PPD for efficient MH sampling. Appendix E describes a method of creating unbiased samples of a target distribution from a collection of biased samples.

Chapter 2

Trans-dimensional joint inversion of seabed scattering and reflection data

This chapter examines joint inversion of acoustic scattering and reflection data to resolve seabed interface roughness parameters (spectral strength, exponent, and cutoff) and geoacoustic profiles. Trans-dimensional (trans-D) Bayesian sampling is applied with both the number of sediment layers and the order (zeroth or first) of auto-regressive parameters in the error model treated as unknowns. A prior distribution that allows fluid sediment layers over an elastic basement in a trans-D inversion is derived and implemented. Three cases are considered: scattering-only inversion, joint scattering and reflection inversion, and joint inversion with the trans-D auto-regressive error model. Including reflection data improves the resolution of scattering and geoacoustic parameters. The trans-D auto-regressive model further improves scattering resolution and correctly differentiates between strongly and weakly correlated residual errors.

2.1 Introduction

Ocean acoustic reverberation modelling and sonar performance predictions in shallow water require estimates of scattering parameters defining seafloor roughness. Direct measurement of roughness parameters (e.g., stereoscopic photography or laser imaging) is time consuming and expensive. Hence, the estimation of *in-situ* seabed rough-

ness from remote acoustic measurements is a problem of practical interest, but has received little attention to date. This chapter develops a trans-dimensional (trans-D) Bayesian inversion approach to estimate seabed scattering parameters and a layered geoacoustic model, as well as data error parameters, from multi-frequency acoustic scattering and/or reflection data. The approach is applied in simulations based on existing measurement techniques to evaluate the ability of the two data types, inverted separately and jointly, to resolve the seabed parameters.

Simulations represent an important initial step in developing effective measurement/inversion approaches, in that the true model is known and may be used to evaluate parameter estimates and uncertainties. Further, error processes can be controlled such that the physics of the acoustic measurements can be examined within specific assumptions (such as lateral homogeneity), without potential complicating factors which may arise in specific experiments. Realistic simulations presented here are based on a geoacoustic test bed located on the Malta Plateau in the Straits of Sicily. In particular, the true geoacoustic profile contains layering at a variety of scales, including fine scales below the resolution of the acoustic data. A flat basement layer (limestone) which supports shear waves is included. Errors include correlations, with both variance and covariance varying with frequency and, for reflection data, with angle.

Bayesian inversion estimates model parameters and uncertainties by quantifying the information content of data and prior, and has been applied widely to geoacoustic inverse problems.[7,35–40] Bayesian inversion is based on formulating the posterior probability density (PPD) which combines both data information, expressed in terms of a likelihood function, and prior information.[7,20,49,75,76] Joint inversion of independent data sets (e.g., scattering and reflection data) is accommodated naturally by formulating a joint likelihood function as the product of the individual likelihoods. An appropriate model parametrization (e.g., number of seabed layers resolved by the data) is generally not known in practice; this is addressed here by trans-D inversion[68,77] which provides an effective automated approach to Bayesian model selection[35,43,44,78,79] that has been applied to several problems in geophysics[26,28,32,80] and geoacoustics.[53,69] Trans-D inversion samples a set of models (which may vary in dimension) according to the support by the data and prior. In particular, partition modelling and the reversible jump Markov chain Monte Carlo (rjMCMC) algorithm are applied here for trans-D sampling over the number of seabed layers.[69] Including elastic (shear) parameters in the basement makes this layer distinct from the overlying

(fluid) sediment layers. Treating layers with different numbers or types of parameters has not been considered previously in trans-D inversion, and requires a novel formulation of the partition prior, which is developed here. Frequency-dependent residual error statistics, including variance and first-order auto-regressive [AR(1)] parameters to model covariance[33,53] of unevenly spaced data, are also sampled (marginalized) in the inversion. The AR(1) coefficients are estimated by trans-D inversion providing an efficient data-driven process to include AR(1) parameters at frequencies where they are required (typically low frequencies) but to omit them where not required to avoid over- or under-parameterizing the error model. The combination of trans-D sampling and hierarchical error modelling provides a rigorous and general inversion approach.

Seabed acoustic scattering data are dependent on both the two-dimensional (2D) seafloor roughness and seabed reflectivity, which is itself dependent on the sub-bottom geoacoustic profile. Since little quantitative work on scattering inversion has been reported, it is of interest to examine to what extent realistic scattering data can resolve roughness parameters and geoacoustic profiles, with resolution characterized here in terms of marginal probability distributions. Possible improvements in parameter resolution in joint scattering/reflection inversion over scattering-only inversion are also of interest. Joint inversion with trans-D error modelling is shown to improve resolution of scattering parameters, even though the reflection data provide no scattering information, by reducing the uncertainty of geoacoustic parameters.

2.2 Forward Models and Data

Two forward models are used here to compute mono-static scattering and spherical-wave reflection data, and are applied to a seabed model consisting of a layered half-space as shown in Fig. 2.1. The top (zeroth) layer is seawater and is assumed to be homogeneous and isotropic with known properties. The seabed is a series of j flat homogeneous layers, terminated by a homogeneous semi-infinite basement ($j+1$ layers with j interfaces in all). Layer properties include interface depth z (the lower boundary of a layer), sound speed c, density ρ, and attenuation α. As there are $j+1$ layers and only j interfaces, the jth layer does not have an associated interface depth; the depth of the interface between the jth layer and the basement is denoted z_b and considered an attribute of the basement (an important distinction, addressed in Sec. 2.4). In addition, the basement is assumed to be elastic with a shear-wave speed c_s and attenuation α_s. The only difference between the seabed model for scattering and

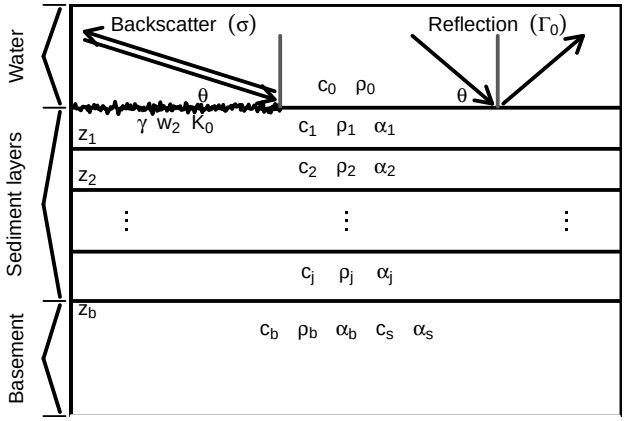

Figure 2.1: Schematic diagram of the seabed model. Parameters are defined in the text.

reflection is that the water-sediment interface is assumed to be rough for scattering and planar for reflection.

The parametrization and forward models for calculating simulated data are consistent with measurements which have been made at the Malta plateau.[12,55] The geoacoustic parameters of the true model are chosen to represent sand layers over a limestone basement. The layering structure of the true geoacoustic model, shown in Fig. 2.2, is more complicated than the data can resolve (∼100 layers) to allow a meaningful evaluation of the trans-D procedure (this figure also includes the optimal profile for an inversion discussed later). The true shear-wave speed and attenuation of the basement are $c_s = 2200$ m/s and $\alpha_s = 0.1$ dB/m/kHz.

In addition to the geoacoustic model it is also necessary to define the residual-error distribution for both scattering and reflection data. The data residuals are assumed to be multivariate Gaussian distributed (as observed for measured reflection-coefficient data[38,41,69]). Residuals at different frequencies are assumed to be independent; however, residuals at the same frequency are not assumed independent over angle.

2.2.1 Scattering kernel

The scattering kernel[17] considered here defines the mono-static acoustic backscatter (as a function of angle θ and frequency f) from a single rough interface between two

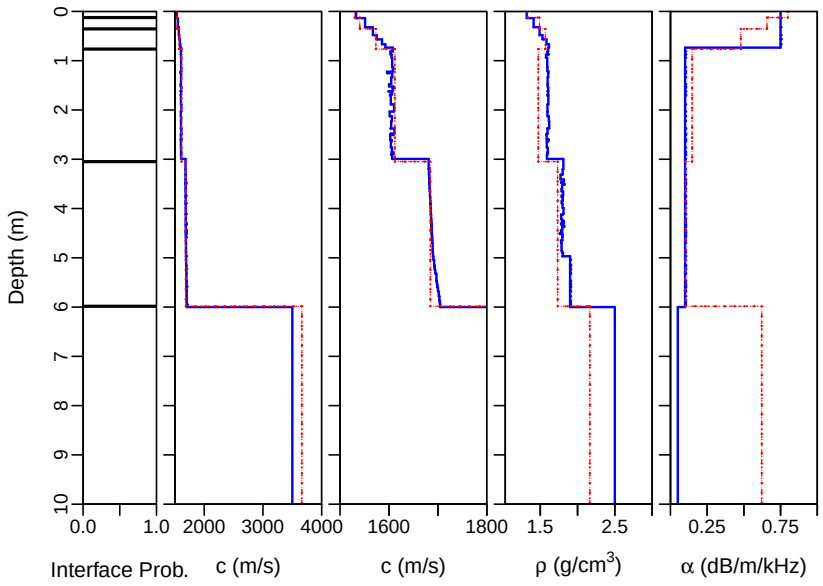

Figure 2.2: True geoacoustic profile (solid line) and maximum *a posteriori* model profile (dotted line) for trans-D auto-regressive inversion (sound speeds shown at two scales).

fluid layers (water and first sediment layer) over a layered medium and is given by

$$\sigma'(\theta, f) = \frac{k_0^4 \left|1 + R(\theta, f)\right|^4}{4} W(2\mathbf{K}) \quad \times \\ \left| 1 - \left(\frac{k_1}{k_0}\right)^2 \frac{\rho_0}{\rho_1} + \left(1 - \frac{\rho_0}{\rho_1}\right)\left(\cos^2\theta + \frac{\rho_0}{\rho_1}\sin^2\theta\right)\left(\frac{1 - R(\theta, f)}{1 + R(\theta, f)}\right)^2 \right|^2, \tag{2.1}$$

where k_0 and k_1 are the wavenumbers in the water and first sediment layers, and R is the plane-wave reflection coefficient for the $j+1$ layer seabed, which is evaluated recursively and accounts for the elastic basement.[14] In Eq. (2.1) W defines the 2D spatial-roughness power spectrum of the water-sediment interface given by

$$W(\mathbf{K}) = w_2 \left(|\mathbf{K}|^2 + K_0^2\right)^{-\gamma/2}, \tag{2.2}$$

where γ, w_2, and K_0 are known as the spectral exponent, spectral strength, and spectral cutoff, respectively, and $\mathbf{K}$ is the transverse component of the wave vector with magnitude $|\mathbf{K}| = k_0 \cos\theta$. Backscatter is considered in decibels,

$$\sigma(\theta, f) = 10 \log_{10} \sigma'(\theta, f) \quad \text{(dB)}. \tag{2.3}$$

Table 2.1: True residual error parameters for scattering (top) and reflection (bottom) data.

Frequency (Hz)	600	900	1200	1800	2400	3600
$\mathbf{S}_s$ (dB)	1.5	1.5	1.5	2.5	1.5	1.5
$\mathbf{a}_s$	0.6	0.5	0.3	0.2	0.1	0
Frequency (Hz)	630	800	1000	1600	2500	4000
$\mathbf{S}_r^L$	0.1	0.1	0.1	0.1	0.05	0.05
$\mathbf{S}_r^H$	0.03	0.03	0.05	0.03	0.03	0.03
$\mathbf{a}_r$	0.7	0.6	0.5	0.2	0.1	0

The scattering errors are assumed to be correlated such that the correlation is non-negative and decreases exponentially with angular separation. To avoid the computational expense of repeatedly taking inverses and determinants of the covariance matrix in the inversion (see Sec. 2.3), the residual correlation structure is modelled using a first-order AR(1) process[33,53,81] given by

$$r_i = {a_s}^{\Delta\theta_i} r_{i-1} + e_i, \tag{2.4}$$

where the r_i are the residuals (indexed over angle), a_s is the AR(1) coefficient, $\Delta\theta_i = \theta_i - \theta_{i-1}$, and the e_i are the total residuals, which are identical independently-distributed (IID) Gaussian random variables with zero mean and standard deviation S_s (frequency subscripts in Eq. (2.4) are omitted for simplicity).

For the simulated data (Fig. 2.3), the true scattering parameters values are selected to be consistent with measurements:[1] $\gamma = 3.15$, $w_2 = 0.002$, and $K_0 = 1.5$ 1/m. Simulated backscatter data are generated at frequencies of 600–3600 Hz and grazing angles of 6–19° (the angular range is intentionally limited to reduce the effects of sub-surface scatterers and allow seafloor interface scattering to be isolated). Data errors are Gaussian at each frequency and the AR(1) parameters decrease with frequency (as often observed) such that only errors at 600 and 900 Hz are significantly correlated. The standard deviations and AR(1) parameters of the true model are listed in Table 2.1.

2.2.2 Spherical-wave reflection coefficient

Reflection-coefficient data are modelled here using a spherical-wave reflection model to accommodate significant penetration depths.[42] Spherical-wave reflection coefficients

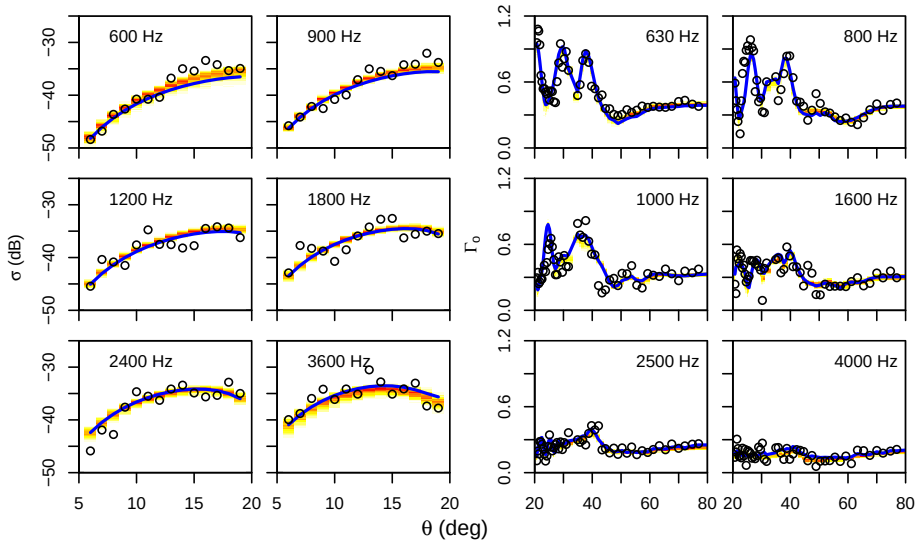

Figure 2.3: Simulated noisy scattering (left) and reflection-coefficient (right) data (∘) and marginal predicted data from joint trans-D auto-regressive inversion (shading). Solid lines indicate noise-free simulated data.

$\Gamma_\circ$ for an arbitrary N-layer half-space are computed as a superposition of plane waves (the Sommerfeld integral)[13]

$$\Gamma_\circ(\theta, f) = \frac{ik_0\sqrt{x^2+z^2}}{\exp\left(ik_0\sqrt{x^2+z^2}\right)} \int_0^{\pi/2-i\infty} J_0(k_0 x\cos\theta') \exp\left(ik_0 z\sin\theta'\right) R(\theta', f)\cos\theta' d\theta'. \tag{2.5}$$

In Eq. (2.5) x and z are the horizontal offset (range) and vertical offset, respectively, and J_0 is the zeroth-order Bessel function. The integral is computed numerically using Simpson's rule.[82] The complex exponential and Bessel function are environment independent, and are pre-computed for an array of argument values.

The covariance structure of the reflection data is similar to that of the scattering data,

$$r_i = {a_r}^{\Delta\theta_i} r_{i-1} + e_i, \tag{2.6}$$

where r_i are the reflection residuals and a_r is the AR(1) parameter. However, in this case the e_i are not IID in that their assigned standard deviations change at an angle $\theta_C = 50°$ (approximately the critical angle), since errors commonly change structure at the critical angle. Thus, if $\theta_i < \theta_C$ a low-angle standard deviation S_r^L is used, otherwise a high-angle value S_r^H is used. The data are third-octave band averages from 630–4000 Hz with an angular range of 20–85° (non-uniformly spaced as obtained

in practical measurements[12]). Errors are Gaussian and correlations decrease with frequency (Table 2.1, Fig. 2.3)

2.3 Bayesian inversion

This section provides a brief overview of Bayesian inference as applied to trans-D geoacoustic inversion; for a more complete description of the approach, see Refs. [26, 28,32,69]. In Bayesian inversion, model parameters are considered random variables with distributions that evolve, with the addition of data, from the prior distribution to the PPD. Let $\mathbf{d}$ be the vector of N observed data (comprising the scattering and/or reflection measurements at multiple frequencies and angles) and let $\mathcal{J}$ be a countable set (indexed by j) specifying the choice of parametrization (e.g., number of seabed sediment layers), with a set of parameter values denoted $\mathbf{m}_j$ (comprising all unknown roughness, geoacoustic, and error parameters). Using Bayes' rule the trans-D PPD can be written[68]

$$P(j,\mathbf{m}_j|\mathbf{d}) = \frac{\pi(\mathbf{m}_j)\mathcal{L}(\mathbf{m}_j)}{\mathcal{Z}} = \frac{P(j)P(\mathbf{m}_j|j)P(\mathbf{d}|j,\mathbf{m}_j)}{\sum_{j'\in\mathcal{J}}\int P(j')P(\mathbf{m}'_{j'}|j')d\mathbf{m}'_{j'}P(\mathbf{d}|j',\mathbf{m}'_{j'})d\mathbf{m}'_{j'}}\,, \quad (2.7)$$

where $\pi(\mathbf{m}_j) = P(j)P(\mathbf{m}_j|j)$ is the prior distribution of j and $\mathbf{m}_j$, $\mathcal{L}(\mathbf{m}_j) = P(\mathbf{d}|j,\mathbf{m}_j)$ is the likelihood of the parameter vector [$P(\mathbf{d}|j,\mathbf{m}_j)$ interpreted as a function of $\mathbf{m}_j$ for a fixed $\mathbf{d}$], and $\mathcal{Z}$ is the total evidence of the ensemble of models. Here the data residuals are assumed to be Gaussian distributed, leading to the likelihood function

$$\mathcal{L}(\mathbf{m}_j) = \frac{1}{(2\pi)^{N/2} \mid \mathbf{C}_d \mid^{1/2}} \exp\left(-\frac{1}{2}\left[\mathbf{d}-\mathbf{d}(\mathbf{m}_j)-\mathbf{d}(\mathbf{a})\right]^{\top}\mathbf{C}_d^{-1}\left[\mathbf{d}-\mathbf{d}(\mathbf{m}_j)-\mathbf{d}(\mathbf{a})\right]\right), \quad (2.8)$$

where $\mathbf{C}_d$ is a diagonal covariance matrix with ith diagonal element equal to the variance of the ith total residual [e_i from Eqs. (2.4) and (2.6) for scattering and reflection data, respectively].[33] In Eq. (2.8) the vector $\mathbf{d}(\mathbf{m}_j)$ represents the data predicted by the forward models [Eqs. (2.3) and (2.5) for scattering and reflection data, respectively] for parameters $\mathbf{m}_j$. The vector $\mathbf{d}(\mathbf{a})$ represents the AR(1) process; i.e., for a given data type and frequency $d_i(\mathbf{a}) = a^{\Delta\theta_i} r_{i-1}(\mathbf{m}_j)$, where $\mathbf{r}(\mathbf{m}_j) = \mathbf{d} - \mathbf{d}(\mathbf{m}_j)$.

The trans-D PPD is sampled using the rjMCMC algorithm, which creates a Markov chain that converges to the PPD.[49] Let $\mathbf{m}_j$ be the current Markov chain state and $Q(\mathbf{m}'_{j'}|\mathbf{m}_j)$ be the proposal distribution by which a new state $\mathbf{m}'_{j'}$ is gener-

ated. The proposed model can represent a perturbation of the parameters of $\mathbf{m}_j$ or a change (jump) in dimension of $\mathbf{m}_j$, i.e., $j' \neq j$. The proposed state $\mathbf{m}'_{j'}$ is accepted with probability

$$A = \min\left[1, \frac{\pi(\mathbf{m}'_{j'})}{\pi(\mathbf{m}_j)} \frac{\mathcal{L}(\mathbf{m}'_{j'})}{\mathcal{L}(\mathbf{m}_j)} \frac{Q(\mathbf{m}_j|\mathbf{m}'_{j'})}{Q(\mathbf{m}'_{j'}|\mathbf{m}_j)} |\mathbf{J}|\right], \tag{2.9}$$

where $\mathbf{J}$ is the Jacobian of the diffeomorphism between the parameter spaces associated with $\mathbf{m}_j$ and $\mathbf{m}'_{j'}$. For the common case of fixed-dimensional (fixed-D) inversion with uniform prior and symmetric proposal, Eq. (2.9) simplifies to the likelihood ratio

$$A = \min\left[1, \frac{\mathcal{L}(\mathbf{m}'_{j'})}{\mathcal{L}(\mathbf{m}_j)}\right]. \tag{2.10}$$

2.4 Prior information

In denoting the parametrization of a model as $\mathbf{m}_j$, the subscript indicates that the number of parameters of the model depends on j, the number of sediment layers. The model vector is a list of parameter vectors, $\mathbf{m}_j = (j, \mathbf{z}_j, \boldsymbol{\beta}, \boldsymbol{\Sigma}, \boldsymbol{\chi}_j, \mathbf{S}_s, \mathbf{S}_r, \mathbf{a}_s, \mathbf{a}_r)$, where $\mathbf{z}_j$ represents the sediment partition; $\boldsymbol{\beta}$ represents the basement parameters c_b, α_b, c_s, α_s, and ρ_b; $\boldsymbol{\Sigma}$ represents the scattering parameters γ, w_2, and K_0; $\boldsymbol{\chi}_j$ contains the three vectors of sediment parameters $\mathbf{c}_j$, $\boldsymbol{\rho}_j$, and $\boldsymbol{\alpha}_j$; $\mathbf{S}_s$ and $\mathbf{S}_r = \left(\mathbf{S}_r^L, \mathbf{S}_r^H\right)$ contain the standard deviations for scattering and reflection-coefficient data, respectively; and $\mathbf{a}_s$ and $\mathbf{a}_r$ contain the AR(1) parameters for scattering and reflection data.

The prior distribution $P(\mathbf{m}_j)$ can be written as a hierarchical distribution $P(\mathbf{m}_j) = P(\mathbf{z}_j|j, \boldsymbol{\beta})P(\boldsymbol{\chi}_j|j)P(\boldsymbol{\Sigma})P(\boldsymbol{\beta})P(j)P(\mathbf{S}_s)P(\mathbf{S}_r)P(\mathbf{a}_s)P(\mathbf{a}_r)$. The independent distributions are assumed to be uniform over some interval of width $\Delta* = *_U - *_L$ ($*$ represents an arbitrary parameter); the upper and lower parameter limits used here are given in Table 2.2. The conditional prior distribution for the physical parameters of the sediment layers is thus $P(\boldsymbol{\chi}_j|j) = (\Delta c\, \Delta\alpha\, \Delta\rho)^{-j}$ when the parameters are within the prior bound and zero otherwise. A collection of laboratory and *in-situ* measurements[83,84] is used to create joint priors for ρ and c for both the sediment and basement layers, as shown in Fig. 2.4.

The interpretation of the conditional prior distribution for the partition $P(\mathbf{z}_j|j, \boldsymbol{\beta})$ is more complicated and requires a novel formulation developed here. Let $\mathbf{z}'_j = \mathbf{z}_j/z_b$, then P is assumed to be a Dirichlet distribution, $P(\mathbf{z}'_j|j, \boldsymbol{\beta}) = \mathrm{Dir}(\mathbf{z}'_j|p_1, p_2, ..., p_j)$,

Table 2.2: Lower and upper prior bounds (LB and UB) for basement, sediment, scattering, and error parameters.

Sediment			Scattering		
Parameter	LB	UB	Parameter	LB	UB
c (m/s)	1450	2100	γ	2	4
α (dB/m/kHz)	0	1	w_2	10^{-5}	10
ρ (g/cm^3)	1.20	2.25	K_0 (1/m)	10^{-5}	32
Basement			Error		
Parameter	LB	UB	Parameter	LB	UB
z_b (m)	0	10	s_s (dB)	0	6
ρ_b (g/cm^3)	1.20	3.00	a_s	0	1
c_b (m/s)	1500	6000	s_s^L	0	0.2
α_b (dB/m/kHz)	0	1	s_s^H	0	0.2
c_s (m/s)	0	$c_b/\sqrt{2}$	a_r	0	1
α_s (dB/m/kHz)	0	1			

where $p_1, \ldots, p_j$ are the Dirichlet parameters. The Dirichlet distribution is a generalization of the binomial distribution that describes the probability of the partition of one unit into j parts.[20] The distribution $P(\mathbf{z}_j|j, \boldsymbol{\beta})$ is found by making a variable transform of $\mathbf{z}'_j$ to $\mathbf{z}_j$. In addition, it is assumed that all possible partitions are equally probable, which is equivalent to assuming that all Dirichlet parameters are unity. Thus $P(\mathbf{z}_j|j, \boldsymbol{\beta}) = (j-1)! z_b^{-(j-1)}$.

In other partition inversion work[28,49,69,80] an alternative method of describing the prior, referred to as the "grid trick," is used. The grid trick assumes a discrete set of possible partition locations corresponding to an underlying grid, such that $P(\mathbf{z}_l|l) = (l!(G-l)!)/G!$, where G is the number grid points and l is the number of layers (because the interface depth of the basement is interpreted differently in the two methods the definition of the number of layers is also different, $l = j - 1$). It is then found that the number of grid points cancels out in the acceptance probability, Eq. (2.9), and consequently an explicit value for G is not required. However the grid trick cannot be used here since the elastic basement is distinct from overlying fluid layers. The Dirichlet prior developed above precludes the need for a fictitious grid, and even if the maximum depth of the partition is known, it has the advantage that it allows for the prior (and consequently the posterior) probability of a model to be evaluated explicitly (not normally possible with the grid trick). This is advantageous in application which require selecting a "best" (i.e., most-probable) model.

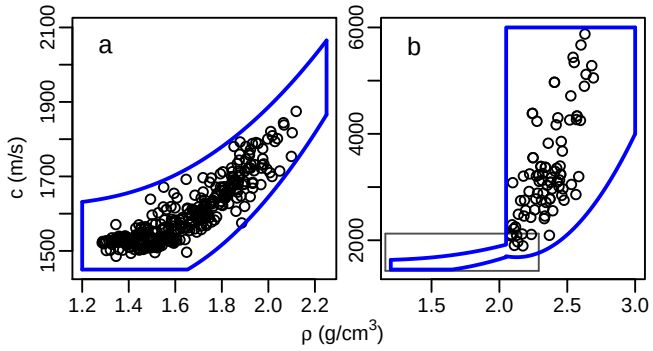

Figure 2.4: Joint prior bounds for c and ρ for (a) sediments, and (b) basement. Empirical [83,84] data are plotted as circles. The grey rectangle in (b) shows the extent of the sediment prior in the basement.

2.5 Sampling scheme

In trans-D inversion, the proposal distribution Q randomly selects between three types of moves at each step: perturb, Q_P; birth, Q_B; and death, Q_D. If Q_P is selected, all parameters in the model are updated. If Q_B is selected, an interface is proposed at a random depth, and the geoacoustic parameters of the proposed layer are defined by perturbing the previous values at that depth. Finally, if Q_D is selected, a randomly-chosen sediment layer is removed. The acceptance probability, Eq. (2.9), can thus be rewritten as

$$A = \min\left[1, \frac{\pi(\mathbf{m}'_{j'})}{\pi(\mathbf{m}_j)} \frac{\mathcal{L}(\mathbf{m}'_{j'})}{\mathcal{L}(\mathbf{m}_j)} \frac{Q_P(\mathbf{m}_j|\mathbf{m}'_{j'}) + Q_B(\mathbf{m}_j|\mathbf{m}'_{j'}) + Q_D(\mathbf{m}_j|\mathbf{m}'_{j'})}{Q_P(\mathbf{m}'_{j'}|\mathbf{m}_j) + Q_B(\mathbf{m}'_{j'}|\mathbf{m}_j) + Q_D(\mathbf{m}'_{j'}|\mathbf{m}_j)}\right]. \quad (2.11)$$

Consider first the perturbation step. Here all parameters are updated sequentially using a symmetric proposal distribution so the proposal ratio is unity. Since the prior distribution of interface depths for any $\mathbf{m}_j$ is not uniform, even given j, the priors do not generally cancel, and the prior ratio must be evaluated. The prior ratio for a perturbation move is

$$\left(\frac{\pi(\mathbf{m}'_{j'})}{\pi(\mathbf{m}_j)}\right)_P = \left(\frac{z_b}{z'_b}\right)^{(j-1)} \quad (2.12)$$

and the acceptance probability for a perturbation step can be written as

$$A_P = \min\left[1, \left(\frac{z_b}{z'_b}\right)^{(j-1)} \frac{\mathcal{L}(\mathbf{m}'_j)}{\mathcal{L}(\mathbf{m}_j)}\right]. \quad (2.13)$$

Both the prior and proposal ratios for the birth step do not cancel and must be evaluated. The prior ratio for the birth step is

$$\left(\frac{\pi(\mathbf{m}'_{j'})}{\pi(\mathbf{m}_j)}\right)_B = \left(\frac{H_c}{\Delta c\,\Delta\alpha\,\Delta\rho}\frac{j}{z_b}\right), \tag{2.14}$$

where H_c is the product of the uniform priors for ρ and c divided by the total area of their joint prior. The proposal ratio for the birth step is

$$\left(\frac{Q_D(\mathbf{m}_j|\mathbf{m}'_{j'})}{Q_B(\mathbf{m}'_{j'}|\mathbf{m}_j)}\right)_B = \frac{z_b}{j}\frac{1}{Q_n(\boldsymbol{\chi}'|\boldsymbol{\chi})}, \tag{2.15}$$

where $\boldsymbol{\chi}'$ and $\boldsymbol{\chi}$ are vectors of geoacoustic parameters for the new layer and the previous values at the depth of the proposed interface, and here Q_n is a multi-variate Gaussian distribution centered at $\boldsymbol{\chi}$. The acceptance probability for the birth step is

$$A_B = \min\left[1, \frac{H_c}{\Delta c\,\Delta\alpha\,\Delta\rho}\frac{1}{Q_n(\boldsymbol{\chi}'|\boldsymbol{\chi})}\frac{\mathcal{L}(\mathbf{m}'_{j'})}{\mathcal{L}(\mathbf{m}_j)}\right]. \tag{2.16}$$

The acceptance probability of the death step can be found in a similar way to be

$$A_D = \min\left[1, \frac{(\Delta c\,\Delta\alpha\,\Delta\rho)}{H_c}Q_n(\boldsymbol{\chi}'|\boldsymbol{\chi})\frac{\mathcal{L}(\mathbf{m}'_{j'})}{\mathcal{L}(\mathbf{m}_j)}\right]. \tag{2.17}$$

In addition to inverting for the geoacoustic parameters trans-dimensionally the same framework can be applied to the AR(1) parameters such that the parameters can be added or removed to avoid over- or under-parameterizing the error model. There are again three moves: birth, death, and perturbation, with acceptance probabilities (a_B, a_D, and a_P) and proposal distributions (q_B, q_D, and q_P), respectively. To describe the trans-D AR(1) procedure, consider a case with only one AR(1) parameter and a fixed number of other model parameters. The model subscript is taken to refer to the status of the AR(1) parameter with $\mathbf{m}_0$ representing a model without the AR(1) and $\mathbf{m}_1$ represents a model with the AR(1) parameter. A model without the AR(1) parameter always proposes a birth and the new value is sampled from the prior distribution, $q_B = p(\mathbf{m}'_1|\mathbf{m}_0) = \Delta a^{-1}$. Conversely, a model with the AR(1) parameter randomly proposes a death or perturbation move: $q_D = p(\mathbf{m}'_0|\mathbf{m}_1) = 0.5$ and $q_P = 0.5\tilde{q}_P(\mathbf{m}'_1|\mathbf{m}_1)$, where $\tilde{q}_P$ is a Gaussian proposal distribution centered at the current AR(1) value. Using these definitions the acceptance probability for birth,

death, and perturbation moves, respectively, are

$$a_B = \min\left[1, \frac{1}{2}\frac{\mathcal{L}(\mathbf{m}_1')}{\mathcal{L}(\mathbf{m}_0)}\right], \tag{2.18}$$

$$a_D = \min\left[1, 2\frac{\mathcal{L}(\mathbf{m}_0')}{\mathcal{L}(\mathbf{m}_1)}\right], \tag{2.19}$$

$$a_P = \min\left[1, \frac{\mathcal{L}(\mathbf{m}_1')}{\mathcal{L}(\mathbf{m}_1)}\right]. \tag{2.20}$$

The convergence of rjMCMC algorithms can be slow. To improve mixing (speed of convergence) population methods can be used. These are based on drawing samples from the product of multiple distributions, at least one of which is the PPD.[53,57] Markov chains are allowed not only to wander within a given distribution but also to interchange (swap) with chains in other distributions.

The choice of additional distributions used here is

$$P_T(\mathbf{m}_j) = \frac{\pi(\mathbf{m}_j)\mathcal{L}(\mathbf{m}_j)^{1/T}}{\sum_{j'\in\mathcal{J}}\int_{\mathcal{M}}\pi(\mathbf{m}_{j'}')\mathcal{L}(\mathbf{m}_{j'}')^{1/T}d\mathbf{m}_j'}, \tag{2.21}$$

where T is known as the sampling temperature. Population-based sampling using this collection of distributions is called parallel tempering.[85,86] Equation (2.21) can be interpreted as the standard PPD with the likelihood raised to the power $1/T$. If $T > 1$ the significance of the data relative to the prior is diminished; if $T < 1$ the significance of the data is exaggerated. Models selected from distributions with $T > 1$ tend to under-fit the data and have j values lower than those for models sampled from the PPD; models sampled from distributions with $T < 1$ tend to over-fit the data and have higher j values.

The acceptance probability for a parallel tempering swap move is[57]

$$A_S = \min\left[1, \left(\frac{\mathcal{L}(\mathbf{m}_{j'}')}{\mathcal{L}(\mathbf{m}_j)}\right)^{(1/T-1/T')}\right]. \tag{2.22}$$

Equation (2.22) assumes that swapping partners are selected such that the probability of one chain picking another chain as its partner is the same as the reverse; if this condition is met, any system for selecting partners is allowed. Inversions in this chapter were run in parallel on 31 groups each consisting of 10 Markov chains. Only Markov chains within the same group interact. Each group had one chain with $T = 1$,

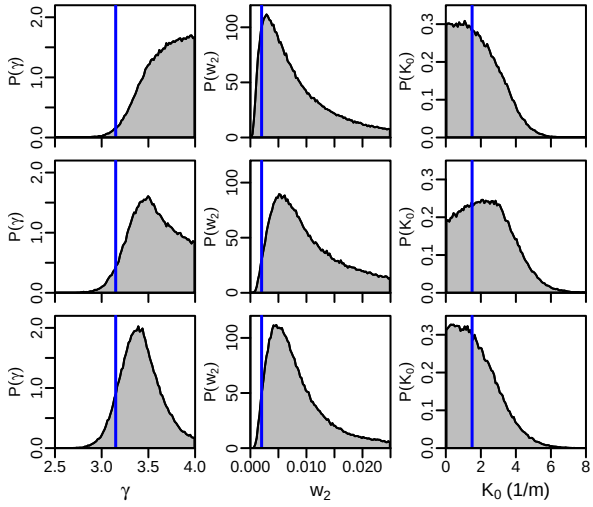

Figure 2.5: Marginal posterior distributions of the scattering parameters for top: scattering-only inversion; middle: joint inversion; bottom: joint TDAR inversion. Vertical lines indicate the true values.

two chains with $T < 1$, and seven chains with $T > 1$. The T values form a geometric progression $T_{i+1}/T_i = 1.175$.

2.6 Inversion results

Three inversions of the simulated data are considered here. The first uses only the scattering data, the second uses both scattering and reflection data, and the third uses both data sets with trans-D sampling of the AR(1) parameters. These inversions are referred to as scattering-only, joint, and joint trans-dimensional auto-regressive (TDAR), respectively. In each case, approximately 2,000,000 samples were collected at $T = 1$ and ∼400,000 samples from the start of the chains were deleted (burn-in). The remaining samples were chain thinned by a factor of 8 to reduce the autocorrelation.

2.6.1 Posterior probability density

The marginal posterior distributions of the scattering parameters for the three inversions are shown in Fig. 2.5. The marginals for the scattering-only inversion for γ and K_0 are not well determined within their prior bounds; the data are unable to differentiate γ from the prior bound of 4 or K_0 from 0. The joint inversion improves the

Table 2.3: Standard deviation about true values for scattering parameters.

Inversion	γ	w_2	K_0 (1/m)
Scattering-Only	0.562	0.0139	1.278
Joint	0.465	0.0188	1.655
TDAR	0.342	0.0120	1.277

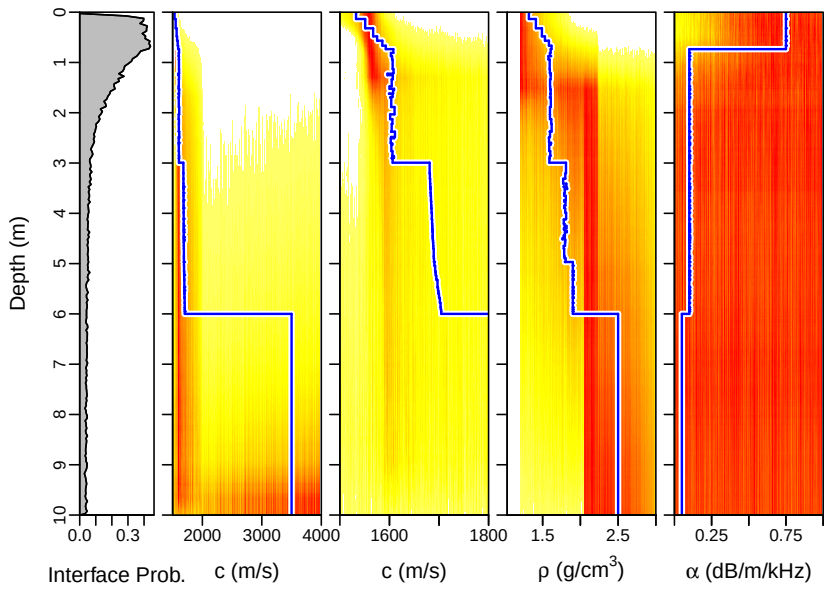

Figure 2.6: Marginal posterior geoacoustic profiles from scattering-only inversion (sound speeds shown at two scales). Solid lines indicate true profiles. Probability values are normalized independently for each depth for display purposes.

resolution of γ; the marginal now has a clear mode centered near the true value. However, the joint inversion appears to slightly degrade the resolution of w_2. The TDAR inversion further improves the resolution of γ. Table 2.3 gives the standard deviation about the true solution for scattering parameters; the TDAR inversion provides the highest accuracy for all parameters.

Figures 2.6 and 2.7 show the marginal profiles for geoacoustic parameters from scattering-only and TDAR inversions, respectively (the joint-inversion profile is indistinguishable from the TDAR profile). For the scattering-only inversion, the near-surface speed and density values are well determined and there is an indication of the positive gradient over the top $\sim$1 m. Below 3-m depth the scattering data are unable to resolve any structure. High attenuation is indicated to 0.5 m depth; below this there is essentially no information.

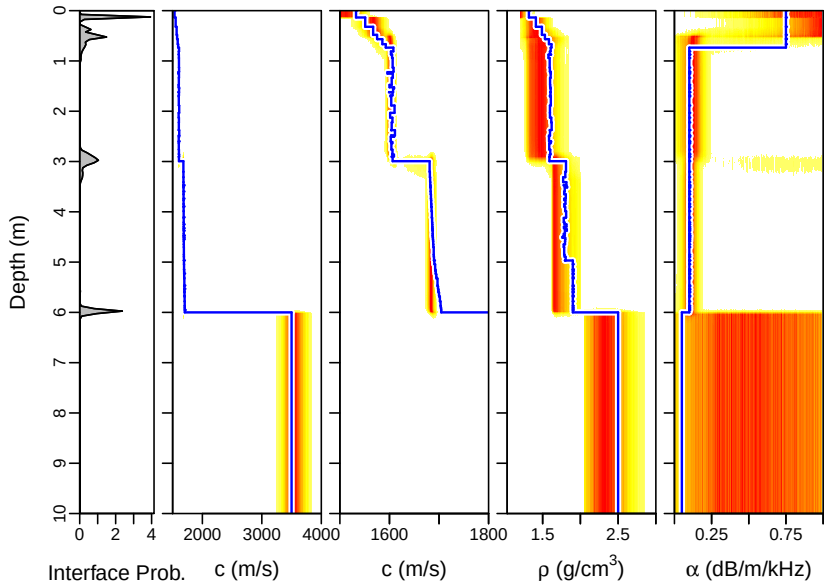

Figure 2.7: Marginal posterior geoacoustic profiles from joint TDAR inversion (joint inversion results are essentially identical). Solid lines indicate true profiles. Probability values are normalized independently for each depth for display purposes.

The joint inversions (Fig. 2.7) have much smaller posterior uncertainty than the scattering-only inversion, particularly at depth. The joint inversions resolve c and ρ for the entire profile depth; α is resolved to the depth of the basement (6 m). In particular, the joint inversions follow the gradient in c from 0–0.9 m, and find the discontinuities at 3 and 6 m. Speed uncertainties are small (relative to their prior bounds). The density and attenuation profiles agree well with the true model but have larger uncertainties than the speed profile. The joint inversions resolve the basement interface depth but not the slight gradient in c from 5-6 m depth.

The marginal posterior distributions of the basement parameters for the joint TDAR inversion are shown in Fig. 2.8 (scattering-only inversion results are not shown as parameters are unresolved). The shear-wave speed is well determined (marginal width $\sim$75 m/s) with significant probability at the true value. The compressional-wave speed is less well determined (marginal width $\sim$800 m/s), but does also have significant probability at the true value. Since high c_b and c_s values are resolved, the data clearly identify the basement as limestone rather than sediment. The better resolution of c_s than c_b results from the smaller contrast between the sediment sound speed (c_j) and c_s than between c_j and c_b. In cases where the ratio between c_j and c_s is greater than between c_j and c_b, c_b is observed to be better resolved than the c_s.[76]

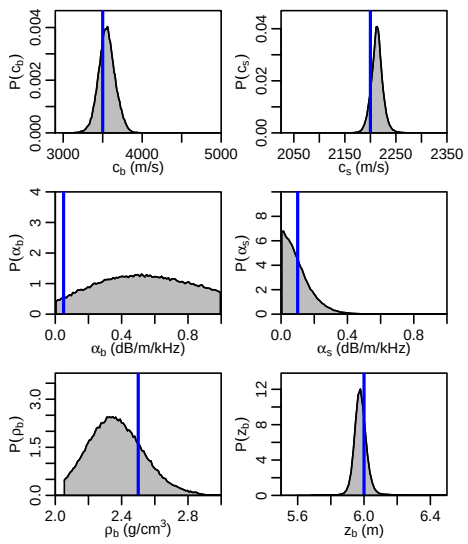

Figure 2.8: Marginal posterior distributions of basement geoacoustic parameters from joint TDAR inversion. Vertical lines indicate true values.

An inversion that treated the basement as fluid, rather than an elastic solid, strongly biased the compressional-wave speed and density estimates (not shown).

The results so far are presented in terms of posterior marginal distributions; however, some applications may require an estimate of the optimal model. Here the maximum *a posteriori* (MAP) model is used. This also highlights the significance of being able to explicitly evaluate the prior distribution as such an estimate is not be possible otherwise. The MAP geoacoustic profile for the TDAR inversion is given in Fig. 2.2; all seabed parameter are given in Table 2.6.1. The MAP scattering parameters are close to their true values. Figure 2.2 shows that the MAP seabed sound-speed profile closely follows the true model, although with less structure (fewer layers) that the true model. The MAP density and attenuation profiles are generally good representations over the sediment layers, but differs more from the true model in the basement (where the marginal probability profiles in Fig. 2.7 indicates larger uncertainties). The MAP model has interfaces very close to both of the large discontinuities in the true profile.

2.6.2 Data Analysis

The fit to the scattering data is shown in terms of marginal predicted data in Fig. 2.9 for the scattering-only inversion and Fig. 2.3 for the TDAR inversion. These figures

Table 2.4: The scattering and geoacoustic parameter values of $\hat{\mathbf{m}}_j$, the model with maximum posterior probability for the TDAR inversion.

Sediment			
z (m)	c (m/s)	ρ (g/cm^3)	α (dB/m/kHz)
0.124	1530	1.330	0.8002
0.355	1540	1.488	0.6573
0.764	1573	1.568	0.4793
3.048	1612	1.472	0.1473
	1684	1.733	0.1026

Basement		Scattering	
z_b (m)	5.984	γ	3.126
c_b (m/s)	3662	w_2	0.0017
c_s (m/s)	2209	K_0 (1/m)	1.507
ρ (g/cm^3)	2.165		
α_b (dB/m/kHz)	0.6218		
α_s (dB/m/kHz)	0.0296		

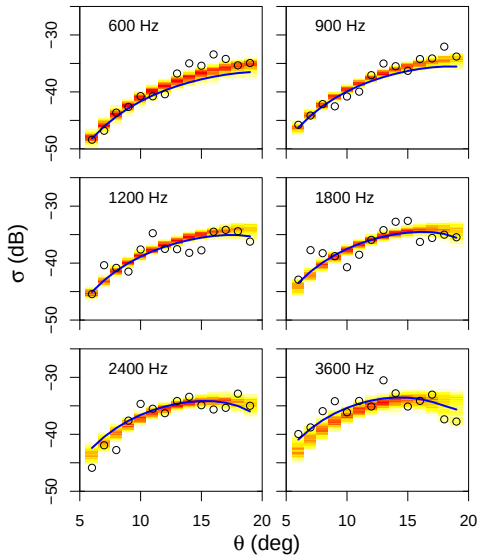

Figure 2.9: Simulated noisy scattering data (∘) and marginal predicted data (shading) from scattering-only inversion. Solid lines indicate noise-free simulated data

also show the observed (noisy) data which were inverted and the noise-free data. The agreement between predicted and noise-free data is good. The introduction of the reflection data reduces the variance of the predicted data about the noise-free data at all frequencies. In addition, the TDAR inversion further reduces the variance at

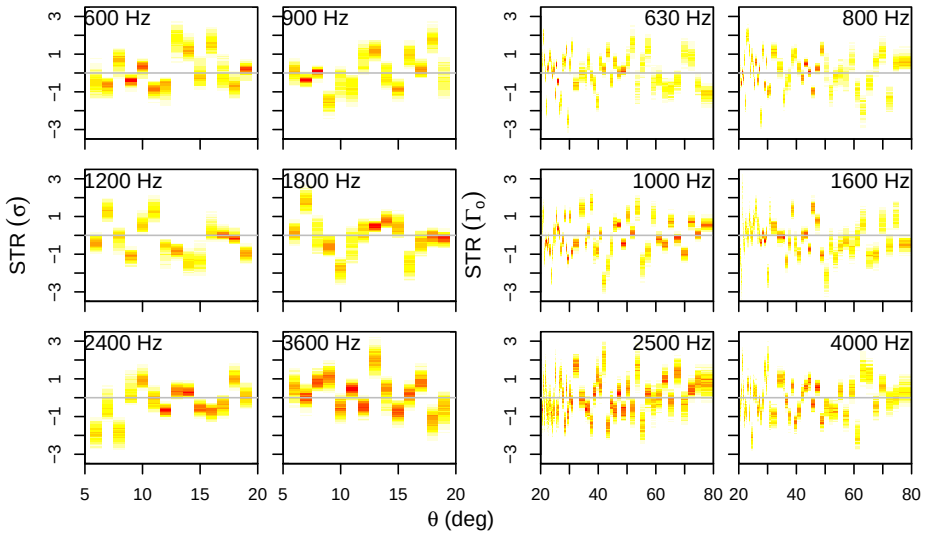

Figure 2.10: Marginal standardized total residuals, $\hat{\mathbf{e}}$, from joint TDAR inversion of scattering data (left) and reflection-coefficient data (right).

high frequencies. These comparisons are, of course, not possible in practical cases when the noise-free data are unknown. In such cases it is natural to consider the data residuals.

The standardized total residuals ($\hat{\mathbf{e}}$) are defined as the total residuals divided by their standard deviations; i.e.

$$\hat{e}_i = \frac{e_i}{S_i} = \frac{d_i - d_i\left(\mathbf{m}_j\right) - d_i\left(\mathbf{a}\right)}{S_i}, \tag{2.23}$$

where the data type and frequency subscripts are omitted for clarity. The standardized total residuals for the scattering and reflection data for the TDAR inversion are shown in Fig. 2.10. The scattering residuals appear stationary, uncorrelated, and unbiased. The residuals for the reflection data also appear homostochastic although there appears to be some residual structure at 630 and 1000 Hz. This is not unexpected as the true geoacoustic profile was selected so it could not be fully resolved by the inversion. Such theory error (the true model is not in the parameter space of the inversion) commonly results in correlated residuals.[33,43,44]

Figures 2.11 and 2.12 show the marginal distributions for the residual standard deviation and AR(1) parameters from the joint inversion. To evaluate whether residuals with significant correlations were correctly identified by the joint inversion, a location test is applied to all AR(1) parameters.[33] The true distributions of the pos-

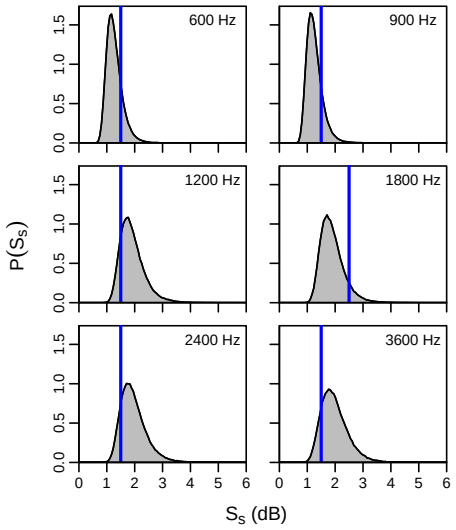

Figure 2.11: Marginal posterior distributions of the scattering data standard errors $\mathbf{S}_s$ for the joint inversion. Vertical lines indicate true values.

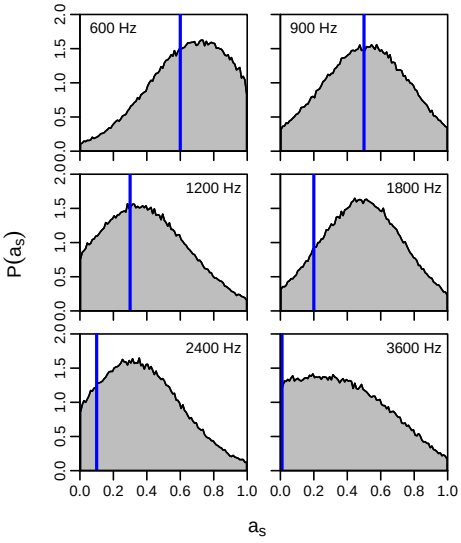

Figure 2.12: Marginal posterior distributions of the scattering data AR(1) parameters $\mathbf{a}_s$ for the joint inversion. Vertical lines indicate true values.

terior means of the AR parameters under the null hypotheses (H_0: $\mathbf{a_s} = \mathbf{a_r} = \mathbf{0}$) are not known analytically; however, they can be approximated using Monte Carlo simulation. This simulation consists of drawing sets of uncorrelated Gaussian random variables and inverting for the AR model [Eq. (2.4) or (2.6)]. The distribution (over many random draws) of the sampled means is used as the null distribution. The lo-

Table 2.5: Location test results for AR(1) parameter means and p-values for scattering (top) and reflection (bottom) errors from joint inversion

Frequency (Hz)	600	900	1200	1800	2400	3600
Scattering means	0.621	0.518	0.408	0.491	0.381	0.404
p-value	0.005	0.023	0.092	0.033	0.124	0.095
Frequency (Hz)	630	800	1000	1600	2500	4000
Reflection means	0.609	0.537	0.505	0.223	0.166	0.177
p-value	0.000	0.000	0.000	0.036	0.108	0.088

cations tests quantifies evidence (in terms of a p-value) against the null hypothesis H_0 that the AR(1) sample means are drawn from the null distribution. The results of the location tests are given in Table 2.5. A p-value > 0.05 is interpreted as no significant evidence against H_0, i.e., no evidence the residuals are serially correlated. A p-value < 0.05 is interpreted as significant evidence against H_0, indicating the residuals are correlated. The inversion procedure correctly identifies the strongly correlated residuals ($a_s, a_r \geq 0.5$) from the weakly correlated ones ($a_s, a_r \leq 0.1$). Moderately-correlated residuals ($0.2 \leq a_s, a_r \leq 0.3$) are not identified.

Generally, AR(1) parameters found to be insignificant could be excluded to avoid over-parametrizing the error model and the inversion repeated. However, this approach is not entirely satisfactory as removing AR(1) parameters is equivalent to fixing their value at zero, and there is insufficient evidence to make such a strong assumption. The TDAR inversion samples AR(1) parameters according to their support by the data and scales well with the number of parameters (unlike hypothesis tests). The uncertainty resulting from including/excluding AR(1) parameters is accounted for in the PPD. The TDAR inversion can also be used to evaluate the necessity of the AR(1) parameters; Table 2.6 gives the proportion of samples with AR(1) parameters for both the scattering and reflection data. The results of the TDAR inversion are similar to the location tests and correctly differentiate the strongly and weakly correlated residuals at the various frequencies.

2.7 Summary and conclusions

The rigorous estimation of *in-situ* seabed scattering (roughness) parameters and their uncertainties from remote acoustic measurements is a problem of practical interest which has received little attention to date. This chapter developed a Bayesian in-

Table 2.6: Proportion of samples with AR(1) parameter for scattering (top) and reflection (bottom) errors from TDAR inversion.

Frequency (Hz)	600	900	1200	1800	2400	3600
Scattering	0.9112	0.7858	0.5369	0.7469	0.5520	0.4474
Frequency (Hz)	630	800	1000	1600	2500	4000
Reflection	0.9995	1.0000	0.9997	0.4592	0.2426	0.2233

version approach to examine resolving a power-law representation of the 2D spatial roughness spectrum parameterized in terms of the spectral strength, exponent, and cutoff. The primary measurement of interest is the backscatter strength as a function of angle and frequency; however, since the scattering kernel also depends on the geoacoustic profile, the improvement in scattering spectral resolution from joint inversion of (simulated) backscatter and spherical-wave reflection-coefficient data was also examined and quantified. Basing initial inversion studies such as this on (realistic) simulations allows comparisons of results with the true model and control over error processes such that the information content of the physics of the acoustic measurements can be quantified. It is, of course, important to follow such studies with inversions of measured data, which represents future work.

The simulations considered here are based on observed scattering parameters and a realistic geoacoustic profile consisting of a large number of sediment layers ($\sim$100 layers over 6 m) with fine structure below the resolution limit of the data. An elastic half-space with compressional and shear properties representative of limestone comprised the basement. The number of sediment layers is unknown in the inversion and treated with a trans-dimensional approach which samples over differing numbers of layers according to their support by the data and prior.

Realistic errors were added to the scattering and (unevenly-spaced) reflection data, including variances which changed with frequency (and near the critical angle for reflection data) and serial correlations which decreased with frequency such that only the low-frequency errors were significantly correlated. The error statistics were treated as unknown in the inversion and parameterized in terms of variances and first-order auto-regressive parameters. Given that error correlations were significant at only some frequencies, to avoid either over- or under-parameterizing the error model in the inversion, a new trans-D sampling approach was developed for the AR(1) parameters.

To treat a geoacoustic profile in which the number of parameters was not constant

with depth (i.e., fluid sediment layers over an elastic basement) required developing a new trans-D partition prior distribution which is not based on a fictitious grid (the standard approach). This enhanced formulation also allows the explicit evaluation of the prior and posterior probabilities (rather than just relative probabilities). In addition, the prior used here includes empirical information on the relationship between compressional-wave speed and density in sediments and limestone. This decreases the total volume of the parameter space without excluding plausible models, which is important in trans-D inversion where parsimony is based on the trade-off between data misfit and prior volume.

Three cases were considered: scattering-only inversion, joint scattering and reflection inversion, and joint inversion with the trans-D auto-regressive error model. The ability of the inversions to resolve the model was evaluated using marginal posterior probability distributions and profiles. The resolution of the scattering parameters (particularly the spectral exponent) and geoacoustic parameters were improved by the introduction of the reflection data. The scattering-only inversion showed some sensitivity to the near surface geoacoustics, however, resolution decreased dramatically with depth. Both joint inversions resolved the geoacoustics at all depths. In the basement, the high shear-wave speed was well resolved (better than compressional-wave speed), highlighting the importance of including elasticity in an inversion when required by the data.

Auto-regressive models were found to effectively account for residual correlation in the scattering-only and joint inversion. However, including AR(1) parameters at all frequencies over-parameterized the error model and increased the variance of scattering parameters. The trans-D AR(1) scheme (i.e., the TDAR inversion) mitigated the loss of PPD resolution while accounting appropriately for the residual correlation at all frequencies.

The results in this chapter indicate significant potential for Bayesian inversion of acoustic data to determine seabed scattering properties and uncertainties, particularly for joint inversion of scattering and reflection data.

Chapter 3

Seabed roughness parameters from joint backscatter and reflection inversion at the Malta Plateau

This chapter presents estimates of seabed roughness and geoacoustic parameters and uncertainties on the Malta Plateau, Mediterranean Sea, by joint Bayesian inversion of mono-static backscatter and spherical-wave reflection-coefficient data. The data are modelled using homogeneous fluid sediment layers overlying an elastic basement. The scattering model assumes a randomly rough water-sediment interface with a von Karman roughness power spectrum. Scattering and reflection data are inverted simultaneously using a population of interacting Markov chains to sample roughness and geoacoustic parameters as well as residual error parameters. Trans-dimensional sampling is applied to treat the number of sediment layers and the order (zeroth or first) of an auto-regressive error model (to represent potential residual correlation) as unknowns. Results are considered in terms of marginal posterior probability profiles and distributions, which quantify the effective data-information content to resolve scattering/geoacoustic structure. Results indicate well-defined scattering (roughness) parameters in good agreement with existing measurements, and a multi-layer sediment profile over a high-speed (elastic) basement, consistent with independent knowledge of sand layers over limestone.

3.1 Introduction

Ocean sonar performance predictions require knowledge of the reverberation levels characteristic of a particular environment. An important cause of acoustic reverberation is seabed roughness; hence, determining roughness parameters is required to model scattering and reverberation. The use of direct techniques (e.g., stereoscopic photography, laser imaging) to measure seafloor roughness over small scales[1,87] and long-range reverberation methods to infer roughness over large scales have been carried out.[19] However, the characterization of seafloor roughness on the meso-scale (patch size of order 10^2 m with sensitivity to seabed fluctuations in the vertical and horizontal of order 10^{-1} m) requires attention. This chapter applies trans-dimensional (trans-D) Bayesian inversion to measured backscatter and reflection data to estimate meso-scale interface roughness parameters and geoacoustic profiles together with a fully nonlinear uncertainty analysis. This provides an acoustic remote-sensing alternative to direct measurements for scattering and geoacoustic parameters. Previous scattering inversion work has been based on optimization techniques to provide a best-fit model but with no indication of uncertainties.[46,47] To our knowledge, this chapter represents the first quantitative estimation (with rigorous uncertainties) of seafloor roughness spectra via inversion of measured acoustic scattering data.

The data considered here consist of mono-static backscatter and spherical-wave reflection coefficients as a function of angle and frequency measured at a geoacoustic test bed located on the Malta Plateau in the Straits of Sicily. Two inversions are conducted, the first using only the backscatter data and the second using both the backscatter and reflection data. The value of joint versus scattering-only inversion was considered previously in a simulation study[71] which found the inclusion of reflection data improved the accuracy (but not the posterior uncertainty) of scattering parameters.

The inversion problem considered here is solved using a Bayesian approach which provides rigorous evaluation of parameter uncertainties and inter-relationships. Bayesian inversion has been applied widely to other geoacoustic inverse problems.[7,35–40] The inversion is conducted trans-dimensionally, a relatively new method that has been applied recently to several problems in geophysics[26,28,32–34] and geoacoustics.[69] Trans-D inversion samples over model dimension (number of unknowns) and intrinsically account for model selection uncertainty in parameter estimate uncertainty.[68,77] As in the previous simulation study,[71] both the number of seabed sediment layers and the

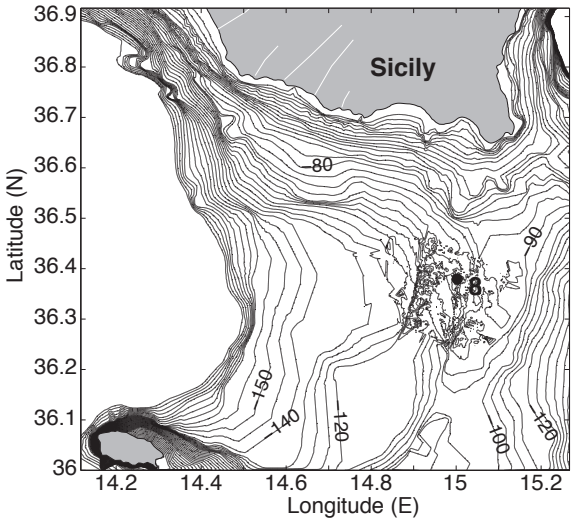

Figure 3.1: Bathymetry (in meters) and experiment site 8 location on the Malta Plateau.

number of data sub-sets (distinct frequencies for backscatter and reflection data) requiring a first-order auto-regressive [AR(1)] processes to model error correlation are treated as unknowns. Bayesian sampling is conducted using the reversible jump Markov chain Monte Carlo (rjMCMC) algorithm[49,68] with parallel tempering.[53,85,86]

The trans-D Bayesian inversion methodology applied here was developed and validated using synthetic scattering and reflection data in an earlier paper.[71] That work indicated good resolution of seafoor roughness and geoacoustic parameters given the idealized physics and controlled statistical errors employed in simulations. This chapter carries out inversion of measured data, and hence quantifies the actual ability to resolve roughness and geoacoustic parameters in practice, including the effects of theory errors (e.g., due to the assumed physics of the forward problem, simplified seabed parameterization, and idealized error statistics) which cannot be evaluated in simulation. Importantly, the parameter uncertainties estimated for measured data quantify the resolution of seabed roughness and geoacoustic parameters that can be achieved in real experimental conditions–a fundamentally important result.

3.2 Data collection

The two types of data inverted here, mono-static backscatter (σ) and spherical-wave reflection coefficients (Γ_0), were measured at site 8 on the Malta Plateau in the Straits

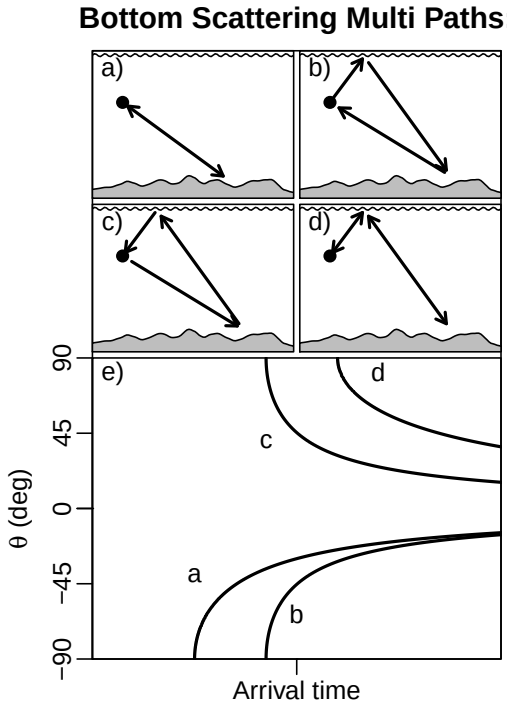

Figure 3.2: (a)-(d): Bottom scattering multi-paths with one seafloor interaction. (e): Vertical arrival angle versus arrival time for each of the multi-paths (after Ref. [12]). The horizontal length scale in (e) is arbitrary depending on the experiment geometry.

of Sicily, as shown in Fig. 3.1. Both measurements sample localized areas of similar scale (∼500 m of the seabed).[12,55] The scattering measurements employed a sound source at the bottom of a vertical array of hydrophones to transmit an impulsive signal which scatters from the seafloor and is recorded at the array.[12] Possible acoustic paths from the source to the receiver are depicted in Fig. 3.2(a)-(d). Beamforming and time windowing are applied to the received signal to separate it into these paths, as shown in Fig. 3.2(e). In this chapter, only data corresponding to the single bottom backscattering path shown in Fig. 3.2(a) are used, as this is sufficient to define the mono-static scattering kernel. Scattering data were collected at frequencies of 600, 900, 1200, 1800, 2400, and 3600 Hz and an angular range of 5–60°. The scattering data are shown in Fig. 3.3 (left); the angular range is restricted to 5–25° for the inversion to minimize the effects of subsurface scatterers or other heterogeneities.

The reflection data were collected using an impulsive acoustic source (seismic boomer) towed past a moored receiver to provide reflections with a range of inci-

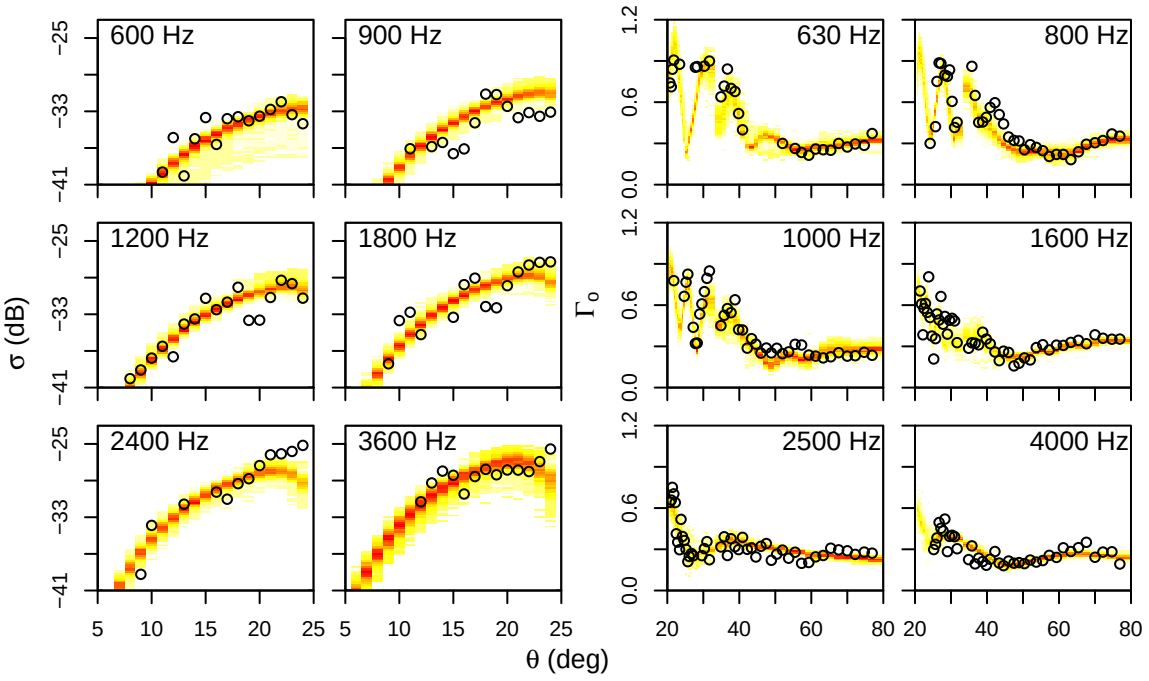

Figure 3.3: Measured (○) and predicted scattering data (left) and reflection data (right) for joint inversion.

dent angles at the seabed.[55] The height of the receiver was selected such that the bottom-bounce and direct-path arrivals were separated in time. The source was located close to the surface so the surface-bounce and direct paths combined to form a single arrival. Source directivity is accounted for using the ratio of direct and bottom-reflected arrivals with the same takeoff angle (i.e., arrivals with the same directivity term, which then cancels in the reflection coefficient, instead of arrivals from the same source transmission) with angular interpolation applied to closely match angles. This method is appropriate because the acoustic source is more repeatable than omnidirectional.[16] The data were recoded over a wide range of frequencies; however, only reflection data in the same frequency range as the scatting data are used: 630, 800, 1000, 1600, 2500, and 4000 Hz. The angular range of the reflection-coefficient data is 20–78°. Figure 3.3 (right) displays the reflection data.

3.3 Forward models

The forward models used here to predict the scattering[17] and reflection[13] data are applied to a seabed model consisting of a layered half-space as shown in Fig. 3.4. The

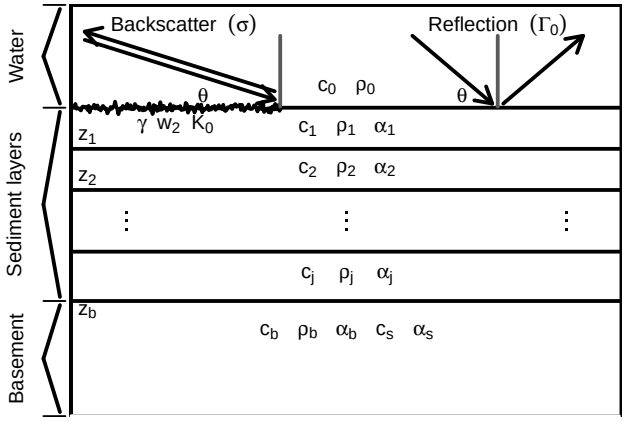

Figure 3.4: Schematic diagram of the seabed model. Parameters are defined in the text.

top (zeroth) layer is seawater and is assumed to be homogeneous and isotropic with known properties. The seabed is a series of j flat homogeneous layers, terminated by a homogeneous semi-infinite basement. Layer properties include interface depth z, sound speed c, density ρ, and attenuation α. In addition, the basement is elastic with a shear-wave speed c_s and attenuation α_s. The scattering model also assumes a randomly rough water-sediment interface with a von Karman roughness power spectrum

$$W(\mathbf{K}) = w_2 \left(|\mathbf{K}|^2 + K_0^2\right)^{-\gamma/2}, \tag{3.1}$$

where γ, w_2, and K_0 are known as the spectral exponent, spectral strength, and spectral cutoff, respectively, and $\mathbf{K}$ is the transverse component of the wave vector with magnitude $|\mathbf{K}| = k_0 \cos\theta$ (k_0 is the wavenumber in the water).[1] The scattering kernel, in decibels, is given by[17]

$$\begin{aligned}\sigma(\theta, f) = \;& 10 \log_{10}\left[W(2\mathbf{K}) k_0^4 \left|1 + R(\theta, f)\right|^4\right] + \\ & 10 \log_{10}\left[\frac{1}{4}\left|1 - \left(\frac{k_1}{k_0}\right)^2 \frac{\rho_0}{\rho_1} + \left(1 - \frac{\rho_0}{\rho_1}\right)\left(\cos^2\theta + \frac{\rho_0}{\rho_1}\sin^2\theta\right)\left(\frac{1 - R(\theta, f)}{1 + R(\theta, f)}\right)^2\right|^2\right],\end{aligned} \tag{3.2}$$

where k_1 is the wavenumber in the first sediment layer, R is the total plane-wave reflection coefficient for the ($j + 1$ layer) sediment stack (evaluated recursively and accounting for the elastic basement).[13]

The scattering errors are assumed to be independent over frequency but potentially correlated over angle such that the correlation is non-negative and decreases

exponentially with angular separation. The residual correlation structure is modelled using an AR(1) process[33,53,71,81] given by

$$r_i = {a_s}^{\Delta\theta_i} r_{i-1} + e_i, \tag{3.3}$$

where the r_i are the residuals (indexed over angle with $r_0 = 0$), a_s is the AR(1) coefficient, $\Delta\theta_i = \theta_i - \theta_{i-1}$, and the e_i are identical independently-distributed (IID) Gaussian random variables with zero mean and standard deviation S_s (the frequency index is omitted here for clarity).

The spherical-wave reflection coefficients $\Gamma_\circ$ for an arbitrary layered half-space are computed by plane-wave decomposition[13] with the Sommerfeld integral computed numerically using Simpson's rule.[82]

The reflection data residuals are also modelled with an AR(1) process over angle, similar to Eq. (3.3), with AR(1) parameter a_r. The reflection data are not considered IID in that their standard deviations are allowed to differ above and below a cut off angle θ_c which approximates the critical angle (since it is commonly observed that reflection errors change at this point). Thus, if $\theta_i < \theta_c$ a low-angle standard deviation S_r^L is used, otherwise a high-angle value S_r^H is used. In previous work,[71] θ_c was assigned arbitrarily; here is is treated as an unknown in the inversion.

3.4 Bayesian inversion

A complete description of Bayesian inference as applied to trans-D geoacoustic inverse problems can be found in Refs. [26,28,32,69] and the specific implementation used here is described in Ref. [71]. To summarize briefly, let $\mathbf{d}$ be a random vector of observed data and let $\mathbf{m}_j$ represent a set of model parameters where j indexes the different choices of parameterizations [e.g., number of seabed sediment layers and/or non-zero AR(1) coefficients]. Using Bayes' rule the trans-D PPD can be written[68]

$$P(\mathbf{m}_j|\mathbf{d}) = \frac{\pi(\mathbf{m}_j)\mathcal{L}(\mathbf{m}_j)}{\mathcal{Z}}, \tag{3.4}$$

Table 3.1: The prior distribution as defined by lower and upper bounds (LB and UB) for basement, sediment, and scattering parameters.

Basement			Sediment		
Parameter	LB	UB	Parameter	LB	UB
z_b (m)	0	10	c (m/s)	1450	2100
ρ_b (g/cm^3)	1.20	3.00	α (dB/m/kHz)	0	1
c_b (m/s)	1500	6000	ρ (g/cm^3)	1.20	2.25
α_b (dB/m/kHz)	0	1	Scattering		
c_s (m/s)	0	$c_b/\sqrt{2}$	Parameter	LB	UB
α_s (dB/m/kHz)	0	1	γ	2	4
			w_2 (m$^{(4-\gamma)}$)	10^{-5}	10
			K_0 (m^{-1})	10^{-5}	32

where $\pi(\mathbf{m}_j)$ is the prior distribution of j and $\mathbf{m}_j$, $\mathcal{Z}$ is the evidence of the ensemble of models, and $\mathcal{L}(\mathbf{m}_j)$ is the likelihood of the parameter vector

$$\mathcal{L}(\mathbf{m}_j) = \frac{1}{(2\pi)^{N/2} \mid \mathbf{C}_d \mid^{1/2}} \exp\left(-\frac{1}{2} \left[\mathbf{d} - \mathbf{d}(\mathbf{m}_j) - \mathbf{d}(\mathbf{a})\right]^\top \mathbf{C}_d^{-1} \left[\mathbf{d} - \mathbf{d}(\mathbf{m}_j) - \mathbf{d}(\mathbf{a})\right] \right), \tag{3.5}$$

where $\mathbf{d}(\mathbf{m}_j)$ represents the data predicted for $\mathbf{m}_j$, $\mathbf{d}(\mathbf{a})$ is the AR(1) process prediction, and $\mathbf{C}_d$ is a diagonal covariance matrix where the ith diagonal element corresponds to the unknown variance of the ith total residual e_i (i.e., an element of $\mathbf{S}_s$, $\mathbf{S}_s^L$, or $\mathbf{S}_r^H$).[33,71]

The priors for the roughness and geoacoustic parameters consist of bounded uniform distributions constraining the parameters to physically meaningful values, as listed in Table 3.1. The relationship between sound speeds and densities for the sediments and basement are also constrained by 2D bounds[71] derived from a large set of measurements.[84]

Direct analytic interpretation of the trans-D PPD is not possible; consequently, the PPD is approximated using a sampling algorithm and inference is conducted on the samples. Here the reversible jump Markov chain Monte Carlo algorithm[49,68] with parallel tempering[53,85,86] is applied. Approximately 1,000,000 samples were collected; the first 250,000 samples from the start of the chain (referred to as the burn-in) were discarded to prevent the arbitrary starting model from impacting the sample distribution. The remainder was chain thinned to 200,000 samples to reduce sample autocorrelation and allow for more convenient storage and analysis.

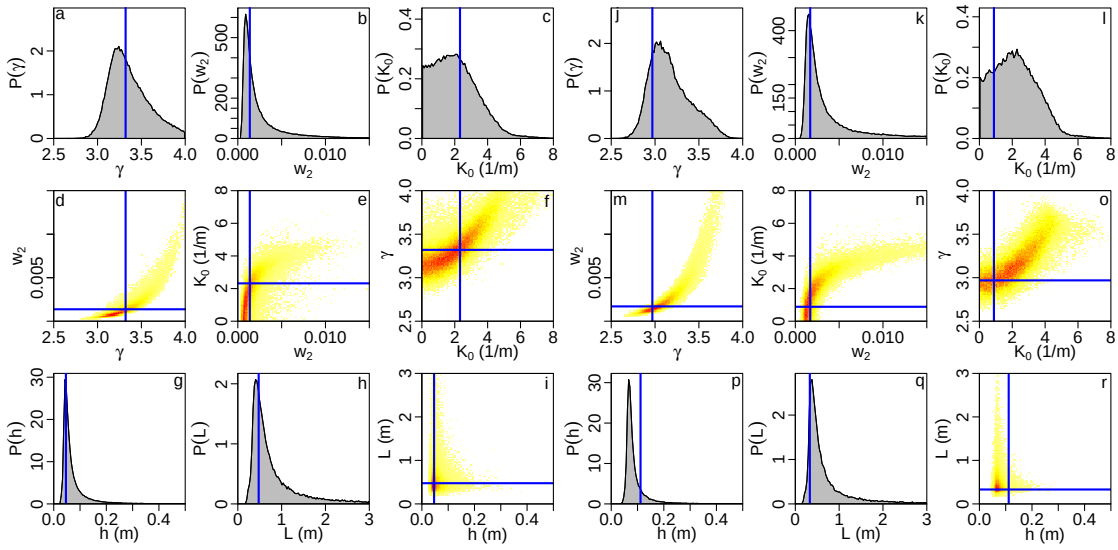

Figure 3.5: Marginal and joint marginal posterior probability distributions for scattering parameters from scattering-only (a–i) and joint (j–r) inversions. Solid lines indicate the MAP model parameters.

3.5 Inversion results

Marginal posterior probability distributions for the scattering parameters estimated via Scattering-only inversion are shown in Fig. 3.5(a)–(f); the equivalent marginals for the joint inversion are shown in Fig. 3.5(j)–(r). All of these distributions are uni-modal with peaks near $\gamma = 3$, $w_2 = 0.003$ m$^{(4-\gamma)}$, and $K_0 = 2$ m^{-1}. All three parameters are reasonably well resolved with only small differences between the joint and scattering-only inversions. The maximum *a posteriori* (MAP) model scattering parameters are indicated in Fig. 3.5 and given, with 95% credibility bounds, in Table 3.2. This represents the most probable model given the prior and data information. The MAP parameter values are generally near the modes of the marginal posterior distributions (although the two need not coincide in nonlinear problems). The agreement between scattering-only and joint inversion results indicates that the two data sets contain consistent geoacoustic information. The joint inversion is taken as the preferred result, despite its slightly larger parameters uncertainties (Table 3.2). The simulation study in Ref. [71] showed that posterior distribution from the joint inversion had a smaller expected difference between an sample of the posterior and the true value. This results from the ability of strongly correlated residuals to shift parameter estimates away from the true value. The inclusion of a second data set is expected to protect

Table 3.2: The MAP scattering parameters with a 95% credibility interval (CI) for the scattering-only (top) and joint (bottom) inversions.

	Parameter	95% CI	MAP
Scattering-only	γ	[3.01, 3.90]	3.23
	w_2 ($\mathrm{m}^{(4-\gamma)}$)	[5.71×10^{-4}, 1.36×10^{-2}]	1.37×10^{-3}
	K_0 (1/m)	[0.11, 5.00]	2.32
Joint	γ	[2.79, 3.80]	2.97
	w_2 ($\mathrm{m}^{(4-\gamma)}$)	[6.55×10^{-4}, 1.58×10^{-2}]	1.66×10^{-3}
	K_0 (1/m)	[1.56×10^{-4}, 5.09×10^{0}]	8.84×10^{-1}

against this.

Fig. 3.5(j)–(l) shows that the spectral exponent γ is distinct from both its upper and lower prior bounds (see Table 3.1), spectral strength w_2 has a sharp peak but a long tail to high values, and spectral cutoff K_0 has an upper limit of $\sim$5 m^{-1} but its lower limit is not distinct from the prior bound of 0 m^{-1}. The mode in γ indicates that scattering strength increases roughly linearly with frequency [for $\gamma = 3$, $W(2\mathbf{K})$ is approximately proportional to k_0^{-3} by Eq. (3.1), which, when substituted into Eq. (3.2), results in the scattering cross section proportional to k_0]. There is a nonlinear relation between γ and w_2 [the joint marginal, Fig. 3.5 (d), has a curved shape]; this is consistent with the observation that γ and $\log(w_2)$ would be multi-variate Gaussian distributed if all other parameters where fixed and $K_0 = 0$ in Eq. (3.1).

Figure 3.6 shows a comparison of γ and $\log(w_2)$ estimates from joint inversion with direct measurements reported in a compilation of roughness parameters using stereophotogrammetric measurements on continental-shelf seafloors at a variety of locations.[1] The probability density from acoustic inversion is located well within the scatter of the direct measurements, and the precision of the inversion result is sufficient to differentiate it from most of the measured values (i.e., its uncertainty is small compared to the scatter of the measured values). The measured values do not share the same correlation structure as the inverted parameters. This can be explained as the inversion correlation structure is a result of the scattering physics and not due to the geological/oceanographic processes that create seabed roughness. Note that measured γ and w_2 values in Fig. 3.6 were originally reported in terms of 1D spectra parameters (γ_1 and w_1), and are converted here to their equivalent 2D

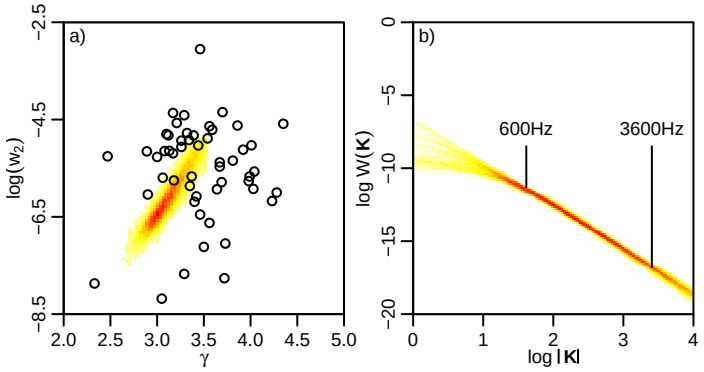

Figure 3.6: (a) Joint marginal of γ and $\log(w_2)$ from acoustic inversion compared to parameters measured with stereoscopic camera. (b) Marginal for the von Karman (power-law) spectrum.

spectrum values (assuming isotropy) by[87]

$$\gamma = \gamma_1 + 1 \tag{3.6}$$

and

$$w_2 = w_1 \frac{(2\pi)^{(\gamma-2)}\,\Gamma\left(\gamma/2\right)}{\left(\sqrt{\pi}\right)\Gamma\left(\left[\gamma-1\right]/2\right)}, \tag{3.7}$$

where Γ is the gamma function. The spectral cutoff (K_0) for both the 1D and 2D spectra are equal under isotropy; the direct measurements assumed $K_0 = 0\ \mathrm{m}^{-1}$. The validity of this comparison can be evaluated by considering Fig. 3.6(b), which shows the posterior marginal roughness power spectrum from the scattering inversion. This spectrum is approximately linear for the range of the data considered here (600–3600 Hz), and consequently the γ and w_2 estimates are not significantly affected by including a non-zero K_0 (although it is expected that the posterior distribution of w_2 would shift slightly towards lower values if K_0 where to be fixed at zero). This, along with the marginal distribution of K_0 in Fig. 3.5(l) and the linear joint marginal for γ and $\log(w_2)$ in Fig. 3.6(a), indicate that K_0 is neither statistically nor practically discernible from zero by the data.

For some applications it is useful to characterize the seabed roughness random process with an alternative parametrization of root-mean-square (RMS) roughness h and correlation length L. The von Karman spectrum can be converted analytically

to an expression for RMS roughness by[1]

$$h = \left[\frac{2\pi w_2}{(\gamma - 2)\, K_0^{(\gamma-2)}} \right]^{\frac{1}{2}} . \tag{3.8}$$

The mathematical definition of L used here assumes $B_1(r) \approx \tilde{B}_1(r) = h^2 \exp(-|r|/L)$, where $B_1(r)$ is the 1D spatial roughness covariance function, which is the inverse Fourier transform of $W_1(k_r)$, the 1D spatial power spectrum (r is lag distance and k_r is the horizontal wavenumber in the direction of lag). Let $\tilde{W}_1(k_r)$ be the Fourier transform of $\tilde{B}_1(r)$:

$$\tilde{W}_1(k_r) = \frac{1}{2\pi} \int_{-\infty}^{\infty} h^2 \exp\left(-|r|/L\right) \exp\left(ik_r r\right) dr = \frac{h^2}{\pi} \frac{L}{L^2 k_r^2 + 1}. \tag{3.9}$$

L is found numerically by minimizing the sum of $(\tilde{W}_1(k_r) - W_1(k_r))^2$ over an array of k_r values. The 1D and 2D marginal posterior densities for the RMS roughness and correlation length are shown in Fig. 3.5(p)–(r). The RMS roughness mode is centered near 7 cm and the correlation-length mode near 40 cm. The joint marginal shows that the h and L estimates are uncorrelated.

Figure 3.7 (top) shows marginal posterior profiles for the geoacoustic parameters from the joint inversion (as in Ref. [71] the scattering-only inversion was only able to resolve the geoacoustic structure of top meter of sediments and is not shown here). As expected, sound speed is resolved better than density and density better than attenuation. Resolution of all parameters generally decreases with depth. The sound speed, density and attenuation values for the top 6 m are consistent with sand, which is known (from underwater video footage) to be present at the experiment site. The high attenuation close to the water sediment interface indicates the surficial sediments also likely include some silt. There is a sharp jump to a high-speed layer at ∼6 m depth. Marginal probability distributions for the basement (including shear parameters) are shown in Fig. 3.8. The geoacoustic parameter values indicated (c_b ∼3500 m/s, c_s ∼2100 m/s) are consistent with limestone, which is known to out-crop in this region. Further, a strong reflector is present at ∼6 m depth in a high-resolution seismic section recorded at the experimental site, shown in Fig. 3.9 which is consistent with limestone at this depth. The multi-modality of the sound speed marginal in Fig. 3.7 for layers from 4–6 m depth may be a result of the roughness of the sediment-basement interface. The large uncertainty of the compressional speed (c_b)

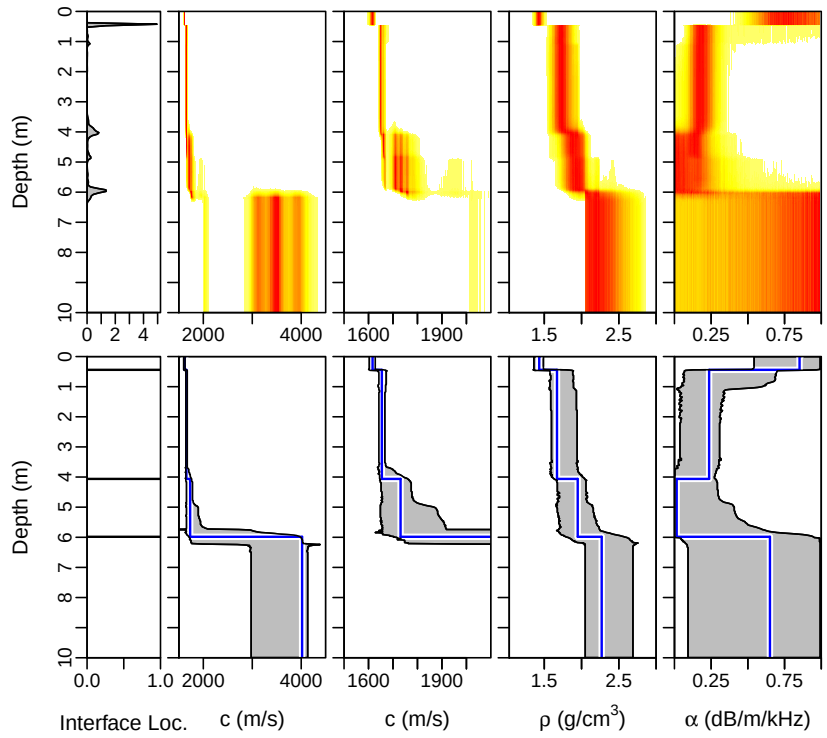

Figure 3.7: Top: Marginal posterior geoacoustic profiles from joint inversion (sound speeds shown at two scales). Probability values are normalized independently for each depth for display purposes. Bottom: MAP geoacoustic profile (solid line) with 95% credibility interval (shaded region).

in Fig. 3.8(a) results from the much smaller contrast between the sediment sound speed (c_j) and shear-wave speed (c_s) than between c_j and c_b; consequently, changes in c_b have little effect on either σ or Γ_0. It is also interesting to note that the small mode near $c_b = 2150$ m/s in Fig. 3.8(a) corresponds closely to the shear-wave speed mode in Fig. 3.8(b).

Figure 3.10(a) shows the posterior marginal distribution for the number of sediment layers, j. The inversion favors three layer models although from 2-6 layers are sampled. Three layers over a basement halfspace can be seen clearly on the profile plot in Fig. 3.7 (at 0–0.5 m, 0.5–4 m, and 4–6 m).

The MAP model profile is shown in Fig. 3.7 (bottom). The MAP basement parameters with 95% credibility interval bounds are listed in Table 3.3. The MAP has $j = 3$ layers, which is consistent with the most probable j from Fig. 3.10(a). The low-speed mode for c_b is excluded from the credibility interval. Also note the MAP does not always correspond to the peak of the marginal profiles.

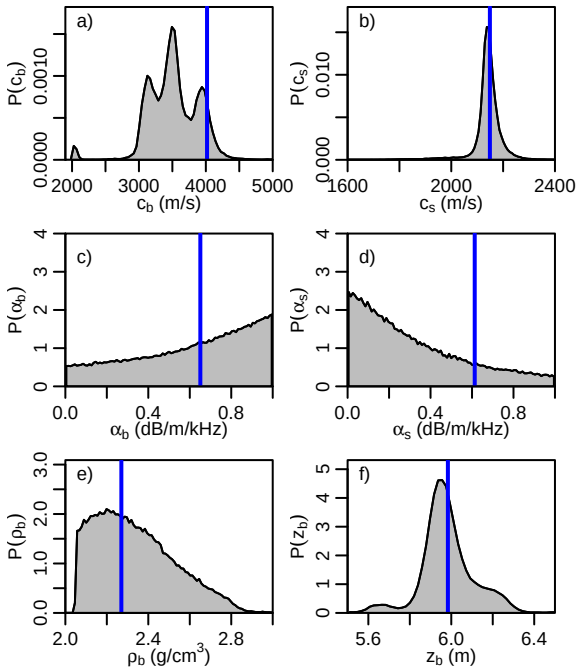

Figure 3.8: Posterior marginal probability distributions for the basement parameters for the joint inversion. Solid lines indicate the MAP model.

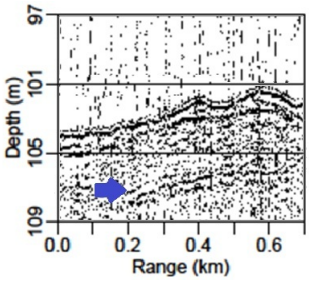

Figure 3.9: High-resolution seismic section at the experiment site showing a strong reflector at ∼6 m sub-bottom (indicated by arrow). The depth scale assumes a sound speed of 1500 m/s.

3.6 Data fit and error model

The fit to the measured data is shown in terms of marginal predicted data in Fig. 3.3. The agreement between predicted and observed data is good; however, correlated residuals are indicated at certain frequencies. A possible source for the correlated residuals is the range dependant bathymetry at the experimental site (shown in Fig. 3.9). The standardized total residuals (STR), denoted $\hat{\mathbf{e}}$, are defined as the

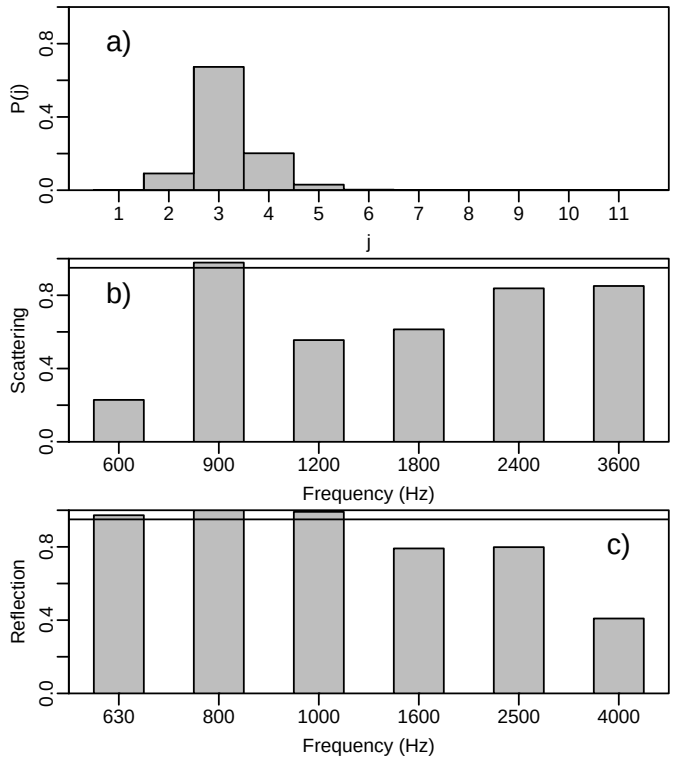

Figure 3.10: (a) Marginal probability distribution for number of sediment layers, j. (b) and (c) Proportion of models with non-zero AR(1) parameter for scattering and reflection data, respectively, at indicated frequencies. Horizontal lines indicates 0.95 level. All results are for the joint inversion.

total residuals [e_i from Eqs. (3.3)] divided by their standard deviations; i.e.,

$$\hat{e}_i = \frac{e_i}{S_i} = \frac{d_i - d_i\left(\mathbf{m}_j\right) - d_i\left(\mathbf{a}\right)}{S_i}, \tag{3.10}$$

where the data type and frequency subscripts are omitted for clarity. The STR for the scattering and reflection data are shown in Fig. 3.11. The scattering residuals in general appear stationary, uncorrelated, and unbiased; only STR at 900 Hz appear significantly auto-correlated. The STR for the reflection data also appear generally homostochastic although there appears to be some residual structure in the higher grazing angles at 1000 and 2500 Hz.

Figure 3.12 shows the marginal distributions for the residual standard deviation and AR(1) parameters. The standard deviations $\mathbf{S}_s$, $\mathbf{S}_r^L$, and $\mathbf{S}_r^H$ have distributions which appear approximately χ–distributed. The $\mathbf{S}_s$ (scattering) marginals indicate that the residual error is approximately 1–2.5 dB. The $\mathbf{S}_r^L$ (low angle reflection) val-

Table 3.3: The MAP basement parameters with a 95% credibility interval (CI) for joint inversion.

Basement		
Parameter	95% CI	MAP
z_b (m)	[5.60, 6.31]	5.98
ρ_b (g/cm^3)	[2.05, 2.70]	2.27
c_b (m/s)	[2960, 4170]	4020
α_b (dB/m/kHz)	[9.04×10^{-2}, 1.00×10^{0}]	6.50×10^{-1}
c_s (m/s)	[1980, 2270]	2150
α_s (dB/m/kHz)	[3.32×10^{-6}, 8.41×10^{-1}]	6.13×10^{-1}

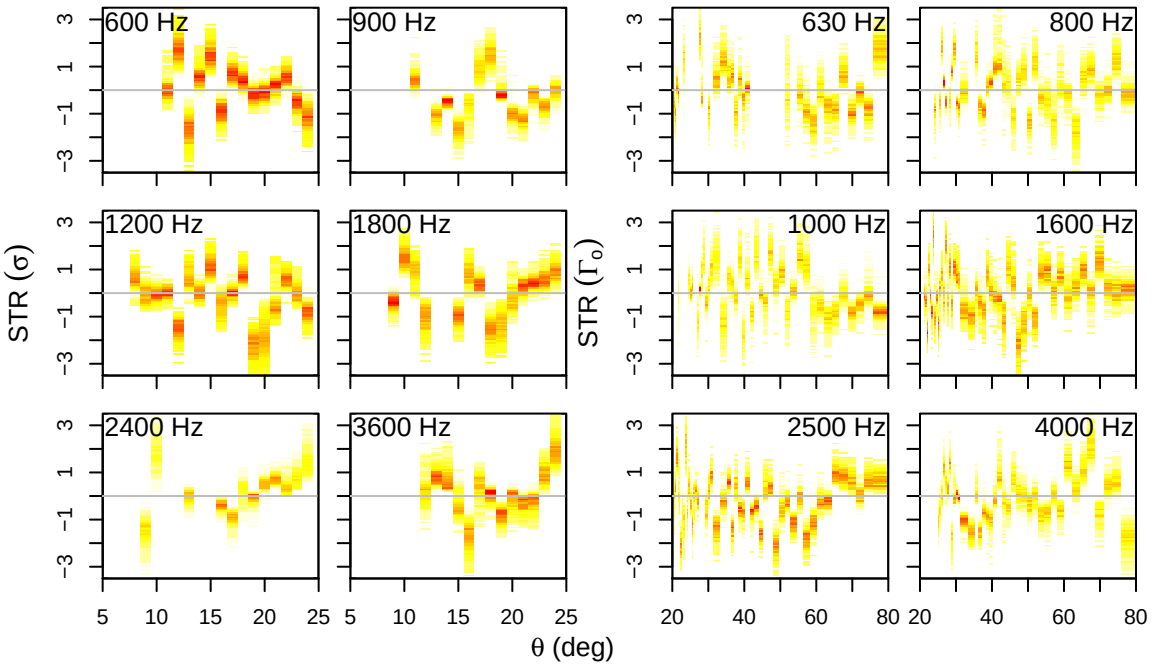

Figure 3.11: Marginal standardized total residuals for scattering data (left) and reflection data(right) for joint inversion.

ues are in general larger than the $\mathbf{S}_r^H$ (high angle reflection) values; this reflects the increased uncertainty below the critical angle. Figure 3.10(b) and (c) give the proportion of samples with non-zero AR(1) parameters for the scattering and reflection data, respectively. This figure shows $\geq$95% of the models require AR(1) parameters for 900-Hz scattering data and for 630-, 800-, and 1000-Hz reflection data; consequently only these frequencies require (at 0.95 level) lag-one auto-correlated residuals. For other frequencies it is ambiguous if the AR(1) parameters are necessary (for

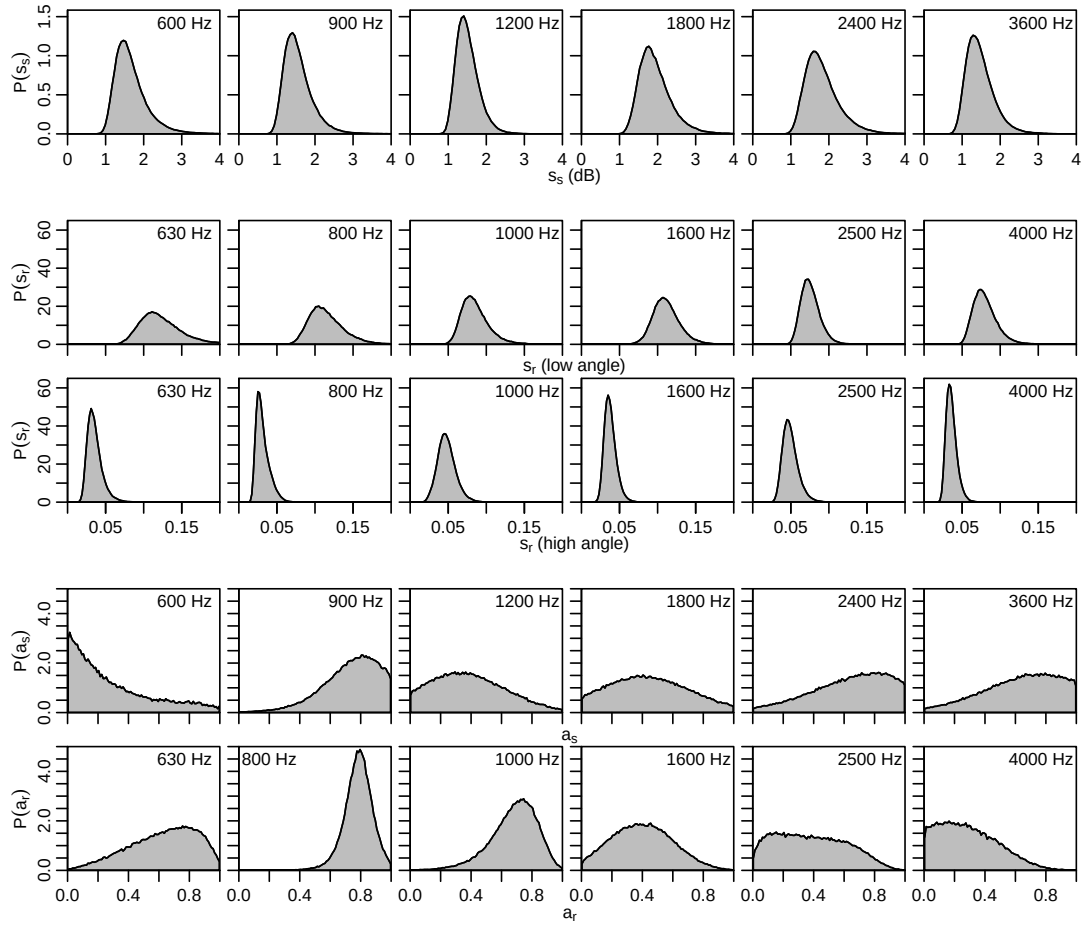

Figure 3.12: Marginal distributions for residual error parameters [Standard errors for scattering and high/low angle reflection and AR(1) for scattering and reflection] for joint inversion.

no frequency is the proportion of models with non-zero AR(1) less than 0.05). This highlights the value of the trans-D AR(1) procedure as it is not necessary to choose *a priori* if an AR(1) parameter is required in the inversion and thus avoids over- or under-parameterizing the error model.[71]

3.7 Summary and conclusions

Estimation of *in-situ* seabed scattering (roughness) parameters and uncertainties from remote acoustic measurements is a problem of practical interest for ocean acoustic reverberation modelling and sonar performance predictions. This chapter applied a Bayesian inversion approach to measured acoustic backscatter data to estimate

parameters for a von Karman (power-law) representation of the 2D seafloor spatial roughness spectrum at a test bed on the Malta Plateau. The experimental procedure probed seafloor roughness on the meso-scale which has been under-represented to date. This also reduces the effects of ocean/seabed variability and uncertainty relative to long-range acoustic measurements while still maintaining a high resolution for seabed variability. The scattering data were augmented with reflection-coefficient data (measured at the same experimental site) to help constrain the roughness and geoacoustic parameters.

An appropriate seabed model parametrization (e.g., number of layers resolved by the data) was addressed by trans-dimensional inversion using the rjMCMC algorithm. Data residual correlations were modelled using a trans-D auto-regressive error process. This allowed the data to determine the level of complexity of the residual model without over- or under-parametrizing the error model. Trans-D inversion also enables uncertainty about the number of layers and AR order (zeroth or first) to be accounted for in the roughness/geoacoustic uncertainties. Sampling over two levels of dimensionality (i.e., seabed layers and AR order) slowed convergence; this was addressed using parallel tempering.

Both a scattering-only and a joint scattering reflection inversion were conducted; the joint inversion is taken as the preferred result. The inversion results indicate that the seabed at the experimental site consisted of $\sim$6 m of sediment layers over a high-speed basement; the parameter values were consistent with sand over limestone. These results are in agreement with video observation of the seafloor at the experimental site and with a high-resolution seismic section. The parameters of the von Karman representation of seabed interface roughness spectrum were well resolved with a posterior mode near $\gamma = 3$, $w_2 = 0.003$ m$^{(4-\gamma)}$, and $K_0 = 2$ m^{-1}. These values can also be interpreted as a root-mean-square roughness of $\sim$7 cm with a correlation-length of $\sim$40 cm. The roughness parameter estimates were consistent with reported values of direct measurements at a variety of other sites and the posterior uncertainties of the scattering parameters are sufficiently small to differentiate the roughness properties of the experiment site from those observed at other locations.

Chapter 4

Determining the dominant seafloor scattering mechanism

This chapter presents a quantitative and objective approach to classify the dominant seabed acoustic scattering mechanism(s) and estimate seabed scattering and geoacoustic parameters from measured backscatter data. The classification system is based on trans-dimensional sampling and the deviance information criterion. Scattering is modeled using first-order perturbation theory as due to one of three mechanisms: interface scattering from a rough seafloor, volume scattering from a heterogeneous sediment layer, or mixed scattering combining both interface and volume scattering. The classification system is applied to six simulated test cases where it correctly identifies the true dominant scattering mechanism as having greater support from the data in five cases; the remaining case is indecisive. The approach is also applied to measured backscatter data from the Malta Plateau where volume scattering is determined as the dominant scattering mechanism. This conclusion and the scattering/geoacoustic parameters estimated in the inversion are consistent with properties from previous inversions and with core measurements from the site.

4.1 Introduction

One of the primary uses of sonar is to detect and classify targets such as submarines, mines, and marine animals. In such cases, acoustic scattering from the seabed is a source of interference which degrades target signals and reduces sonar performance. An accurate model of seabed scattering in a particular environment can help miti-

gate this performance reduction.[1] Seabed scattering is caused by irregularities such as roughness at the water-sediment interface and heterogeneities or discrete inclusions (e.g., shells or stones) within the volume of a sediment layer. To model seabed scattering it is necessary to identify the observable and/or dominant scattering mechanism(s) (DSM) at the location of interest for the relevant frequency range. This chapter develops a Bayesian approach to classifying the DSM as either surface scattering, volume scattering, or both surface and volume scattering. The classification procedure is carried out for both simulated and measured direct-path backscatter data. These data are preferred here for scattering inversion over long-range reverberation data due to a higher information content, and the ability to resolve meso-scale scattering structure (10^2 m patch size with sensitivity to seabed fluctuations in the vertical and horizontal of order 10^{-2} m).

The inversion methodology used here represents an extension to previous work on interface scattering;[71,72] acoustic scattering data are modeled using first-order perturbation theory and the posterior probability density (PPD) is approximated by trans-dimensional (trans-D) Bayesian sampling. Specifically, scattering data are predicted using a simplified version of Jackson et al.'s GABIM model,[18] which allows scattering at the water-sediment interface and/or within the volume of sediment layers. The seafloor roughness and sediment-layer heterogeneities are represented as random processes that follow von-Karman power-law spectra. The trans-D Bayesian sampling is conducted using the reversible jump Markov chain Monte Carlo (rjMCMC) algorithm,[26,28,32,68,69,71] sampling over scattering, geoacoustic, and residual-error parameters as well as the unknown number of geoacoustic layers and the number of frequency bands with auto-correlated data residuals. Scattering mechanisms are compared using the deviance information criterion (DIC),[67] which trades off data fit versus model complexity to determine the most parsimonious and, thus preferred, model.

The acoustic inversion approach proposed here represents an *in-situ*, remote-sensing alternative to laborious direct measurements of scattering parameters (e.g., stereoscopic photography and core X-radiography)[1] and can address both interface and volume scattering with a single acoustic data set. The Bayesian approach extends scattering inversion from point estimation based on optimization techniques[15,19,46,47] to a probabilistic solution with fully nonlinear uncertainty analysis.

Six simulated data sets (three representing a muddy sediment regime and three representing a sandy regime) are considered and inverted in a simulation study of the DSM classifier. For both sediment regimes data sets are created for each of interface

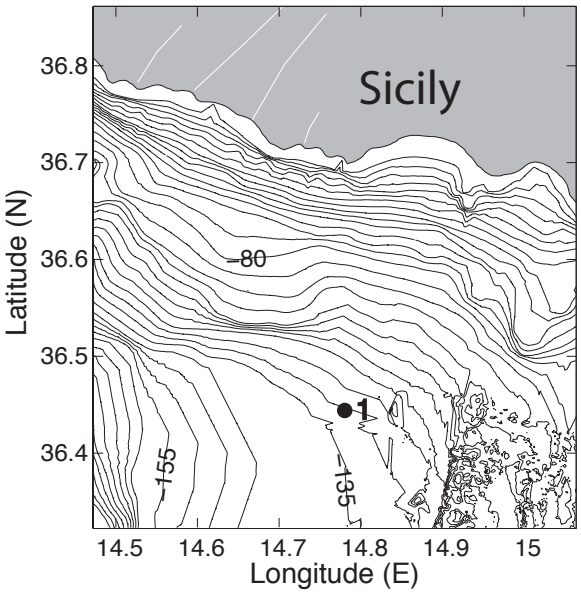

Figure 4.1: Bathymetry (in meters) and experiment Site 1 location on the Malta Plateau.

scattering, volume scattering, and mixed scattering (both interface and volume scattering). The classification procedure is found to correctly identifies the DSM in five inversions. No statistically significant difference is found for the other inversion.

The measured data were recorded at a NATO experiment site on the Malta Plateau, Mediterranean Sea. As with previous scattering inversion work,[72] the measured scattering data are jointly-inverted with reflection-coefficient data collected at the same site and cover roughly the same patch size. The classification procedure finds that volume scattering dominates, with the most probable location of the scatterers found to be just above and within a high-sound-speed, high-density geoacoustic layer. Scattering and geoacoustic parameter estimates are consistent with previous inversions[51,88] and with measurements on a core extracted at the site.[15]

4.2 Data collection

The backscatter and reflection-coefficient data used here were measured at Site 1 on the Malta Plateau in the Strait of Sicily, as shown in Fig. 4.1. Previous geoacoustic inversions of reflection-coefficients from active-source measurements and from ambient-noise data at this site have been reported in Refs. [88] and [51], respectively. Both a $\sim$0.8 m gravity core and $\sim$3.0 m piston core of the seabed sediments were

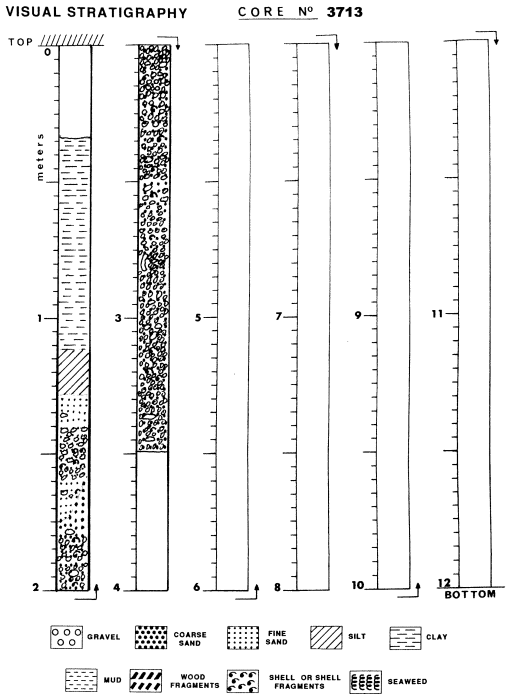

Figure 4.2: Split piston core stratigraphy, the core diameter is 10 cm and the core slumped 35 cm.

extracted at the site.[15] Figure 4.2 shows the stratigraphy of the piston core. Note that the core slumped 0.35 m; i.e., the true position of a given layer is 0.35 m less than shown in the stratigraphy. Features of the core include a silty-clay layer ($\sim$1.0 m thick) over a sandy layer (0.7 m thick) above a pebble/shell layer which extends to the end of the core. Occasional larger stones ($\sim$0.05–0.1 m in diameter) are present from $\sim$0.8 m to the end of the core (not shown in the stratigraphy).

Mono-static backscatter data (σ) are used as observables from which the seabed DSM is classified. As backscatter data are relatively insensitive to geoacoustic structure, spherical-wave reflection-coefficient (Γ_0) data collected at the site are also included in the inversion.[15] Both measurements sample areas of similar scale ($\sim$500 m lateral patch of the seabed).[12,55] The scattering measurements employed a sound source at the bottom of a vertical array of hydrophones to transmit an impulsive signal which scatters from the seafloor and is recorded at the array.[12] Beamforming and time windowing are applied to the received signal to isolate the single bottom-

backscattering path. Backscatter data were collected at frequencies of 900, 1800, and 3600 Hz and an angular range of 6–40° and are shown in Fig. 4.3

The reflection data were collected using an impulsive acoustic source (seismic boomer) towed past a moored receiver to provide reflections with a range of incident angles at the seabed.[55] The height of the receiver was selected such that the bottom-bounce and direct-path arrivals were separated in time. The source was located close to the surface so the surface-bounce and direct paths combined to form a single arrival. Source directivity is accounted for using the ratio of direct and bottom-reflected arrivals with the same takeoff angle (i.e., arrivals with the same directivity term, which then cancels in the reflection coefficient, instead of arrivals from the same source transmission) with angular interpolation applied to closely match angles.[16] The reflection data are Gaussian band averaged (1/15 octave band width) with center frequencies at 400, 504, 635, 800, 1008, 1270, 1600, and 2016 Hz with angular range 12–80° (Fig. 4.3).

4.3 Forward models

The forward models used to predict the scattering[18] and reflection[13] data are applied to a layered model of the ocean environment (Fig. 4.4). The top (zeroth) layer is seawater and is assumed to be homogeneous and isotropic with known properties. The seabed is a series of j flat-lying layers, terminated by a homogeneous semi-infinite basement. Layer properties include interface depth z, sound speed c, density ρ, and attenuation α. In addition, the basement is elastic with a shear-wave speed c_s and attenuation α_s. While the seabed layers are considered homogeneous for reflection, they are allowed to be heterogeneous for scattering. Specifically, the scattering model allows a randomly rough water-sediment interface and/or a layer of volume heterogeneities at depth. Note that, for generality, the scattering layer does not necessarily correspond to the geoacoustic layers (i.e., scattering layer boundaries are independent of the geoacoustic layer boundaries).

The roughness of the water-sediment interface is described as a random process governed by a von Karman roughness power spectrum[1]

$$W(\mathbf{K}) = w_\mathrm{I} \left(|\mathbf{K}|^2 + K_\mathrm{I}^2\right)^{-\frac{1}{2}\gamma_\mathrm{I}}, \tag{4.1}$$

where γ_I, w_I, and K_I are known as the interface spectral exponent, spectral strength,

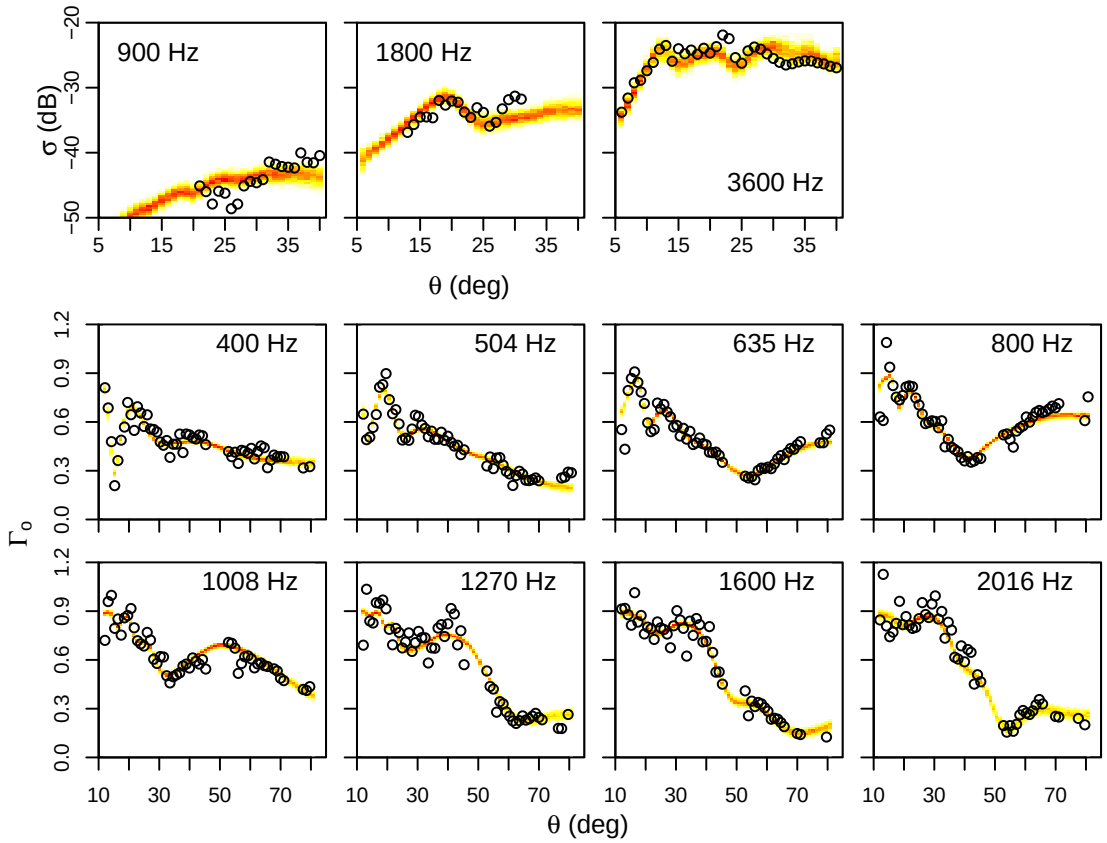

Figure 4.3: Measured (∘) and predicted scattering data (top row) and reflection data (bottom two rows) at indicated frequencies.

and spectral cutoff, respectively, and $\mathbf{K}$ is the transverse component of the wave vector with magnitude $|\mathbf{K}| = k_0 \cos\theta$ (k_0 is the wavenumber in the water).[1] The volume heterogeneities are also modeled as a random process with a density heterogeneity power spectrum[18]

$$W_{\rho\rho}(\mathbf{K}) = w_{\mathrm{V}} \left(|\mathbf{K}|^2 + k_z^2/\lambda^2 + K_{\mathrm{V}}^2\right)^{-\frac{1}{2}\gamma_{\mathrm{V}}}, \tag{4.2}$$

compressibility heterogeneity spectrum

$$W_{cc}(\mathbf{K}) = \left(u^2 + v^2\right) W_{\rho\rho}(\mathbf{K}), \tag{4.3}$$

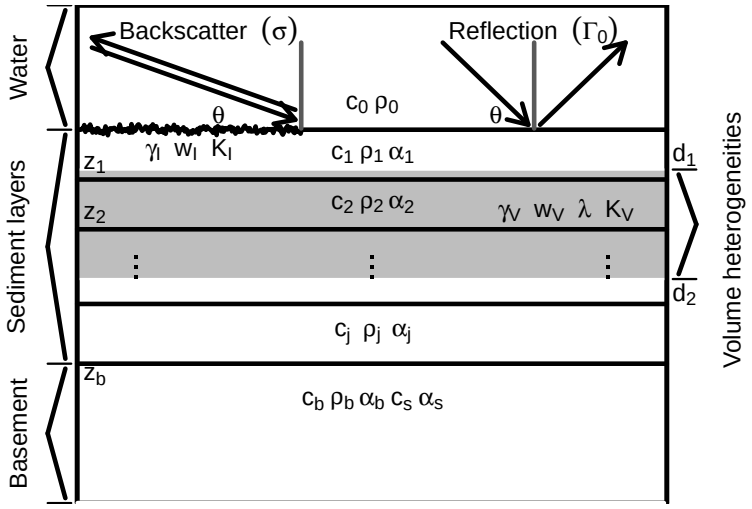

Figure 4.4: Schematic diagram of the seabed model. Parameters are defined in the text.

and cross-heterogeneity spectrum

$$W_{c\rho}\left(\mathbf{K}\right) = v\, W_{\rho\rho}\left(\mathbf{K}\right). \tag{4.4}$$

The parameters γ_{V}, w_{V}, λ, and K_{V} are the volume spectral exponent, spectral strength, vertical anisotropy, and spectral cutoff, respectively. The weights u and v determine the relative importance of the spectra. The interface and volume spectra as statistically independent.

The scattering kernel used here is a simplified version of the GABIM model[18]

$$\sigma(\theta, f) = 10 \log_{10}\left[\sigma_I(\theta, f) + \sigma_V(\theta, f)\right] \quad \text{(dB)}. \tag{4.5}$$

The interface scattering contribution σ_{I} is

$$\sigma_{\mathrm{I}}(\theta, f) = \frac{k_0^4 \left|1 + R(\theta, f)\right|^4}{4} W(2\mathbf{K}) \quad \times \\ \left| 1 - \left(\frac{k_1}{k_0}\right)^2 \frac{\rho_0}{\rho_1} + \left(1 - \frac{\rho_0}{\rho_1}\right)\left(\cos^2\theta + \frac{\rho_0}{\rho_1}\sin^2\theta\right)\left(\frac{1 - R(\theta, f)}{1 + R(\theta, f)}\right)^2 \right|^2, \tag{4.6}$$

where k_1 is the wavenumber in the first sediment layer, and R is the total plane-wave reflection coefficient for the ($j + 1$ layer) sediment stack.[13] The volume-scattering

contribution σ_{V} is

$$\sigma_{\mathrm{V}}(\theta, f) = \sum_{l=1}^{j} \sigma_{\mathrm{V}_l}(\theta, f), \tag{4.7}$$

where σ_{V_l} is the contribution for geoacoustic layer l calculated as

$$\begin{aligned} \sigma_{\mathrm{V}_l}(\theta, f) = \; & \frac{\pi\rho_0^2}{2\rho_l^2} \sum_{i_1=1}^{2}\sum_{i_2=1}^{2}\sum_{i_3=1}^{2}\sum_{i_4=1}^{2} a_{i_1} a_{i_2}^{\star} a_{i_3} a_{i_4}^{\star} F\left(i[g_{i_1} + g_{i_2} - g_{i_3}^{\star} - g_{i_4}^{\star}]\right) \times \\ & \left(|k_l|^4 W_{cc}(\mathbf{K}) + (|\mathbf{K}| + g_{i_1} g_{i_3})(|\mathbf{K}| + g_{i_2} g_{i_4})^{\star} W_{\rho\rho}(\mathbf{K}) + \right. \\ & \left. \left[(|\mathbf{K}| + g_{i_1} g_{i_3})(k_l^{\star})^2 + (|\mathbf{K}| + g_{i_2} g_{i_4})^{\star}(k_l)^2\right] W_{c\rho}(\mathbf{K})\right), \end{aligned} \tag{4.8}$$

where k_l is the wavenumber in the lth layer. In Eq. (4.8) the gs are the vertical wavenumber in the lth layer (k_{z_l}), with the subscript indicating sign (i.e., $g_1 = -k_{z_l}$ and $g_2 = k_{z_l}$). The values a_1 and a_2 are the magnitudes of the down- and up-going waves in the lth layer, [i.e., $a_1 = B$ and $a_2 = A$ if the field in the lth layer is $p_l = Ae^{-kz(z-z_l)} + Be^{kz(z-z_l)}$]. Finally, the function F is defined as $F(x) = (\exp(xh_l) - 1)/x$, where h_l is the thickness of the lth layer. Note that Eq. (4.8) is slightly different from the analogous Eq. (49) in Ref. [18], as a result of the simplifying assumption used here (that all geoacoustic layers within the scattering layer have the same volume spectra parameters) and due to the correction of several typographical errors that appeared in the original paper.

The scattering data residuals (difference between measured and predicted data) are assumed to be independent over frequency but potentially correlated over angle such that the correlation is non-negative and decreases exponentially with angular separation. The residual correlation structure is modeled using a first-order auto-regressive [AR(1)] process[33,53,71,81] given by

$$r_i = {a_{\mathrm{scat}}}^{\Delta\theta_i} r_{i-1} + e_i, \tag{4.9}$$

where the r_i are the residuals (indexed over angle), a_{scat} is the AR(1) coefficient, $\Delta\theta_i = \theta_i - \theta_{i-1}$, and the e_i, referred to as total residuals, are identical independently-distributed (IID) Gaussian random variables with zero mean and standard deviation S_{scat} (the frequency index is omitted here for clarity). Error parameters at all frequencies $\mathbf{a}_{\mathrm{scat}}$ and $\mathbf{S}_{\mathrm{scat}}$ are treated as unknowns and sampled in the inversion. Further, as in Ref. [72], the number of frequency bands with non-zero AR(1) coefficients a_{scat} is considered unknown and marginalized trans-dimensionally with rjMCMC sampling.

This allows the data to determine if auto-regressive parameters are required or not, and avoids over- or under-fitting the residual error model.

The spherical-wave reflection coefficients $\Gamma_\circ$ for an arbitrary layered half-space are computed by plane-wave decomposition[13] with the Sommerfeld integral computed numerically using Simpson's rule.[82] The reflection data residuals are also modeled with trans-D AR(1) processes over angle, similar to Eq. (4.9), with (unknown) AR(1) parameters a_{ref}. The reflection data are not considered IID in that their standard deviations are allowed to differ above and below a cut-off angle θ_c which approximates the critical angle (since it is observed that reflection errors often change at this point[71,72]). Thus, if $\theta_i < \theta_c$ a low-angle standard deviation S^L_{ref} is used, otherwise a high-angle value S^H_{ref} is used. Error parameters $\mathbf{a}_{\text{ref}}$, $\mathbf{S}^L_{\text{ref}}$, $\mathbf{S}^H_{\text{ref}}$ and θ_c are treated as unknowns in the inversion.[72]

4.4 Bayesian inversion

A complete description of Bayesian inference as applied to trans-D geoacoustic inverse problems can be found in Refs. [26,28,32,69] and the specific implementation used here is given in Ref. [71]; this section provides a brief overview. In Bayesian inversion, model parameters are considered random variables; let $\mathbf{m}_j$ represent an arbitrary model where j represents a choice of model complexity [e.g., indicates the number of sediment layers and frequency bands with non-zero AR(1) coefficients]. Using Bayes' rule, the trans-D PPD can be written[68]

$$P(\mathbf{m}_j|\mathbf{d}) = \frac{\pi(\mathbf{m}_j)\mathcal{L}(\mathbf{m}_j)}{\mathcal{Z}}\,, \tag{4.10}$$

where $\pi(\mathbf{m}_j)$ is the prior distribution of $\mathbf{m}_j$, $\mathcal{Z}$ is the total evidence of the ensemble of models (i.e., the probability of the data integrated and summed over all possible parameter values and parameterizations), and $\mathcal{L}(\mathbf{m}_j)$ is the likelihood of the parameter vector (i.e., the probability of the data given the model). Assuming Gaussian-distributed total residuals, the likelihood can be written

$$\mathcal{L}(\mathbf{m}_j) = \frac{1}{(2\pi)^{N/2} \mid \mathbf{C}_d \mid^{1/2}} \exp\left(-\frac{1}{2}\left[\mathbf{d}-\mathbf{d}(\mathbf{m}_j)-\mathbf{d}(\mathbf{a})\right]^\top \mathbf{C}_d^{-1}\left[\mathbf{d}-\mathbf{d}(\mathbf{m}_j)-\mathbf{d}(\mathbf{a})\right]\right), \tag{4.11}$$

where $\mathbf{d}(\mathbf{m}_j)$ is the data predicted by the forward model, $\mathbf{d}(\mathbf{a})$ is the AR(1) prediction [from Eq. (4.9)], and $\mathbf{C}_d$ is a diagonal data covariance matrix.[33,71]

The trans-D PPD is approximated by rjMCMC sampling.[49,68] Let $\mathbf{m}_j$ be the current Markov chain state and $Q(\mathbf{m}'_{j'}|\mathbf{m}_j)$ be the proposal distribution by which a new state $\mathbf{m}'_{j'}$ is generated. The proposed model can represent a perturbation of the parameters of $\mathbf{m}_j$ or a change (jump) in dimension of $\mathbf{m}_j$, i.e., $j' \neq j$. The proposed state $\mathbf{m}'_{j'}$ is accepted with probability

$$A = \min\left[1, \frac{\pi(\mathbf{m}'_{j'})}{\pi(\mathbf{m}_j)} \frac{\mathcal{L}(\mathbf{m}'_{j'})}{\mathcal{L}(\mathbf{m}_j)} \frac{Q(\mathbf{m}_j|\mathbf{m}'_{j'})}{Q(\mathbf{m}'_{j'}|\mathbf{m}_j)} |\mathbf{J}|\right], \tag{4.12}$$

where $\mathbf{J}$ is the Jacobian of the diffeomorphism between the parameter spaces associated with $\mathbf{m}_j$ and $\mathbf{m}'_{j'}$. To improve sampling efficiency (particularly the acceptance rate of dimension jumps), parallel tempering,[53,71,85,86] a population-based sampling scheme, is applied.

4.5 Deviance information criterion

The deviance information criterion is used here to differentiate between models defined by choice of scattering mechanism. The DIC is a measure of model support based on a trade-off of data fit versus model complexity (parsimony) similar to other more common model-selection criteria such as the Bayesian information criterion (BIC).[43,70,89] However, the DIC is calculated from samples drawn from the PPD (rather than from a point-estimate of the best-fit model in BIC), and consequently has the advantages that it accounts for the prior distribution, parameter correlations, and the general non-Gaussianity of the PPD.[67] The DIC trades off the data fit of a characteristic model against a complexity term and is defined as

$$\mathrm{DIC} = D\left(\hat{\mathbf{m}}_{\hat{j}}\right) + 2P_D\,, \tag{4.13}$$

where D is the posterior deviance defined as

$$D\left(\mathbf{m}_j\right) = -2\log\mathcal{L}\left(\mathbf{m}_j\right) \tag{4.14}$$

and $\hat{\mathbf{m}}_{\hat{j}}$ is a central or characteristic model (e.g., the mean, median, mode, or maximum *a posteriori*, MAP, model). The term P_D in Eq. (4.13) is the effective number of focused parameters (parameters that are not marginalized out of the posterior before the DIC is calculated). The effective number of parameters is evaluated as

$P_D = \overline{D(\mathbf{m}_j)} - D(\hat{\mathbf{m}}_{\hat{j}})$, where $\overline{D(\mathbf{m}_j)}$ is the mean of the posterior deviance. Here the characteristic model $\hat{\mathbf{m}}_{\hat{j}}$ is taken to be the MAP model to accommodate the trans-D structure of the PPD.

4.6 Classifying the dominant scattering mechanism

In this chapter acoustic scattering data are classified by their DSM using the DIC. Three possible scattering mechanisms are considered including: a rough water-sediment interface (interface scattering), heterogeneities within the volume of a sediment layer (volume scattering), and both interface roughness and volume heterogeneities (mixed scattering). Additional scattering mechanisms, such as rough interfaces between sediment layers, could also be considered; however, these are expected to be less important here and are not easily conformable to the geoacoustic parameterization used in this chapter. The classifier conducts three inversions of a data set, one for each of the three scattering mechanisms described above. Once all inversions are completed the DIC for each mechanism is calculated and used to measure the support by the data; the mechanism with the lowest DIC has the greatest support and is taken as the DSM for the data set. DSMs with a difference in DIC of 5–10 are generally considered to have significantly different levels of data support and differences of 10 or more are considered definitive.[67]

In this section, the classification approach is validated using six simulated-data inversions; these consist of three cases representing geoacoustic parameters for a muddy seabed and three cases representing a sandy seabed; both regimes use the same scattering parameter values. The true geoacoustic and scattering parameter values are given in Table 4.1. For each sediment type data sets are computed for volume scattering, interface scattering, and mixed scattering. Random errors are added to the true predicted data; the error for all inversions are uncorrelated Gaussian distributed with standard deviations of 1 dB. Note the residual parameters (including possible residual correlations) are treated as unknowns in the inversions. The simulated data have similar angular and frequency ranges as the measured data (5–40° and frequency bands centered at 900, 1800, and 3600 Hz). The simulations only include scattering data, not reflection-coefficient data, to reduce the computational cost of the simulation study.

Each of the six data sets is inverted once assuming each of the three types of scattering, and the DIC is calculated for each of the 18 inversions (Table 4.2). The DIC

Table 4.1: True parameter values for simulated inversions.

Mud Regime					
Layer	1	2	3	Basement	
z (m)	0.1	1.0	–	5.0	
c (m/s)	1500	1475	1550	1590	
ρ (g/cm^3)	1.35	1.5	1.6	1.65	
α (dB/m/kHz)	0.1	0.1	0.1	0.1	
Sand Regime					
Layer	1	2	3	Basement	
z (m)	0.1	1.0	–	5.0	
c (m/s)	1600	1650	1750	1690	
ρ (g/cm^3)	1.5	1.7	1.75	1.65	
α (dB/m/kHz)	0.8	0.1	0.1	0.1	
σ_{I} Parameters		σ_{V} Parameters			
w_{I}	0.015	u	0.0	w_{V}	0.015
K_{I} (1/m)	3.0	v	-1.5	K_{V} (1/m)	10.0
γ_{V}	3.25	d_1 (m)	1.0	γ_{V}	3.5
		d_2 (m)	2.5	λ	5.0

Table 4.2: DIC values for simulated scattering inversions. Column names are the assumed scattering mechanism, row names are the true scattering mechanism. Bold values indicate lowest DIC in the group. The correct choices are on the main diagonal.

Mud				Sand			
	Volume	Interface	Mixed		Volume	Interface	Mixed
Volume	**417.5**	556.7	469.1	Volume	**413.4**	861.8	612.3
Interface	495.7	**427.4**	431.4	Interface	418.9	418.5	**416.5**
Both	491.6	517.2	**441.1**	Both	499.7	511.38	**481.8**

correctly identifies the DSM in five out of six cases. Four of the five correct classifications have definitive DIC differences (≥ 10). The remaining correct classification has a border-line significant DIC difference of 4 (interface scattering over mixed scattering for a muddy seabed). The mis-identified case (mixed scattering chosen over the correct mechanism of interface scattering for the sand seabed) has an insignificant DIC difference of ≤ 2 (i.e., the classification is inconclusive).

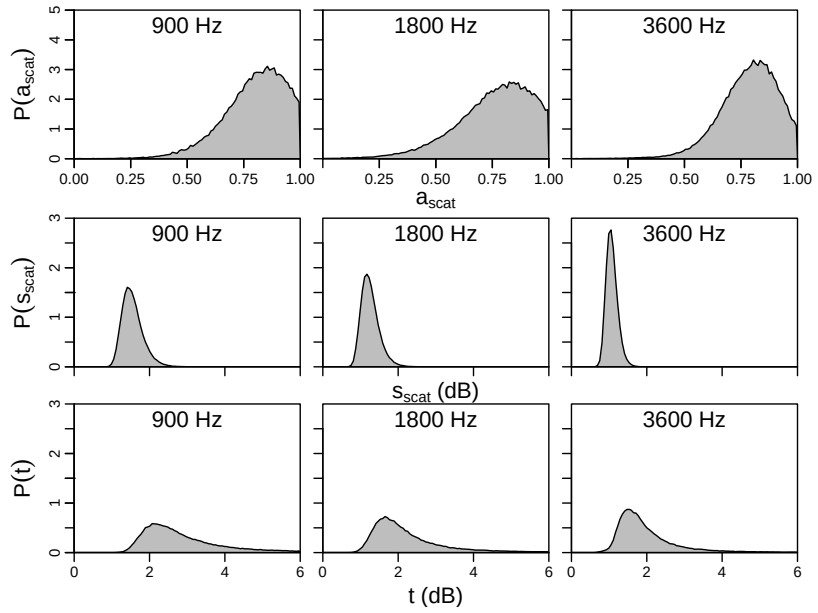

Figure 4.5: Marginal posterior distributions for the scattering residual parameters a_{scat}, s_{scat}, and correlated standard deviation t.

4.7 Malta Plateau scattering inversion results

In this section the DSM classifier is applied to the joint inversion[72] of measured scattering and reflection-coefficient data at Site 1 on the Malta Plateau (Fig 4.1). The DIC values for volume, interface, and mixed scattering are -1014, -841, and -1003, respectively. Thus, volume scattering is definitively preferred over interface and mixed scattering (DIC ≥ 11). Note the negative DIC values result from the small ($\ll 1$) reflection data residual standard errors. The remainder of this chapter will consider only the volume scattering results.

Figure 4.3 shows the fit of the volume scattering inversion to the measured data. The data predictions follow the trend of both the scattering and reflection data. The cutoff angle θ_c (where the standard deviations of the reflection data residuals changes from $S_{\text{ref}}^{\text{L}}$ to $S_{\text{ref}}^{\text{H}}$, see Sec. 4.3) is estimated to be $\sim 15°$. This accounts for the heterostochasticity in the reflection data residuals (particularly visible in the 800, 1008, 1270, and 2016 Hz frequency bands of Fig. 4.3). Figure 4.5 shows the marginal posterior probability densities of the scattering residual AR(1) and standard deviation parameters. All three frequency bands have only ~ 0.5 % of samples with uncorrelated ($a_{\text{scat}} = 0$) residuals; i.e., correlated errors are clearly indicated. The residual standard deviation parameters represent the uncorrelated or conditional standard deviation;

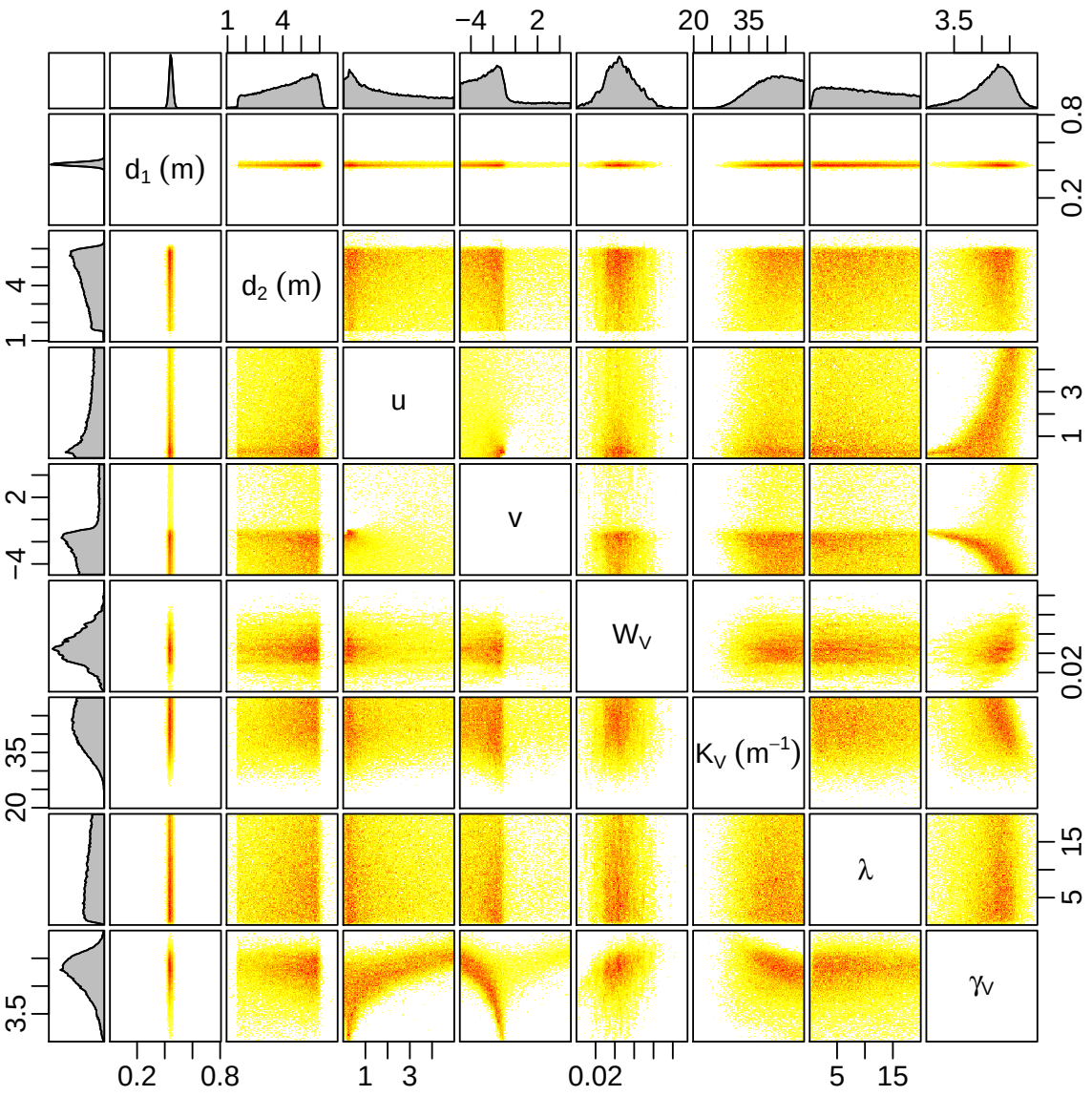

Figure 4.6: Marginal and joint marginal posterior probability distributions for volume scattering parameters.

however, the correlated/unconditional standard deviation (denoted t) gives a more natural description of the data fit and is calculated as[81]

$$t = \sqrt{\frac{s^2_{\text{scat}}}{1 - a^2_{\text{scat}}}}. \tag{4.15}$$

The posterior medians of t for the three scattering frequency bands (900, 1800, and 3600 Hz) are 2.6, 2.0, and 1.8 dB, respectively (Fig 4.5).

Table 4.3: The MAP model, 95% central credibility interval (CI), and prior bounds for volume-scattering parameters and coefficient of variation (assuming volume scattering as the DSM). Note where prior bounds are dependent on other parameters the minimum/maximum possible value is given.

Parameter	Prior Bounds	95% CI	MAP
d_1 (m)	[0, 10]	[0.418, 0.472]	0.441
d_2 (m)	[0, 10]	[1.8, 6.1]	5.57
u	[0, 10]	[0.00, 4.60]	0.42
v	[-5, 5]	[−5.00, 3.4]	−2.26
w_{V} ($\mathrm{m}^{(5-\gamma_{\mathrm{V}})}$)	[0, 1]	[1.35×10^{-2}, 8.02×10^{-2}]	4.31×10^{-2}
K_{V} (1/m)	(0, 50]	[31.7, 50.0]	38.1
λ	(0, 20]	[0.6, 18.9]	1.04
γ_{V}	[3, 5]	[3.36, 4.48]	4.07
Σ (%)	–	[1.4, 124.68]	12.0

Posterior marginal and joint marginal densities for all volume scattering parameters are shown in Fig. 4.6; the MAP and 95% central credibility intervals are given in Table 4.3. The distribution for the depth at which volume heterogeneities begin, d_1, is well defined with a mode near 0.44 m and only a few centimeters of uncertainty. The distribution for d_2, where volume heterogeneities end, is less well defined extending from 1.5–6.0 m. The precision scatterer onset is unexpected and the depth of 0.44 m is slightly shallower than indicated by the piston core (Fig. 4.2) where visible scatterers starting at ∼0.8 m. This discrepancy may be a result of the radically different lateral sampling scales of the core (0.1 m in diameter) and the acoustic data (500-m patch size).[15] The posterior distributions of u and v have modes near 0 and −1.5, respectively, and are consistent with reported values.[18] The spectral strength, spectral cutoff, vertical isotropy, and spectral exponent (w_{V}, K_{V}, λ, and γ_{V}) values are consistent with those reported in the Orcas and San Fransico experiments (Ref. [1], Table 14.1) obtained from analysis of small-diameter, short core samples which were analyzed in the vertical coordinate only.

It is of even more interest to compare the scattering parameters with those derived from reverberation measurements in the same general area.[19] For the comparison, it is necessary to point out that the reverberation measurements average over a patch size of 10^7 m^2, which is 4–5 orders of magnitude larger than the scattering-measurement patch size for the data here. At the 10^7 m^2 scale, there is some evidence for spatial stationarity[19] which indicates that the values of γ_{V} and K_{V} do not change drastically

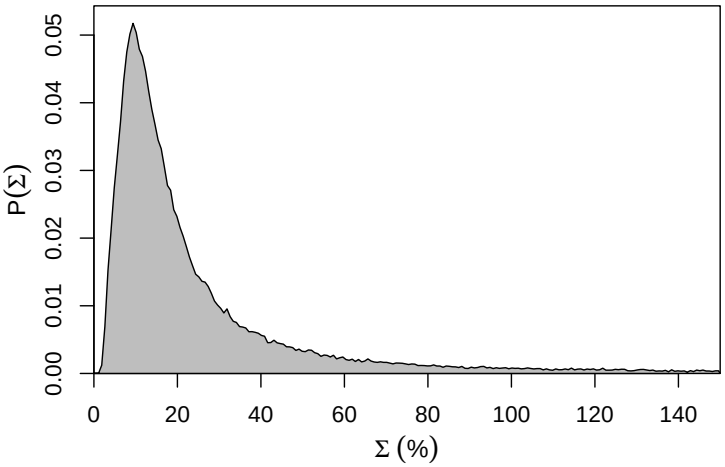

Figure 4.7: Marginal distribution of the coefficient of variation.

for various positions on the shelf. There is no hard evidence that both the scattering mechanism and the scattering characteristics in the small patch (from direct-path backscatter) are the same as for the large patch (from reverberation). This being the case, the comparison is still of interest. In the reverberation measurements, data from 1–3.2 kHz were interpreted as arising from volume scattering with $\gamma_{\mathrm{V}} = 3.9$ and $K_{\mathrm{V}} = 25$ m^{-1} (no uncertainty estimates), whereas the scattering inversion finds $\gamma_{\mathrm{V}} = 4.1 \pm 0.3$ and $K_{\mathrm{V}} = 38 \pm 5$ m^{-1}. These values are remarkably close given the completely different methods, and suggest that the scattering mechanism is similar between the two patch sizes.

Another way to interpret the scattering parameters is with the coefficient of variation (Σ), defined as the standard deviation divided by the mean of the random process of interest, converted to percent. For the density heterogeneity spectrum, Σ can be expressed as[18,90]

$$\Sigma = 100\% \times \pi^{3/4} \sqrt{\frac{w_{\mathrm{V}} \lambda \Gamma\left[\left(\gamma_{\mathrm{V}-3}\right)/2\right]}{K_{\mathrm{V}}^{(\gamma_{\mathrm{V}}-3)} \Gamma(\gamma_{\mathrm{V}}/2)}}, \tag{4.16}$$

where Γ is the standard Gamma function. Figure 4.7 shows the posterior distribution of Σ. The values are considerably higher than those reported from direct measurements at other sites;[1] however, this is likely due to the experiment site considered here which, unlike the sites considered in Ref. [1], has distinct scatterers (cobbles and shells). The 95% CI and MAP values for Σ are given in Table 4.3.

Figure 4.8 shows marginal posterior profiles for the geoacoustic parameters. The sound speed, density and attenuation profiles are consistent with previous inversions

at this site.[51,88] In addition, the profile matches the geoacoustic properties of the two cores taken at the site. Sound speed values are measured to ∼0.8 m depth from the gravity core and density values to ∼2.6 m from a piston core (shown in Fig. 4.2).[15] The posterior sediment sound-speed profile from inversion is in excellent agreement with the gravity core sound speed measurements. The density profile is in excellent agreement with the first 1.6 m of the piston core, but there is significant disagreement over the bottom 0.80 m of the core. This may be due to disturbance (compaction) in the bottom of the core caused by the piston coring process. The location of the geoacoustic layer interfaces is consistent with the visible interfaces of the piston core. The fourth panel in Fig. 4.8 (volume probability) shows the probability that heterogeneities (volume scatterers) are at the indicated depth. Almost all (∼ 99.997%) of the sampled models place scatterers across the sediment layer with high speed and density values that extends from approximately 1–1.5 m depth where the largest concentration of scatterers are visible in Fig. 4.2. The slow monotonic decrease in the probability of scatterers with depth suggests that scatterers and/or data sensitivity decrease with depth.

4.8 Summary and conclusions

This chapter considered three mechanisms for acoustic scattering at the seabed: scattering from a randomly rough water-sediment interface, scattering from volume heterogeneities within a sediment layer, and mixed scattering from both interface roughness and volume heterogeneities. These mechanism are modeled using first-order perturbation theory. A classification system based on the deviance information criterion is proposed for determining the dominant scattering mechanism for a measured backscatter data set. The DIC is similar to other model selection criteria in that it trades off data fit versus model complexity. However, it has the advantage of accounting for nonlinearity, prior distribution, and the effective number of parameters of the inverse problem. The classification system is validated using inversions of six simulated scattering data sets where it correctly identifies the true DSM as better than the other scattering mechanisms in five inversions; the remaining inversion case was unable to distinguish between interface scattering (the true scattering mechanism) and mixed scattering. The results of the simulations likely depend on the true scattering parameters assigned in each case, but these results indicate the general utility of the DSM classification approach.

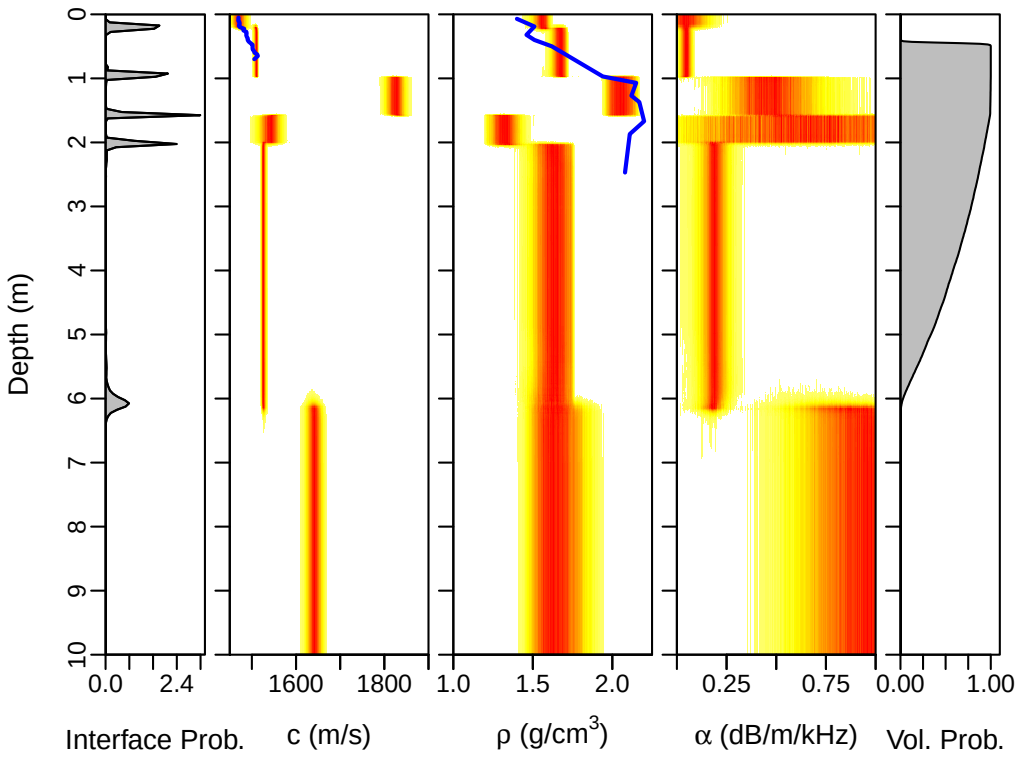

Figure 4.8: Marginal posterior geoacoustic profiles, solid line indicates core sample values. (Probability values for geoacoustic parameters are normalized independently for each depth for display purposes.)

Classification of the DSM is carried out for a measurement site on the Malta Plateau via joint inversion of backscatter and reflection-coefficient data. The DSM is found to be volume scattering with scatterers located from approximately 0.45–5.5 m in depth. The parameters of the density von Karman heterogeneity spectrum are well resolved. These heterogeneities are also interpreted as a 12% coefficient of variation. The seabed at the experimental site is found to consist of five sediment layers over a semi-infinite basement, with geoacoustic profiles in close agreement with previous geoacoustic inversions at the site. In addition, the geoacoustic parameters of the first 2–3 layer are in excellent agreement with the gravity and piston cores extracted at the site. Most of the layers and basement have relatively low sound-speed and density values; however, the inversion also finds a high sound-speed and density layer at approximately 1–1.6 m depth. This layer is also found to have a high probability of containing scatterers. This result along with the classification of volume scattering as the DSM is consistent with the piston core which shows a high density layer at a

similar depth range with 5–10 cm diameter stone scatterers.

Chapter 5

A trans-dimensional polynomial-spline parameterization for gradient-based geoacoustic inversion

This chapter presents a polynomial spline-based parameterization for trans-dimensional geoacoustic inversion. The parameterization is demonstrated for both simulated and measured data and shown to be an effective method of describing sediment geoacoustic profiles dominated by gradients. Specifically, the spline parameterization is compared to the standard stack-of-homogeneous-layers parameterization for the inversion of bottom-loss data measured at a muddy seabed experiment site on the Malta plateau. Inversion results for both parameterizations are in good agreement with measurements on a sediment core extracted at the site. However, the spline parameterization more accurately resolves the power-law like structure of the core density profile. In addition the spline parameterization is found to be more parsimonious, and hence preferred, according to the deviance information criteria. The trans-dimensional polynomial spline approach is general, and applicable to any inverse problem for gradient-based profiles.

5.1 Introduction

Many ocean acoustic modeling and sonar applications require knowledge of seabed geoacoustic properties and their associated uncertainties. Of particular interest in this chapter are the geoacoustic properties of marine muds which have received relatively little attention despite their common presence. Geoacoustic profiles of soft muddy sediments often involve parameter variations with depth in terms of continuous gradients rather than as discontinuous changes between homogeneous layers as is common for harder seabeds. This chapter develops a new and general approach to profile parameterization in terms of a polynomial spline with an unknown number of nodes. This approach can provide a more efficient and effective model than a stack of homogeneous layers for gradient-based problems.

Bayesian inference provides a rigorous framework to estimate parameters and quantify their uncertainties in geoacoustic inversion.[7,35–40] In particular, trans-dimensional (trans-D) Bayesian inversion provides an effective, automated approach to model selection (e.g., determining the number of sediment layers consistent with the resolving power of the data), such that the parameter estimate uncertainties account for model selection uncertainties.[26,28,32–34,52,69,72] However, to date trans-D Bayesian inversion has only been used to account for levels of complexity within a given type of parameterization (e.g., discontinuous layered models) not to compare different types of parameterizations. As the concept of model parsimony is linked to the choice of model parameterization type, the posterior uncertainty of parameter estimates are still affected by *a priori* parameterization selection decisions. This work highlights the importance of model parameterization in Bayesian inversion approaches by presenting a new type of parameterization for geoacoustic structure based on trans-D polynomial splines, and compares it with the common, stack-of-homogeneous-layers parameterization. The comparison is illustrated here using two examples of geoacoustic inversion of bottom-loss (BL) data. The first considers simulated data allowing for a meaningful evaluation of the inversion procedure; the second considers measured data and shows the practical value of the new parameterization. The preferred parameterization is selected using the deviance information criterion[67] (DIC), which trades off model complexity (accounting for the effective number of parameters) with data fit to determine the most parsimonious, and hence preferred, model.

For both inversion examples considered here the spline parameterization is found to be superior (lower DIC) to the stack-of-homogeneous-layers parameterization. For

the measured data both parameterizations are found to adequately fit the data and produce roughly similar geoacoustic profiles. However, the spline parameterization has fewer effective parameters and smaller posterior uncertainties, and its geoacoustic profile is in closer agreement with core measurement taken at the experiment site.

5.2 Theory

5.2.1 Parametrization

This work considers two approaches to parametrization of geoacoustic profiles. The first is the ubiquitous description of the sediment as a stack of j homogeneous layers over a basement half-space with discontinuities in parameter values at layer boundaries. This parameterization is referred to here as the "pancake" model. In the geoacoustic inversion considered here, each layer has an associated depth z, sound speed c, density ρ, and attenuation α. The second parametrization, developed and applied to geoacoustic inversion for the first time here, describes the sediment structure above the basement in terms of splines consisting of cubic polynomial segments[49] and is referred to as the spline model. Each geoacoustic property (c, ρ, or α) is described by an independent spline to prevent over- or under-parameterizing one property due to structure in another. This choice may not be obvious from a geological standpoint where sediment properties might be expected to change as a group at interfaces instead of individually. However, acoustic data vary in their sensitivity to different geoacoustic properties such that differing amounts of structure can generally be resolved for different properties. Consequently, linking the geoacoustic properties in the inversion may result in under- or over-parameterizing the model, which can be avoided by independent profiles.

The splines for each property have one node at the water-sediment interface (0-m depth) and one at the sediment-basement interface, and may have additional nodes at intermediate depths as required to fit the data. Each node has a depth z, property value x (corresponding to c, ρ, α), and property derivative or slope x'. The splines are created by using a third-order interpolation polynomial to link each adjacent pair of nodes while maintaining the property value and slope (x and x') at each node. The inversion is carried out treating both the number of nodes and node parameters for each geoacoustic property as unknowns which are sampled using Markov chain Monte Carlo methods (described in Sec. 5.2.3). However, to apply the forward theory

to predict data for candidate models, the splines must be discretized and transformed into a stack of homogeneous layers, as described in Sec. 5.2.2. To be clear, the inversion scheme samples node parameters; these are transformed into spline parameters which define the continuous-gradient property profiles. These spline profiles are in turn transformed into homogeneous-layer profiles for data prediction via the forward model.

For a given set of depths, properties, and derivatives for adjacent nodes i and $i+1$, the vector of coefficients $\mathbf{b}_i$ of the interpolation polynomials are found by solving the linear system of equations

$$\begin{pmatrix} 1 & z_i & {z_i}^2 & {z_i}^3 \\ 1 & z_{i+1} & {z_{i+1}}^2 & {z_{i+1}}^3 \\ 0 & 1 & 2z_i & 3{z_i}^2 \\ 0 & 1 & 2z_{i+1} & 3{z_{i+1}}^2 \end{pmatrix} \times \begin{pmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} x_i \\ x_{i+1} \\ x'_i \\ x'_{i+1} \end{pmatrix}. \tag{5.1}$$

If $\mathbf{A}_i$ is the sensitivity matrix in Eq. (5.1) and vector $\mathbf{x}_i$ lists the property values and derivatives, the solution is $\mathbf{b}_i = \mathbf{A}_i^{-1}\mathbf{x}_i$. The nodes are allowed to be arbitrarily close to represent discontinuous changes; this results in the possibility of $\mathbf{A}_i$ being singular or ill-conditioned. To avoid numerical instability the matrix is inverted using singular value decomposition.[91]

5.2.2 Discretization of spline models

The discretization of each interpolation spline for forward modeling is accomplished through an iterative procedure of splitting the spline into segments (note these do not correspond to the nodes used to create the spline). A segment is split if the difference between the maximum and minimum values of the spline are greater than Δx (there is one Δx for each spline property c, ρ, α and they are treated as unknowns which are sampled in the inversion) and if it is longer than a pre-specified minimum length (3 cm in depth is applied in the BL inversions considered here). When the procedure starts, the spline is considered to be in one segment; it continues until either no segment needs to be split or it reaches a pre-assigned maximum number of homogeneous layers (20 is used in this work). At each iteration of the discretization process the segment with the largest difference between the maximum and minimum values is split at its midpoint (equidistant from the two extrema). The algorithm starts with c, then ρ, and finally α; once a split is present in one spline it is used in all other splines. After

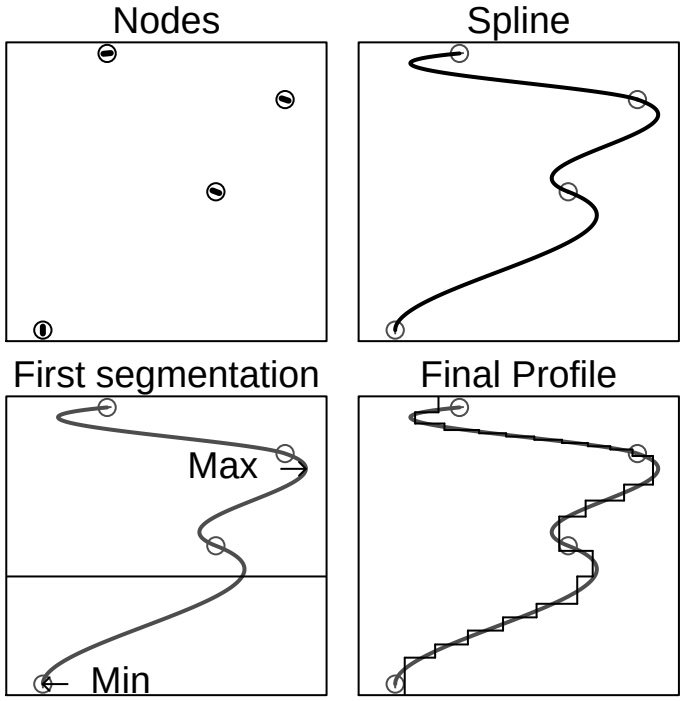

Figure 5.1: Example of the process of transforming node parameter values into a layered profile.

the splitting process is complete, each segment is assigned a uniform property value taken as the mean value of the spline within the segment (determined by integration). Figure 5.1 illustrates the steps of the splitting algorithm: The process starts with only nodes. The nodes are splined together. The splines are then segmented to produce a layered model for which the data can be predicted.

5.2.3 Bayesian inversion

A complete description of Bayesian inference as applied to trans-D geoacoustic inverse problems can be found in Refs. [26,28,32,69] and the specific implementation used here for the pancake parametrization is given in Ref. [71] This section gives the implementation for the new spline parametrization, and highlights the differences between the two parameterizations. Let $\mathbf{m_j}$ represent an arbitrary model where $\mathbf{j}$ represents a choice of model complexity [e.g., $\mathbf{j} = (j_c, j_\rho, j_\alpha)$ gives the number of internal nodes in each of the c, ρ, and α splines]. Using Bayes' rule, the trans-D

posterior probability density (PPD) can be written[68]

$$P(\mathbf{m_j}|\mathbf{d}) = \frac{\pi(\mathbf{m_j})\mathcal{L}(\mathbf{m_j})}{\mathcal{Z}}\,, \tag{5.2}$$

where $\pi(\mathbf{m_j})$ is the prior distribution of $\mathbf{m_j}$, $\mathcal{Z}$ is the total evidence of the ensemble of models, and $\mathcal{L}(\mathbf{m_j})$ is the likelihood of the parameter vector. Assuming Gaussian-distributed errors, the likelihood can be written

$$\mathcal{L}(\mathbf{m_j}) = \frac{1}{(2\pi)^{N/2} \mid \mathbf{C}_d \mid^{1/2}} \exp\left(-\frac{1}{2} \left[\mathbf{d} - \mathbf{d}(\mathbf{m_j}) - \mathbf{d}(\mathbf{a})\right]^\top \mathbf{C}_d^{-1} \left[\mathbf{d} - \mathbf{d}(\mathbf{m_j}) - \mathbf{d}(\mathbf{a})\right] \right), \tag{5.3}$$

where $\mathbf{d}(\mathbf{a})$ is a first-order auto-regressive process [AR(1)] prediction, and $\mathbf{C}_d$ is a diagonal data covariance matrix.[33,71] A more complete description of the residual statistics is given in Sec. 5.3.

When the pancake parameterization is applied, priors are generally taken to be bounded uniform distributions. For the spline model, the prior for the node slopes x' is taken to be a standard Cauchy distribution. This is equivalent to assuming that any node angle in the range $(-\pi/2, \pi/2)$ is equally likely; the equivalence follows from conducting a change of variables from angle to slope and noting that the cumulative density function of a Cauchy distribution is equivalent to the arc-tangent function.[20] In addition, the property values of the homogeneous layers from the discretized spline must satisfy empirical constraints on the joint prior for c and ρ given in Ref. [71]; this creates a non-uniform prior distribution. An important subtlety of the spline parametrization is that although the splines for all geoacoustic properties are independent, they share the same basement-depth parameter z_b: this results in the partition prior given by extension of the Dirichlet prior derived in Ref. [71] as

$$P(\mathbf{z}|\mathbf{j}, z_b) = \frac{(j_c - 1)!(j_\rho - 1)!(j_\alpha - 1)!}{z_b^{j_c + j_\rho + j_\alpha}}\,. \tag{5.4}$$

As is commonly the case in non-linear Bayesian inversion, analytic interpretation of the PPD is not possible and numeric sampling is used to approximate it. Sampling is conducted here using the reversible jump Markov chain Monte Carlo algorithm[49,68] with dimension jumps conducted independently for the splines of each geoacoustic property. Parallel tempering,[53,71,85,86] a population-based sampling scheme, is also used to improve sampling efficiency.

5.2.4 Deviance information criterion

There are several ways to determine which parametrization is best supported by the data. Ideally the evidence [$\mathcal{Z}$ from Eq. (5.2)] would be calculated for competing parametrizations to determine the preferred model.[64] However, evaluation of evidence is computationally expensive and intractable in practice for the present problem due to the large number of unknowns.[92] Instead, the deviance information criterion[67] (DIC) is used here. The DIC is a measure of parsimony similar to other more common model-selection criteria such as the Bayesian information criterion[70] which has been used previously in geoacoustic inversion.[43,89] However, the DIC has the advantages that it accounts for the prior distribution, parameter correlations, and the general non-Guassiananity of the PPD.[67] The DIC trades off the data fit of a characteristic model against a complexity term and is defined

$$\mathrm{DIC} = D\left(\hat{\mathbf{m}}_{\hat{\mathbf{j}}}\right) + 2P_D\,, \tag{5.5}$$

where D is the posterior deviance defined as

$$D\left(\mathbf{m_j}\right) = -2\log\mathcal{L}\left(\mathbf{m_j}\right) \tag{5.6}$$

and $\hat{\mathbf{m}}_{\hat{\mathbf{j}}}$ is a central or characteristic model. The term P_D in Eq. (5.5) is the effective number of what are referred to as focused parameters (parameters that are not marginalized out of the posterior before the DIC is calculated). In the cases considered here no nuisance- or hyper-parameters are marginalized out of the comparison because both parametrization have similar hierarchical structures.[67] The effective number of parameters is evaluated as $P_D = \overline{D\left(\mathbf{m_j}\right)} - D(\hat{\mathbf{m}}_{\hat{\mathbf{j}}})$, where $\overline{D\left(\mathbf{m_j}\right)}$ is the mean of the posterior deviance. Here the characteristic model $\hat{\mathbf{m}}_{\hat{\mathbf{j}}}$ is taken to be the maximum *a posteriori* (MAP) model (the choice of the MAP model instead of other possible characteristic models such as the mean or median models is made to accommodate the trans-D structure of the PPD). As with other similar criteria, the model with the lowest DIC has the greatest support from the data and is preferred.

5.3 Forward model

Acoustic bottom-loss (BL) data are defined as

$$\mathrm{BL}(\theta, f) = -20 \log_{10} |R(\theta, f)|, \tag{5.7}$$

where θ is grazing angle, f is the frequency, and R is the plane-wave reflection coefficient for a layered medium.[14] Both BL[51,76,93] and reflection coefficients[42,72] have been applied in geoacoustic inversion. The choice to use BL instead of $|R|$ is made here because BL is more sensitive to the small reflection amplitudes near the angle of intromission for low sound-speed sediments.[94]

In addition to modeling the physics of BL, it is also necessary to define a statistical model for the data residuals $\mathbf{r} = \mathbf{d} - \mathbf{d}(\mathbf{m_j})$ (observed minus predicted data). Here the residuals for each frequency band are modelled as first-order auto-regressive [AR(1)] heterostochastic zero-mean Gaussian processes (over angle), i.e.,

$$r_i = a^{\Delta\theta_i} r_{i-1} + e_i, \tag{5.8}$$

where a is the AR(1) parameter and e_i is referred to as the total residual.[71] The e_i are assumed to be independent Gaussian random variables; the heterostochasticity results from including two unknown variance parameters, low-angle and high-angle parameter, S_L^2 and S_H^2, respectively. The change in variance is included since reflection-coefficient or BL errors are commonly observed to change structure near the critical or intromission angle.[72] The cut-off angle θ_c used to determine the angular range of each variance is treated as an unknown and sampled in the inversion. The exponent ($\Delta\theta_i = \theta_i - \theta_{i-1}$) in Eq. (5.8) is used to account for uneven data spacing which is common in measured reflection data.

5.4 Simulation Study

This section describes the inversion of a simulated BL data set using both the spline and pancake models. The simulated data were generated using the true model shown in Fig. 5.2, which was designed to have both gradients and large discontinuities. Data were created at a range of frequencies and intensity averaged over third-octave bands centered at 1000, 1300, 1600, 2050, 2500, 3250, and 4000 Hz. The data have an angular range of 7–36° with ∼130 measurements per frequency band. Uncorrelated

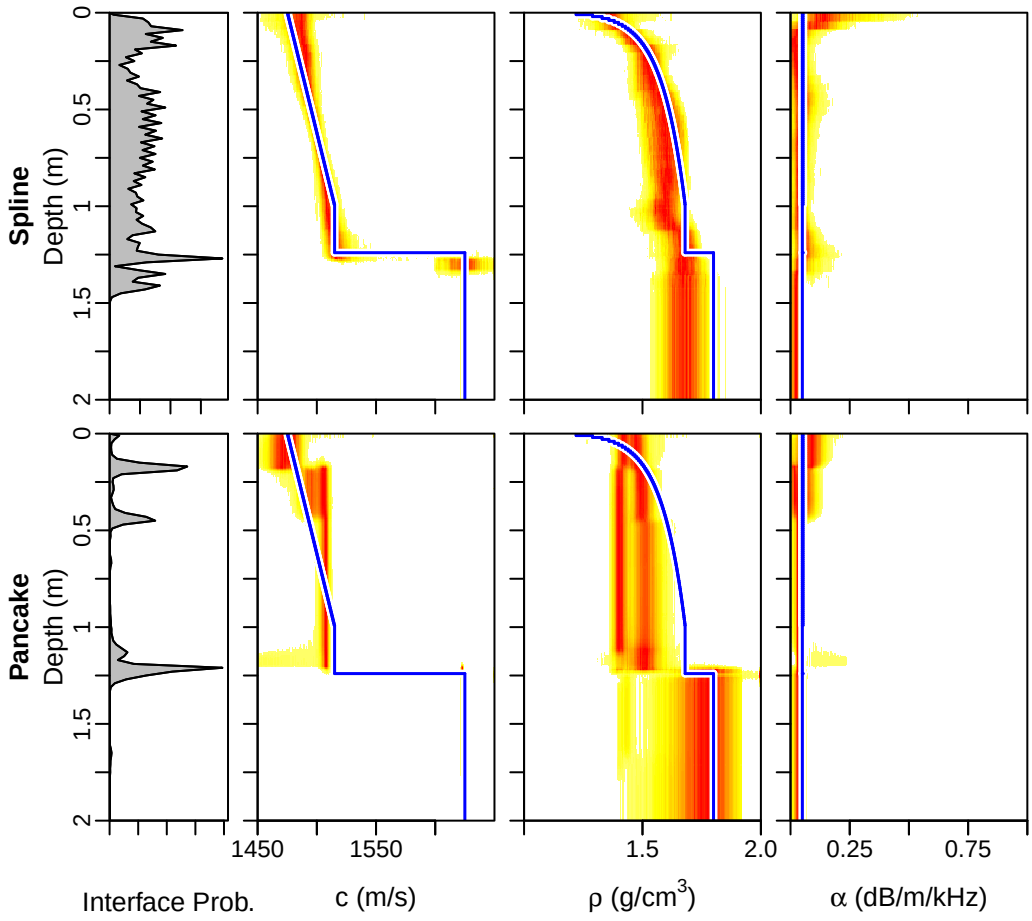

Figure 5.2: Marginal posterior geoacoustic profiles from spline (top) and pancake (bottom) inversions for simulated data. Probability values are normalized independently at each depth for display purposes. Solid lines indicate true geoacoustic parameterizations.

Gaussian random errors where added to the true data to produce noisy simulated data; the standard deviations for all frequencies are $S_L = 1.0$ dB and $S_H = 1.5$ dB with $\theta_c = 22°$. The data are shown in Fig. 5.3.

Trans-D Bayesian inversion was applied to the BL data in Fig. 5.5 using both the standard pancake and the new spline parameterizations. For both inversions (spline and pancake) approximately 3,000,000 samples were collected; the first 1,000,000 samples were discarded to prevent the arbitrary starting model from impacting the sample distribution. The remainder was chain-thinned to 200,000 samples to reduce sample autocorrelation and allow for more convenient storage and analysis.

Figure 5.2 displays the marginal posterior probability profiles for the three geoa-

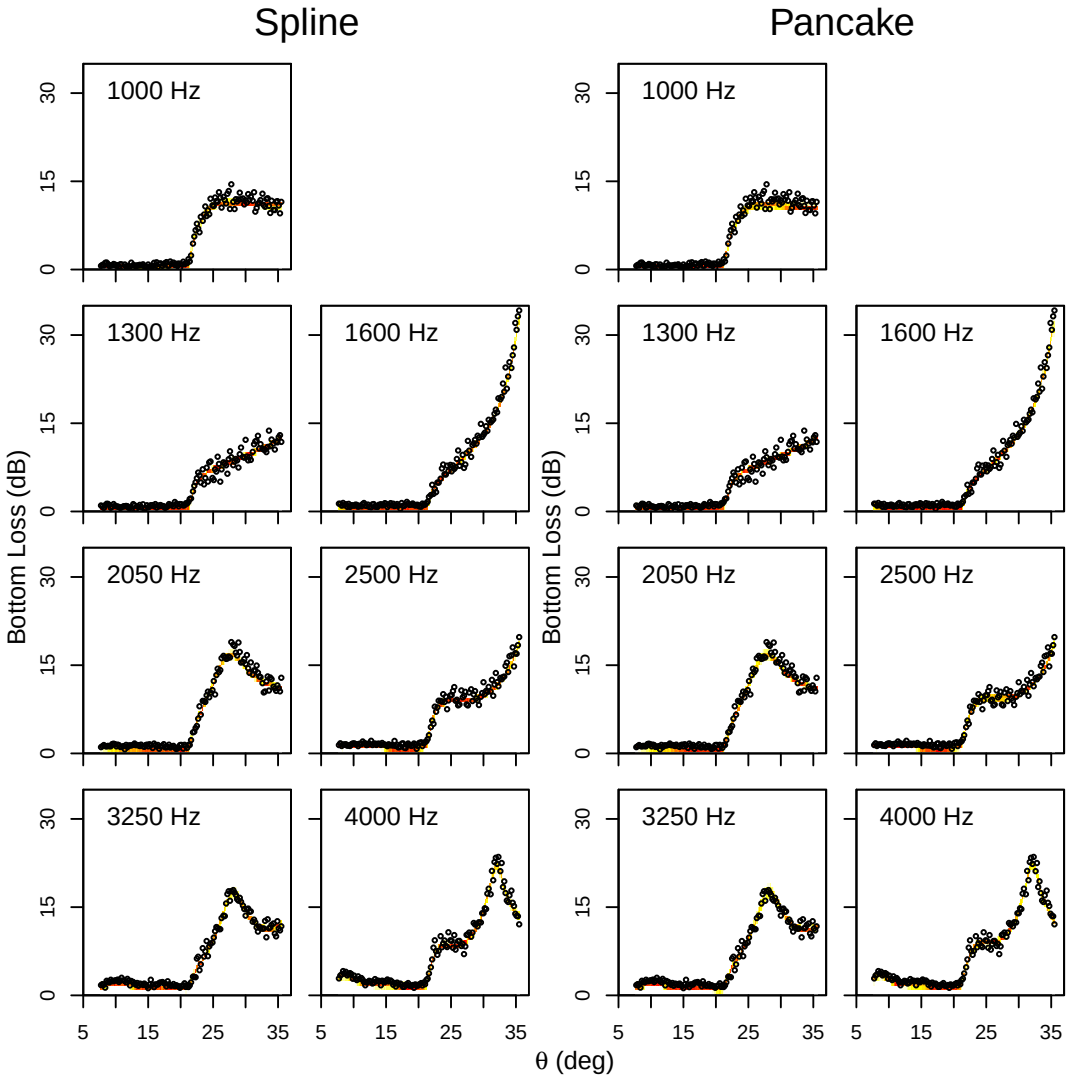

Figure 5.3: Simulated (∘) and predicted bottom loss data at frequencies indicated for the spline (left) and pancake (right) parametrization.

coustic properties for both the spline and pancake inversions. The spline inversion more accurately represents the gradient in sound speed over the top 1 m. Both inversions capture the sound speed discontinuity at 1.25 m well. Note also that both inversions estimate the basement sound speed to high accuracy such that the marginal distribution is hidden by the line indicating the true value. For density, the spline model more accurately captures the power-law like gradient structure. However, the pancake model better localizes the discontinuity in ρ, although the discontinuity is substantially over estimated by the pancake model. Both models are generally close

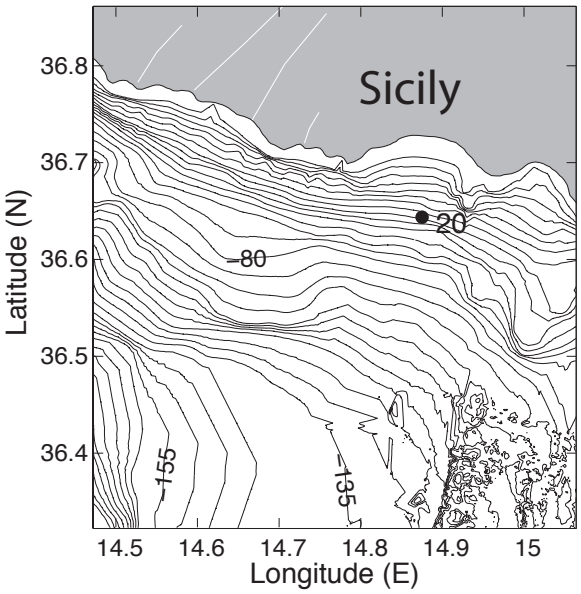

Figure 5.4: Bathymetry (in meters) and experiment site 20 location on the Malta Plateau.

to the true attenuation profile, although both indicate somewhat high α values for the surficial sediments. The DIC values for the spline and pancake models are ∼2598 and 2612, respectively, a difference of 14. In interpreting the DIC a difference of more than 10 is considered definitive;[67] hence, the spline model is clearly preferred over the pancake model in this simulation example.

5.5 Malta Plateau inversion

The measured data analyzed in this chapter were collected at a test bed (site 20) on the Malta Plateau in the Strait of Sicily, shown in Fig. 5.4. The sediment at the test site is known to consist of soft mud from a gravity core extracted at the site. Bottom-loss (BL) data were collected using a fixed receiver and a ship-towed impulsive source; full details of the experiment procedure are given in Ref. [55]. Data processing included time windowing to remove the effect of geoacoustic structure below approximately 2.0 m. The measured data shown in Fig. 5.5 have the same angular range and frequency bands as the simulated data considered in the previous example.

The same trans-D Bayesian inversion scheme used for the simulation example was applied to the measured BL data. The posterior fit to the data for the two inversions

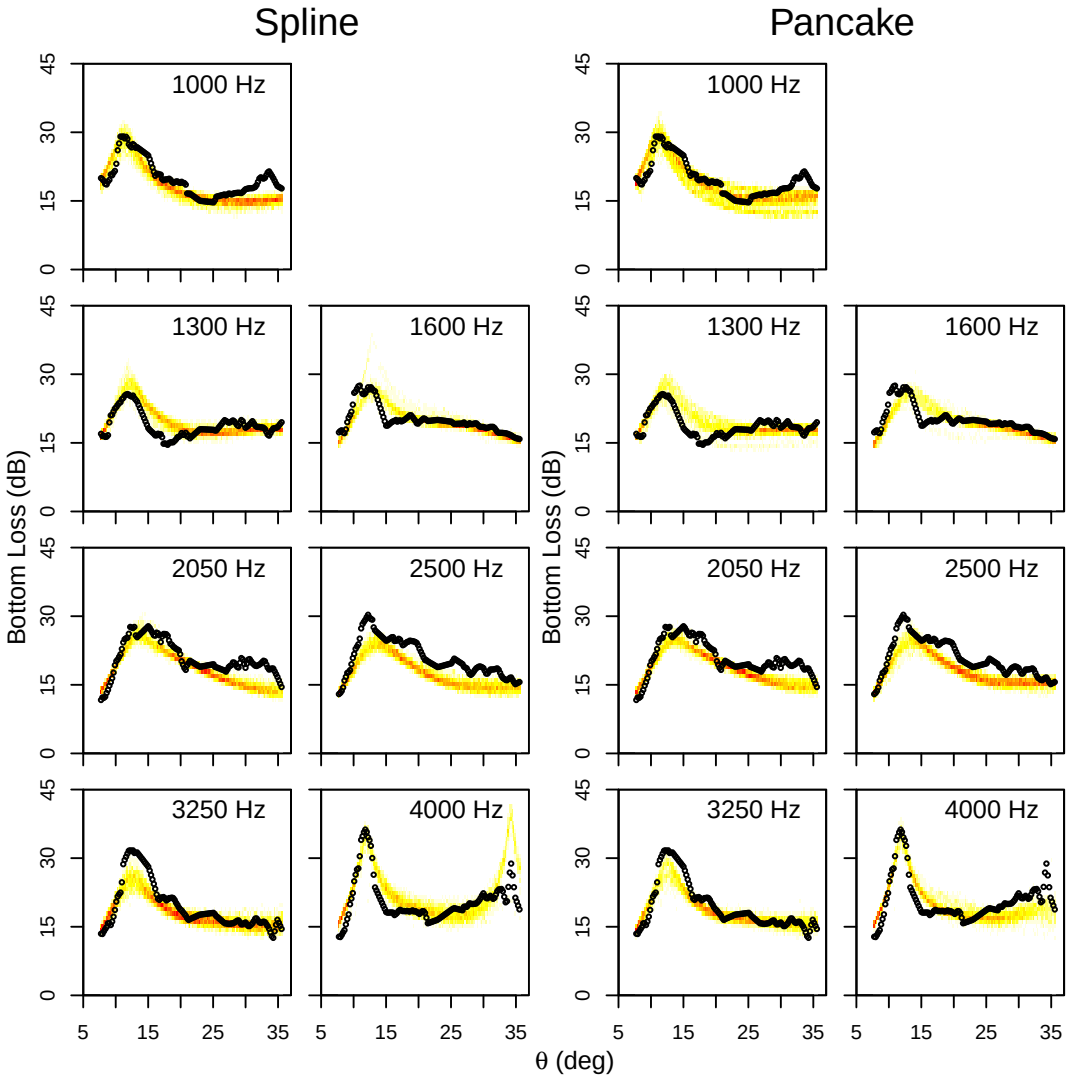

Figure 5.5: Measured (∘) and predicted bottom loss data at frequencies indicated for the spline (left) and pancake (right) parametrization.

is shown in Fig. 5.5. Both parameterizations appear to fit the general structure of the data quite well; the fits of both inversions for most frequency bands are almost indistinguishable. The spline model appears to fit better at the higher angles of the 1000 Hz band, while the pancake model fits better at the lower angles of the 3250 Hz band. Only the spline inversion caught the secondary peak in BL at 4000 Hz, although it does over estimate it. Detailed structure (fluctuations) in the data are generally not fit by either model predictions. The residual structure (or auto-correlation) is accounted for through the AR(1) processes included in the inversion.

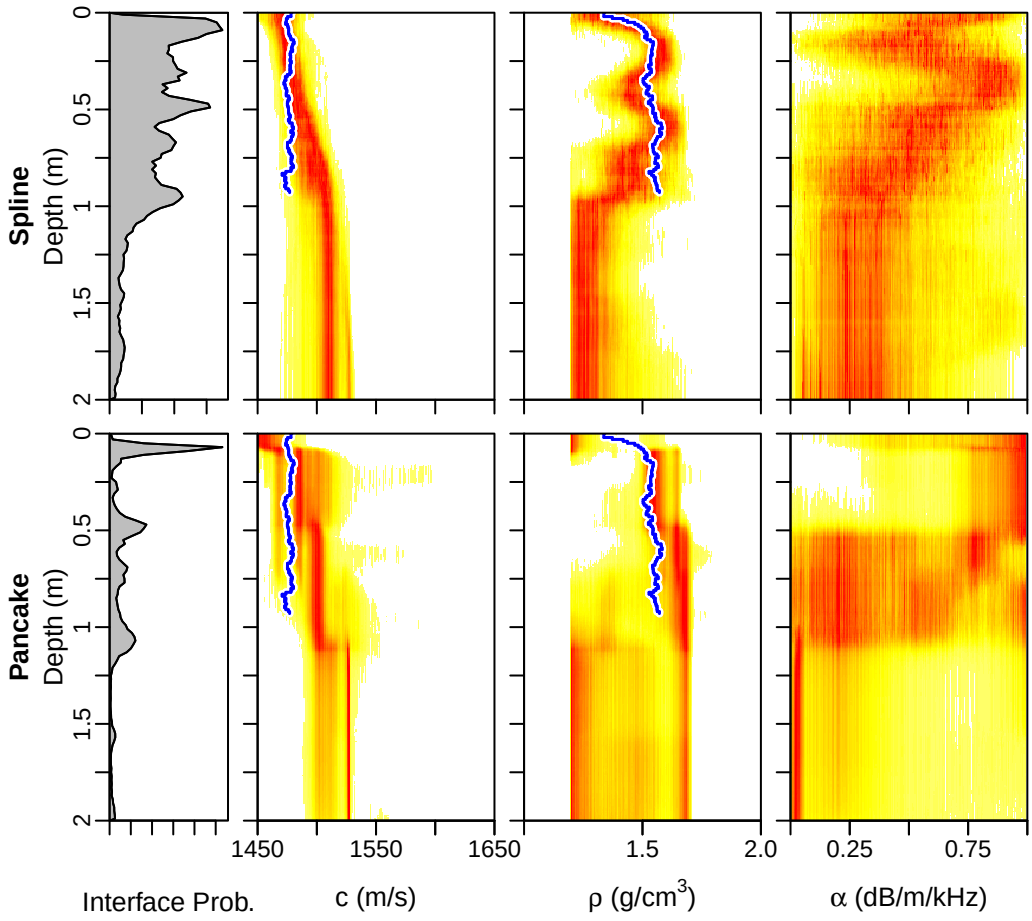

Figure 5.6: Marginal posterior geoacoustic profiles from spline (top) and pancake (bottom) inversions. Probability values are normalized independently at each depth for display purposes. Solid lines indicate values measured for a core extracted at the site.

The marginal geoacoustic probability profiles for the two inversions are shown in Fig. 5.6 and 95% central credibility intervals and median values are shown in Fig. 5.7. Both inversion results indicate generally similar sediment structure: c increases with depth over the first meter and there is a sharp decrease in ρ at approximately 1 m (unfortunately the core did not penetrate deep enough into the sediment to verify the existence of this feature). For both models the marginal profiles are in reasonable agreement with the core measurements which extend over the top 0.9 m. In particular, the spline inversion is in excellent agreement with the core density profile. The power-law like gradient over the top $\sim$0.1 m is well reproduced, and there appears to be some sensitivity to variations in ρ between 0.3–0.9 m. Both the spline and pancake

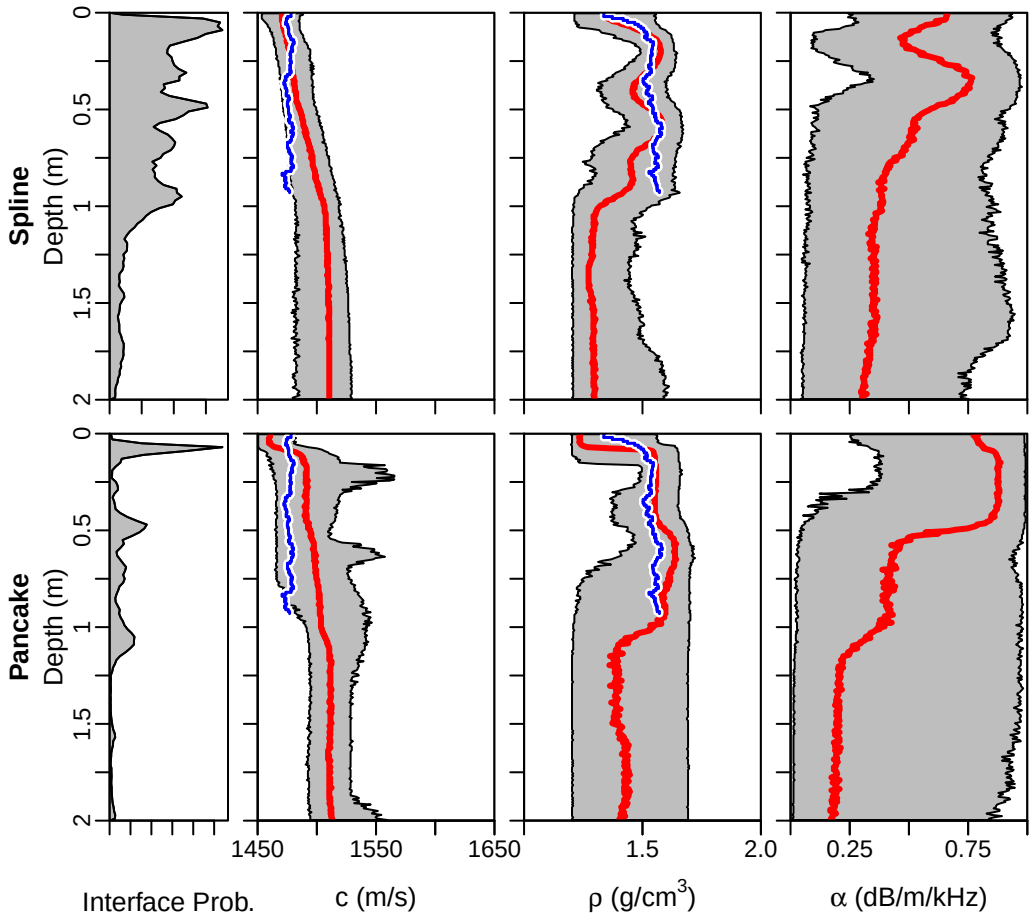

Figure 5.7: 95% central credibility interval (shaded regions) with median model (heavy lines) for the geoacoustic profiles from spline (top) and pancake (bottom) inversions. Medium solid lines indicate values measured for a core extracted at the site.

inversions indicate an increase in c over the top 1 m which is not present in the core although the spline is closer overall to the core values. The α values for both inversions are 1 or 2 orders of magnitude too high for intrinsic attenuation in muddy sediments. Careful attenuation measurements of this same muddy fine-grained layer in the same frequency range showed values in the range 0.006–0.012 dB/m/kHz.[95] Though the two measurements are separated by roughly 10 km, seismic data clearly shows that the upper 10 m at both sites are of identical Holocene origin. One difference between the two sites is that the Ref. [95] measurements are an average over the upper 10 m of the mud, whereas this samples only the upper meter. However, attenuation gradients in muddy sediments are expected to be very small.[96] Thus, the inversion results

represent an effective attenuation, i.e., intrinsic plus some other loss mechanism. The peak of the bottom loss at the angle of intromission (where the reflection coefficient is nearly 0) is very sensitive to attenuation. Modeling (not shown) indicates that very small-scale sound speed and density fluctuations affect (reduce) the peak in the same way as an increase in attenuation. Since the information content in the data is not sufficiently high enough to capture these small scale fluctuations, the attenuation is biased high in the inversion mimicking the effect of the fluctuations.

Overall there is less uncertainty with the spline model for c and ρ than with the pancake model (see credibility interval widths in Fig. 5.7). The DIC values for the two parametrization are $\sim$1206 for the spline model and 1235 for the pancake model. Thus the spline model has a definitively lower DIC (by 29) and is the clearly preferred model. The P_D values for the two models are 31.1 and 42.3 and the total absolute misfits $[\Sigma\,|d_i - d_i\,(\mathbf{m_j})|]$ are 2232 and 2352 dB; thus, the spline parameterization has about 11 fewer effective parameters while still fitting the data better than the pancake model.

5.6 Summary and conclusions

This work considered two distinctly different types of parameterizations for describing geoacoustic profiles of marine sediment in geoacoustic inversion. These are the common discrete stack-of-homogeneous-layers parameterization (referred to here as the pancake model) and the new polynomial-spline parameterization (spline model). The spline model describes each sediment property with its own trans-dimensional polynomial spline. The splines are then discretized to be compatible with existing forward codes. The spline model naturally parametrizes environments with geoacoustic profiles dominated by smooth transitions or gradients (although discontinuities can be represented by large transitions over very thin layers). Alternatively, the pancake model more naturally parametrizes geoacoustic profiles with discrete discontinuities. The pancake model links the complexity of the different geoacoustic properties (e.g., if there is a change in ρ at some depth then a new c and α value must also be estimated at that depth), whereas the spline allows the complexities of the description of each property to be independent of the others. Consequently, the two models have different interpretations of parsimony and their parameter uncertainties are affected by model selection uncertainties differently. These two models do not, of course, represent an exhaustive list of possible parametrizations: indeed one of the goals of

this work is to demonstrate the value of exploring alternative parametrizations for trans-D inversion.

The deviance information criterion is introduced here to geoacoustic inversion as a method of determining which of several parameterizations is best supported by the data. The DIC is similar to other model selection criteria in that it trades data fit versus complexity. However, it has the advantage that it accounts for non-linearity, prior distribution, and the effective number of parameters of the inverse problem.

Geoacoustic inversions of bottom-loss data collected on the Malta Plateau were conducted using both the spline and pancake models. For both inversions the estimated geoacoustic profiles are in good agreement with the independent information about the site, i.e., the inversion results are consistent with a muddy sediment and geoacoustic values similar to those measured for a core sample. In particular, the density profile from the spline inversion closely matches the power-law like structure of the core. The sound speed is found to increase with depth, attenuation is found to decrease with depth, and density has more complex profile. The spline model was found to be definitively superior to the pancake parameterization in terms of DIC value, and results in a profile closer to that of the core.

It should be noted that inversions conducted here were *a priori* believed to have profiles dominated by gradient features. For problems dominated by large discrete discontinuities the pancake parameterization could be a superior choice. In cases where no *a priori* information exists, inversions should be performed with both parameterizations and the DIC employed to determine which is most appropriate. Finally the inversion approach developed here is general and could be applied to any inverse problem.

Chapter 6

Conclusions

The pursuit of innovative techniques for measuring shallow-water seabed parameters has motivated the ocean acoustic community to pursue research on reflection[89,97] and matched-field[5,52,53] geoacoustic inversion. However, little work on rigorous scattering inversion has been conducted. Scattering inversion is important because it allows the relaxation of idealized assumptions for acoustic environments (e.g., interfaces are planar and layers are homogeneous). These assumptions can be problematic in many regimes, e.g., ripples in sand, bubbles in mud, or shelly layers. In addition, the scattering mechanisms contain information about the geological and biological processes that create them.

The development of bottom-scattering databases is currently of great interest to the naval survey community. The approach developed in chapters 2–4 represents an important step as it allows for an automated inversion procedure to remotely estimate seabed scattering properties. This is a problem of practical interest for sonar applications which has received little attention to date.

Two types of acoustic data are used in the inversion procedure, representing measures of backscatter and reflectivity. The physics used to model the data are first-order perturbation theory for the scattering data and spherical-wave reflection coefficient for the reflectivity data. The acoustic scattering model accounts for scattering from a rough water-sediment interface and/or from sediment-layer volume heterogeneities, as well as depth-dependent geoacoustic profiles. Statistically independent von-Karman (power-law) spatial spectra are used to govern the interface roughness and volume-heterogeneity random processes. The primary measurement of interest is the backscatter strength as a function of angle and frequency; however, since the scattering kernel also depends on the geoacoustic profile, the reflection-coefficient data

are included to provide additional information about these parameters.

The thesis addresses model selection, an important issue in quantitative inversion, through trans-dimensional sampling and the deviance information criterion. Trans-D sampling is applied over the number of sediment layers and the number of data frequency bands with auto-correlated residuals. Trans-D sampling is conducted using the reversible jump Markov chain Monte Carlo algorithm with parallel tempering. This process allows the data to determine the level of complexity of both the geoacoustic and residual models without risk of over- or under-parameterization. To accommodate the depth-dependent number of parameters of the geoacoustic profile (i.e., fluid sediment layers over an elastic basement) required the development of a new trans-D partition prior distribution based on the Dirichlet distribution. Unlike existing partition priors for trans-D inversion, the Dirichlet prior allows for the explicit evaluation of the prior and posterior probabilities.

The DIC is introduced to geoacoustic inversion in this thesis. It is used to choose between parameterizations where trans-D sampling would be impractical to implement as the dimension jumps would involve adding or removing large numbers of parameters making acceptance unlikely. In particular, the DIC is used to determine the dominant scattering mechanism (interface or volume scattering) or the preferred parameterization (polynomial-spline or stack-of-homogeneous-layers).

Several realistic simulated data inversion studies are conducted to compare estimated parameter values with the true model. The simulations are based on observed scattering and geoacoustic parameters/profiles and contain finer scale structure than is resolvable by the data. Realistic (auto-correlated and heterostochastic) errors are added to the scattering and reflection data. In general the simulated inversions presented here are able to adequately resolve both the scattering and geoacoustic parameters.

The thesis also considers inversions of data sets recorded at three sites on the Malta Plateau (Mediterranean Sea). The data from the first site are used to conduct both a scattering-only and a joint scattering/reflection inversion. The joint inversion is taken as the preferred result. The inversion assumed that interface roughness is the dominant scattering mechanism for the site. The parameter estimates indicated that the seabed at the experiment site consisted of $\sim$6 m of sediment layers over a high-speed basement, with values consistent with sand layers over limestone. These results are in agreement with video observation of the seafloor and a high-resolution seismic section at the site. The parameters of the von-Karman spatial interface rough-

ness spectrum were well resolved and agreed with a collection of values derived from seafloor stereoscopic photography.

Data from the second site are used to demonstrate the procedure for identifying the dominant scattering mechanism using the DIC, which found volume scattering to be dominant. The scattering parameters and geoacoustic profile are well resolved. The volume heterogeneities (scatterers) are found to extend from 0.4–6.0 m depth. The geoacoustic profile indicates a high density and sound speed-layer at $\sim$1 m depth. The geoacoustic parameters and preference for volume scattering are consistent with a core extracted at the site showing large (0.1 m) stones located below approximately 1 m of mud at the seafloor.

In addition to scattering inversion, this thesis presents a trans-D polynomial-spline based parameterization which is useful for geoacoustic profiles based on gradients. The spline model is compared to the common discrete stack-of-homogeneous-layers parameterization, and is shown to be preferred (i.e., has both a lower data misfit and a more parsimonious parameterization) for regimes with smoothly-varying depth-dependent geoacoustic parameters. Comparisons are made using both simulated and measured data; the preferred model is determined using DIC. The polynomial-spline based parameterization is general and could be applied to any geoacoustic or geophysical inverse problem.

The successful use of this parameterization highlights the limitations to the objectivity of the posterior uncertainty. The choice of parameterization is inevitably somewhat arbitrary, i.e., an experanced user is unlikely to select a parameterization that is not *a priori* plausible; however, there are no formal criteria on how to chose an appropriate parameterization. The choice of parameterization then impacts the posterior uncertainty. Thus, although trans-D inversion improves over fixed-dimensional inversion by incorporating model selection uncertainty into parameter estimate uncertainty, it is still, like other statistical paradigms, limited by user choices. This leads to the final caveat of this thesis, that those intending to use the procedures described here should ensure that all levels of assumptions made in the inversion procedure are applicable, particularly those that do not seem like assumptions.

Appendix A

Joint empirical prior for sound speed and density

This appendix provides a detailed derivation of the joint empirical prior for sound speed (c) and density (ρ) which is used through out this thesis. It is generally accepted that sound speed and density of marine sediments have a strong inter-dependence; for example, a quadratic relationship of c on ρ was found empirically by Bachman.[98] Hence, it can be assumed *a priori* that sediments will not have a low sound speed and a high density or vice versa, even if the individual sound speed and density values are reasonable. Thus, independent priors for c and ρ can over-estimate the region of parameter space with plausible values for both parameters.

A.1 Sediment layer priors

More-constraining prior bounds which account for the relationship between c and ρ are derived here for sediment layers using a large compilation of sediment data presented by Hamilton.[83,84] The first step in estimating the joint bounds of c and ρ is to estimate their joint distribution. This accomplished by finding a conditional distribution $f_c(c|\rho)$ (a regression model) and multiplying it by the distribution of $f_\rho(\rho)$. For the Hamilton data, the residuals from regressions of c onto ρ are found empirically to be approximately t-distributed and the distribution of $\log \rho$ resembles

a truncated normal distribution. Thus an applicable joint distribution is

$$f_{c,\rho}(c,\rho) = \frac{\Gamma\left(\frac{q_1+1}{2}\right)\left(\sqrt{2q_1}\pi q_6 q_4\right)^{-1}}{\Gamma\left(\frac{q_1}{2}\right)\left(\Phi\left(\frac{q_2-q_5}{q_6}\right)-\Phi\left(\frac{q_3-q_5}{q_6}\right)\right)} \frac{\exp\left(-\frac{1}{2}\frac{(\log(\rho)-q_5)^2}{q_6^2}\right)}{\rho\left(1+\frac{(c+q_7+q_8\rho+q_9\rho^2)^2}{q_1 q_4^2}\right)^{\frac{q_1+1}{2}}}, \tag{A.1}$$

where Γ is the gamma function and Φ is the standard Gaussian cumulative density function. The parameters of the distribution, $q_1, \ldots, q_9$, are estimated to maximize the probability of the ρ and c pairs reported by Hamilton.[83,84]

The desired joint prior distribution for c and ρ is defined to be uniform within some bounds and zero outside these bounds, so the second step in creating the prior is to define bounds for the acceptable region of $f_{c,\rho}(c,\rho)$. A natural choice for these bounds is a closed isoprobability contour. Such a bound is convenient as it does not require the evaluation of the normalizing constant [the first term in Eq. (A.1) containing the gamma and cumulative density functions], as unnormalized and normalized isoprobability contours are coincident. The contour level is selected such that a desired cumulative probability mass of $f_{c,\rho}(c,\rho)$ is contained within. In the work presented here, the contour is selected to contain 99.5 percent of the probability mass. The shape of the contour is described as a function mapping ρ to upper and lower bounds of c:

$$\left.\begin{matrix} c^+ \\ c^- \end{matrix}\right\} = \pm\sqrt{q_1 q_4^2\left(\exp\left[-\left(\frac{(\ln\rho - q_5)^2}{2q_6^2}+\log\rho+\log P\right)\left(\frac{2}{q_1+1}\right)\right]-1\right)}-q_7-q_8\rho-q_9\rho^2. \tag{A.2}$$

The unnormalized probability of the contour is P. Where these joint bounds exceed the individual parameter prior bounds they are replaced with the individual prior bound values. The joint prior bounds for ρ and c developed for the sediment layers are shown in Fig. A.1(a).

A.2 Basement prior

Since it is may be unknown *a priori* if the basement layer is sediment or a consolidated material such as limestone, the joint prior bounds for basement sound speed c_b and density ρ_b must span a larger range than those of the overlying sediment layers. This is accomplished by augmenting the prior bounds of the sediment layers with prior bounds for limestone. The limestone prior is again derived using empirical data from

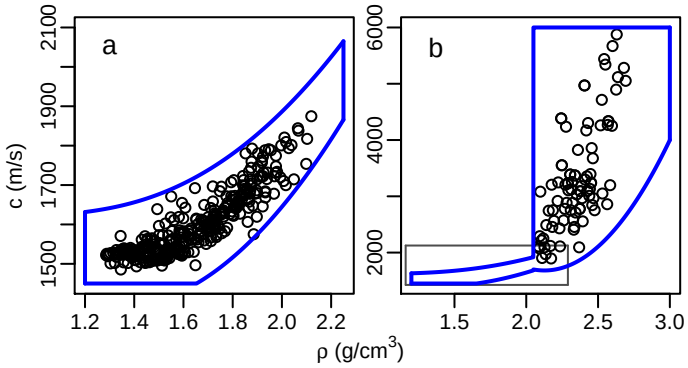

Figure A.1: Joint prior bounds for c and ρ for sediments (a) and basement (b). Hamilton [83,84] data are plotted as circles. The grey rectangle in (b) shows the extent of the sediment prior in the basement.

Hamilton.[83,84] As the individual parameter prior bounds are more restrictive than the joint prior upper bounds, only the lower bound of the joint prior for c_b and ρ_b in limestone needs to be defined. In this case the bound is quadratic, $c^- = q_0 + q_1\rho + q_3\rho^2$. The joint prior bounds are shown in Fig. A.1(b).

Appendix B

Equivalence of exponential-form parametric covariance matrix and a first-order auto-regressive process

This section shows that the use of a first-order auto-regressive process denoted AR(1),

$$r_i = ar_{i-1} + e_i, \tag{B.1}$$

to model data residuals

$$\mathbf{r} = \mathbf{d} - \mathbf{d}(\mathbf{m}_j) \tag{B.2}$$

is equivalent to using the analogous exponential residual covariance matrix. The residual covariance matrix can be written as $\mathbf{C}_r = \sigma_r^2 \mathbf{R}$ where σ_r^2 is the unconditional variance of the data residuals and $\mathbf{R}$ has element $\mathbf{R}_{ij} = a^{|i-j|}$. The proof is presented here as its absence from the geoacoustic and geophysical literature has resulted in some speculation the equivalence of the two representations. The log likelihoods for the two definitions [using $\mathbf{C}_r$ and AR(1)] are

$$\log \mathcal{L}_{\mathrm{C}}(\mathbf{m}) = -\frac{1}{2}\left(\mathbf{r}^T \mathbf{C}_r^{-1} \mathbf{r} + \log |\mathbf{C}_r| + \log (2\pi)\right) \tag{B.3}$$

and

$$\log \mathcal{L}_{\mathrm{AR}}(\mathbf{m}) = -\frac{1}{2}\left(\sum_{i=1}^{N} \frac{e_i^2}{\sigma_{e_i}^2} + \sum_{i=1}^{N} \log \sigma_{e_i}^2 + \log (2\pi)\right), \tag{B.4}$$

respectively, where $e_i = r_i - ar_{i-1}$ and $\sigma_{e_i}^2$ are the variances of the e_i.

An important step in establishing the equivalence of Eqs. (B.3) and (B.4) is showing the relationship between σ_r^2 and the $\sigma_{e_i}^2$. This can be found by considering the variance (V) of an arbitrary residual r_i. First, write r_i in terms of the e's,

$$r_i = ar_{i-1} + e_i = a\left(ar_{i-2} + e_{i-1}\right) + e_i = \sum_{j=1}^{i} a^{i-j} e_j = a^{i-1} e_1 + \sum_{j=2}^{i} a^{i-j} e_j \,. \tag{B.5}$$

Note this assumes that $r_1 = e_1$. Taking the variance of the left and right hand sides of Eq. (B.5)

$$\begin{aligned} V\left(r_i\right) &= V\left(a^{i-1} e_1 + \sum_{j=2}^{i} a^{i-j} e_j\right) \\ \Rightarrow \sigma_r^2 &= a^{2(i-1)} \sigma_{e_1}^2 + \sum_{j=2}^{i} a^{2(i-j)} \sigma_{e_j}^2, \end{aligned} \tag{B.6}$$

since e_i is independent of e_j, $i \neq j$. There are two straightforward ways to proceed: the first is to assume r_1 represents the start of the random process; the second is that the r_i represent only a section of an infinite random process (i.e., r1 is correlated with previous realizations which are not part of the observed section). The first assumption is equivalent to $\sigma_{e_i}^2 = \sigma_e^2 \; \forall \, i$; the second is $\sigma_{e_1}^2 = \sigma_r^2$ and $\sigma_{e_i}^2 = \sigma_e^2 \; \forall \, i \geq 2$. Here the second assumption is used:

$$\sigma_r^2 = a^{2(i-1)} \sigma_r^2 + \sigma_e^2 \sum_{j=2}^{i-1} a^{2(j-1)}. \tag{B.7}$$

Rearranging terms leads to

$$\begin{aligned} \left(1 - a^{2(i-1)}\right) \sigma_r^2 = \sigma_e^2 \sum_{j=2}^{i-1} a^{2(j-1)} &= \sigma_e^2 \frac{1 - a^{2(i-1)}}{1 - a^2} \\ \Rightarrow \left(1 - a^2\right) \sigma_r^2 &= \sigma_e^2 \end{aligned} \tag{B.8}$$

Using this relationship and choice of $\sigma_{e_1}^2 = \sigma_r^2$ Eq. (B.4) can be rewritten as

$$\log \mathcal{L}(\mathbf{m}) = -\frac{1}{2}\left[\frac{e_1^2}{\sigma_r^2} + \frac{\sum_{i=2}^{N} e_i^2}{\left(1 - a^2\right)\sigma_r^2} + N \log \sigma_r^2 + (N-1) \log\left(1 - a^2\right) + \log\left(2\pi\right)\right]. \tag{B.9}$$

The matrix $\mathbf{R}$ can be written as $\mathbf{R} = \mathbf{L}\mathbf{L}^T$ where $\mathbf{L}$ is the lower triangular matrix

$$\mathbf{L} = \begin{bmatrix} 1 & 0 & 0 & \cdots \\ a & \sqrt{1-a^2} & 0 & \cdots \\ a^2 & a & \sqrt{1-a^2} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}. \tag{B.10}$$

Thus $|\mathbf{R}| = |\mathbf{L}|^2$ which is the product of the squared elements of the main diagonal,[99] i.e., $|\mathbf{R}| = (1-a^2)^{N-1}$. Hence Eq. (B.3) can be written as

$$\log \mathcal{L}_{\mathrm{C}}(\mathbf{m}) = -\frac{1}{2}\left[\mathbf{r}^T\mathbf{C}_r^{-1}\mathbf{r} + N\log\sigma_r^2 + (N-1)\log\left(1-a^2\right) + \log(2\pi)\right]. \tag{B.11}$$

Thus only

$$\mathbf{r}^T\mathbf{C}_r^{-1}\mathbf{r} = \frac{e_1^2}{\sigma_r^2} + \frac{\sum_{i=2}^{N} e_i^2}{(1-a^2)\,\sigma_r^2} \tag{B.12}$$

needs to be shown to establish the equivalence of Eqs. (B.3) and (B.4). Staring with the left side of Eq. (B.12), the inverse of $\mathbf{C}_r$ is

$$\mathbf{C}_r^{-1} = \frac{1}{(1-a^2)\,\sigma_r^2}\begin{bmatrix} 1 & -a & 0 & \cdots & 0 & 0 \\ -a & 1+a^2 & -a & \cdots & 0 & 0 \\ 0 & -a & 1+a^2 & \cdots & 0 & 0 \\ & & & \vdots & & \\ 0 & 0 & 0 & \cdots & 1+a^2 & -a \\ 0 & 0 & 0 & \cdots & -a & 1 \end{bmatrix}. \tag{B.13}$$

This allows for the evaluation of $\mathbf{r}^T\mathbf{C}_r^{-1}\mathbf{r}$ as

$$\begin{aligned} &\mathbf{r}^T\mathbf{C}_r^{-1}\mathbf{r} = \\ &\frac{r_1\left(r_1 - ar_2\right)}{(1-a^2)\,\sigma_r^2} + \frac{\sum_{i=2}^{N-1}\left[\left(1+a^2\right)r_i - a\left(r_{i-1}+r_{i+1}\right)\right]}{(1-a^2)\,\sigma_r^2} + \frac{r_N\left(r_N - ar_{N-1}\right)}{(1-a^2)\,\sigma_r^2}. \end{aligned} \tag{B.14}$$

The right side of Eq. B.14 can be grouped such that

$$\mathbf{r}^T\mathbf{C}_r^{-1}\mathbf{r} = \frac{1}{(1-a^2)\,\sigma_r^2}\left[r_1^2 + r_N^2 + \left(1+a^2\right)\sum_{i=2}^{N-1} r_i^2 - 2a\sum_{i=1}^{N-1} r_i r_{i+1}\right]. \tag{B.15}$$

Returning to the right hand side of Eq. (B.12) and substituting r_1 for e_1 and $r_i - ar_{i-1}$

for e_i given $i \geq 2$.

$$\begin{aligned} \frac{e_1^2}{\sigma_r^2} + \frac{\sum_{i=2}^{N} e_i^2}{(1-a^2)\,\sigma_r^2} &= \frac{r_1^2}{\sigma_r^2} + \frac{\sum_{i=2}^{N} (r_i - a r_{i-1})^2}{(1-a^2)\,\sigma_r^2} \\ &= \frac{r_1^2}{\sigma_r^2} + \frac{\sum_{i=2}^{N} r_i^2 - 2a \sum_{i=2}^{N} r_i r_{i-1} + a^2 \sum_{i=2} r_{i-1}^2}{(1-a^2)\,\sigma_r^2} \\ &= \frac{1}{(1-a^2)\,\sigma_r^2} \left[r_1^2 + r_N^2 \left(1+a^2\right) \sum_{i=2}^{N-1} r_i^2 - 2a \sum_{i=1}^{N-1} r_i r_{i+1} \right] \end{aligned} \tag{B.16}$$

Equations (B.16) and (B.15) are equal; thus, the two sides of Eq. (B.12) are equal proving that the likelihoods for the two definitions are equivalent.

Appendix C

Chain Thinning Approaches

In order for a Markov chain to converge in distribution to its target it may be necessary to collect a very large number of samples. Ideally, all samples would be recorded; however, due to practical restrictions it is not always possible or desirable to keep all samples. A popular strategy is to thin the Markov chain by only recording every nth sample. This has the advantage that as n gets large the samples become independent. However, it also has the disadvantage that a large amount of computational effort is ignored. This appendix proposes two novel methods of chain thinning; these methods use the wasted computational effort of conventional thinning to more efficiently sample the target distribution. Both methods utilize an inclusion probability; that is, once a Markov chain moves to a state or model, that state is not automatically recorded. Instead an inclusion probability is generated for each model which define how likely it is that a model will be included in the record.

C.1 Temperature thinning

Temperature thinning (TT) is based on importance reweighing (also called high temperature sampling).[100] The TT method consists of Metropolis-Hastings (MH) sampling from a target distribution which is raised to the power $1/T^\star$ where $T^\star$ referred to as the sampling temperature ($T^\star \geq 1$); the importance weights for each sample are then used as the inclusion probability.

The acceptance probability A of a Metropolis step from $\mathbf{m}$ to $\mathbf{m}'$ for target distribution P at $T^\star$ is

$$A = \min\left(1, \left[\frac{P(\mathbf{m}')}{P(\mathbf{m})}\right]^{1/T^\star}\right). \tag{C.1}$$

The importance weight (α) for sample $\mathbf{m}'$ is

$$\alpha = \left(\frac{P(\mathbf{m}')}{\beta P(\hat{\mathbf{m}})} \right)^{(1-1/T^\star)}, \tag{C.2}$$

where $\hat{\mathbf{m}}$ is the an approximation to the most probable model and $\beta \geq 1$ is a constant chosen for numeric stability. Thus for a sample to be included in the record it must first be accepted with probability A and then recorded with probability α. All samples in the record are assigned equal importance. Note that if a Markov chain fails to move the same value may be recorded several times.

The model $\hat{\mathbf{m}}$ must be a highly probable model; the restriction is that no model $\mathbf{m}'$ may be sampled such that $P(\mathbf{m}') > \beta P(\hat{\mathbf{m}})$. In general, $\hat{\mathbf{m}}$ is found by an initial optimization carried out before the sampling is conducted. Ideally, $\hat{\mathbf{m}}$ would be the most probable model as sampling efficiency is improved if β is small (close to 1) as this increases α and result in less wasted sampling.

Consider Eqs. (C.1) and (C.2), as $T^\star \to 1$, $\alpha \to 1$ and the process converges in to unthinned Metropolis sampling. As $T^\star \to \infty$, $A \to 1$ and the process becomes standard rejection sampling.[59]

Figure C.1 shows the samples recorded and resulting marginal distributions for Metropolis-Hastings sampling from a tri-modal distribution. The proposal distribution is purposely selected such that sampling is inefficient. Each row of Fig. C.1 displays samples collected with a different sampling temperature $T^\star = 10, 3, 1$, in all cases $\beta = 1$ as the true maximum probability model is know. For each temperature approximately 400,000 samples were drawn; as the number of samples removed by TT is random the remaining samples where further thinned down to 40,000 for ease of comparison between the temperatures. To compare accuracy of the sampling consider the total misfit,

$$\int_{-\infty}^{\infty} |f(m) - \hat{f}(m)| dm \approx \sum_{i=1}^{N} \frac{|f(m_i) - \hat{f}_{\text{Bin}}(m_i)| \Delta m}{N}, \tag{C.3}$$

where f is the target distribution, $\hat{f}$ is the sampled distribution, and $\hat{f}_{\text{Bin}}(m_i)$ is the sampled distribution approximated by a histogram at the ith bin. The total misfits for each sampling temperature are 0.067 at $T^\star = 10$, 0.073 at $T^\star = 3$, and 0.189 at $T^\star = 1$. Thus high temperature thinning is a more efficient use of computation for this problem then conventional chain thinning.

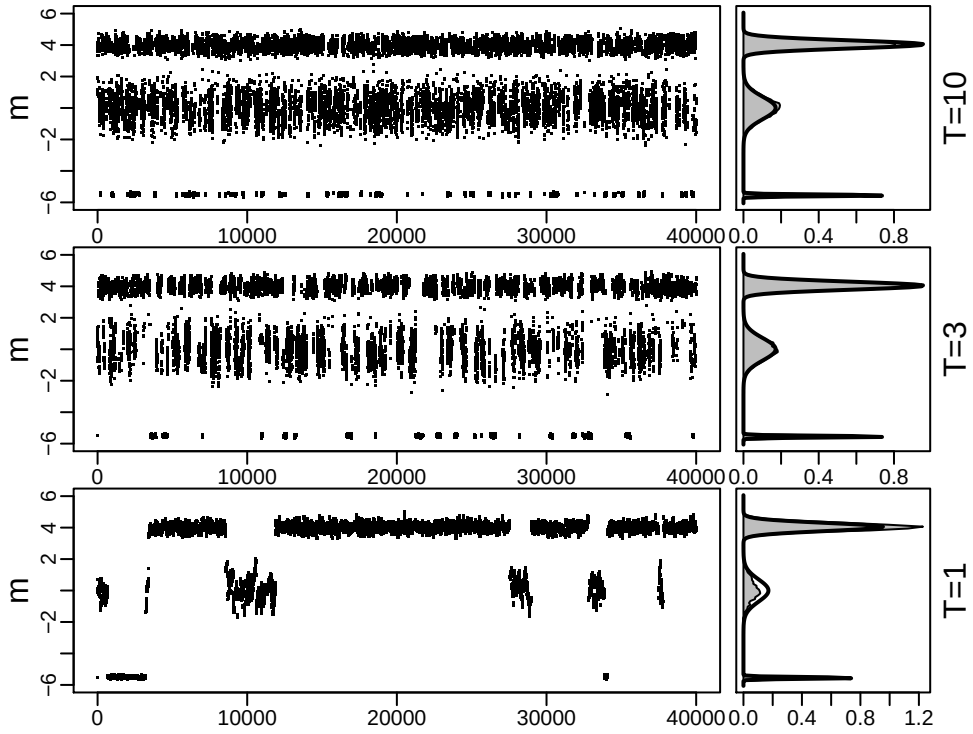

Figure C.1: Samples and marginal distribution for by the MH algorithm using the temperature thinning. Records shown on each row are drawn using a different sampling temperature (T).

C.2 Fixed length thinning

Fixed length thinning (FLT) is a method of chain thinning in which the desired number of recorded samples is set before the commencement of sampling. Samples are added and removed from the record list such that the probabilities of any two samples being in the final record are equal, regardless of the total number of samples collected. This is desirable because it causes more thinning to occur the longer a Markov chain runs.

The FLT algorithm has three steps: First, a new sample ($\mathbf{m}_i$) is drawn from the PPD. Second, the inclusion probability α_i (defined below) is computed. Third, a randomly-selected previously-recorded sample is overwritten by $\mathbf{m}_i$ with probability α_i. The probability α_i depends only on the number of samples drawn so far (i) and the number of samples to be kept (N). The inclusion probability for the ith sample

is

$$\begin{cases} \alpha_i = 1 & i \leq N \\ \alpha_i = \frac{\alpha_{i-1} N}{\alpha_{i-1}+N} & i > N \end{cases}. \tag{C.4}$$

The probability for each sample being included in the record ($\mathcal{R}$) after t iterations of the FLT algorithm is constant over all samples, e.g., $P_t(\mathbf{m}_i \in \mathcal{R}) = P_t(\mathbf{m}_j \in \mathcal{R})\ \forall i, j$, where $P_t(x)$ is the probability of x at iteration t. To demonstrate this, consider two samples $\mathbf{m}_i$ and $\mathbf{m}_{i+1}$. The probability of their inclusion in the record at the $i+k$th iteration of the algorithms are $P_{i+k}(\mathbf{m}_i \in \mathcal{R})$ and $P_{i+k}(\mathbf{m}_{i+1} \in \mathcal{R})$, respectively. If the two probabilities are equal then it is clear by induction that all probabilities are equal. The probability that a sample is in the record can be evaluated by considering the probability that it was initially included in $\mathcal{R}$ times the probability that it is not overwritten by later models, i.e.,

$$P_{i+k}(\mathbf{m}_i \in \mathcal{R}) = P_{i+k}(\mathbf{m}_{i+1} \in \mathcal{R}) \tag{C.5}$$

is equivalent to

$$\alpha_i \left(1 - \frac{\alpha_{i+1}}{N}\right) \prod_{j=2}^{k} \left(1 - \frac{\alpha_{i+k}}{N}\right) = \alpha_{i+1} \prod_{j=2}^{k} \left(1 - \frac{\alpha_{i+k}}{N}\right). \tag{C.6}$$

The product term can be cancelled from both sides of Eq. (C.6) resulting in

$$\alpha_i \left(1 - \frac{\alpha_{i+1}}{N}\right) = \alpha_{i+1} \tag{C.7}$$

Solving Eq. (C.7) for α_{i+1} results in

$$\frac{\alpha_i N}{\alpha_i + N} = \alpha_{i+1}. \tag{C.8}$$

Equation (C.8) is equivalent to Eq. (C.4); thus the definition of the inclusion probabilities α_i results in all adjacent pairs of models having equal probability of being in final record. Since all adjacent pairs of models have equal probability of being in the final record, all models have equal probability of being there.

Figure C.2 shows an example of the FLT on samples drawn from a normal distribution using the MH algorithm. The proposal distribution of the sampling is chosen to be smaller than that of target distribution to help display how FLT reduces sample autocorrelation. The record length is $N = 5000$ and t indicates the number of

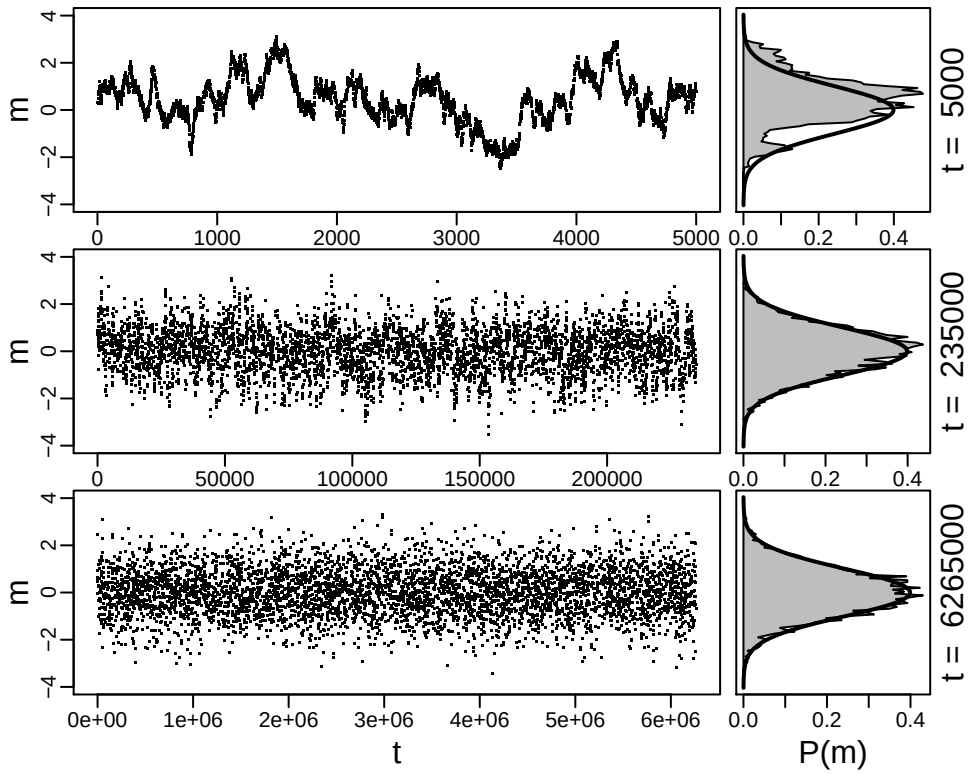

Figure C.2: Samples trajectory and marginal distribution for samples collected using MH algorithm with FLT targeting a Gaussian distribution at three iteration numbers (t). Note t is the unthinned sample index for all iterations, only the 5,000 models remaining after thinning are plotted.

MH steps. The sample autocorrelation for the three times are 0.99, 0.75, and 0.10 respectively. The total misfit [Eq. (C.3)] of the approximation for each time is 0.38, 0.10, and 0.09.

Appendix D

Adaptive Metropolis Hastings Sampling

Metropolis-Hastings (MH) sampling can be an efficient way to sample from complicated distributions. One of the key limitations of MH sampling is that a user must *a priori* chose a proposal distribution whereas the ideal proposal distribution is not known until after sampling is conducted. Adaptive Markov-chain Monte Carlo (AMCMC) methods attempt to mitigate this limitation by utilizing proposal distributions that change/evolve over the parameter space and/or the sampling index.[101–105] These methods allow for the sampling process to improve the proposal distribution without *a priori* instruction from the user. Currently the a popular paradigm is creating a weakly non-Markovian sampling procedure[101,102] where previously drawn samples are used to calculate a proposal covariance matrix. This approach has been combind with other sampling schemes such as delayed rejection[103] and regional proposal methods.[104]

This section describes an alternative paradigm that uses local gradient information for creating an adaptive proposal distribution to be used in MH sampling. To describe the adaptive Metropolis Hastings (AMH) algorithm, consider a model $\mathbf{m}$ representing the current location of a Markov chain in the parameter space. The next step of the Markov-chain will update m_j, the jth element of $\mathbf{m}$. Let M_j be the this one-dimensional conditional parameter subspace. Local gradient information at m_j along M_j is used to create a parametric approximation of a conditional target distribution $f_j(M_j)$; this approximation is then used as the proposal distribution Q for next MH step. Note that once a candidate model $\mathbf{m}'$ is found, a proposal distribution for the

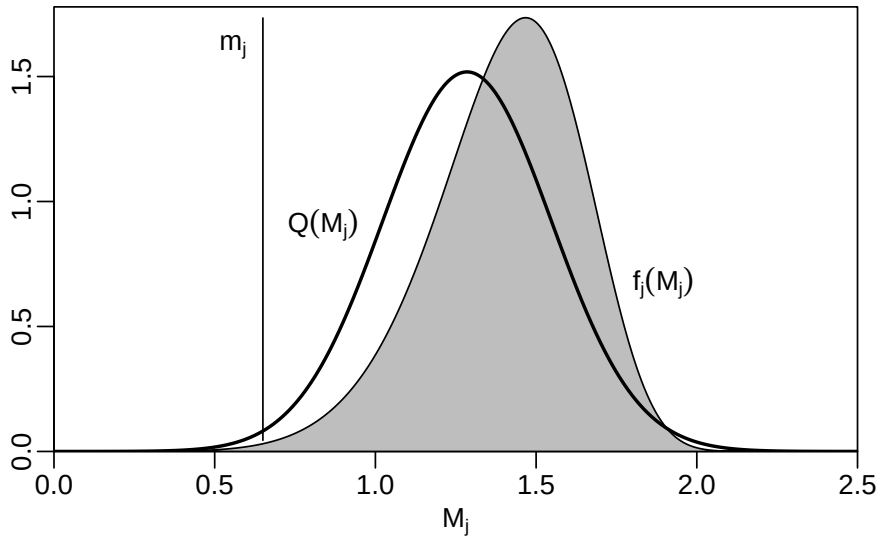

Figure D.1: The proposal distribution Q calculated at the model m_j in the conditional parameter space M_j to approximate the conditional target distribution f_j.

reverse step is also necessary and must be calculated. Figure D.1 shows an example of a proposal distribution Q calculated at the point m_j.

To create Q (assuming that the proposal is definable by two parameters), it is necessary to calculate the functional value and first and second derivatives (v_0, v_1, and v_2, respectively) of the target distribution at $\mathbf{m}$. These values are then equated with their equivalent for Q, i.e., $v_0 = Q(m_j)$, $v_1 = Q'(m_j)$, and $v_2 = Q''(m_j)$. The third equation is required because it is possible (and common) that f_j will not integrate to unity but rather to some value a. Thus, Q can be conceptualized as $Q = aQ^\star$, where $Q^\star$ has the desired functional form. A strategy that is effective for many types of proposal distributions is to use the ratios of v_1 to v_0 and v_2 to v_0, that is

$$\begin{array}{l} \frac{v_1}{v_0} = \frac{Q'(m_j)}{Q(m_j)} \\ \frac{v_2}{v_0} = \frac{Q''(m_j)}{Q(m_j)} \end{array} . \tag{D.1}$$

If $Q^\star$ is assumed to be Gaussian distributed with mean μ and variance σ^2 then Eq. (D.1) can be solved for

$$\mu = \frac{v_0 v_1}{v_1^2 - v_0 v_2} + m_j \tag{D.2}$$

and

$$\sigma^2 = \frac{(\mu - m_j)^2}{v_0 v_1} \left(1 - (\mu - m_j)\frac{v_2}{v_1}\right)^{-1} . \tag{D.3}$$

It is possible to find an analytical expression for the parameters of other distributions such as the β distribution; however, their expressions are not as compact. It is important to observe that the mean of the proposal distribution is generally not centred at m_j, as in Fig. D.1, thus the proposal can closely resemble the target distribution regardless of the location of $\mathbf{m}$ and result in the AMH sampling that approximates Gibbs[20,23] sampling (a Bayesian sampling method that always accepts proposed models, which are drawn the conditional target distribution[20]).

The rest of this section presents an example of AMH sampling applied to a synthetic inverse problem known to have a highly correlated posterior probability density. The problem is inverting plane-wave reflection-coefficient data from a fluid-solid interface for the geoacoustic parameters. Unknown parameters including densities and sound-speeds of both layers (ρ_0, ρ_1, c_0, c_1), the compressional attenuation, shear-wave speed, and shear attenuation in the second layer (α_1, c_s, α_s), and the data residual standard error σ. The observed data are modelled as[14]

$$d_i = \left| \frac{z_{tot} - z_0}{z_{tot} + z_0} \right| + e_i, \tag{D.4}$$

where $z_{tot} = z_1 \cos^2(2\theta_S) + z_S \sin^2(2\theta_S)$, $z_1 = \hat{c}_1\rho_1/\sin(\theta_1)$, $z_2 = \hat{c}_2\rho_2/\sin(\theta_2)$, and $z_S = \hat{c}_S\rho_2/\sin(\theta_S)$. The value $\hat{c}_\star$ is the complex sound speed which includes $\alpha_\star$, i.e., $\hat{c}_\star = (1/c_\star + i\alpha_\star/K)^{-1}$, where $K = 40,000\pi/\log(10)$ is constant used to convert the attenuation in dB/m/kHz to Np. Snell's law is used to find θ_1, and θ_S in terms of θ_0, the incident angle. Finally, the e_i are assumed to be independent and identically distributed Gaussian random variables with mean equal to zero and standard deviation σ.

The inversion is conducted using AMH sampling with the starting model selected through an initial optimization. Approximately 5,000,000 samples are collected; these are then chain thinned down to 50,000. The samples as a function of iteration number are shown in Fig. D.2. The one- and two-dimensional posterior marginal distributions are shown in Fig. D.3. The sampling scheme is able to resolve the parameters despite the differing levels of sensitivity without requiring the specification of proposal distributions. However, the process does not appear to handle parameter correlation as well as sampling in rotated space.[7] In more complex problems it may be useful to use both procedures.

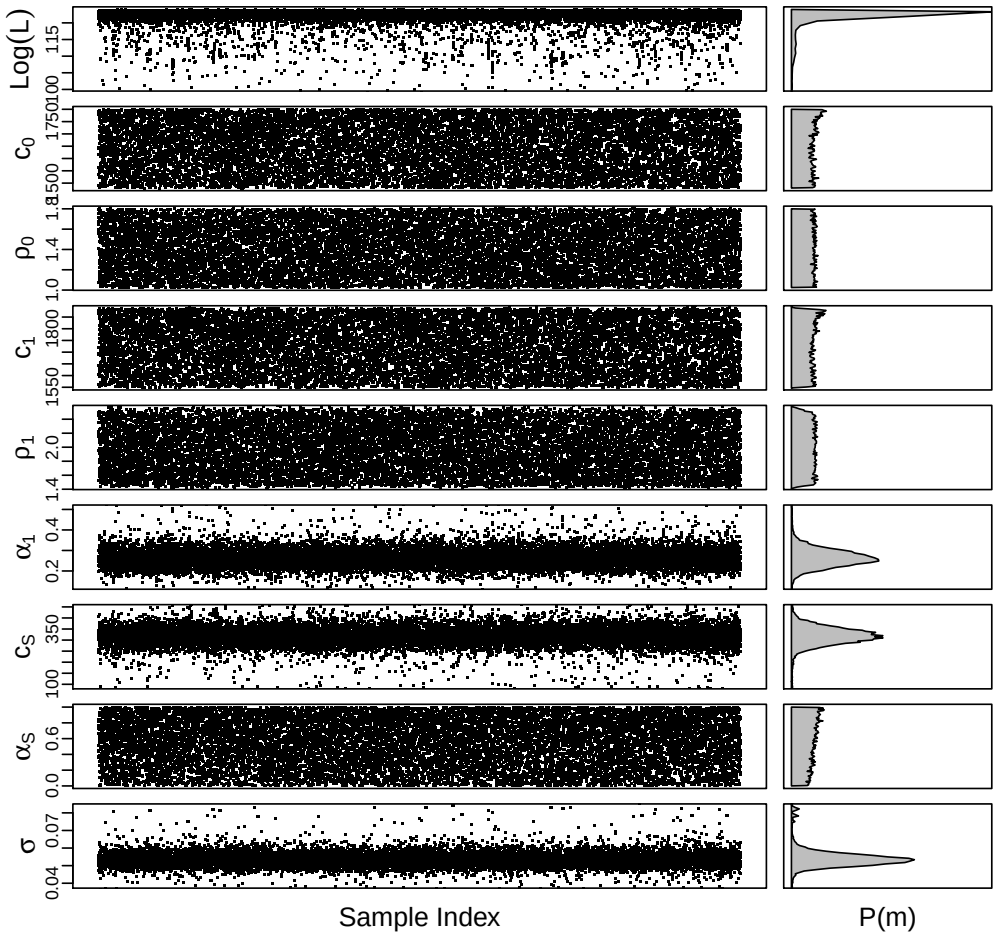

Figure D.2: The samples (dots) and marginal distributions (shaded area) of the model parameters.

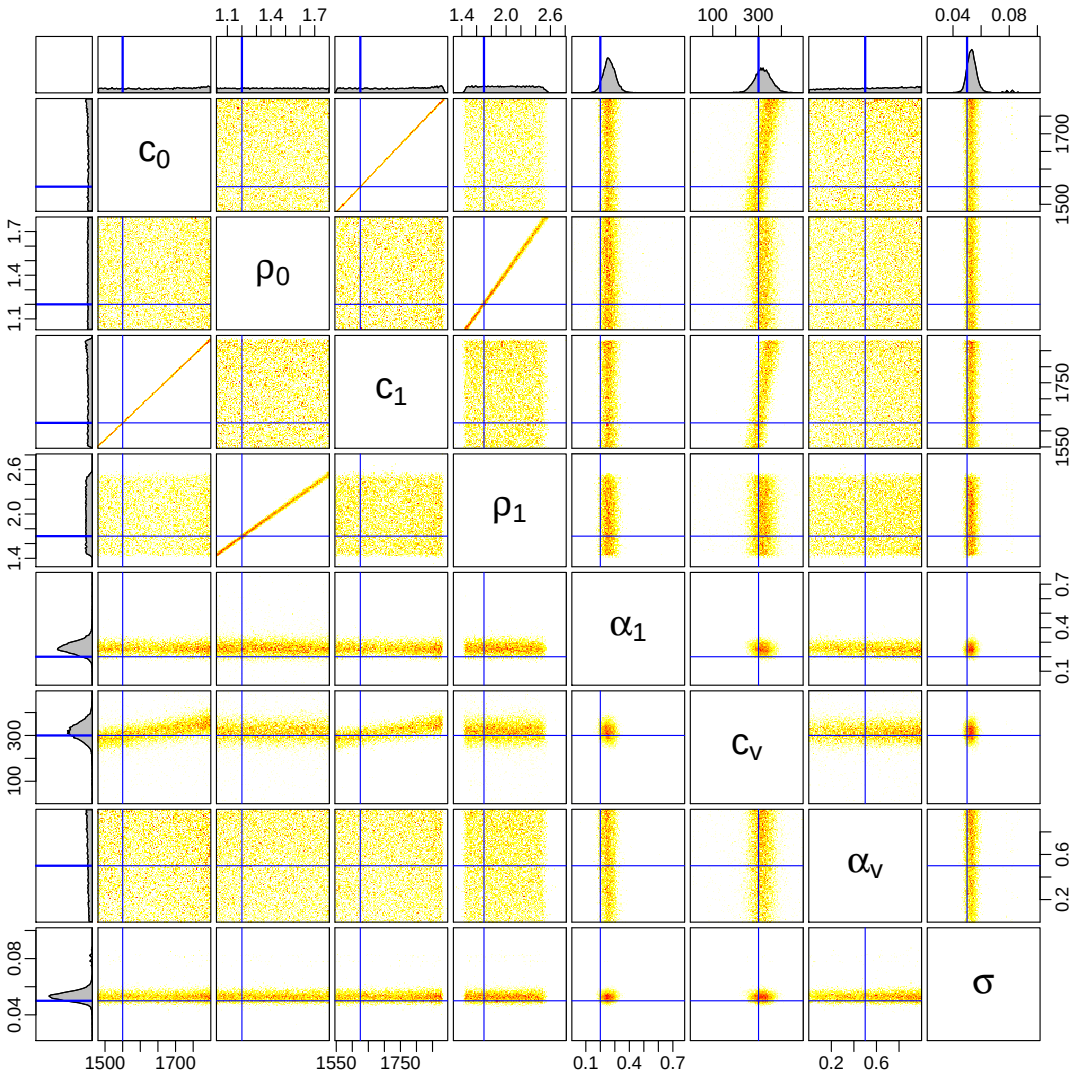

Figure D.3: A matrix of one- and two- dimensional marginal distributions of the model parameters.

Appendix E

Correcting Biased Sampling

This section describes a novel resampling-based approach for producing unbiased samples of a target distribution from biased samples of the same distribution. This approach could be used to transform optimization runs (e.g., simulated anealing or genetic algorithims) into unbiased uncertainty estimates, or to correct for non-ergodic or Markovian sampling of a target distribution.

The process works in several phases. First, a biased sample of the target distribution is acquired. The only two restriction on this sample is that it must span the parameter space and no sample may appear more than once. Next a weight is assigned to each sample. Finally the samples are re-sampled (with replacement) with probabilities for each sample proportional to the weights. If $\mathbf{m}_i$ is the ith sample, its weight w_i is calculated as $w_i = \mathcal{L}(\mathbf{m}_i) V_i$, where V_i is the volume of the parameter space closer to $\mathbf{m}_i$ than any other $\mathbf{m}_j$. That is, if D is a Vorinoi partition of the parameter space with centroids $\mathbf{m}_1$–$\mathbf{m}_n$, V_i is volume of the Vorinoi cell centred on $\mathbf{m}_i$.[52,106] The computation of D is the limiting component of this procedure. Creating a Vorinoi partition for n centroids in N-dimensional space is $O(n^{[2N-1]/N})$.[107]

An example of this process is shown in Fig. E.1. Simulated annealing is used to find the peak of a bi-variate Gaussian distribution. The steps of the annealing run are used to approximate the posterior probability density (PPD). As a result of the annealing schedule the PPD is over sampled near the peak and far tails but under-sampled on its "bell". After resampling the bias is corrected and the sampled PPD closely matches the target density.

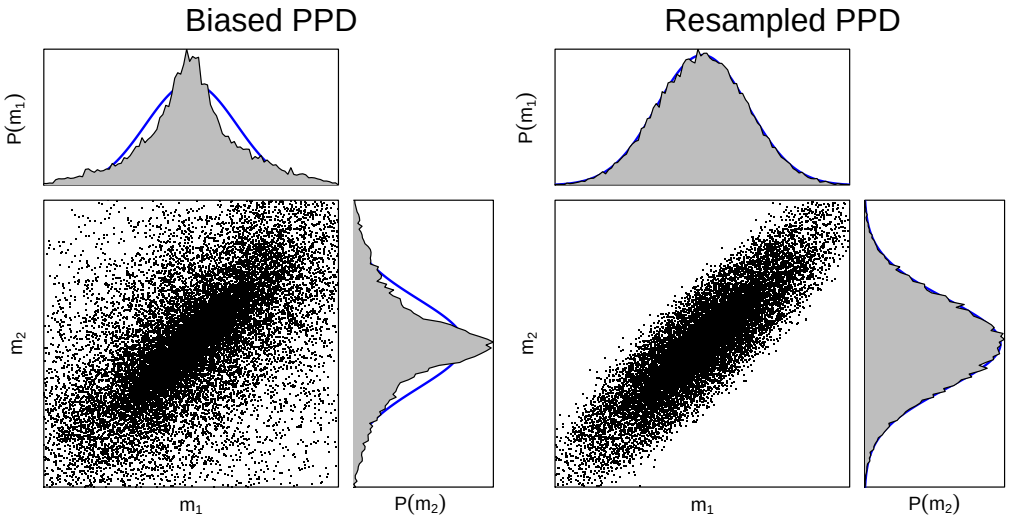

Figure E.1: Left: samples (.) of a bi-variate Gaussian distribution collected by simulated annealing with 1D marginal distributions. Right: samples (.) of a bi-variate Gaussian distribution collected by a resampling from the biased samples.

Bibliography

[1] Darrell R. Jackson and Michael D. Richardson. *High-Frequency Seafloor Acoustics*, pages 178–200. Springer, New York, 1st edition, 2007.

[2] M. K. Sen and P. L. Stoffa. *Global Optimization Methods in Geophysical Inversion*. Elsevier, Amsterdam, 1995.

[3] M. Sambridge and K. Mosegaard. Monte Carlo methods in geophysical inverse problems. *Reviews of Geophysics*, 40:3–29, 2002.

[4] G. Beylkin. The inversion problem and applications of the generalized radon transform. *Comm. Pure Appl. Math.*, 37:579–599, 1984.

[5] M. D. Collins, W. A. Kuperman, and H. Schmidt. Nonlinear inversion for ocean-bottom properties. *J. Acoust. Soc. Am.*, 93:2770–2783, 1992.

[6] S. E. Dosso and M. J. Wilmut. Quantifying data information content in geoacoustic inversion. *IEEE J. Ocean. Eng.*, 27:296–304, 2002.

[7] S. E. Dosso. Quantifying uncertainty in geoacoustic inversion. I. A fast Gibbs sampler approach. *J. Acoust. Soc. Am.*, 111:129–142, 2002.

[8] R. L. Parker. *Geophysical Inverse Theory*. Princeton University Press, Princeton, 1994.

[9] M. R. Fallat and S. E. Dosso. Geoacoustic inversion for the workshop 97 benchmark test cases using simulated annealing. *J. Computat. Acoust.*, 6:29–43, 1998.

[10] M. R. Fallat and S. E. Dosso. Geoacoustic inversion via local, global, and hybrid algorithms. *J. Acoust. Soc. Am.*, 105:3219–3230, 1999.

[11] S. E. Dosso and C. W. Holland. *Acoustic Sensing Techniques for the Shallow Water Environment*. Inversion Methods and Experiments. Springer, New York, 2006.

[12] C. W. Holland, R. Hollett, and L. Troiano. Measurement technique for bottom scattering in shallow water. *J. Acoust. Soc. Am.*, 108:997–1011, 2000.

[13] L. M. Brekhovskikh and O. A. Godin. *Acoustics of Layered Media II: Point source and bounded beams*, pages 1–33. Wave Phenomena. Springer, New York, 1st edition, 1992.

[14] F. B. Jensen, W. A. Kuperman, M. B. Porter, and H. Schmidt. *Computational Ocean Acoustics*, pages 1–64. Series in Modern Acoustic and Signal Processing. AIP Press, New York, 1993.

[15] C. W. Holland. Coupled scattering and reflection measurements in shallow water. *IEEE J. Ocean. Eng.*, 27:454–470, 2002.

[16] C. W. Holland. Seabed reflection measurement uncertainty. *J. Acoust. Soc. Am.*, 114:1861–1873, 2003.

[17] J. E. Moe and D. R. Jackson. First-order perterbation solution for rough surface scattering cross section including the effects of gradients. *J. Acoust. Soc. Am.*, 96:1748–1754, 1994.

[18] D. R. Jackson, R. I. Odom, M. L. Boyd, and A. N. Ivkin. A geoacoustic bottom interaction model. *IEEE J. Ocean. Eng.*, 35:603–617, 2010.

[19] C. W. Holland. Evidence for a common scale O(0.1) m that controls seabed scattering and reverberation in shallow water. *J. Acoust. Soc. Am.*, 132:2232–2238, 2012.

[20] A. Gelman, J. B. Carlin, H. S. Stern, and D. B. Rubin. *Bayesian Data Analysis*, pages 73–100. Chapman, New York, 2004.

[21] T. J. Ulrych, M. D. Sacchi, and A. Woodbury. A Bayes tour of inversion: A tutorial. *Geophysics*, 66:55–69, 2001.

[22] A. F. M. Smith. Bayesian computational methods. *Phil. Trans. R. Soc. Lond.*, 337:369–386, 1991.

[23] A. F. M. Smith and G. O. Roberts. Bayesian computation via the Gibbs sampler and related Markov Chain Monte Carlo methods. *J. R. Statist. Soc.*, 55:3–23, 1993.

[24] A. Malinverno and W. S. Leaney. Monte-Carlo Bayesian look-ahead inversion of walkaway vertical seismic profiles. *Geophys. Prosp.*, 53:689–703, 2005.

[25] A. Malinverno. A Bayesian criterion for simplicity in inverse problem parametrization. *Geophys. J. Int.*, 140:267–285, 2000.

[26] A. Malinverno. Parsimonious Bayesian Markov chain Monte Carlo inversion in a nonlinear geophysical problem. *Geophys. J. Int.*, 151:675–688, 2002.

[27] K. Mosegaard and M. Sambridge. Monte Carlo analysis of inverse problems. *Inverse Problems*, 18:R29–R54, 2002.

[28] M. Sambridge, K. Gallagher, A. Jackson, and P. Rickwood. Trans-dimensional inverse problems, model comparison and the evidence. *Geophys. J. Int.*, 167:528–542, 2006.

[29] M. Sambridge. Geophysical inversion with the neighbourhood algorithm – i. searching a parameter space. *Geophys. J. Int.*, 138:479–494, 1999.

[30] M. Sambridge. Geophysical inversion with the neighbourhood algorithm – ii. appraising the ensemble. *Geophys. J. Int.*, 138:727–746, 1999.

[31] M. K. Sen and P. L. Stoffa. Bayesian inference, Gibbs' sampler and uncertainty estimation in geophysical inversion. *Geophys. Prosp.*, 44:313, 1996.

[32] T. Bodin and M. Sambridge. Seismic tomography with the reversible jump algorithm. *Geophys. J. Int.*, 178:1411–1436, 2009.

[33] J. Dettmer, S. Molnar, G. A. M. W. Steininger, S. E. Dosso, and J. F. Cassidy. Trans-dimensional inversion of microtremor array dispersion data with hierarchical autoregressive error models. *Geophys. J. Int.*, 188:719–734, 2011.

[34] A. Ray and K. Key. Bayesian inversion of marine csem data with a trans-dimensional self parametrizing algorithm. *Geophys. J. Int.*, 191:1135–1151, 2012.

[35] D. J. Battle, P. Gerstoft, W. S. Hodgkiss, W. A. Kuperman, and P. L. Nielsen. Bayesian model selection applied to self-noise geoacoustic inversion. *J. Acoust. Soc. Am.*, 116:2043–2056, 2004.

[36] Y. Jiang, N. R. Chapman, and H. A. DeFerrari. Geoacoustic inversion of broadband data by matched beam processing. *J. Acoust. Soc. Am.*, 119:3707–3716, 2006.

[37] C.-F. Huang, P. Gerstoft, and W. S. Hodgkiss. Uncertainty analysis in matched-field geoacoustic inversions. *J. Acoust. Soc. Am.*, 119:197–207, 2006.

[38] J. Dettmer, S. E. Dosso, and C. W. Holland. Joint time/frequency-domain inversion of reflection data for seabed geoacoustic profiles. *J. Acoust. Soc. Am.*, 123:1306–1317, 2008.

[39] D. Tollefsen and S. E. Dosso. Bayesian geoacoustic inversion of ship noise on a horizontal array. *J. Acoust. Soc. Am.*, 124(2):788–795, 2008.

[40] Y. Caglar, P. Gerstoft, and W. S. Hodgkiss. Tracking of geoacoustic parameters using Kalman and particle filters. *J. Acoust. Soc. Am.*, 125(2):746–760, 2009.

[41] J. Dettmer, S. E. Dosso, and C. W. Holland. Uncertainty estimation in seismo-acoustic reflection travel-time inversion. *J. Acoust. Soc. Am.*, 122:161–176, 2007.

[42] J. Dettmer, S. E. Dosso, and C. W. Holland. Full wave-field reflection coefficient inversion. *J. Acoust. Soc. Am.*, 122:3327–3337, 2007.

[43] J. Dettmer, S. E. Dosso, and C. W. Holland. Model selection and Bayesian inference for high resolution seabed reflection inversion. *J. Acoust. Soc. Am.*, 125:706–716, 2009.

[44] J. Dettmer, C. W. Holland, and S. E. Dosso. Analyzing lateral seabed variability with Bayesian inference of seabed reflection inversions. *J. Acoust. Soc. Am.*, 126:56–69, 2009.

[45] J. Dettmer, C. W. Holland, and S. E. Dosso. Spatial variability of density gradients in the transition layer from Bayesian inference of seabed reflection data., 2009.

[46] D. Tang, G. Jin, D. R. Jackson, and K. L. Williams. Analyses of highfrequency bottom and subbottom backscattering for two distinct shallow water environments. *J. Acoust. Soc. Am.*, 96:2930–2936, 1994.

[47] R. J. Soukup and R. F. Gragg. Backscatter from a limestone seafloor at 2-3.5 khz: Measurements and modeling. *J. Acoust. Soc. Am.*, 113:2501–2514, 2003.

[48] G. Casella and R. Berger. *Statistical Inferance*, pages 1–45. Thomson Learning Group, Oasific Grove, 2002.

[49] D. G. T. Denison, C. C. Holmes, B. K. Mallick, and A. F. M. Smith. *Bayesian Methods for Nonlinear Classification and Regression*, pages 1–277. Wiley, New York, 2002.

[50] A. Gelman. Objections to bayesian statistics. *Bayesian Anal.*, 3:445–450, 2008.

[51] J. E. Quijano, S. E. Dosso, J. Dettmer, L. M. Zurk, and M. Siderius. Trans-dimensional geoacoustic inversion of wind-driven ambient noise. *J. Acoust. Soc. Am.*, 133, 2013. EL47.

[52] J. Dettmer and S. E. Dosso. Probabilistic two-dimensional water-column and seabed inversion with self-adapting parameterizations. *J. Acoust. Soc. Am.*, 133:2612–2623, 2013.

[53] J. Dettmer and S. E. Dosso. Trans-dimensional matched-field geoacoustic inversion with hierarchical error models and interacting markov chains. *J. Acoust. Soc. Am.*, 132:2239–2250, 2012.

[54] C. E. Lindsay and N. R. Chapman. Matched field inversion for geoacoustic model parameters using adaptive simulated annealing. *IEEE J. Ocean. Eng.*, 18:224–231, 1993.

[55] C. W. Holland and J. Osler. High-resolution geoacoustic inversion in shallow water: A joint time- and frequency-domain technique. *J. Acoust. Soc. Am.*, 107:1263–1279, 2000.

[56] K. Gallagher, K. Charvin, S. Nielsen, M. Sambridge, and J. Stephenson. Markov chain Monte Carlo (MCMC) sampling methods to determine optimal models, model resolution and model choice for earth science problems. *Marine and Petrol. Geol.*, 26:525–535, 2009.

[57] A. Jasra, D. A. Stephens, and C. Holmes. Population-based reversible jump markov chain monte carlo. *Biometrika*, 94:787–807, 2007.

[58] I. Murray. *Advances in Markov chain Monte Carlo methods*. PhD thesis, Gatsby Computational Neuroscience Unit, University College London, London, 2007.

[59] J. S. Liu. *Monte Carlo strategies in scientific computing*. Springer, 1st edition, 2001.

[60] N. Metropolis, A. Rosenbluth, M. Rosenbluth, and A. Teller A. E. Teller. Equations of state calculations by fast computing machines. *Journal of Chemical Physics*, 21:1087–1092, 1953.

[61] D. J. C. MacKay. *Information Theory, Inference, and Learning Algorithms*, pages 343–386. Cambridge University Press, Cambridge, 2003.

[62] S. Chib. Marginal likelihood from the Gibbs output. *J. Am. Stat. Assoc.*, 90:1313–1321, 1995.

[63] R. M. Neal. Annealed importance sampling. *Stat. Comp.*, 11:125–139, 2001.

[64] J. Dettmer, S. E. Dosso, and John C. Osler. Bayesian evidence computation for model selection in geoacoustic inversion. *J. Acoust. Soc. Am.*, 128:3406–3415, 2010.

[65] J. Skilling. Nested sampling for general Bayesian computation. *Bayesian Analysis*, 1:833–860, 2006.

[66] T. Jasa1 and N. Xiang. Nested sampling applied in bayesian room-acoustics decay analysis. *J. Acoust. Soc. Am.*, 132:3251–3262, 2012.

[67] D J. Spiegelhalter, N. G. Best, B. P. Carlin, and A. van der Linde. Bayesian measures of model complexity and fit. *J. R. Statist. Soc.*, 64:583–639, 2002.

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

[69] J. Dettmer, S. E. Dosso, and C. W. Holland. Trans-dimensional geoacoustic inversion. *J. Acoust. Soc. Am.*, 128:3393–3405, 2010.

[70] G. Schwartz. Estimating the dimension of a model. *Ann. Statist.*, 6:461–464, 1978.

[71] G. Steininger, J. Dettmer, S. E. Dosso, and C. W. Holland. Trans-dimensional joint inversion of seabed scattering and reflection data. *J. Acoust. Soc. Am.*, 133:1347–1357, 2013.

[72] G. Steininger, C. W. Holland, S. E. Dosso, and J. Dettmer. Seabed roughness parameters from joint backscatter and reflection inversion at the malta plateau. *J. Acoust. Soc. Am.*, 134:1833–1842, 2013.

[73] G. Steininger, C. W. Holland, S. E. Dosso, and J. Dettmer. Determining the dominant seabfloor scattering mechanism. *J. Acoust. Soc. Am.* Submitted.

[74] G. Steininger, C. W. Holland, S. E. Dosso, and J. Dettmer. A trans-dimensional polynomial-spline parameterization for gradient-based geoacoustic inversion. *J. Acoust. Soc. Am.*, 2013. Submitted.

[75] S. E. Dosso and P.L. Nielsen. Quantifying uncertainty in Geoacoustic Inversion. II. Application to broadband, shallow-water data. *J. Acoust. Soc. Am.*, 111:143–159, 2002.

[76] S. E. Dosso and C. W. Holland. Geoacoustic uncertainties from visco-elastic inversion of seabed reflection data. *IEEE J. Ocean. Eng.*, 31:657–671, 2006.

[77] P. J. Green. *Highly Structured Stochastic Systems*, chapter Trans-dimensional Markov chain Monte Carlo, pages 179–198. Oxford Statistical Science Series. Oxford University Press, Oxford, 2003.

[78] C. F. Mecklenbräuker, P. Gerstoft, J. F. Bhme, and P.-J. Chung. Hypothesis testing for geoacoustic environmental models using likelihood ratio. *J. Acoust. Soc. Am.*, 105:1738–1748, 1999.

[79] T. Jasa and N. Xiang. Using nested sampling in the analysis of multi-rate sound energy decay in acoustically coupled rooms. *AIP Conference Proceedings*, 803(1):189–196, 2005.

[80] N. P. Agostinetti and A. Malinverno. Receiver function inversion by trans-dimensional monte carlo sampling. *Geophys. J. Int.*, 181:858–872, 2010.

[81] R.H. Shumway and D.S. Stoffer. *Time Series Analysis and Its Applications*, pages 89–212. Springer, New York, 2000.

[82] C. W. Holland, P. L. Nielsen, J. Dettmer, and S. E. Dosso. Resolving meso-scale seabed variability using reflection measurements from an autonomous underwater vehicle. *J. Acoust. Soc. Am.*, 131:1066–1078, 2012.

[83] E. L. Hamilton. Geoacoustic modeling of the sea floor. *J. Acoust. Soc. Am.*, 68:1313–1340, 1980.

[84] E. L. Hamilton and R. T. Bachman. Sound velocity and related properties of marine sediments. *J. Acoust. Soc. Am.*, 72:1891–1904, 1982.

[85] D. J. Earl and M. W. Deem. Parallel tempering: Theory, applications, and new perspectives. *Phys. Chem. Chem. Phys.*, 7:3910–3916, 2005.

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

[87] Darrell R. Jackson and Kevin B. Briggs. High-frequency bottom backscattering: Roughness versus sediment volume scattering. *J. Acoust. Soc. Am.*, 92:962–977, 1992.

[88] J. Dettmer, C. W. Holland, and S. E. Dosso. Transdimensional uncertainty estimation for dispersive seabed sediments. *Geophysics*, 73:1–14, 2013.

[89] J. E. Quijano, S. E. Dosso, J. Dettmer, L. M. Zurk, M. Siderius, and C. Harrison. Bayesian geoacoustic inversion using wind-driven ambient noise. *J. Acoust. Soc. Am.*, 131:2658–2667, 2011.

[90] D. R. Jackson, R. I. Odom, M. L. Boyd, and A. N. Ivkin. A correction to a geoacoustic bottom interaction model. *IEEE J. Ocean. Eng.*, 36:373, 2011.

[91] W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery. *Numerical Recipes in Fortran 77*, pages 1–933. Cambridge University Press, 2nd edition, 1997.

[92] J. Botts. Nested sampling in practice. *J. Acoust. Soc. Am.*, 134:3575, 2013.

[93] J. Yang, D. R. Jackson, and D. Tang. Mid-frequency geoacoustic inversion using bottom loss data from the shallow water 2006 experiment. *J. Acoust. Soc. Am.*, 131:1711–1721, 2012.

[94] C. W. Holland, J. Dettmer, and S. E. Dosso. Remote sensing of sediment density and velocity gradients in the transition layer. *J. Acoust. Soc. Am.*, 118:163–177, 2005.

[95] C. W. Holland and S. E. Dosso. Mid frequency shallow water fine-grained sediment attenuation measurements. *J. Acoust. Soc. Am.*, 134:131–143, 2013.

[96] F. A. Bowles. Observations on attenuation and shear-wave velocity in fine-grained, marine sediments. *J. Acoust. Soc. Am.*, 101:3385–3397, 1997.

[97] J. Dettmer, S. E. Dosso, and C. W. Holland. Sequential trans-dimensional Monte Carlo for range-dependent geoacoustic inversion. *J. Acoust. Soc. Am.*, 129:1794–1806, 2011.

[98] R. T. Bachman. Acoustic and physical property relationships in marine sediment. *J. Acoust. Soc. Am.*, 78:616–621, 1985.

[99] T. Bodin. *Transdimensional approaches to geophysical inverse problems*. PhD thesis, College of Physical and Mathematical Sciences, Australian National University, Canberra, Australia, 2010.

[100] B. A. Brooks and L. N. Frazer. Importance reweighting reduces dependence on temperature in Gibbs samplers: an application to the coseismic geodetic inverse problem. *Geophys. J. Int.*, 161:12–20, 2005.

[101] H. Haario, E. Saksman, and J. Tamminen. An adaptive metropolis algorithm. *Bernoulli*, 7:223–242, 2001.

[102] H. Haario, E. Saksman, and J. Tamminen. Componentwise adaptation for high dimensional mcmc. *Stat. Comput.*, 20:265–274, 2005.

[103] H. Haario, M. Laine, A. Mira, and E. Saksman. Dram: Efficient adaptive mcmc. *Stat. Comput.*, 16:339–354, 2006.

[104] H. Haario, E. Saksman, and J. Tamminen. Componentwise adaptation for high dimensional mcmc. *J. Computational and Graphical Stats.*, 18:349–367, 2012.

[105] C. Andrieu and J. Thoms. A tutorial on adaptive mcmc. *Stat. Comput.*, 18:343–373, 2008.

[106] M. G. Voronoi. Nouvelles applications des parametres continus a la theorie des formes quadratiques (new applications of continuous parameters to the theory of quadratic forms). *J. Rein. Angew. Math.*, 134:198–287, 1908.

[107] D. F. Watson. Computing the n-dimensional delaunay tessellation with application to voronoi polytopes. *The Computer Journal*, 24:167–172, 1980.