

OPTIMIZATION AND BAYESIAN APPROACHES TO  
GEOACOUSTIC INVERSION

by

Anna-Liesa Salome Lapinski

UNIVERSITY OF VICTORIA

# Optimization and Bayesian Approaches to Geoacoustic Inversion

by

Anna-Liesa Salome Lapinski  
B.Sc., Brock University, 1999

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

MASTER OF SCIENCE

in the

SCHOOL OF EARTH AND OCEAN SCIENCE

We accept this thesis as conforming  
to the required standard

[REDACTED]  
Dr. S. E. Dosso, Supervisor (School of Earth and Ocean Science)

[REDACTED]  
Dr. N. R. Chapman, Co-Supervisor (School of Earth and Ocean Science)

[REDACTED]  
Dr. M. J. Wilmut, Departmental Member (School of Earth and Ocean Science)

[REDACTED]  
Dr. M. Lesperance, Outside Member (Department of Mathematics and Statistics)

[REDACTED]  
Dr. G. H. Brooke, External Examiner (Anteon Corp., Victoria, BC)

© Anna-Liesa Salome Lapinski, July 1, 2002  
UNIVERSITY OF VICTORIA

All rights reserved. This thesis may not be reproduced in whole or in part,  
by photocopy or other means, without the permission of the author.


Supervisor: Dr. S. E. Dosso

### **Abstract**

This thesis presents the results of applying two geoacoustic inversion methods, an hybrid optimization algorithm and a Bayesian sampling algorithm, to data from environments with range independent and range dependent geophysical and geometric properties. The hybrid algorithm combines the local method of downhill simplex with the global method of simulated annealing, in an adaptive algorithm. The Bayesian inversion algorithm uses a Gibbs sampler to estimate properties of the posterior probability density (PPD), such as mean and maximum *a posteriori* parameter estimates, marginal probability distributions, highest posterior density intervals, and the model covariance matrix can be calculated. The methods were applied to noise-free and noisy benchmark data for several shallow ocean environments. An appropriate model parameterization is unknown for many of the environments, which increases the difficulty of the problems. An under-parameterized approach was applied to determine the optimal parameterization. The model solutions were estimated well given the varying sensitivities of the parameters. The Bayesian inversion method provided complete solutions including quantitative uncertainty estimates to the inversion problems, while the hybrid inversion method provided parameter estimates in a fraction of the computation time.


Examiners:

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

  
Dr. N. R. Chapman, Co-Supervisor (School of Earth and Ocean Science)

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

  
Dr. M. Lesperance, Outside Member (Department of Mathematics and Statistics)

  
Dr. G. H. Brooke, External Examiner (Anteon Corp., Victoria, BC)

---

## Table of Contents

<b>Acknowledgments</b>	<b>ix</b>
<b>1 Introduction</b>	<b>1</b>
<b>2 Theory</b>	<b>9</b>
2.1 Adaptive Simplex Simulated Annealing (ASSA) . . . . .	9
2.1.1 Downhill Simplex . . . . .	9
2.1.2 Simulated Annealing . . . . .	11
2.1.3 ASSA Algorithm . . . . .	13
2.2 Fast Gibbs Sampler (FGS) . . . . .	15
2.2.1 Bayesian Formulation of Inverse Theory . . . . .	16
2.2.2 Importance Sampling . . . . .	18
2.2.3 Gibbs Sampler (GS) . . . . .	19
2.2.4 FGS Algorithm . . . . .	19
2.3 Error Function . . . . .	21
2.4 Unknown Model Parameterization . . . . .	23
2.5 Sound Propagation Models . . . . .	25
2.5.1 Normal Modes . . . . .	25
2.5.2 Parabolic Equations . . . . .	27
<b>3 Range Independent Inversion</b>	<b>32</b>
<b>4 Range Dependent Inversion</b>	<b>45</b>
4.1 Introduction . . . . .	45
4.1.1 Workshop Data . . . . .	46
4.1.2 ASSA Error Function . . . . .	48

---

4.1.3	Range Dependent Propagation Modeling . . . . .	49
4.2	ASSA Results . . . . .	52
4.2.1	Test Case 1 . . . . .	52
4.2.2	Test Case 2 and Calibration Case . . . . .	60
4.3	FGS Results . . . . .	65
4.3.1	Test Case 1 . . . . .	66
4.3.2	Test Case 2 and Calibration Case . . . . .	73
<b>5</b>	<b>Range Dependent Inversion for Noisy Synthetic Data</b>	<b>80</b>
5.1	Test Case 1 . . . . .	80
5.2	Test Case 2 . . . . .	84
<b>6</b>	<b>Conclusions</b>	<b>92</b>
	<b>References</b>	<b>95</b>

## List of Tables

6.1	Mismatch comparison for data presented. . . . .	93
-----	---	----

---

## List of Figures

2.1	Downhill Simplex method. . . . .	10
3.1	Schematic representation of the NL test case environment and model parameterization. . . . .	33
3.2	NLa mismatch and variation. . . . .	35
3.3	NLa convergence of mismatch and parameters for $L = 5$ sediment layers. . . . .	37
3.4	NLa Cauchy-distribution scaling factors, $\delta m_i$ , for $L = 5$ . . . . .	38
3.5	NLa parameter profiles of the ASSA solution, for $L = 5$ . . . . .	39
3.6	NLa sensitivity analysis for $L = 5$ . . . . .	40
3.7	NLa parameter profiles of sub-optimal ASSA solution, for $L = 3$ & 8. . . . .	43
3.8	NLb and NLc parameter profiles of the ASSA solutions, for $L = 4$ . . . . .	44
4.1	TC0 environment and model parameterization. . . . .	46
4.2	TC1 environment and model parameterization. . . . .	47
4.3	TC2 environment and model parameterization. . . . .	47
4.4	PECan parameter study. . . . .	51
4.5	TC1 mismatch and variation. . . . .	53
4.6	TC1 convergence of mismatch and parameters for $L = 3$ sediment layers. . . . .	54
4.7	TC1 parameter profiles of the ASSA solution, for $L = 3$ . . . . .	55
4.8	TC1 sensitivity analysis for $L = 3$ . . . . .	56
4.9	TC1 parameter profiles of sub-optimal ASSA solution, for $L = 2$ & 4. . . . .	58
4.10	TC1 prediction of data using best model. . . . .	59
4.11	TC1 parameter profiles of the three frequency ASSA solution, for $L = 3$ . . . . .	59
4.12	TC2 mismatch and variation. . . . .	61
4.13	TC2 parameter profiles of the ASSA solution, for $L = 2$ sediment layers. . . . .	62

---

4.14	TC2 sensitivity analysis for $L = 2$ . . . . .	63
4.15	TC2 prediction of data using best model. . . . .	64
4.16	TC0 parameter profiles of the ASSA solution, for $L = 2$ sediment layers. . . . .	64
4.17	TC1 cumulative fractional energy, $\eta_P$ , for principal component analysis. . . . .	67
4.18	TC1 estimated marginal PPDs. . . . .	69
4.19	TC1 parameter profiles of the FGS MAP. . . . .	70
4.20	TC1 model correlation matrix. . . . .	72
4.21	TC2 cumulative fractional energy, $\eta_P$ , for principal component analysis. . . . .	74
4.22	TC2 estimated marginal PPDs. . . . .	75
4.23	TC2 parameter profiles of the FGS MAP. . . . .	76
4.24	TC2 model correlation matrix. . . . .	77
4.25	TC0 parameter profiles of the FGS MAP. . . . .	78
5.1	Noisy TC1 mismatch and variation. . . . .	81
5.2	Noisy TC1 parameter profiles of the best ASSA solution, for $L = 2$ . . . . .	82
5.3	Noisy TC1 sensitivity analysis for $L = 2$ . . . . .	83
5.4	Noisy TC1 estimated marginal PPDs. . . . .	85
5.5	Noisy TC1 parameter profiles of the FGS MAP. . . . .	86
5.6	Noisy TC2 parameter profiles of the best ASSA solution, for $L = 2$ . . . . .	87
5.7	Noisy TC2 sensitivity analysis for $L = 2$ . . . . .	88
5.8	Noisy TC2 estimated marginal PPDs. . . . .	89
5.9	Noisy TC2 parameter profiles of the FGS MAP. . . . .	90

## Acknowledgments

I would like to thank Dr. Stan Dosso. It was his guidance, unhesitating help, patience, deep understanding of the problem, affirmations, enthusiasm, and geniality that made this work possible. Thank-you Stan for guiding me through these three years and for always being there when I needed any kind of help. I would like to thank Dr. Ross Chapman, who, as the SEOS graduate advisor in late 1998/early 1999, asked me if I'd ever thought about studying ocean acoustics. I hadn't. I'd also like to thank him for always being a kind, supportive and generous supervisor. I would like to thank Dr. Michael Wilmot and the other members of the ocean acoustics group, past and present, for their unhesitating help and constant amiability. I would like to thank my parents, Barb and Stan Lapinski, for their infinite support, encouragement and love. Mom, I could write a dissertation on how much you've helped me during my time at UVic. And Dad, I could feel your strength, and your unwavering belief and pride in me everywhere I've gone in the world. Thank-you Mom and Dad for always being at the other end of the phone line and my e-mails. I would also like to thank my other family and friends: thank-you for the e-mails, phone calls, care packages, monthly dinners, movies, celebrations, and Italy; thank-you for your encouragement and for my life outside of school. Finally, I give my reverent deep thanks to the higher power in the universe for helping me in every part of my life.

Financial support for this research was provided for by the Natural Sciences and Engineering Research Council (NSERC) of Canada and the Office of Naval Research (ONR).

## Chapter 1

### Introduction

In recent years, a great amount of effort has been applied to developing accurate and efficient methods of estimating geoacoustic properties using measured ocean acoustic fields. Acoustic energy propagates as a pressure disturbance,  $p$ , that evolves as it travels through an environment. The propagation characteristics are governed by the acoustic wave equation. The acoustic field at any given point has been influenced by the medium it has traveled through. Geoacoustic inversion uses the measured acoustic fields to determine unknown geoacoustic properties that influenced the fields.

To consider an inversion problem, the corresponding forward problem must first be formulated. Given a model  $\mathbf{m}$  representing a physical system, the forward problem involves determining the data  $\mathbf{d}$  that will be measured when a physical process interacts with the system. If  $\mathbf{d}$  cannot be determined, then there can be no inverse problem. The inverse problem is to determine  $\mathbf{m}$ , given  $\mathbf{d}$ . Geoacoustic inversion belongs to the class of nonlinear geophysical inverse problems that have no direct solution. The inverse problem can be formulated by postulating a model  $\mathbf{m} = \{m_i, i = 1, M\}$ , to represent the unknown geoacoustic parameters. Each parameter,  $m_i$ , is assigned lower and upper bounds,  $m_i^- < m_i < m_i^+$ , which define the parameter space for the problem. Matched-field inversion (MFI) is a technique that can be used to determine a likely  $\mathbf{m}$ . In MFI, measured (true) acoustic fields are compared with (replica) fields generated for a proposed  $\mathbf{m}$  using a numerical propagation model. The difference or mismatch is quantified by defining an objective function (or error function),  $E(\mathbf{m})$ . When the difference between the true data and replica data is minimal, the proposed model is taken to represent the sensed environment.

The primary goal of many inversion algorithms is to minimize  $E(\mathbf{m})$ . This is

accomplished by searching the parameter space of the problem, which can be a non-trivial exercise. For example, the strong non-linearity of geoaoustic inverse problems results in objective functions with many local minima; therefore, isolating the global minimum becomes more difficult. Also, if the objective function has narrow oblique valleys due to correlated parameters, then the search difficulty is increased. In addition, the size of the search space increases geometrically with the number of model parameters. Therefore, the greater the number of parameters, the more challenging the search will be. Widely varying parameter sensitivities also increase the difficulty. Parameters that vary laterally (i.e., a range dependent environment) will increase the computational difficulty of an inverse problem. To add to the challenge of geoaoustic inversion, an appropriate model parameterization is commonly unknown.

A variety of inversion techniques are used in geoaoustic inversion to locate the global minimum of  $E(\mathbf{m})$ . The grid search method computes the mismatch at all points in a discretized search space. It can accommodate local minima, correlated parameters, and widely varying sensitivities in the objective function. However, for more than a small number of parameters or a fine grid spacing, the technique becomes prohibitively time consuming. Several modified grid search techniques have been used recently (e.g., Tolstoy 1998; Ainslie *et al.* 2000; Dizaji *et al.* 2002). Local inversion methods use local gradient information from the objective function to iteratively improve upon a starting model. While this technique moves efficiently downhill, the algorithm typically converges in a local minimum near the starting model. Local methods include the Gauss-Newton method, conjugate gradient method and the downhill simplex (DHS) method. Local methods have been successfully applied to geoaoustic inversion problems by Rajan *et al.* (1987) and Zala and Ozard (1998). Global inversion methods widely search the parameter space by repeatedly perturbing the models using a random process. This method includes the ability to move uphill in the objective function and therefore allows the algorithm to escape local

minima. The model perturbations are generated without using gradient information; therefore, the algorithm is inefficient at navigating downhill and into the narrow oblique valleys generated by correlated parameters. Two popular global methods in geoacoustic inversion are Genetic Algorithm (GA) (e.g., Gerstoft 1994; Ratilal *et al.* 1998; Heard *et al.* 1998; Siderius *et al.* 2000; Taroudakis and Markaki 2000) and Simulated Annealing (SA) (e.g., Collins *et al.* 1992; Fallat and Dosso 1998).

When an inversion problem has more than a few parameters, or parameters with varying sensitivities, one attempt to improve the efficiency of standard inversion methods is to simplify the inversion problem by dividing the parameter space into subspaces, i.e., to group parameters together and invert for each group in a preferential way during the algorithm. The subspaces have a lower dimensionality than the original model space, which will reduce the difficulty of the problem. However, *a priori* knowledge of parameter correlations is needed so that the model groupings are appropriate (e.g. Ratilal *et al.* 1998; Ainslie *et al.* 2000; Taroudakis and Markaki 2000).

Another way to improve upon standard inversion methods is found in hybrid inversion methods. Hybrid inversion methods combine local and global methods so as to retain the good points of both methods and overcome their respective weaknesses. The local component enables the algorithm to effectively move downhill while the global component enables the algorithm to escape local minima. The method, therefore, is appropriate for inverse problems with complicated or unknown objective function topographies. Some hybrid algorithms applied to geoacoustic problems include GA combined with Gauss-Newton (Gerstoft 1995), GA combined with DHS (Musil *et al.* 1999) and Fast Simulated Annealing (FSA) combined with DHS (Fallat and Dosso 1999). Adaptive Simplex Simulated Annealing (ASSA), a hybrid algorithm which combines an adaptive FSA with DHS (Dosso, Wilmut and Lapinski 2001) is described in Chapter 2 of this work with results presented in Chapters 3 through 5.

While great effort has been applied to finding optimal solutions to inverse problems, comparatively little effort has been applied to rigorously determining the uncertainties or correlations of the parameters. For linear inverse problems with Gaussian-distributed random errors on the data, a complete, straightforward, error analysis can be conducted. Gaussian data uncertainties imply Gaussian parameter uncertainties in linear inversion. For nonlinear inverse problems, such as geoacoustic inversion, Gaussian data uncertainties do not imply Gaussian parameter uncertainties. A frequentist approach can be used to put approximate error bounds on the parameter estimates. If multiple data sets are available for a given environment, for example the measured pressure fields for a continuous source in time are divided into snapshots (Siderius *et al.* 2000), then inversions for each data set will produce parameter estimates. The set of parameter estimates can be used to determine a mean solution and the standard deviation of the parameter estimates can be used to estimate the uncertainty of the mean solution. However, it is unclear how meaningful these estimates are.

Bayesian inference theory offers a method to estimate parameters and their uncertainties for nonlinear inverse problems. Bayesian inference theory treats parameters as random variables, which allows probabilities to be assigned to parameter values (Olson 1987). Bayesian methods are based on sampling the posterior probability density (PPD) which combines prior information about the model with information from a measured data set. The PPD embodies the solution to the inverse problem. The Bayesian formalism has previously been applied through GA by Gerstoft (Gerstoft 1994; Gerstoft and Mecklenbräuker 1998). An often used early method (Gerstoft 1994; Ratilal *et al.* 1998; Taroudakis and Markaki 2000) is based on an empirically normalized data likelihood function resulting in an inaccurate PPD (Gerstoft and Mecklenbräuker 1998). In a later work, Gerstoft and Mecklenbräuker (1998) developed PPD estimation for geoacoustic inversion using a likelihood-based approach. In their work, fast estimation of the PPD was considered more important than accuracy; therefore, the final generations of GA inversions were used to generate the PPD. This

introduced unknown bias into the estimates due to the unknown sampling distribution of GA. Dosso (2002) applied the likelihood-based formulation of Gerstoft and Mecklenbräuker (1998) in a Gibbs sampling technique. Unlike GA, a Gibbs sampler (GS) has a known sampling distribution. Gibbs sampling in its basic form can be computationally slow; therefore, modifications were applied to form a Fast Gibbs Sampler (FGS). The FGS has been compared with GS and Monte Carlo methods for standard benchmark geoacoustic test cases, and found to produce identical results in orders of magnitude less computation time (Dosso 2002). It has also been successfully applied to measured data (Dosso and Nielsen 2002; Reidel *et al.* 2002). Bayesian theory and the application of the FGS are presented in Chapter 2 and Chapters 4 and 5 of this work, respectively.

In addition to applying an appropriate inversion algorithm, a proper formulation of an inversion problem is very important. For example, choosing an appropriate parameterization is necessary. If the parameterization does not adequately represent the true environment, then the best model found will be meaningless. If the model parameterization is unknown, using either an under-parameterized or an over-parameterized approach can be used to determine the model structure. An under-parameterized approach, as discussed later in this work, finds the model with the least amount of structure that resolves the acoustic data. An over-parameterized approach can also be used, in which a penalty for structure is included in the objective function being minimized.

Choosing proper parameter bounds is also necessary. The bounds should include the true solution, however, the true solution is typically not known. Therefore, information is needed about the environment so that realistic parameter bounds can be applied. In addition, the parameter space should not be allowed to become prohibitively large by choosing unnecessarily wide parameter bounds.

In MFI, the forward problem may be solved tens of thousands of times, therefore it is important to ensure that the calculation of the replica fields as efficient as possible.

To calculate the replica fields, an appropriate propagation model must be chosen with respect to the environment and the inversion algorithm. For some propagation models, such as those used in this work, the computational effort required to calculate the replica fields increases with the source frequency. Care must be taken when choosing what frequencies to use. However, to determine geometric and geoacoustic properties in both the water column and the ocean bottom, a range of frequencies could be needed. Attenuation increases with frequency; therefore, high frequencies are more sensitive to upper sediment properties, and lower frequencies are more sensitive to deeper structure (Collins *et al.* 1992; Ratilal *et al.* 1998; Taroudakis and Markaki 2000). High frequencies are also more sensitive to source position and water depth (Ratilal *et al.* 1998; Taroudakis and Markaki 2000) as well as to small-scale structure. Subspace inversion methods sometimes use this feature to carry out inversions (Ratilal *et al.* 1998; Ainslie *et al.* 2000; Taroudakis and Markaki 2000). Finally, a matched field processor (error function) appropriate to the data available must be chosen (Dosso and Wilmut 2002).

There were two primary goals with respect to the work presented in this thesis. One objective was to establish if appropriate parameterization could be determined for both range independent (i.e., properties vary vertically but not laterally) and range dependent problems using ASSA and the under-parameterized approach. The other objective of the work was to assess the ability of ASSA and the FGS to generate solutions for range dependent inversion problems. Inversions for range independent environments require less computational effort than similar range dependent problems. However, applying the inversion methods to range dependent environments is important because only a subset of ocean environments can be approximated as range independent environments. A successful application of the FGS includes generating parameter and uncertainty estimates and other statistical information about the problem by sampling the PPD. A successful application of ASSA includes the

algorithm converging to a good model estimate that approximates the true environment. Uncertainty estimates from FGS and, in a limited way, a relative sensitivity analysis from ASSA can be used to judge how well model parameters are determined for each inversion method.

An international workshop was held in Vancouver, BC, Canada, in 1997 to provide a forum to compare current geoacoustic inversion methods (Tolstoy *et al.* 1998). A suite of benchmark ocean acoustic data representing range independent, shallow water environments was made available. The benchmark data allowed for comparisons between inversion methods. Data for six test cases, each with three subcases, were provided (the subcases had the same model parameterization, but different parameter values). Calibration data with known parameters were also made available so that computation errors could be removed from the inversion codes. In 2001 a similar international workshop was held in Gulfport, MS, USA (Chapman *et al.* 2001). A focus of this workshop was to compare inversion techniques applied to data from range dependent shallow water environments. Four synthetic data test cases were supplied, one being a calibration test case. Except for the calibration case, the model parameterization was unknown for each test case, prior to the workshop. In this work, the applications of ASSA and the FGS to test case data from both workshops are analyzed, with particular attention focused on determining appropriate model parameterizations.

This thesis is divided as follows: Chapter 2 presents the theory for this work. In Sec. 2.1 the hybrid optimization algorithm ASSA is described. In Sec. 2.2, the theory and development of the FGS is presented. In the subsequent sections of Chapter 2, error functions, acoustic propagation models and the under-parameterized approach to determining model structure are discussed with respect to the work presented in the remaining chapters. Chapters 3, 4 and 5 present the results of applying Bayesian and/or optimization inversion methods to data generated for range independent and

---

range dependent environments. In Chapter 3, ASSA is applied to noise-free benchmark data for a range independent environment with unknown model parameterization. In Chapter 4, ASSA and FGS are applied to noise-free benchmark data for three range dependent environments, two of which have unknown model parameterization. In Chapter 5, ASSA and FGS are applied to noisy data. The data used were generated by adding noise to two of the data sets used in Chapter 4. Chapter 6 presents general conclusions from the work.

---

## Chapter 2

### Theory

This chapter outlines the theory behind the work presented in this thesis. ASSA and the FGS are first developed and described, followed by discussions about error functions, unknown model parameterization and acoustic propagation models.

#### 2.1 Adaptive Simplex Simulated Annealing (ASSA)

In this section, the minimization technique for a highly non-linear objective function  $E(\mathbf{m})$  is described, where  $E(\mathbf{m})$  is the mismatch between the observed (measured) acoustic field data,  $\mathbf{d}^{\text{obs}}$ , and the modeled acoustic field data,  $\mathbf{d}(\mathbf{m})$ . ASSA is a hybrid inversion algorithm which combines DHS with an adaptive FSA. The results of applying ASSA to data from both range independent and range dependent environments are presented in Chapters 3 through 5 of this work. In the following sections DHS, SA, FSA, and ASSA are described.

##### 2.1.1 Downhill Simplex

DHS (Nedler and Mead 1965; Press *et al.* 1992) uses a geometric scheme to determine local downhill steps in an objective function  $E(\mathbf{m})$ . For an  $M$ -parameter model,  $\mathbf{m}$ , a simplex of  $M + 1$  models is used. The goal is to improve the model with the highest mismatch through a series of geometric steps including reflections, expansions and contractions. For example, Fig. 2.1(a) represents a simplex with  $M = 3$  parameters. The model with the highest mismatch is first reflected through the face of the simplex (Fig. 2.1(b)). If the new, reflected, model is a better model

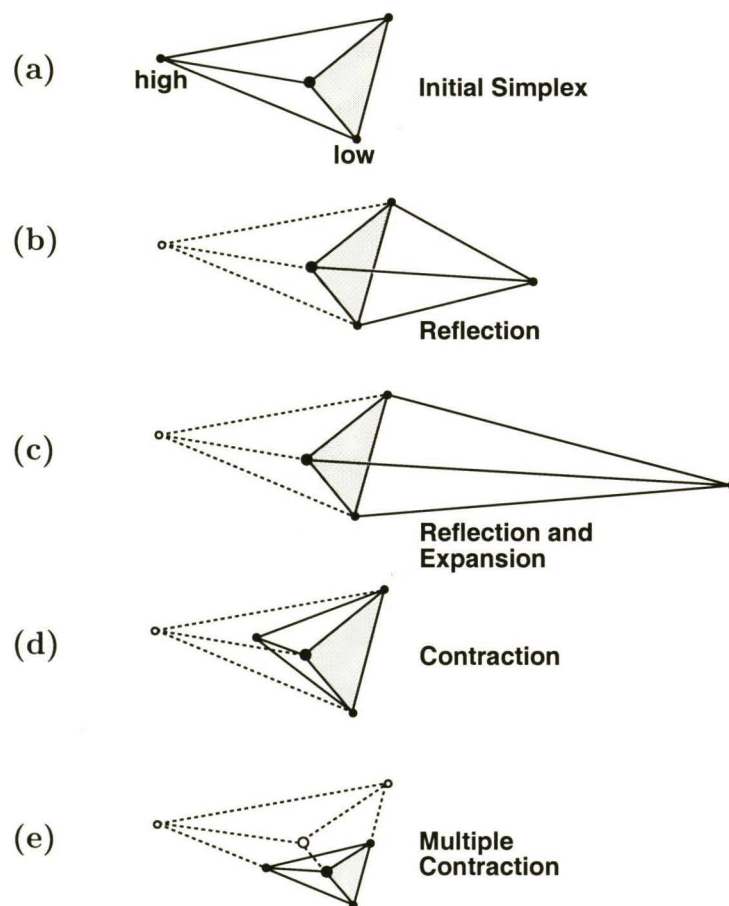


Figure 2.1 Geometric steps of the Downhill Simplex method for a three dimensional model, after Press *et al.* 1992.

(has a lower mismatch) than the original model, then it replaces the original model in the simplex; otherwise it is discarded. If the reflected model has the lowest mismatch in the simplex, then an expansion by a factor of two is attempted in the same direction as the reflection (Fig. 2.1(c)). After the reflection, if the model in the simplex with the second highest mismatch has remained the model with the second highest mismatch, then the model with the highest mismatch is contracted by a factor of two towards the simplex face (Fig. 2.1(d)). If none of these attempts reduce the mismatch below the second highest mismatch, then a multiple contraction is performed, about the model with the lowest mismatch (Fig. 2.1(e)). The algorithm continues until the highest and lowest mismatches satisfy

$$\frac{E_{\text{high}} - E_{\text{low}}}{(E_{\text{high}} + E_{\text{low}})/2} < \epsilon_1, \quad (2.1)$$

where  $\epsilon_1$  is a user specified tolerance level, and  $E_{\text{high}}$  and  $E_{\text{low}}$  are the highest and lowest mismatches, respectively, in the simplex. The criterion is tested prior to each reflection. Each step in the algorithm is straightforward to calculate. However, the method may not be as efficient in finding the local minimum as more precise local methods that require the calculation of partial derivatives and the solution of systems of equations to compute downhill steps. However, for ASSA only a general descent rather than a precise descent is required when DHS is combined with FSA.

### 2.1.2 Simulated Annealing

SA (Kirkpatrick *et al.* 1983; Press *et al.* 1992) is based on an analogy with crystal growth through thermodynamic annealing. SA and GS, discussed in Sec. 2.2.3, have the same statistical origins. In statistical thermodynamics, the Boltzmann or Gibbs' distribution

$$P_G(\mathbf{m}) = \frac{\exp[-E(\mathbf{m})/T]}{\sum \exp[-E(\mathbf{m})/T]} \quad (2.2)$$

gives the probability that a system of many particles is in a state  $\mathbf{m}$  with an energy  $E(\mathbf{m})$  at absolute temperature  $T$ . The sum in the denominator that provides the normalization is over all possible states. To simulate the behavior of a system in equilibrium, an algorithm was proposed (Metropolis *et al.* 1953) in which the system is repeatedly perturbed and the perturbations are accepted if the criterion

$$\xi \leq \exp[-\Delta E/T] \quad (2.3)$$

is satisfied, where  $\xi$  is a uniform random number on  $[0, 1]$  and  $\Delta E$  is the difference between the energies of the proposed and current state. This is the so called Metropolis Algorithm (MA). After a large number of perturbations are accepted at a constant  $T$ , Markov-chain analysis confirms that the MA equilibrium distribution is given by the Gibbs' distribution (2.2) (Sen and Stoffa 1995).

In SA, the state  $\mathbf{m}$  is the model and the energy  $E(\mathbf{m})$  is the error function. The MA is applied and any perturbed model that moves downhill in the objective function is accepted while a model that moves uphill may still be accepted. The temperature  $T$ , a control parameter, is decreased slightly after each iteration. As  $T$  decreases, the rate of accepting uphill steps decreases, which forces the algorithm to search more thoroughly for the minimum of the region it is in.

The method of perturbation is an important aspect of SA. For example, in FSA (Szu and Hartley 1987) the random perturbations are chosen from a Cauchy distribution about the current parameter values while the width of the distribution is reduced linearly with the temperature. A Cauchy distribution has a narrow central peak and heavy tails. This ensures a focus on local sampling with occasional large perturbations. To increase efficiency further, the rate at which the distribution width decreases can be manipulated. For example, successful results were obtained by reducing the distribution width using a square-root dependence with respect to temperature (Fallat and Dosso 1998). ASSA uses an adaptive method to individually reduce the distribution width of each parameter, as described in the following

sub-section.

### 2.1.3 ASSA Algorithm

ASSA combines DHS and SA methods to form an adaptive hybrid algorithm (Dosso, Wilmut and Lapinski 2001). Previously, other non-adaptive DHS-SA hybrid algorithms have been suggested (e.g., Press *et al.* 1992; Pan and Wu 1998; Fallat and Dosso 1999). In ASSA, SA features are applied at each DHS geometric step. ASSA has previously been shown to be an effective optimization algorithm. The following briefly describes the ASSA algorithm.

For an  $M$ -parameter model, the ASSA algorithm begins by randomly generating a simplex of  $M + 1$  models. The DHS algorithm then proceeds, but with the following variations. For every new model,  $\mathbf{m}'$ , generated from a DHS geometric step, random perturbations  $\delta m_i C$  are added to each parameter  $m'_i$ :

$$m''_i = m'_i + \delta m_i C \quad (2.4)$$

where  $C$  is a Cauchy-distributed random variable and  $\delta m_i$  effectively scales the Cauchy-distribution width for parameter  $m_i$ . If  $m''_i$  does not fall within its parameter bounds, then  $m'_i$  is perturbed again. The Metropolis criterion (2.3) is used to either accept or reject  $\mathbf{m}''$ . If accepted,  $\mathbf{m}''$  replaces  $\mathbf{m}$  in the simplex. The algorithm then proceeds to the next DHS geometric step. One note about the algorithm is that the DHS multiple contraction is typically applied only in the final phase of the algorithm due to the computationally expensive nature of the step.

In ASSA, the random variable,  $C$ , is drawn from a distribution such that

$$C = \eta_1 \tan(\eta_2 \pi / 2) \quad (2.5)$$

where  $\eta_1$  and  $\eta_2$  are uniform random variables on  $[-1, 1]$ . Lane 1992, states that this is a Cauchy distribution. For the first  $S$  perturbations, the Cauchy distribution is scaled to the width of the parameter bounds,  $\delta m_i = m_i^+ - m_i^-$  where  $m_i^+$  and  $m_i^-$

are the maximum and minimum allowed parameter values. For the remainder of the algorithm, the Cauchy distribution is scaled to the smaller of either  $m_i^+ - m_i^-$  or  $s\bar{m}_i$ , where  $s$  is a control parameter,  $s > 1$ , and  $\bar{m}_i$  is the mean perturbation size. The mean perturbation size is the running average of the last  $S$  accepted random perturbations for that parameter. Therefore, the allowed perturbation size *adapts* as the algorithm runs and is unique to each parameter. By increasing  $s$  and  $S$ , the parameter search becomes slower and more thorough. It is found that the algorithm may still have difficulty accepting a new model, at times. Incorporated into the algorithm is the criterion that if a large number of consecutive models are rejected, then a further reduction in the Cauchy distribution is applied.

The rate of the temperature decrease is controlled according to

$$T_n = \beta^n T_o, \quad (2.6)$$

where  $T_n$  is the temperature at the  $n^{\text{th}}$  temperature step,  $\beta$  is a control parameter that controls the rate of temperature decrease ( $\beta < 1$ ), and  $T_o$  is the initial temperature. The number of new models to be accepted at each temperature step, the maximum number of temperature steps, and  $\beta$  and  $T_o$ , are all values that are problem specific and must be adjusted with respect to each problem. To ensure that the temperature decreases slowly enough to allow the parameter space to be sufficiently searched,  $\beta$  must be sufficiently close to 1. Also, if  $T_o$  is too high, then computation time will be wasted by accepting a large number of unrealistic models and if  $T_o$  is too low then the parameter space will not be properly searched. An appropriate  $T_o$  allows  $\sim 90\%$  of the models to be accepted initially. Typically, the control parameters do not require a comprehensive analysis to be adequately determined.

A modification to the ASSA algorithm described by Dosso, Wilmut and Lapinski (2001) is the addition of a convergence test prior to the final quenching phase of the ASSA algorithm. The convergence of the simplex is judged by evaluating (2.1), with  $\epsilon_1 = 0.005$ . If the convergence criterion is satisfied, then the algorithm begins its

final phase, as described below. Otherwise, the final phase does not begin until the maximum number of temperature steps has been achieved.

The final phase of the algorithm is a DHS quenching to move directly to the bottom of the minimum that the ASSA algorithm has located. The best model in the current simplex is saved, and the remainder of the simplex is reinitialized with randomly generated models. The DHS method is then applied to the simplex until either the convergence criterion (2.1) is met with  $\epsilon_1 = 0.0001$  or the maximum number of DHS steps is realized. The quenching allows one final search of the parameter space for the lowest mismatch.

The efficiency of the ASSA algorithm, compared to SA (or FSA), has been improved on several fronts. The DHS steps help the algorithm maneuver down the objective function, adapting to parameter correlations. The adaptive nature of the algorithm also helps parameters converge to an optimal value at a rate related to their effect on the mismatch. This, therefore, also helps increase the efficiency of the algorithm. Another attribute of the algorithm is that the simplex always maintains the  $M + 1$  best models encountered providing an effective memory of good solutions. SA and FSA have no memory and a good model could be replaced and never encountered again. The improved efficiency can be seen through the comparison of ASSA and FSA results with respect to the 1997 workshop test cases. ASSA provided significantly better parameter estimates in less computation time than FSA for all benchmark test cases (Fallat and Dosso 1998; Dosso, Wilmut and Lapinski 2001).

## 2.2 Fast Gibbs Sampler (FGS)

In the following sections, inverse theory from a Bayesian perspective, importance sampling and the GS are introduced, based on Dosso (2002), Sen and Stoffa (1995), and Sen and Stoffa (1996). The FGS algorithm, as presented by Dosso (2002), is also

outlined.

### 2.2.1 Bayesian Formulation of Inverse Theory

For data  $\mathbf{d}$  and model  $\mathbf{m}$  that represent random variables, Bayes' rule can be written

$$P(\mathbf{m}|\mathbf{d}) \propto P(\mathbf{d}|\mathbf{m})P(\mathbf{m}), \quad (2.7)$$

where  $P(\mathbf{m}|\mathbf{d})$  is the conditional probability density function (PDF) of  $\mathbf{m}$  given  $\mathbf{d}$ ,  $P(\mathbf{d}|\mathbf{m})$  is the conditional PDF of  $\mathbf{d}$  given  $\mathbf{m}$ , and  $P(\mathbf{m})$  is the PDF of  $\mathbf{m}$  (the *a priori* information). The PDF of observed data,  $\mathbf{d}^{\text{obs}}$ , is referred to as the likelihood function,  $L(\mathbf{d}^{\text{obs}}|\mathbf{m})$ . It contains the *a posteriori* information from the data. When  $\mathbf{d}^{\text{obs}}$  replaces  $\mathbf{d}$  in Eq. (2.7), the relation is rewritten as

$$P(\mathbf{m}|\mathbf{d}^{\text{obs}}) \propto L(\mathbf{d}^{\text{obs}}|\mathbf{m})P(\mathbf{m}), \quad (2.8)$$

where  $P(\mathbf{m}|\mathbf{d}^{\text{obs}})$  is the posterior probability density (PPD). The prior PDF contains the available *a priori* information about the model while the PPD denotes the information content given both the data and the prior information.

The statistical distribution of errors on the data, from both measurement and theory, governs the expression of  $L(\mathbf{d}^{\text{obs}}|\mathbf{m})$ . For example, under the assumption that the measured data contain complex, zero-mean, Gaussian-distributed random errors, the likelihood function is,

$$L(\mathbf{d}^{\text{obs}}|\mathbf{m}) = \frac{1}{\pi^N |\mathbf{C}_D|} \exp \left[ -[\mathbf{d}^{\text{obs}} - \mathbf{d}(\mathbf{m})]^\dagger \mathbf{C}_D^{-1} [\mathbf{d}^{\text{obs}} - \mathbf{d}(\mathbf{m})] \right] \quad (2.9)$$

(e.g., Hammersley and Handscomb 1964; Gerstoft and Mecklenbräuker 1998) where  $\mathbf{C}_D$  is the data covariance matrix and  $\dagger$  refers to the conjugate-transpose. The likelihood function can also be represented as,

$$L(\mathbf{d}^{\text{obs}}|\mathbf{m}) \propto \exp[-E(\mathbf{m})], \quad (2.10)$$

where,

$$E(\mathbf{m}) = [\mathbf{d}^{\text{obs}} - \mathbf{d}(\mathbf{m})]^\dagger \mathbf{C}_D^{-1} [\mathbf{d}^{\text{obs}} - \mathbf{d}(\mathbf{m})]. \quad (2.11)$$

Thus, the PPD can be written

$$P(\mathbf{m}|\mathbf{d}^{\text{obs}}) = \frac{\exp[-E(\mathbf{m})] P(\mathbf{m})}{\int \exp[-E(\mathbf{m}')] P(\mathbf{m}') d\mathbf{m}'}, \quad (2.12)$$

where the integration bounds of the normalization span the model space.

The PPD  $P(\mathbf{m}|\mathbf{d}^{\text{obs}})$  embodies the general Bayesian solution to the inverse problem. The calculation of integral moments of the multi-dimensional PPD, such as the posterior mean model, model covariance matrix and the marginal probability distribution for parameter  $m_i$ , are needed to interpret the PPD solution. These PPD properties are defined, respectively, as

$$\begin{aligned} \langle \mathbf{m} \rangle &= \int \mathbf{m}' P(\mathbf{m}'|\mathbf{d}^{\text{obs}}) d\mathbf{m}', \\ \mathbf{C}_M &= \int (\mathbf{m}' - \langle \mathbf{m}' \rangle)(\mathbf{m}' - \langle \mathbf{m}' \rangle)^T P(\mathbf{m}'|\mathbf{d}^{\text{obs}}) d\mathbf{m}', \\ P(m_i|\mathbf{d}^{\text{obs}}) &= \int \delta(m'_i - m_i) P(\mathbf{m}'|\mathbf{d}^{\text{obs}}) d\mathbf{m}', \end{aligned} \quad (2.13)$$

where  $\delta$  represents the Dirac delta function and  $T$  represents the transpose. These moments can be written in the general form

$$I = \int f(\mathbf{m}') P(\mathbf{m}'|\mathbf{d}^{\text{obs}}) d\mathbf{m}'. \quad (2.14)$$

In addition, the model that maximizes the PPD is known as the maximum *a posteriori* (MAP) solution. For PPDs that are multi-modal or considerably asymmetric, the MAP solution can provide more meaningful parameter estimates than the posterior mean  $\langle \mathbf{m} \rangle$  (2.13). If the MAP solution is the only estimate required, then optimization methods such as ASSA can sometimes be employed and are typically more efficient than Bayesian methods.

### 2.2.2 Importance Sampling

Efficiently evaluating integrals of the form (2.14) is of primary concern in Bayesian inversion. Grid searches and Monte Carlo methods can be used, but are computationally intensive for more than a small number of parameters. A more efficient approach is the technique of importance sampling. In importance sampling, models  $\{\mathbf{m}_j, j = 1, Q\}$ , are drawn from regions of the model space that contribute the most to the integral (i.e., important regions). This reduces the variance of the integral estimator for a given sample size. If the  $Q$  models are drawn from a normalized, non-uniform, sampling distribution,  $g(\mathbf{m})$ , then Eq. (2.14) can be written

$$I = \int \left[ \frac{f(\mathbf{m}') P(\mathbf{m}'|\mathbf{d}^{\text{obs}})}{g(\mathbf{m}')} \right] g(\mathbf{m}') d\mathbf{m}' \approx \frac{1}{Q} \sum_{j=1}^Q \frac{f(\mathbf{m}_j) P(\mathbf{m}_j|\mathbf{d}^{\text{obs}})}{g(\mathbf{m}_j)}. \quad (2.15)$$

Importance sampling provides an unbiased estimate of the integral. However, if an unknown sampling distribution is used so that the estimate cannot be corrected for  $g(\mathbf{m})$  (denominator of Eq. (2.15)), then the estimate will be biased. The formulation of  $g(\mathbf{m})$  that minimizes the variance of the integral estimator is given by

$$g(\mathbf{m}) = \frac{|f(\mathbf{m})|P(\mathbf{m}|\mathbf{d}^{\text{obs}})}{\int |f(\mathbf{m}')|P(\mathbf{m}'|\mathbf{d}^{\text{obs}}) d\mathbf{m}'} \quad (2.16)$$

(e.g., Sen and Stoffa 1995). However, since  $P(\mathbf{m}|\mathbf{d}^{\text{obs}}) \propto \exp[-E(\mathbf{m})]$  as in (2.12), then to evaluate the integral in Eq. (2.16)  $E(\mathbf{m})$  must be evaluated at all points in the sample space. If this is done, then there is no need to use importance sampling. Therefore, in importance sampling, a more practical function  $g(\mathbf{m})$  is used. The GS, as described in the following section, is a sampling method that provides a useful sampling function, without requiring the enumeration of the sample space. It satisfies the normalization requirement and has a known sampling distribution which therefore ensures that the estimates of (2.15) are unbiased. This is in contrast to Gerstoft and Mecklenbräuer (1998), who used GA to estimate the PPD. GA has an unknown sampling distribution, which will in turn bias the evaluations of Eq. (2.15).

### 2.2.3 Gibbs Sampler (GS)

As mentioned previously, the GS and SA have their origins in the MA (2.3). In the MA, Markov-chain analysis confirms that at a constant  $T$  the equilibrium distribution is given by the Gibbs' PDF  $P_G(\mathbf{m})$  (2.2). Therefore, the MA at equilibrium is a GS. If  $T = 1$  and the prior PDF,  $P(\mathbf{m})$ , is uniform between the upper and lower parameter bounds (as is typical for matched-field inversion), then the Gibbs' PDF  $P_G(\mathbf{m})$  and the PPD  $P(\mathbf{m}|\mathbf{d}^{\text{obs}})$  (2.12) are equivalent, except for the sum and integral calculated in their denominators. In this case, at equilibrium, the MA samples directly from the PPD  $P(\mathbf{m}|\mathbf{d}^{\text{obs}})$ . Since  $P_G(\mathbf{m})$  is a normalized non-uniform sampling distribution, it is an appropriate choice for  $g(\mathbf{m})$ . For a GS, Eq. (2.14) can therefore be approximated as

$$I \approx \frac{1}{Q} \sum_{j=1}^Q f(\mathbf{m}_j). \quad (2.17)$$

Thus, the integrals of Eq. (2.13) can be estimated directly from the set of  $Q$  accepted models without solution bias.

### 2.2.4 FGS Algorithm

The GS, though more efficient than other methods such as grid search and Monte Carlo, can still be computationally slow. Modifications can be made to the GS that improve its efficiency without biasing the estimates. The following describes the modifications made to the GS to produce the FGS algorithm (Dosso 2002).

The first stage of the FGS algorithm is a cooling stage that allows the algorithm to find a region of models that are reasonably probable. The cooling stage involves quickly reducing  $T$  from a high temperature to  $T = 1$ . Throughout the cooling, the maximum perturbation size for each parameter shrinks adaptively. Once  $T = 1$  has been achieved, the maximum perturbation size most recently used is retained, and all but the last model are discarded. Following that, after each perturbation the

maximum perturbation size for each parameter is adjusted to be the larger of either the current maximum perturbation size or  $k$  times the last accepted perturbation, while not exceeding the width of the parameter bounds. Typically,  $k = 2$ , however after a large number of accepted perturbations, it can be reduced to  $k = 1.2$ . Thus the maximum perturbation is allowed to grow so as to only sample the important regions.

In an attempt to improve computation time, a modification to the algorithm described by Dosso (2002) was added. During periods when the algorithm has a low rate of accepting a perturbation, the maximum perturbation size is temporarily reduced to help hasten the acceptance of a model. Following the acceptance of a perturbation, the maximum allowed perturbation width returns to its prior width. After observing the behavior of this modification (not shown) it was concluded that it did not adversely influence the algorithm. This modification is only implemented prior to the rotation of the parameter space axes described next.

After a predefined number of accepted perturbations, the parameter space axes are rotated by applying an orthogonal transformation that diagonalizes the model covariance matrix,  $\mathbf{C}_M$  (2.13). The rotation helps improve the sampling efficiency by orienting the parameter space axes along the oblique valleys that result from parameter correlations (e.g., Jaschke and Chapman (1999)). An eigenvector decomposition of  $\mathbf{C}_M$

$$\mathbf{C}_M = \mathbf{A} \mathbf{\Lambda} \mathbf{A}^T \quad (2.18)$$

is required to determine the rotation matrix  $\mathbf{A}^T$  ( $T$  symbolizes the transpose), where  $\mathbf{\Lambda}$  is a diagonal matrix of eigenvalues, and the eigenvectors form the columns of the orthonormal matrix  $\mathbf{A}$ . Each rotated model,  $\tilde{\mathbf{m}}$ , is determined through the orthogonal transformation

$$\tilde{\mathbf{m}} = \mathbf{A}^T \mathbf{m}. \quad (2.19)$$

Following the initial rotation, parameter perturbations are solely applied in the rotated space. The parameter bounds and the maximum perturbation sizes of each rotated parameter must be determined. The bounds are determined by generating a large number (e.g.,  $10^5$ ) of random models from the unrotated sample space. The orthogonal transformation is then applied to these models and the maximum and minimum values found for  $\tilde{m}_i$  are assigned as the bounds of the parameter. Similarly, the maximum parameter perturbations in the rotated space can be approximated. The perturbed models are continually rotated back to the original (physical) parameters

$$\mathbf{m} = \mathbf{A}\tilde{\mathbf{m}} \quad (2.20)$$

so that the forward problem can be solved and the Metropolis criterion (2.3) can be tested.

The FGS algorithm also includes an automated convergence criterion. In the algorithm, the FGS is applied in parallel and two independent samples of accepted models are generated. The two samples are compared periodically and the algorithm is terminated when the difference between the two samples is satisfactorily small. In particular, when the differences between estimated cumulative marginal distributions for each parameter are each below a threshold,  $\epsilon_2$ , the algorithm is terminated. An acceptable threshold is  $\epsilon_2 = 0.1$ . It is expected that when the two independent samples are combined the difference between the estimated marginal distributions and the true marginal distributions will be smaller than the difference between the two samples.

### 2.3 Error Function

The error function,  $E(\mathbf{m})$ , is an important element of MFI. The following briefly describes the derivation of the error function for data generated using multi-frequency sources and multiple independent samples such as that acquired at multiple ranges

from the sources. The error function is a measure of model fitness used in both the optimization techniques described earlier and in sampling. The following follows the approach presented by Dosso and Nielsen (2002) for single-range data but uses somewhat different notation. This approach was originated by Gerstoft and Mecklenbräuker (1998) in geoacoustic inversion.

Let it be assumed that the  $N$  data for each range and frequency contain complex, Gaussian-distributed random errors that are spatially and temporally uncorrelated, and uncorrelated over frequency with an identical spectrum at each sensor. If the phase and amplitude of the source are unknown, then the likelihood function can be maximized with respect to the unknown source information and can be written

$$L(\mathbf{d}^{\text{obs}}|\mathbf{m}) = \prod_{f=1}^F \prod_{a=1}^{N_r} \frac{1}{(\pi \nu_{fa})^N} \exp [-B_{fa}(\mathbf{m}) |\mathbf{d}_{fa}^{\text{obs}}|^2 / \nu_{fa}], \quad (2.21)$$

for  $F$  frequencies and  $N_r$  ranges. In Eq. (2.21),  $\nu_{fa}$  is the data variance and  $B_{fa}(\mathbf{m})$  is the normalized Bartlett mismatch (e.g., Mecklenbräuker and Gerstoft 2000; Dosso and Wilmut 2002), defined

$$B_{fa}(\mathbf{m}) = 1 - \frac{|\mathbf{d}_{fa}^{\dagger}(\mathbf{m}) \mathbf{d}_{fa}^{\text{obs}}|^2}{|\mathbf{d}_{fa}(\mathbf{m})|^2 |\mathbf{d}_{fa}^{\text{obs}}|^2}, \quad (2.22)$$

where  $\mathbf{d}_{fa}(\mathbf{m})$  is the replica field for the proposed model, computed using a numerical propagation model. The range of  $B_{fa}(\mathbf{m})$  is between zero and one. A perfect match (parallel observed and replica data vectors) is represented by zero and a perfect mismatch (orthogonal vectors) by one. The error function is defined as the negative of the exponent of the likelihood function, Eq. (2.21) (and (2.10)), given by

$$E(\mathbf{m}) = \sum_{f=1}^F \sum_{a=1}^{N_r} B_{fa}(\mathbf{m}) |\mathbf{d}_{fa}^{\text{obs}}|^2 / \nu_{fa}. \quad (2.23)$$

Minimizing this function, as in the optimization techniques, will result in the maximum likelihood model for incoherent data (i.e., data where information about the source phase and amplitude are unknown). If the data variance is known, then Eq.

(2.23) can be easily calculated. However, in many cases, an independent estimate of the data variance is not available. Therefore, allowing a variance to be calculated for  $N$  sensors at each range, the variance estimate that results from maximizing the likelihood function (2.21) with respect to  $\nu_{fa}$  can be written

$$\hat{\nu}_{fa} = B_{fa}(\hat{\mathbf{m}}) |\mathbf{d}_{fa}^{\text{obs}}|^2 / N \quad (2.24)$$

where  $\hat{\mathbf{m}}$  is the maximum likelihood solution for data that has unknown variances. By substituting  $\hat{\nu}_{fa} = B_{fa}(\mathbf{m}) |\mathbf{d}_{fa}^{\text{obs}}|^2 / N$  into Eq. (2.21),  $\hat{\mathbf{m}}$  can be determined by minimizing

$$\Phi(\mathbf{m}) = \prod_{f=1}^F \prod_{a=1}^{N_r} B_{fa}(\mathbf{m}), \quad (2.25)$$

using an optimization method. If the assumption that the errors are independent at each sensor is not true because the errors are actually correlated between sensors as defined by the propagating normal modes, then  $N$  in Eq. (2.24) can be replaced by the effective number of independent sensors,  $N_e$ ,

$$\hat{\nu}_{fa} = B_{fa}(\hat{\mathbf{m}}) |\mathbf{d}_{fa}^{\text{obs}}|^2 / N_e. \quad (2.26)$$

The number of propagating modes corresponds to  $N_e$ , and the number can be estimated from a principal component analysis of the data covariance matrix (Neilson *et al.* 1997). By replacing  $\nu_{fa}$  with  $\hat{\nu}_{fa}$  in Eq. (2.23), the error function used in the FGS becomes

$$E(\mathbf{m}) = \sum_{f=1}^F \sum_{a=1}^{N_r} B_{fa}(\mathbf{m}) |\mathbf{d}_{fa}^{\text{obs}}|^2 / \hat{\nu}_{fa}. \quad (2.27)$$

## 2.4 Unknown Model Parameterization

As previously mentioned, not knowing the suitable model parameterization is a common problem in geoacoustic inversion. It is a problem encountered in many of

the inversion problems analyzed in this work. The basis of the under-parameterized approach is to start with a simplistic model of the environment and then repeatedly solve the inversion problem while increasing the complexity of the model parameterization for each inversion run. For example, in the problems considered here, the appropriate number of sediment layers needed to represent the ocean bottom may not be known. The under-parameterized approach described by Dosso, Wilmut and Lapinski (2001) is applied. The inversion algorithm is run a number of times and the number of sediment layers,  $L$ , is increased with each successive run. If the number of sediment layers is too small, then not all the information available in the acoustic data is used and the model will have too little structure. If the number of layers is too large, then excess structure that cannot be resolved by the acoustic data is incorporated into the solution. The optimal number of layers can be determined by analyzing the mismatch and the structure of the model solution for each run of the algorithm.

To quantify the amount of structure in the model, the  $l_1$  norm of variation for the model parameters is used. For a model  $\mathbf{m}$ , the parameters  $m_i$  that represent the same property in different layers are grouped together. For geoacoustic inversion, those properties can include compressional speed, density and attenuation in the sediment. For a property  $q$ , the  $l_1$  norm of variation is defined:

$$V_q = \sum_{l=1}^{L'-1} |q_{l+1} - q_l|, \quad (2.28)$$

where  $L'$  represents the number of model parameters representing a property. The parameters are ordered with respect to depth. The variation is also referred to as the flatness of the model, because it approximates the depth derivative of the model parameters.

## 2.5 Sound Propagation Models

MFI methods, such as ASSA and FGS, require the calculation of replica acoustic fields  $\mathbf{d}(\mathbf{m})$  for a given model  $\mathbf{m}$ . Several types of propagation models have been developed to describe sound propagation in the sea: e.g., spectral or “fast field program”, normal mode, ray, and parabolic equation models, and direct finite-difference or finite-element solutions to the full wave equation (Jensen *et al.* 1994). In this thesis, normal mode and parabolic equation models are used, therefore they are briefly introduced in the following. For further information about the methods presented here, as well as other methods, please see Jensen *et al.* (1994).

### 2.5.1 Normal Modes

The acoustic wave equation that governs sound propagation can be stated

$$\rho \nabla \cdot \left( \frac{1}{\rho} \nabla p \right) - \frac{1}{c^2} \frac{\partial^2 p}{\partial t^2} = 0, \quad (2.29)$$

where  $c$  is sound speed, and  $\rho$  is density. Given a range independent environment and allowing pressure, sound speed, and density to be represented by  $p(r, z)$ ,  $c(z)$  and  $\rho(z)$ , respectively, the two dimensional (2-D) time independent wave equation in cylindrical coordinates that describes the acoustic energy propagating through the ocean from a point source at depth  $z_s$  and range  $r = 0$  can be written as

$$\frac{1}{r} \frac{\partial}{\partial r} \left( r \frac{\partial p}{\partial r} \right) + \rho \frac{\partial}{\partial z} \left( \frac{1}{\rho} \frac{\partial p}{\partial z} \right) + \frac{\omega^2}{c^2} p = -\frac{\delta(r) \delta(z - z_s)}{2\pi r}, \quad (2.30)$$

which is the Helmholtz Equation (HE). The separation of variables technique is applied to the homogeneous form of the equation (i.e., the right hand side is set to zero, implying the absence of the source). In this technique, pressure is redefined as

$$p(r, z) = \sum_m R_m(r) Z_m(z), \quad (2.31)$$

where the solution is a linear combination of all unique solutions. By substituting (2.31) into the homogeneous HE, the equation can be divided into radial and depth

dependent equations. The depth dependent equation reduces to an eigenvalue problem, where the  $m^{\text{th}}$  largest eigenvalue is the radial (horizontal) wave number,  $k_{rm}^2$ , and the  $m^{\text{th}}$  eigenfunction  $Z_m(z)$  is the  $m^{\text{th}}$  mode function or the normal mode. The mode functions are normal (orthogonal) to each other. By re-substituting (2.31) into (2.30),  $R_m(r)$  can be solved with respect to the source. Having solved  $Z_m(z)$  and  $R_m(r)$ , the solution to Eq. (2.30) can now be constructed using (2.31) which is a superposition of all possible modal solutions.

Using normal modes, the wave equation can be solved accurately for range independent environments. There are several methods in use to solve for  $Z_m(z)$ , such as finite-difference, layer methods and shooting methods (Jensen *et al.* 1994). However, each method requires a routine to determine the eigenvalues of the problem. The roots of the characteristic equation must be found. This is a challenging and important facet of normal-mode computation. Jensen *et al.* (1994) suggests bisection, deflation, brute-force searches, and analytic estimates as optional eigenvalue finding methods. For the wave propagation model ORCA (Westwood *et al.* 1996) used here, the eigenvalue problem is expressed in terms of reflection coefficients and the reflection coefficients at a reference depth are utilized to determine the eigenvalues.

The normal mode methods developed thus far have been for range independent environments. Normal mode methods can also be used to approximately solve for a range dependent environment. In the coupled modes method the range of the problem is divided into range independent segments. The field is then calculated within each segment, and boundary conditions at the interface of each segment connect the solutions together. This is a computationally intensive method, and therefore is typically only used as a reference solution for other approximate methods (Jensen *et al.* 1994). One approximate method is the method of adiabatic normal modes which is only appropriate in the farfield (i.e.,  $k_{rm}r \gg 1$ ). For adiabatic normal modes, the coupling between different modes at the interface of the segments is ignored. It is assumed that there is no transfer of energy to higher or lower modes (i.e., an adiabatic

approximation).

### 2.5.2 Parabolic Equations

The parabolic equation (PE) method to solving the Helmholtz equation is based on the assumption that the out-going energy dominates over the backscattered energy. This method is often used to determine fields in range dependent environments. It is less computationally intensive than coupled-modes but is not as accurate as normal modes for range independent environments, due to the main assumption and the farfield approximation it uses. However, PEs do not make the adiabatic approximation. To derive the 2-D parabolic equation for a point source, Eq. (2.30) is restated in homogeneous form:

$$\frac{1}{r} \frac{\partial}{\partial r} \left( r \frac{\partial p}{\partial r} \right) + \rho \frac{\partial}{\partial z} \left( \frac{1}{\rho} \frac{\partial p}{\partial z} \right) + k_o^2 n^2 p = 0 \quad (2.32)$$

where  $k_o = \omega/c_o$  is a reference wavenumber,  $c_o$  is a reference sound speed, and  $n(r, z) = c_o/c(r, z)$  is the index of refraction. The solution to the wave equation is assumed to take the form

$$p(r, z) = \Phi(r, z) H_0^{(1)}(k_o r) \quad (2.33)$$

where  $H_0^{(1)}(k_o r)$  is the zeroth order Hankel function and  $\Phi$  is assumed to be slowly varying in range. For the farfield assumption, ( $k_o r \gg 1$ ), the asymptotic solution to  $H_0^{(1)}(k_o r)$  is

$$H_0^{(1)}(k_o r) \simeq \sqrt{\frac{2}{\pi k_o r}} \exp \left[ i \left( k_o r - \frac{\pi}{4} \right) \right]. \quad (2.34)$$

Therefore, for the farfield assumption Eq. (2.33) can be re-written:

$$p(r, z) = \frac{\exp[ik_o r]}{\sqrt{r}} \Psi(r, z) \quad (2.35)$$

where the difference between  $\Phi$  and  $\Psi$  is merely a constant factor. Substituting Eq. (2.35) into Eq. (2.32) gives the elliptic wave equation:

$$\frac{\partial^2 \Psi}{\partial r^2} + 2ik_o \frac{\partial \Psi}{\partial r} + \rho \frac{\partial}{\partial z} \left( \frac{1}{\rho} \frac{\partial \Psi}{\partial z} \right) + k_o^2 (n - 1) \Psi = 0. \quad (2.36)$$

approximation).

### 2.5.2 Parabolic Equations

The parabolic equation (PE) method to solving the Helmholtz equation is based on the assumption that the out-going energy dominates over the backscattered energy. This method is often used to determine fields in range dependent environments. It is less computationally intensive than coupled-modes but is not as accurate as normal modes for range independent environments, due to the main assumption and the farfield approximation it uses. However, PEs do not make the adiabatic approximation. To derive the 2-D parabolic equation for a point source, Eq. (2.30) is restated in homogeneous form:

$$\frac{1}{r} \frac{\partial}{\partial r} \left( r \frac{\partial p}{\partial r} \right) + \rho \frac{\partial}{\partial z} \left( \frac{1}{\rho} \frac{\partial p}{\partial z} \right) + k_o^2 n^2 p = 0 \quad (2.32)$$

where  $k_o = \omega/c_o$  is a reference wavenumber,  $c_o$  is a reference sound speed, and  $n(r, z) = c_o/c(r, z)$  is the index of refraction. The solution to the wave equation is assumed to take the form

$$p(r, z) = \Phi(r, z) H_0^{(1)}(k_o r) \quad (2.33)$$

where  $H_0^{(1)}(k_o r)$  is the zeroth order Hankel function and  $\Phi$  is assumed to be slowly varying in range. For the farfield assumption, ( $k_o r \gg 1$ ), the asymptotic solution to  $H_0^{(1)}(k_o r)$  is

$$H_0^{(1)}(k_o r) \simeq \sqrt{\frac{2}{\pi k_o r}} \exp \left[ i \left( k_o r - \frac{\pi}{4} \right) \right]. \quad (2.34)$$

Therefore, for the farfield assumption Eq. (2.33) can be re-written:

$$p(r, z) = \frac{\exp[ik_o r]}{\sqrt{r}} \Psi(r, z) \quad (2.35)$$

where the difference between  $\Phi$  and  $\Psi$  is merely a constant factor. Substituting Eq. (2.35) into Eq. (2.32) gives the elliptic wave equation:

$$\frac{\partial^2 \Psi}{\partial r^2} + 2ik_o \frac{\partial \Psi}{\partial r} + \rho \frac{\partial}{\partial z} \left( \frac{1}{\rho} \frac{\partial \Psi}{\partial z} \right) + k_o^2 (n - 1) \Psi = 0. \quad (2.36)$$

By factoring this equation into outgoing and incoming waves and assuming that the range dependence of  $n(r, z)$  is relatively weak, the outgoing (one-way) wave equation (i.e., the parabolic wave equation), can be stated:

$$\begin{aligned} \frac{\partial \Psi}{\partial r} &= ik_o \left( -1 + \sqrt{n^2 + \frac{\rho}{k_o^2} \frac{\partial}{\partial z} \left( \frac{1}{\rho} \frac{\partial}{\partial z} \right)} \right) \Psi \\ &= ik_o (-1 + \sqrt{1 + q}) \Psi \end{aligned} \quad (2.37)$$

where

$$q = n^2 - 1 + \frac{\rho}{k_o^2} \frac{\partial}{\partial z} \left( \frac{1}{\rho} \frac{\partial}{\partial z} \right). \quad (2.38)$$

There are many different PE methods used in ocean acoustics. The differences between the methods lie in the approximations made so that Eq. (2.37) can be solved. The difficulty of the problem lies in evaluating the square-root operator  $\sqrt{1 + q}$ . For example, Tappert, Claerbout and Green PEs (Jensen *et al.* 1994) approximate  $\sqrt{1 + q}$  using the general form

$$\sqrt{1 + q} \simeq \frac{a_o + a_1 q}{b_o + b_1 q} \quad (2.39)$$

where the coefficients  $a_o$ ,  $a_1$ ,  $b_o$  and  $b_1$  can be chosen to minimize the error for a given angular aperture with respect to the point source. A Padé series expansion of  $\sqrt{1 + q}$  is another technique, where

$$\sqrt{1 + q} = 1 + \sum_{j=1}^{N_P} \frac{a_{j,N_P} q}{1 + b_{j,N_P} q} + O(q^{2N_P+1}) \quad (2.40)$$

and  $N_P$  is the number of terms kept in the expansion.

The angular aperture of the source which produces the most important outgoing waves with respect to the problem defines which PE should be used. For example, both the Tappert and Claerbout have increasing phase error with increasing angles relative to the primary source direction. Beyond  $\pm 15^\circ$  and  $\pm 40^\circ$  relative to the main propagation direction, respectively, those PEs are typically not appropriate if reasonable accuracy is needed. For  $N_P > 1$ , the Padé expansion method is computationally

more demanding than the other PEs. However, the limiting angle increases as  $N_P$  increases. Almost the full  $\pm 90^\circ$  can be accounted for if enough terms are used. If computation time is a consideration, then determining the aperture of importance for the problem at hand so as to choose an appropriate PE is necessary.

Another PE technique is to approximate the wave propagator rather than the square-root operator; i.e., apply the approximation directly to the acoustic field rather than to the PE. PEs are solved by marching the solution out in range. To determine a field at range  $R$ , the environment is divided into range independent regions. The solution to Eq. (2.37) is then solved for each region by using the solution from one region to calculate the solution for the next region (solution methods will be mentioned shortly). If  $\Psi$  is expanded through a Taylor series about  $r + \Delta r$ , where  $\Delta r$  is a range independent region, then

$$\Psi(r + \Delta r, z) = \exp[\Delta r \partial_r] \Psi(r, z). \quad (2.41)$$

Together, Eqs. (2.37) and (2.41) lead to the relation

$$\Psi(r + \Delta r, z) = \exp \left[ ik_o \Delta r \left( -1 + \sqrt{1 + q} \right) \right] \Psi(r, z). \quad (2.42)$$

In the split-step Padé technique, the wave propagator  $\exp \left[ ik_o \Delta r \left( -1 + \sqrt{1 + q} \right) \right]$  is approximated for each region, rather than the the square-root operator. Therefore, the field is directly approximated rather than indirectly approximated. The propagator is approximated through a Padé expansion similar to Eq. (2.40), where the square-root operator on the left-hand-side of Eq. (2.40) is replaced by the propagator. The split-step Padé method is a wide-angle PE, like the regular Padé expansion of the square-root operator. Because the split-step Padé technique is based on an approximation of the propagator itself, much larger range steps ( $\Delta r$ ) can be used compared to the regular Padé expansion of the square-root operator, which will reduce computation time (Brooke *et al.* 2001).

When choosing a PE algorithm to use, the type of PE which is appropriate to the problem must first be determined: e.g., a narrow angle or a wide angle PE. Next

the solution method must be considered. Two methods are commonly used: the split-step Fourier transform method and the finite-difference/finite-element method. Unlike the split-step Fourier algorithm, the finite-difference (FD) and finite-element (FE) methods can be applied to all PEs. The relative efficiency of the solution methods depend on the problem.

PEs also require an approximation of the starting field to begin range marching the solution. Several examples are given by Jensen *et al.* (1994), which can be classified as either numerical or analytic starters. The choice of a starter depends on the PE used, the aperture needed, the computational nature of the task at hand and the environment of the problem. For example, numerical starters can be more computationally intensive than analytic starters, although some analytic starters are not appropriate for wide-angle PEs.

Finally, the methods used to solve the PE require dividing the environment with respect to range and depth. Finding an appropriate grid size is important. Finer grids provide a more accurate solution but more computation is required; therefore, a balance between the grid size and accuracy is required. Also, in the case of the Padé expansions, a similar balance with respect to the number of expansion terms is required.

To get a more accurate or efficient solution, there are techniques that can be applied within a PE algorithm. For example, due to the use of the one-way wave equation, energy is not properly conserved at the vertical boundaries of the grid. Two boundary conditions need to be satisfied at the vertical interfaces, but the one-way wave equation only allows for one to be satisfied. This becomes a problem for sloping interfaces in the ocean environment. A PE algorithm applied to a range dependent problem should compensate for the energy conservation problem. In PECan, the PE code used in this work, the true energy-conserving condition is approximated by replacing  $\Psi$  with a scaled field  $\Psi^* \equiv \Psi/\gamma$  where  $\gamma = \sqrt{\rho(z) c_o/\mathcal{N}(z)}$  and  $\mathcal{N}(z)$  is the range independent complex index of refraction that handles attenuation (Brooke

*et al.* 2001). Also, the ocean bottom must be terminated at some depth. Typically, an absorbing layer is applied which attenuates the field and prevents any unwanted reflections at the termination of the grid. Another option is to use a non-local boundary condition (Brooke and Thomson 2000; Brooke *et al.* 2001). This method mimics a semi-infinite PE problem using a bounded domain. As a result, the total depth of the grid can be less than for an absorbing layer which will decrease the size of the computational domain and therefore decrease computation time. Non-local boundaries can also be used to mimic a rough ocean surface.

As mentioned earlier, PECan is the PE algorithm used in this work (Brooke *et al.* 2001). PECan is an  $N \times 2$ -D/3-D PE underwater sound propagation model which can march the acoustic field solution in range, depth and azimuth. The user has the option of using either a split-step Padé or regular Padé algorithm. A variety of starting fields can be used. The techniques of energy conservation and non-local boundary conditions can also be applied.

---

## Chapter 3

### Range Independent Inversion

In this chapter, ASSA is applied to data from a range independent environment that has unknown parameterization. The goal of the investigation was to determine if ASSA and the under-parameterized approach could be applied to determine an appropriate parameterization and a good model estimate. Synthetic data generated for the 1997 Geoacoustic Inversion Workshop (Workshop '97) are used. Benchmark acoustic field data were provided at the Workshop for test cases representing shallow ocean environments with range independent properties. The benchmark data were generated using the wavenumber integration model SAFARI (Schmidt 1988). Many inversion techniques have been applied to the data, such as GA (Ratilal *et al.* 1998; Heard *et al.* 1998; Taroudakis and Markaki 2000), FSA (Fallat and Dosso 1998), DHS (Fallat and Dosso 1999), modified grid searches (Tolstoy 1998; Ainslie *et al.* 2000) and hybrid algorithms (Fallat and Dosso 1999). In more recent work, ASSA has been shown to efficiently obtain better results for almost all the test cases compared to the previously published results for the other optimization methods (Dosso, Wilmut and Lapinski 2001). In this section, the results of applying ASSA to the so called NL test case will be presented. These results were originally presented by Dosso, Wilmut and Lapinski (2001).

The NL test case is the only Workshop '97 test case that has unknown model parameterization. The number of sediment layers in the NL subcases (NLa, NLb and NLc), up to a maximum of 50 layers, is unknown, however the parameter bounds were given as part of the benchmark problem. The subcases have the same unknown geophysical and geometric properties, but the data are generated for different environments. Fig. 3.1 shows a schematic representation of the NL environ-

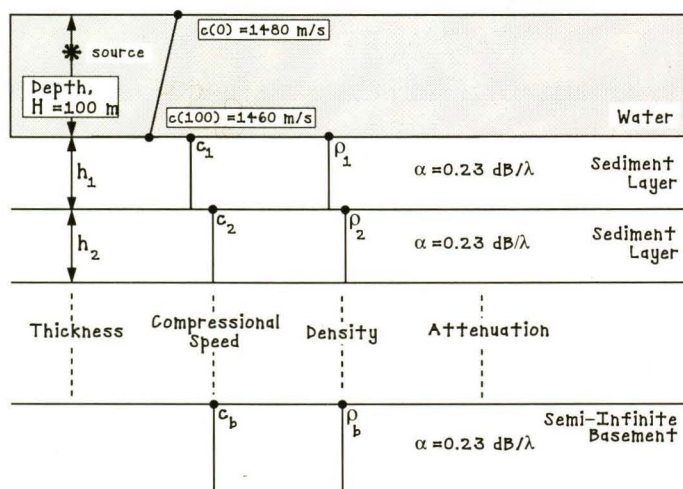


Figure 3.1 Schematic representation of the NL test case environment and model parameterization.

ment and the preliminary model parameterization that was adopted. All but one sediment layer has unknown compressional speed, density, and layer thickness (i.e.,  $\{c_\ell, \rho_\ell, h_\ell; \ell = 1, L - 1\}$ , respectively); the exception being that the semi-infinite basement only has unknown  $c_b$  and  $\rho_b$ . For  $L$  sediment layers, the model has  $3L - 1$  parameters. The remaining geophysical and geometric parameters needed to represent the environments are indicated in Fig. 3.1. The under-parameterized approach described in Sec. 2.4 was applied to the subcase problems to determine the appropriate number of sediment layers needed to describe each ocean bottom environment. In the following, the results of the parameterization analysis for subcase NLa are presented and then the results of similar analyses for subcases NLb and NLc are briefly presented.

A large amount of data are available for the NL test case. In the following results, benchmark data were used representing 100 Hz source frequency data received by 21 receivers at a range of 1 km on a vertical line array (VLA). The receivers were evenly spaced from 5 to 100 m water depth. The source position is at 20 m below the sea

surface.

The error function is an important part of MFI. A variation of the error function (2.23) for one frequency and range was used in the ASSA algorithm. Only the relative value of the error function is of importance in optimization algorithms. The error function can be multiplied or divided by constant factors without affecting the algorithm. Only one frequency and range was used, therefore the factor  $|\mathbf{d}^{\text{obs}}|^2/\nu$  could be removed from (2.23). The new error function becomes the Bartlett mismatch as defined in Eq. (2.22) which is the error function used in the ASSA algorithm for the NL subcases. The normal-mode wave propagation model ORCA (Westwood *et al.* 1996) was used to calculate the replica fields.

The ASSA control parameters were somewhat altered between applications. For subcase NLa, the number of temperature steps in the algorithm varied: 1000 for  $L \leq 2$ , 2000 for  $2 < L \leq 8$  and 2500 for  $L = 9$ . The convergence criterion (2.1) was not included in the ASSA algorithm prior to quenching for the NL subcases. Typically, the algorithm required seven perturbations to be accepted at each temperature step, but this was increased to eleven for  $L = 9$ . The ASSA control parameters also included  $T_o = 0.3$  and  $\beta \approx 0.99$ . For all applications of ASSA presented in this work perturbation scale factors of  $s = 3$  and  $S = 30$  were applied. From  $L = 1$  to  $L = 8$ , the computation time increased from  $\sim 2$  to  $\sim 9$  hours on a 200 MHz computer while  $L = 9$  required  $\sim 8$  hours on a 400 MHz PC.

Fig. 3.2 shows the results of the under-parameterized approach to determining the appropriate number of sediment layers. Figs. 3.2a and b respectively show the mismatch,  $E$ , and variation,  $V$  (2.28), for increasing number of layers  $L$ . In Fig. 3.2a, the mismatch decreases significantly between  $L = 1$  and  $L = 4$ . Beyond  $L = 4$ , the mismatches generally plateau indicating that it is difficult to fit the data any better due to limited precision of the data, the numerical model, and/or an imperfect search. In Fig. 3.2b, the compressional speed and density variations ( $V_c$  and  $V_\rho$ ), which are measures of structure, generally increase with  $L$ . However,  $V_c$  plateaus between

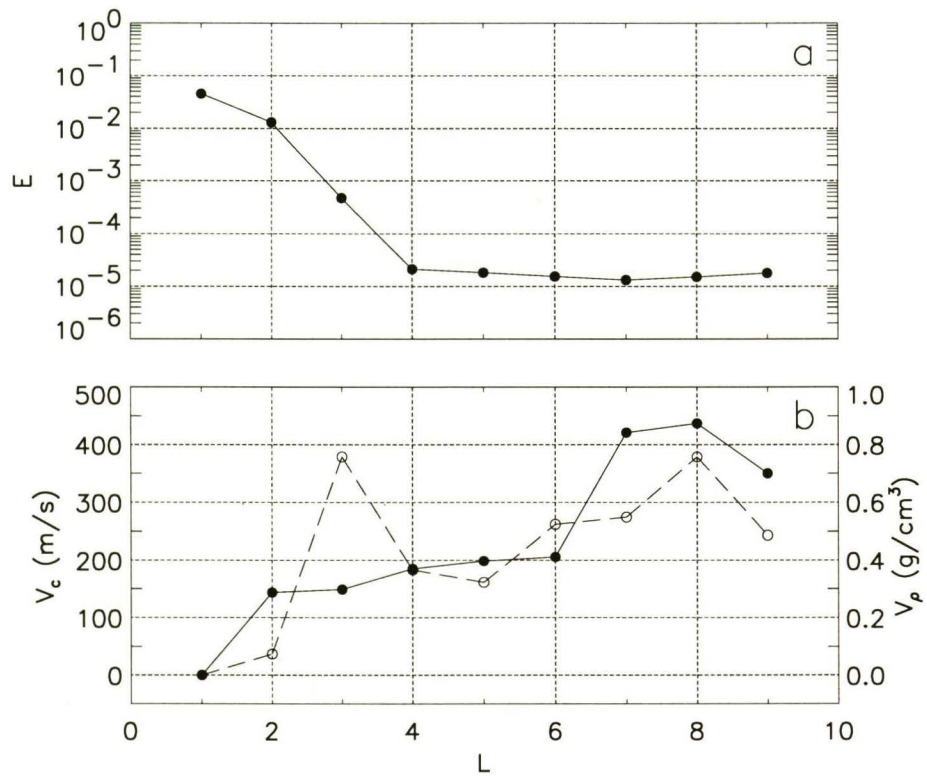


Figure 3.2 Under-parameterized approach to determining the appropriate number of sediment layers,  $L$ , for the NLa subcase: (a) the mismatch  $E$  and (b) the  $l_1$  norm variations of the compressional speed (solid line,  $V_c$ ) and density (dashed line,  $V_\rho$ ).

$L = 4$  and  $L = 6$ . Based on the  $V_c$  plateau, and the behavior of the mismatches in Fig. 3.2a, it was determined that  $L = 4$  through  $L = 6$  allows for models that resolve the geoacoustic structure as well as the acoustic data will permit. The number of sediment layers that best represent the ocean bottom was chosen to be  $L = 5$ .

The convergence of the algorithm is reflected in the plots of mismatch and model parameters of the simplex with respect to temperature step. Fig. 3.3 shows the convergence of the ASSA algorithm for subcase NLa using  $L = 5$ . The convergence of the mismatch  $E$  is shown in Fig. 3.3a. The mismatches decrease with the progression of the algorithm and converge to an exceedingly low value of  $E = 1.8 \times 10^{-5}$ . The final mismatch is lower than the mismatch between the benchmark data and the replica data generated using the true model,  $E = 9.5 \times 10^{-5}$  (the dotted line in Fig. 3.3a). Figs. 3.3b–o show the convergence of the model parameter values. The convergence rate for each parameter qualitatively indicates the relative sensitivity of that parameter. The model parameter values of  $h$  and  $c$  typically converge earlier than  $\rho$ , with  $\rho_b$  converging the slowest of all the parameters. Throughout Fig. 3.3, the abrupt change for all but one model in the simplex at temperature step 2000 is due to the quenching process described in Sec. 2.1.3.

To consider the adaptive properties of the ASSA, Fig. 3.4 illustrates the behavior of the Cauchy-distribution scaling factors,  $\delta m_i$  (Sec. 2.1.3), with respect to temperature step. The scaling factors control the width of the Cauchy-distribution from which the random perturbations are drawn. This figure illustrates the adaptive behavior of the algorithm for each parameter over the course of the inversion. The size of the factors fluctuate widely for the first  $\sim 700$  temperature steps. For  $h$  and  $c$  the factors then rapidly decrease to relatively small values. However, for  $\rho$  the decrease in the average factor size after  $\sim 700$  is more gradual.

The results of the ASSA inversion are given in Fig. 3.5 which shows depth profiles of the compressional speed and density for the  $L = 5$  solution. For comparison, the true model profiles are also included. The proposed profile for the compressional

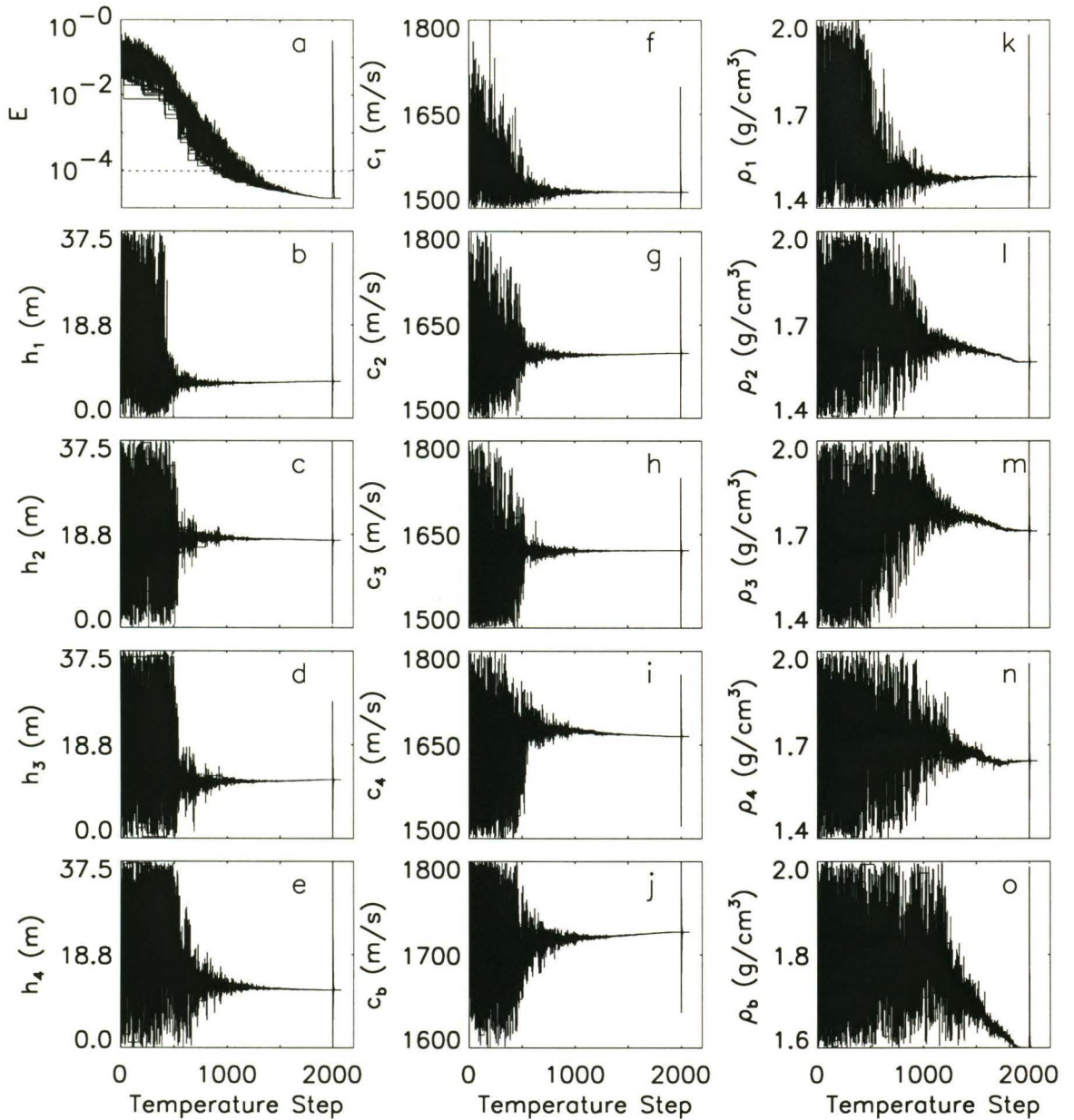


Figure 3.3 Mismatch  $E$  and parameter estimates of the  $L = 5$  ASSA inversion for the NLa subcase. (a) Mismatch for every point in the simplex plotted as a function of temperature step. The dotted line represents the mismatch between the given data and the replica data generated using ORCA and the true model. (b)-(o) Parameter values of each model in the simplex. The parameter bounds used in the inversion are represented by the  $y$ -axis limits for each plot.

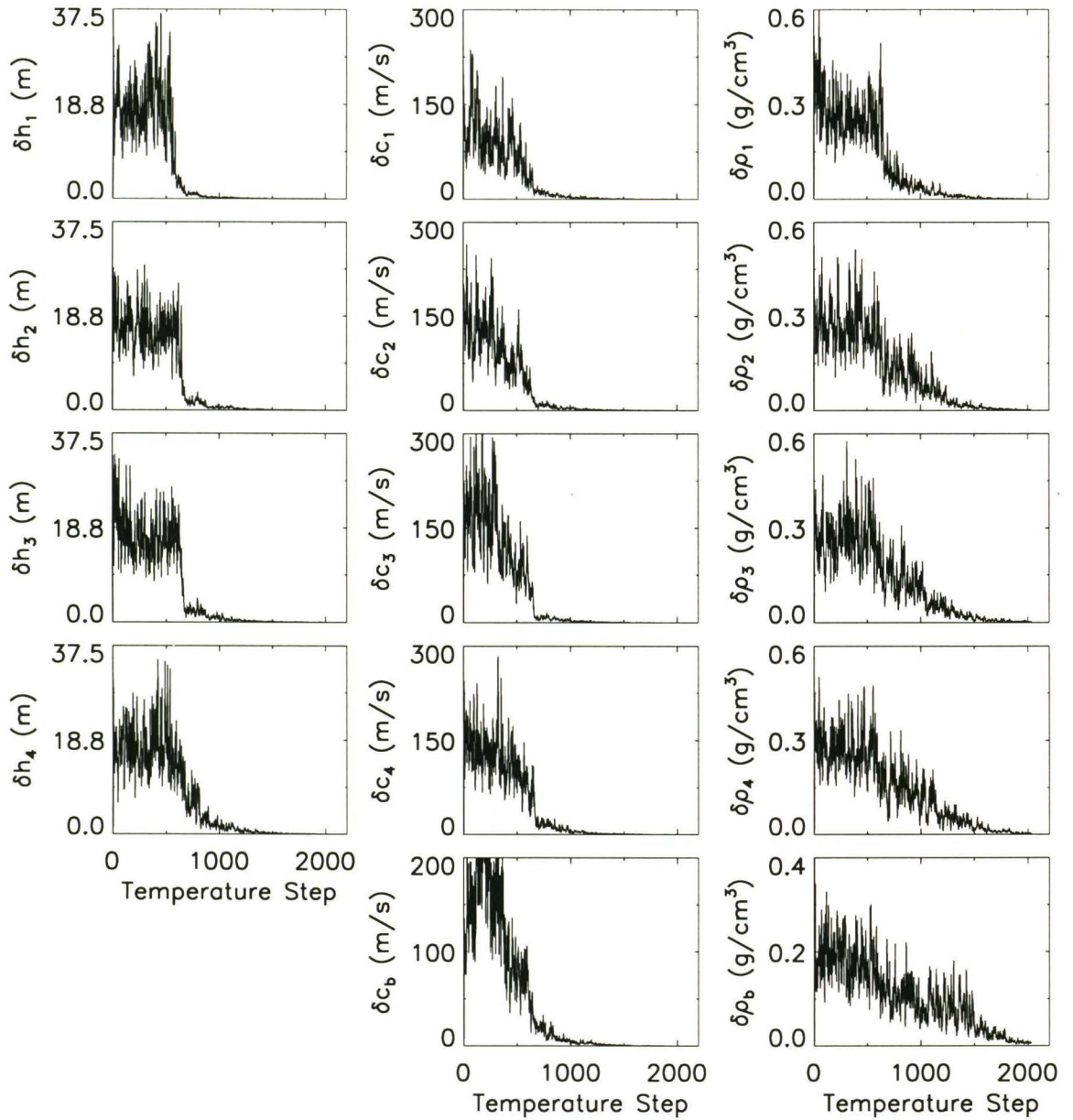


Figure 3.4 Cauchy-distribution scaling factors  $\delta m_i$  for the  $L = 5$  ASSA inversion of subcase NLa.

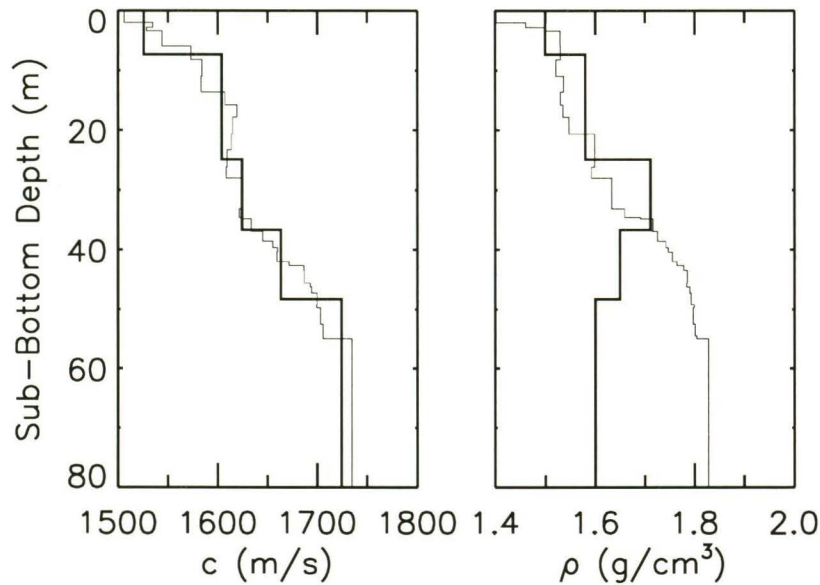


Figure 3.5 Profiles of compressional speed  $c$  and density  $\rho$  for the NLa subcase. The thin line represents the true model and the thicker line represents the ASSA solution for  $L = 5$  sediment layers.

speed approximates the true profile well over the entire sub-bottom depth; however, the small-scale structure on the true model cannot be resolved by the acoustic data. The proposed density profile approximates the true profile reasonably well until a depth of about 40 m.

Determining parameter uncertainties is an important part of using inverse techniques to predict model parameter values. If a true model is not available, as is usually the case, then the uncertainties are a measure of how well the model parameter estimates have been determined. However, determining the uncertainties can also be a time consuming exercise. A relatively quick method to get a general impression of relative parameter uncertainties is to qualitatively determine the sensitivity of the parameters. If a parameter is sensitive, small changes in the parameter value cause large changes in the mismatch; therefore, it can be expected that the parameter will have a small range of uncertainty. If a parameter is insensitive, small changes

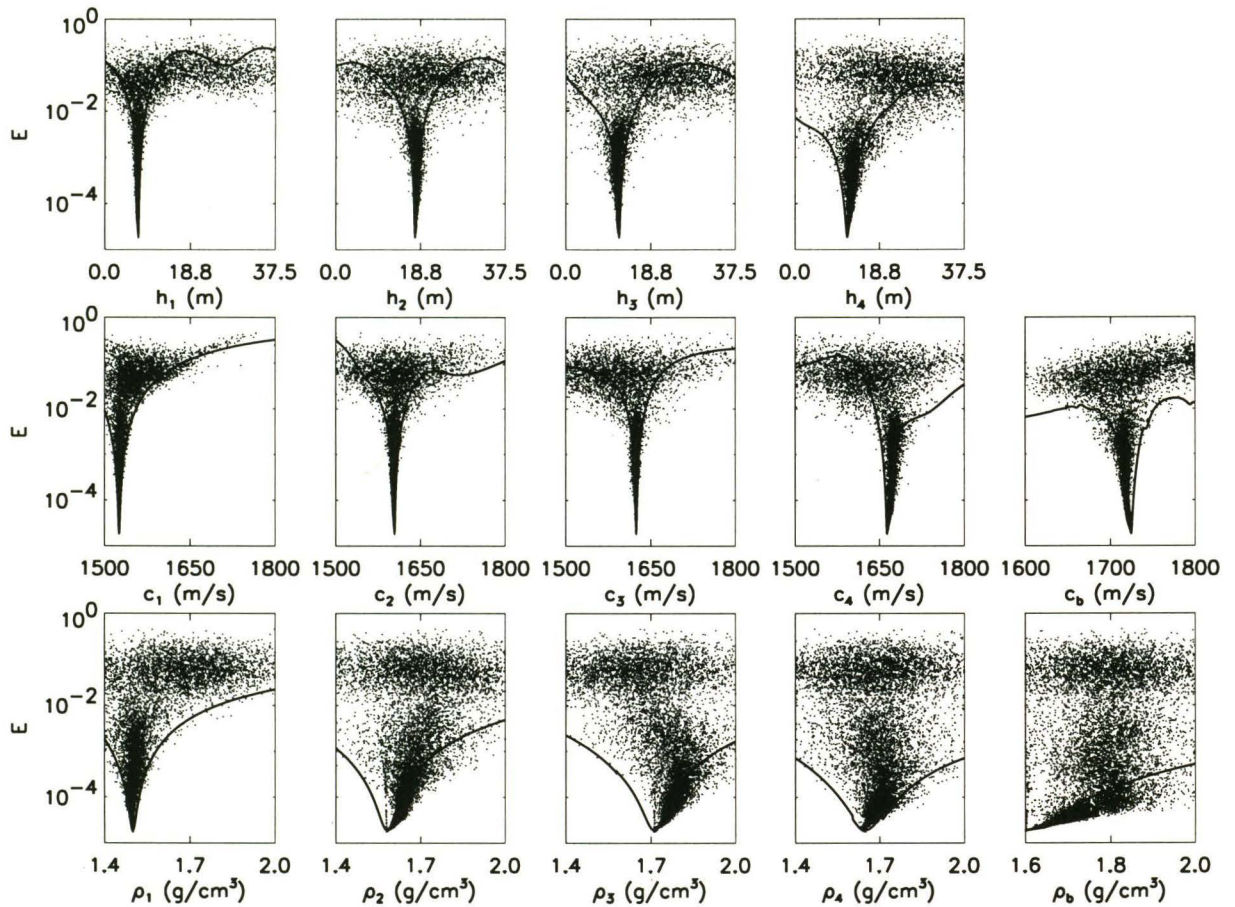


Figure 3.6 Sensitivity analysis for NLa and  $L = 5$  sediment layers. Solid line represents 1-D cross-sections of the objective function. Small dots represent the mismatch  $E$  for every accepted model.

in the parameter value will not cause a significant change in the mismatch and the parameter will likely have a larger range of uncertainty.

Cross-sections of the objective function can be used to predict how well parameters should be expected to be determined through an inversion. To do this, all parameters except the ones under consideration are held constant at their estimated values and the mismatch is calculated with respect to the given fields. Fig. 3.6 shows the one dimensional (1-D) cross-sections for subcase NLa calculated using the  $L = 5$  solution

as the estimated model. 1-D cross-sections can be misleading because they only consider one slice through the  $M$ -dimensional parameter space. The slice may not reflect the true nature of the objective function. In particular, the effect of multi-parameter correlations is not accounted for.

Another approximate method to qualitatively determine parameter sensitivities is to utilize the models accepted while applying the ASSA algorithm. Plotting the mismatch as a function of parameter value for all accepted models (as shown in Fig. 3.6) gives an additional view of the parameter sensitivities. The ASSA algorithm samples the  $M$ -dimensional parameter space (i.e., is not confined to a 1-D slice); therefore, the analysis can sometimes provide alternative sensitivity information. Effects of multi-parameter correlations will be incorporated, to some degree, into the sensitivity plots. However, due to the nature of the algorithm, ASSA does not necessarily sample the entire parameter space. This can be seen in Fig. 3.6 where a 1-D cross-section achieves lower mismatches than the ASSA multi-dimensional plot. For a more complete (though still approximate) sensitivity analysis, it is prudent to combine the multi-dimensional parameter plots with the 1-D cross-sections. However, care should still be exercised because generalizing the objective function into 1-D views cannot properly depict the  $M$ -dimensional parameter space.

Using figures like Fig. 3.6, parameter sensitivities can only be determined in a relative sense. A parameter is considered more sensitive if its distribution width at low energies is narrow compared to the other parameters. Greater sensitivity implies that a parameter should be better determined, with respect to the other parameters. The sensitivity is also dependent on the parameter bounds used. If for additional applications of the algorithm the bounds are changed, then the relative sensitivities may change. Therefore, given an ASSA inversion result, it can be said that certain parameters are more sensitive than other parameters, and that certain parameters are likely to be determined better than other parameters, but qualitative uncertainty bounds for each parameter cannot be determined from the method. The Bayesian

approach introduced in Sec. 2.2 enables parameter uncertainties to be determined. However, knowing the relative sensitivities of the parameters can still give insight into the problem at hand. In Fig. 3.6,  $h$  and  $c$  are more sensitive than  $\rho$  and the sensitivity of  $\rho$  generally decreases with depth. The relative sensitivities reflect how well the solution profiles in Fig 3.5 approximate the true profiles.

While Fig. 3.5 shows the solution for the optimal parameterization, Fig. 3.7 shows sub-optimal solutions for other model parameterizations. Fig. 3.7a shows the parameter profiles for the  $L = 3$  solution, which has a higher mismatch than the  $L = 4$  solution. The proposed compressional speed profile is a reasonably good approximation, but it is apparent from Fig. 3.5 that the acoustic data can provide a more precise fit. Not all the available information in the acoustic data is used in Fig. 3.7a. Fig. 3.7b shows the  $L = 8$  solution. Fluctuations in the structure are evident in both the compressional speed and density profiles, representing features that are not resolved or constrained by the data.

Figs. 3.2–3.7 show a complete analysis of the ASSA inversion of subcase NLa. Similar analyses were carried out for subcases NLb and NLc. For both NLb and NLc,  $L = 4$  was determined to be the optimal number of layers to represent the ocean bottom. Fig. 3.8 shows the parameter profiles for the  $L = 4$  ASSA solutions. Again, the compressional speed profiles approximate the true profiles well, and the density profiles less well. The mismatches of the models are  $5.6 \times 10^{-5}$  and  $8.3 \times 10^{-6}$  for NLb and NLc, respectively.

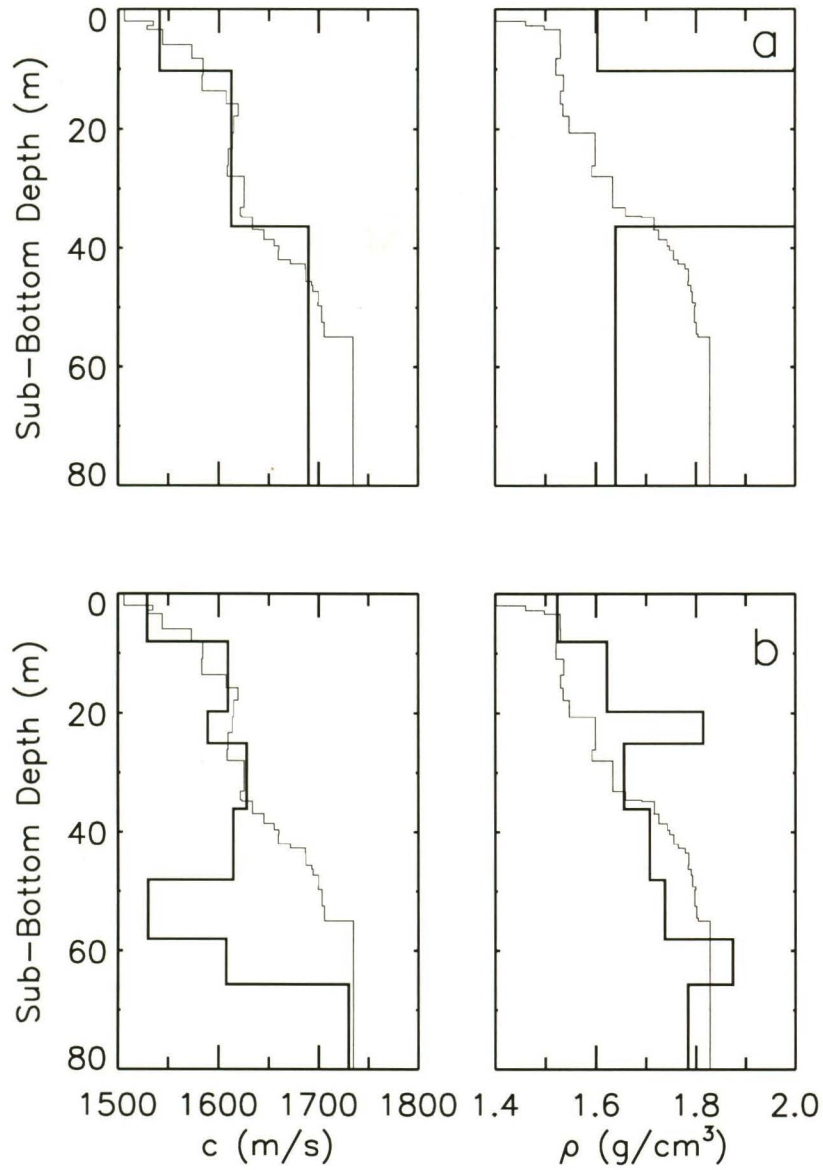


Figure 3.7 Sub-optimal profiles of compressional speed  $c$  and density  $\rho$  for the NLa subcase. The thin line represents the true model and the thicker line represents the ASSA solution for (a)  $L = 3$  and (b)  $L = 8$  sediment layers.

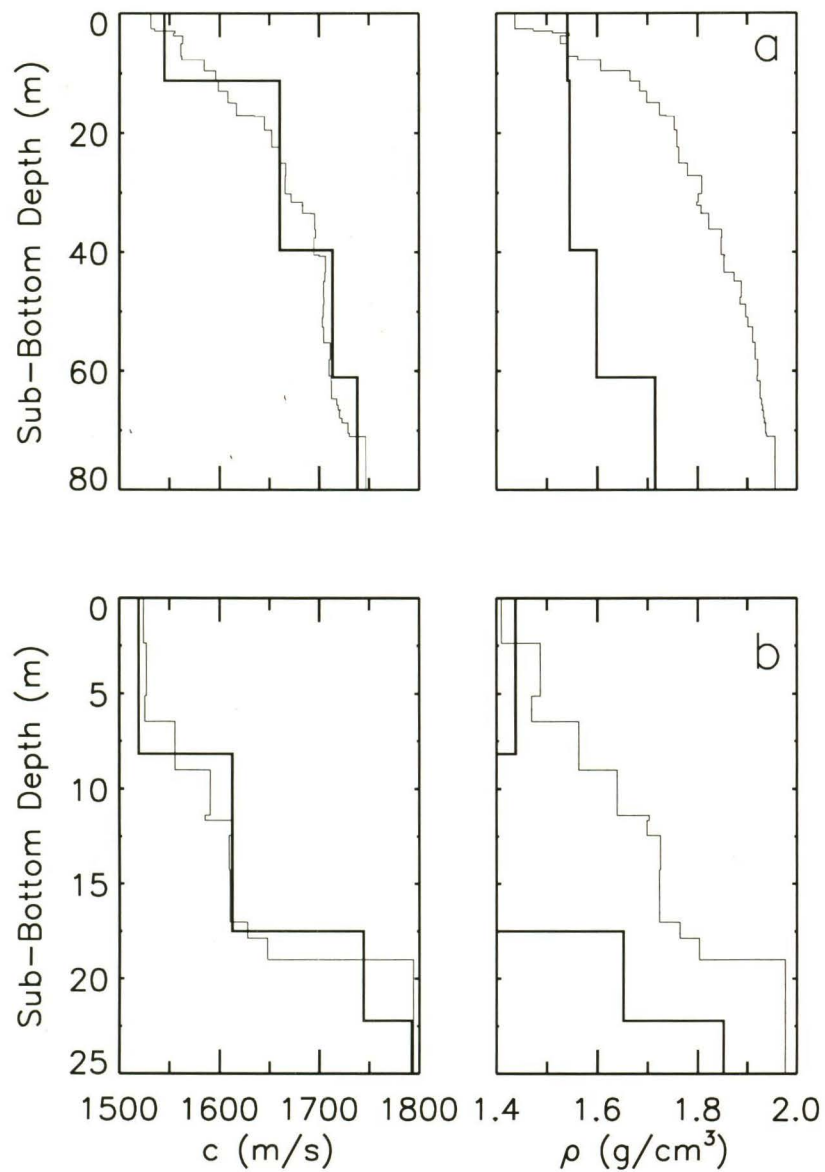


Figure 3.8 Profiles of compressional speed  $c$  and density  $\rho$  for the (a) NLb and (b) NLc subcases. The thin line represents the true model and the thicker line represents the ASSA solution for  $L = 4$  sediment layers.

## Chapter 4

### Range Dependent Inversion

#### 4.1 Introduction

In this chapter, ASSA and the FGS are applied to evaluate their ability to generate solutions for range dependent inversion problems. Range dependence is a common attribute of real ocean environments. Allowing for range dependence makes the inversion problem more general. However, the calculation of the forward problem increases the computational intensity. Due to the number of forward problem calculations that must be performed in an inversion, a coupled mode propagation model would not be appropriate. PE models are less computationally intensive, but still more intensive than, for example, a normal mode propagation model for range independent environments, as used in the previous chapter. Adiabatic normal mode models could also be applied, if coupling between the modes can be ignored. For the range dependent inversions examined here the PE wave propagation model PECan (Brooke *et al.* 2001) was used to solve the forward problems. In addition, when range dependence is introduced, it can become more difficult to determine the model parameters because of a more complicated environment and imprecise field calculations. In this chapter, results from applying ASSA and the FGS to range dependent benchmark data are presented. The benchmark data, the error function used in ASSA and the calibration of PECan are first discussed. In Secs. 4.2.1 and 4.2.2 the results of applying ASSA are presented, and in Secs. 4.3.1 and 4.3.2 the results of applying the FGS are presented.

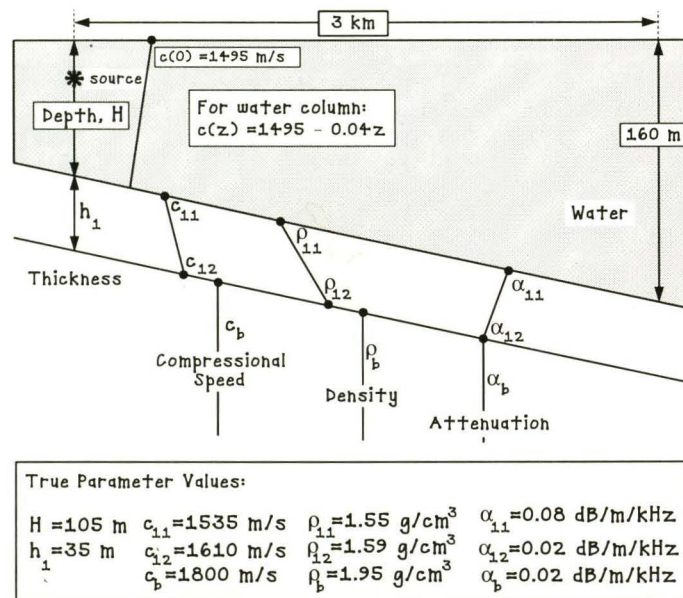


Figure 4.1 Schematic representation of the TC0 environment and model parameterization.

True parameter values are given.

#### 4.1.1 Workshop Data

The 2001 Inversion Techniques Workshop (Chapman *et al.* 2001) provided an opportunity to test geoacoustic inversion techniques using benchmark acoustic data for shallow ocean environments with range dependent properties. The data were generated using COUPLE (Evans 1983) and RAM (Collins 1993). COUPLE is a coupled mode propagation model, while RAM is a range dependent PE propagation model that uses a split-step Padé method to obtain a solution. The data for each environment are grouped into cases. ASSA and the FGS were applied to data from Test Case 1 (TC1), Test Case 2 (TC2), and the Calibration Case (TC0) on a 400 MHz PC.

Figs. 4.1, 4.2, and 4.3 depict the TC0, TC1 and TC2 environments, respectively. The TC0 and TC1 environments consist of an ocean bottom with a monotonic down-slope while the TC2 environment depicts an up-slope shelf break environment. The

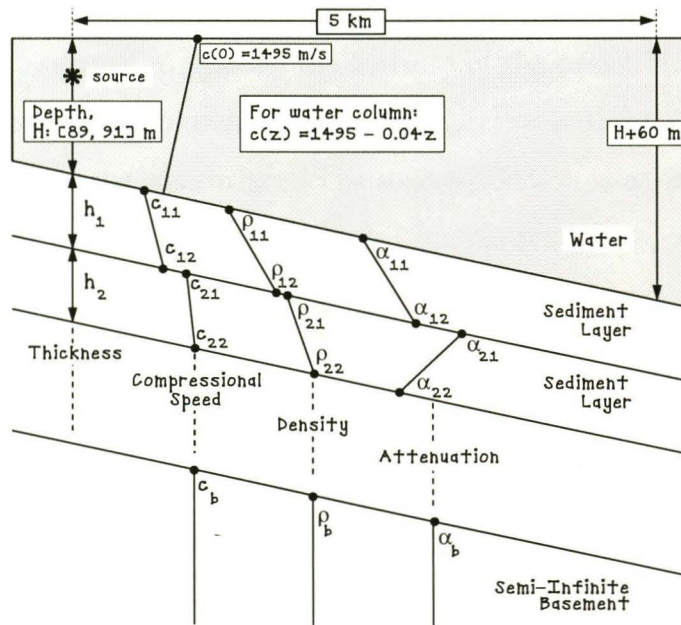


Figure 4.2 Schematic representation of the TC1 environment and model parameterization.

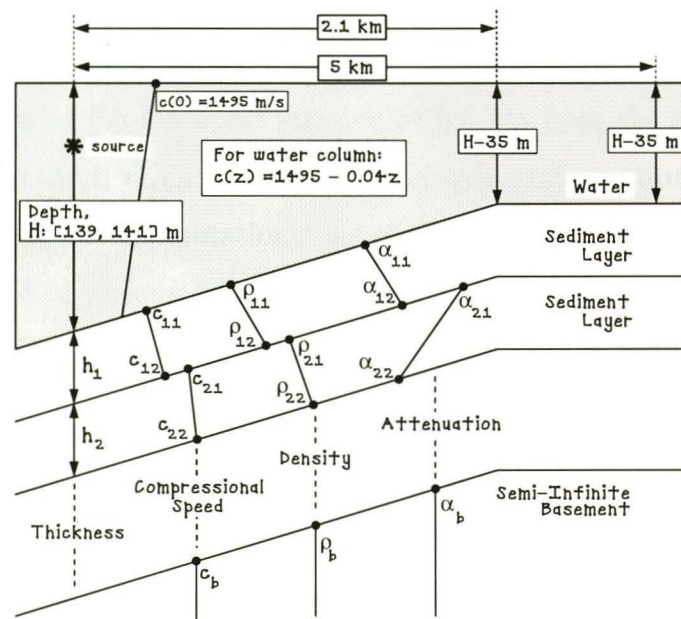


Figure 4.3 Schematic representation of the TC2 environment and model parameterization.

slopes of the ocean bottoms were assumed to be known for the test cases and the sediment layers were assumed to parallel the features of the ocean bottoms. The sound speed profile of each water column is defined by  $c_w(z) = 1495 - 0.04z$ , where  $z$  is the depth in meters below the sea surface. The source position is at 20 m below the sea surface. Parameter bounds were not provided for the test cases, with the exception of the water depth.

The number of sediment layers that is appropriate to represent the ocean bottom is unknown for TC1 and TC2. An under-parameterized approach, as described in Sec. 2.4, was applied to find the optimal number of layers for each environment. The sediment properties included as model parameters were compressional speed  $c$ , density  $\rho$ , attenuation  $\alpha$ , and layer thickness  $h$ . The water depth at the source  $H$ , was also a model parameter. For an  $L$  layer ocean bottom, the model included  $7L - 3$  parameters, where the general model parameterization is as illustrated in Figs. 4.2 and 4.3.

For TC0, TC1 and TC2, a large amount of data are available. Data representing measurements at ten VLAs for TC1 and TC2, and six VLAs for TC0 are available. The VLAs are spaced 0.5 km apart starting at 0.5 km from the source. It was felt that using the first three VLAs would provide adequate information for the inversions while not increasing the computation time of the algorithm prohibitively. For both ASSA and the FGS inversions, 50 and 200 Hz source frequency data were used, unless stated otherwise. The data represented measurements from 21 evenly spaced receivers at 20–80 m depth.

#### 4.1.2 ASSA Error Function

The error function is an essential part of each MFI method. The error function used in the ASSA algorithm is a modified version of Eq. (2.23). For noise-free synthetic data, it can be assumed that there are numerical errors on the data due to

the approximate solutions for both the data and replica fields. For the ASSA under-parameterized analysis, it was assumed that  $\mathbf{d}^{\text{obs}}/\nu_{fa}$  was constant. Therefore, Eq. (2.23) can be rewritten,

$$E(\mathbf{m}) = \frac{1}{\nu} \sum_{f=1}^F \sum_{a=1}^{N_r} B_{fa}(\mathbf{m}). \quad (4.1)$$

Removing or including a constant factor from the error function does not affect the results of an optimization algorithm. Eq. (4.1) was multiplied by  $\nu$  and divided by  $F$  and  $N_r$  to obtain the error function used in the ASSA algorithm,

$$E(\mathbf{m}) = \frac{1}{F \cdot N_r} \sum_{f=1}^F \sum_{a=1}^{N_r} B_{fa}(\mathbf{m}). \quad (4.2)$$

This mismatch function ranges between zero and one.

#### 4.1.3 Range Dependent Propagation Modeling

The range dependent PE propagation model PECAN (Brooke *et al.* 2001) was used to calculate the replica field,  $\mathbf{d}(\mathbf{m})$ , for the Workshop '01 inversions. PECAN provides many options regarding the PE solution method. The analytic Green's starting field, the energy-conservation, the basement non-local boundary condition and the split-step Padé solution method were all utilized in the inversions presented here. The reference sound speed,  $c_o$ , was taken as the average sound speed in the water column at the source position.

A set of PECAN parameters was desired that gave an acceptable mismatch using the least amount of computation time. Using the TC0 data, a thorough study of the appropriate PECAN parameters was undertaken. For the study, a selection of values for the vertical and lateral grid spacing,  $\Delta z$  and  $\Delta r$ , the basement non-local boundary depth, and the number of Padé terms,  $N_P$ , were chosen. Keeping the given geoacoustic parameters constant at their true values (Fig. 4.1), all combinations of the PECAN parameters were tried and the resulting acoustic fields were compared to the

benchmark data using the error function (4.2). Data representing the six VLAs and 21 evenly spaced receivers were used. The mismatches calculated for each combination of the PECAN parameters are shown in Figs. 4.4a–c for 50, 200 and 400 Hz source frequency data.

Using 25 and 100 Hz data, it was first concluded that 350 m sediment depth was an adequate non-local boundary position (not shown). Fig. 4.4a shows that for 50 Hz data an appropriate grid spacing is  $\Delta r = 10$  m and  $\Delta z = 1$  m and an appropriate number of Padé terms is  $N_P = 3$ , giving a mismatch of  $E \sim 6 \times 10^{-4}$ . The best mismatch achieved in the 50 Hz study is of the same order. Fig. 4.4b shows that for 200 Hz data the parameters  $\Delta r = 10$  m,  $\Delta z = 0.2$  m, and  $N_P = 3$  are acceptable, giving a mismatch of  $E \sim 2 \times 10^{-3}$ . The lowest mismatch in the 200 Hz PECAN parameter study is  $E \sim 5 \times 10^{-4}$ . For the 400 Hz data,  $\Delta r = 10$  m,  $\Delta z = 0.1$  m, and  $N_P = 5$  are acceptable, with a mismatch  $E \sim 5 \times 10^{-3}$  (Fig. 4.4c). Using 50 and 200 Hz data from the first three ranges and the above mentioned PECAN parameters, the numerical mismatch between the acoustic data fields and the replica PECAN-generated fields calculated using the TC0 true model is  $1.8 \times 10^{-3}$ . The true model mismatch is not zero because the true and replica fields were calculated using different propagation models. The effect is that the noise-free fields appear to be slightly noisy when compared to PECAN generated fields using the same model. This mismatch set the standard of how well an inversion could be expected to perform.

One minor source of error in the replica field calculations using PECAN is caused by the attenuation. PECAN requires the attenuation to be in dB/ $\lambda$ , where  $\lambda$  is a wavelength. However, for the calibration case, the true model attenuation was given in dB/m/kHz. The conversion from dB/m/kHz to dB/ $\lambda$  involves the sound speed of the sediment. For a given attenuation  $A$  (expressed in dB/m/kHz) and compressional speed  $C$  (m/s), the attenuation in dB/ $\lambda$  is  $A \cdot C/1000$ . To illustrate the problem, imagine that for inversion purposes the compressional speed in a sediment layer is allowed to vary linearly but attenuation is constrained to be constant with units of

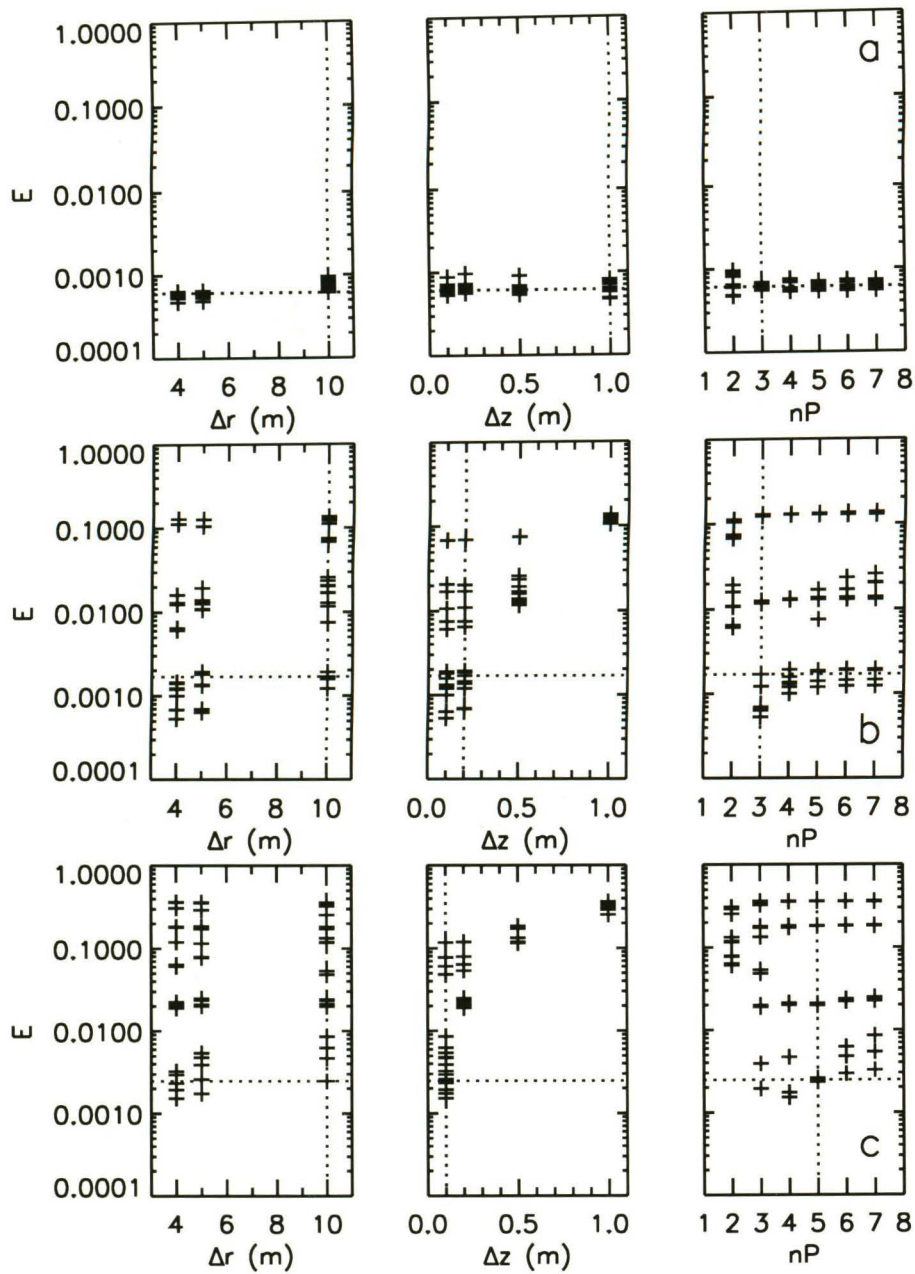


Figure 4.4 PECan parameter study. Mismatch  $E$  is plotted as a function of PECan parameters for a constant geoacoustic model. The study was carried out using (a) 50, (b) 200 and (c) 400 Hz data to determine an acceptable set of PECan parameters for the inversions. The dotted lines indicate the parameter set that was chosen and its mismatch.

dB/m/kHz. If the attenuation profile for that layer is subsequently converted to units of dB/ $\lambda$ , then the attenuation would no longer be constant. Therefore, allowing a layer to have a compressional speed and attenuation that vary linearly will result in profiles that do not have identical slopes when converted from one set of units to the other. However, rather than invert for attenuation in dB/m/kHz and have to recalculate the attenuation in dB/ $\lambda$  at every grid point for each forward model, the unknown attenuation parameters for all inversion results presented here have units dB/ $\lambda$ . It was found, and will be illustrated shortly, that changes in attenuation do not affect the mismatch as much as changes in other parameters; therefore, modeling errors in the attenuation have less importance than errors in the other parameters.

## 4.2 ASSA Results

### 4.2.1 Test Case 1

The inversion results of the TC1 parameterization analysis using ASSA are presented in this section. The control parameters were  $T_o = 1$  and  $\beta = 0.99$  for the inversion algorithm, and four accepted models at each temperature step were required. The number of temperature steps varied with each run of the algorithm due to the convergence criterion (2.1) included prior to quenching. For  $L \geq 3$  the number of temperature steps ranged between 1100 and 1300. For  $L < 3$ , the number of temperature steps was considerably reduced. The algorithm run time ranged from  $\sim 2$  hours for  $L = 1$  to  $\sim 12$  hours for  $L \geq 3$ . An additional constraint included in the ASSA algorithm required that the density at the top of each layer be less than the density at the bottom of the layer.

Fig. 4.5 shows the results of the under-parameterized approach to determining the optimal number of layers. Fig. 4.5a shows the best mismatches obtained with respect

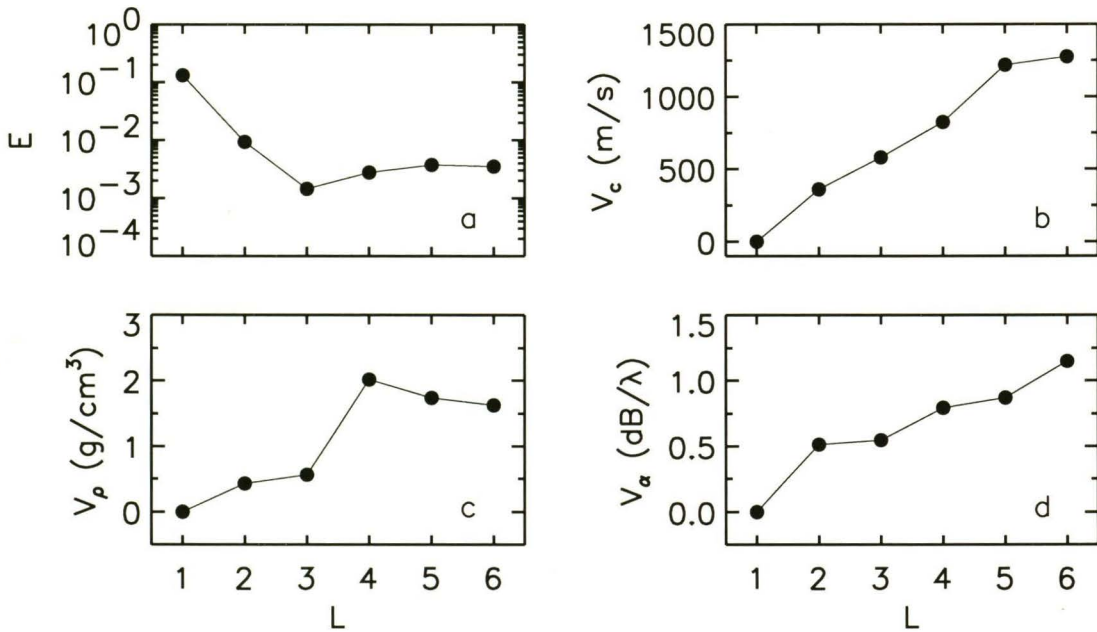


Figure 4.5 Under-parameterized approach to determining the appropriate number of sediment layers,  $L$ , for TC1: (a) the mismatch  $E$  and (b) the  $l_1$  norm variations of the compressional speed  $V_c$ , density  $V_\rho$  and attenuation  $V_\alpha$ .

to  $L$ . Mismatch  $E$  decreases as  $L$  increases to 3, then generally plateaus. Figs. 4.5b–d show the variation (2.28) of the sediment compressional speed  $V_c$ , density  $V_\rho$ , and attenuation  $V_\alpha$ , with respect to  $L$ . The variation of the parameters generally increases as  $L$  increases. The dramatic increase in  $V_\rho$  between  $L = 3$  and  $L = 4$  (Fig. 4.5c) suggests that the density is well constrained until  $L > 3$ . The plateauing of the mismatch beyond  $L = 3$  and the subsequent behavior of the parameter variations suggests that  $L = 3$  is the optimal number of layers to represent the ocean bottom.

The convergence of the ASSA algorithm for  $L = 3$  sediment layers is shown in Fig. 4.6. Fig. 4.6a shows that the mismatch of the accepted models decreases as the algorithm progresses and converges to a value of  $E = 1.4 \times 10^{-3}$ , which is lower than the true model mismatch for TC1 of  $E = 2.3 \times 10^{-3}$ . Figs. 4.6b–s show how the model parameters converge. The parameters that converge the quickest are  $h$ ,  $H$  and  $c$ , not including  $c_b$ . The density parameters do not converge as quickly as  $c$ , but they

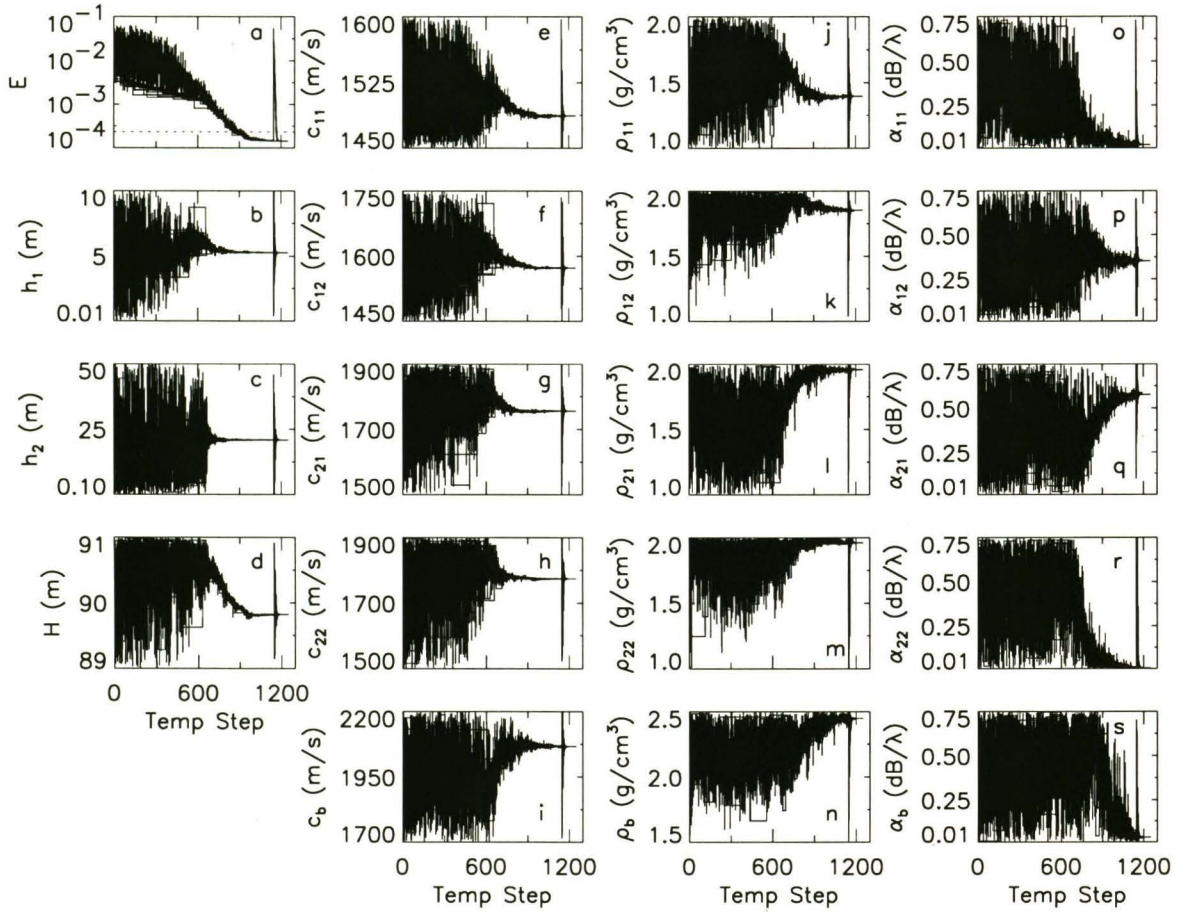


Figure 4.6 Mismatch  $E$  and parameter estimates of the  $L = 3$  ASSA inversion for TC1.

(a) Mismatch for every point in the simplex. The dotted line represents the mismatch between the given data and the replica data generated using PECan and the true model. (b)-(s) Parameter values of each model in the simplex. The parameter bounds used in the inversion are represented by the  $y$ -axis limits for each plot.

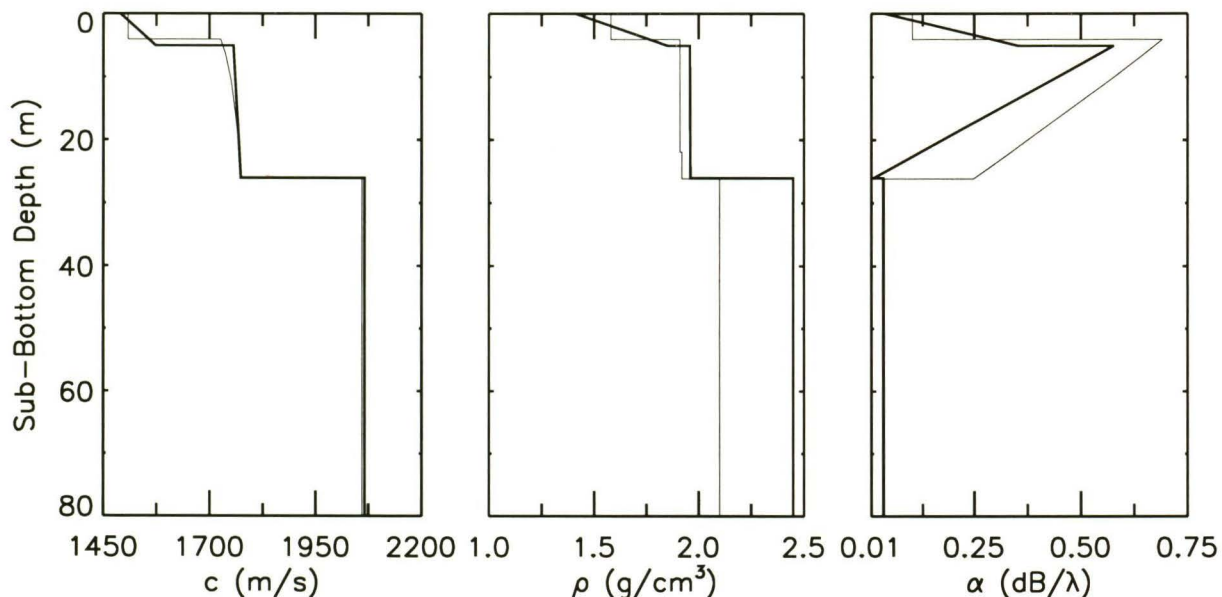


Figure 4.7 Profiles of compressional speed  $c$ , density  $\rho$ , and attenuation  $\alpha$  for TC1. The thin line represents the true model and the thicker line represents the ASSA solution for  $L = 3$ .

converge more quickly than  $\alpha$ . The parameter  $\alpha_b$  converges the slowest.

The profiles of the compressional speed, density and attenuation for the  $L = 3$  ASSA solution are shown in Fig. 4.7 along with the profiles for the true model. The compressional speed profile, including  $h_1$  and  $h_2$ , are well determined. The density profile is well approximated for the top two layers, but not as well as in the compressional speed profile. The attenuation profile is the least well determined, though it still approximates the true profile adequately.

The 1-D sensitivity plots for the  $L = 3$  model are shown in Fig. 4.8. The way the solution profiles approximate the true profiles in Fig. 4.7, relative to each other, is reflected in the relative sensitivity of the parameters determined from Fig. 4.8. The most sensitive parameters are  $h_1$ ,  $h_2$ ,  $H$  and  $c$ ; though  $c_b$  is less so. The ASSA

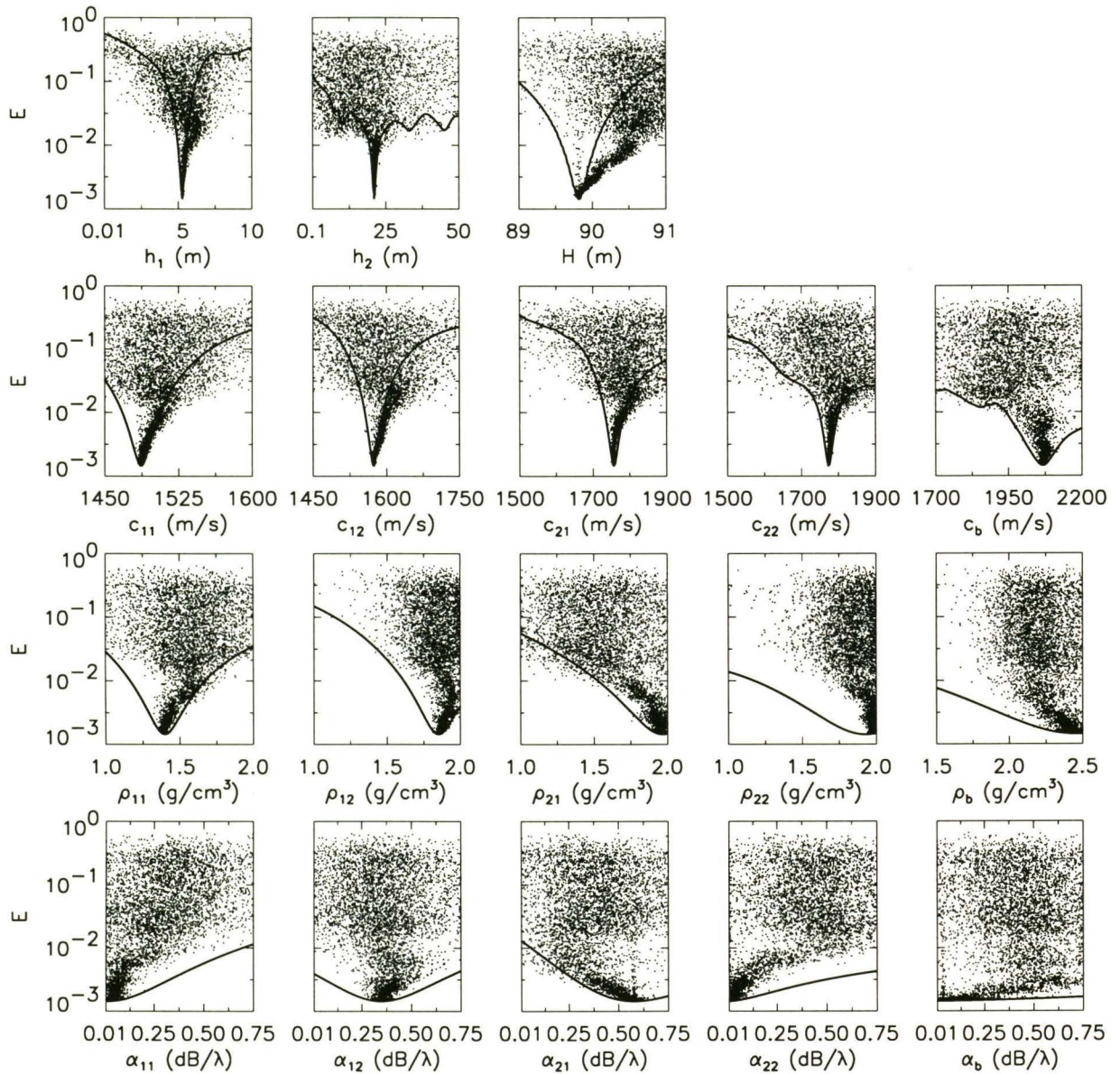


Figure 4.8 Sensitivity analysis for TC1. Solid line represents 1-D cross-sections of the objective function. Small dots represent the mismatch  $E$  for every accepted model.

algorithm required that  $\rho_{\ell 2} \geq \rho_{\ell 1}$  for  $\ell = 1, L - 1$ , which may account for the concentration of large values for  $\rho_{12}$  and  $\rho_{22}$  in the ASSA results. Sensitivity decreases from  $c$  to  $\rho$  to  $\alpha$ . The least sensitive parameter is  $\alpha_b$ ; therefore, the accuracy of the basement attenuation in Fig. 4.7 may be coincidental.

Figs. 4.9a and b show profiles of the sub-optimal results generated using  $L = 2$  and  $L = 4$  sediment layers during the parameterization analysis. The  $L = 2$  solution is too simplistic a model that does not utilize all the information available in the acoustic data. The  $L = 4$  model has excess structure that the acoustic data cannot resolve and does not approximate the profiles for the upper 25 m as well as the  $L = 3$  model.

One measure of the general applicability of a geoacoustic model is how well it fits data that were not included in the inversion. The ASSA inversions used data for three ranges out to 1.5 km. Using the  $L = 3$  solution, Fig. 4.10 shows the mismatches calculated for 50 and 200 Hz and ten ranges out to 5 km. The mismatches at both frequencies are typically at or below the best mismatch found through the ASSA inversion. At long ranges, the 200 Hz replica data have mismatches that are repeatedly an order of magnitude lower than the best mismatch value. In general, the  $L = 3$  solution predicts the data well, relative to previous results.

One way to improve a geoacoustic model is to invert more data, particularly data extending over a greater frequency bandwidth. A model estimation can be improved because with an increase in the number of frequencies, the information content increases, and therefore the optimization solution is better constrained. However, computation time can increase prohibitively when higher frequencies are added. Fig. 4.11 shows the profiles for a three-frequency ASSA inversion. Data from the first three ranges with source frequencies of 50, 200, and 400 Hz were used. For 400 Hz data, the values of  $\Delta r$ ,  $\Delta z$  and  $N_P$  were as stated earlier. The inversion took  $\sim 40$  hours and achieved a mismatch  $E = 1.3 \times 10^{-3}$ . The agreement of the top layer improved considerably compared to the two-frequency model (Fig. 4.7). The compressional

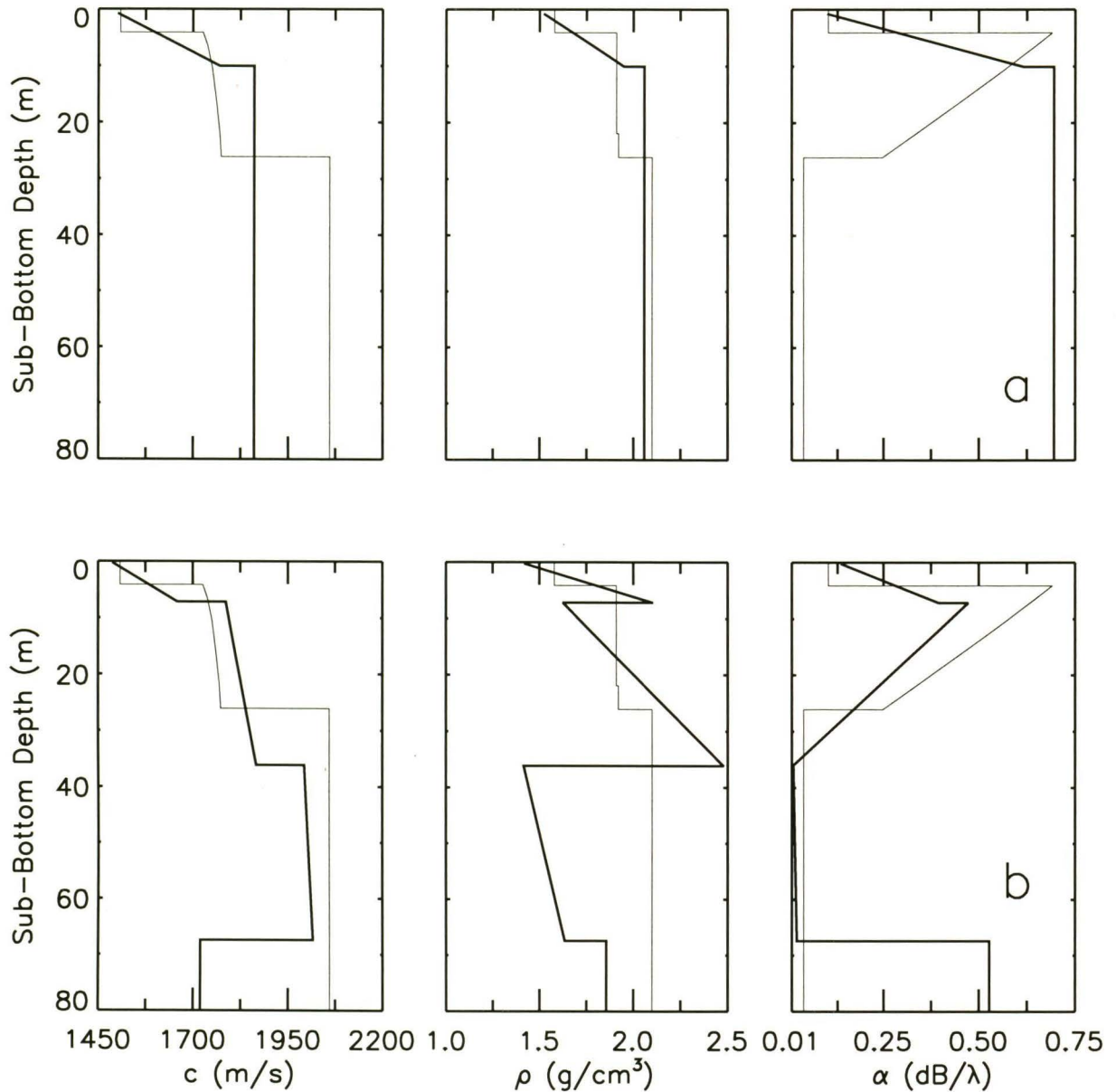


Figure 4.9 Sub-optimal profiles of compressional speed  $c$ , density  $\rho$ , and attenuation  $\alpha$ , for the TC1. The thin line represents the true model and the thicker line represents the ASSA solution for (a)  $L = 2$  and (b)  $L = 4$  sediment layers.

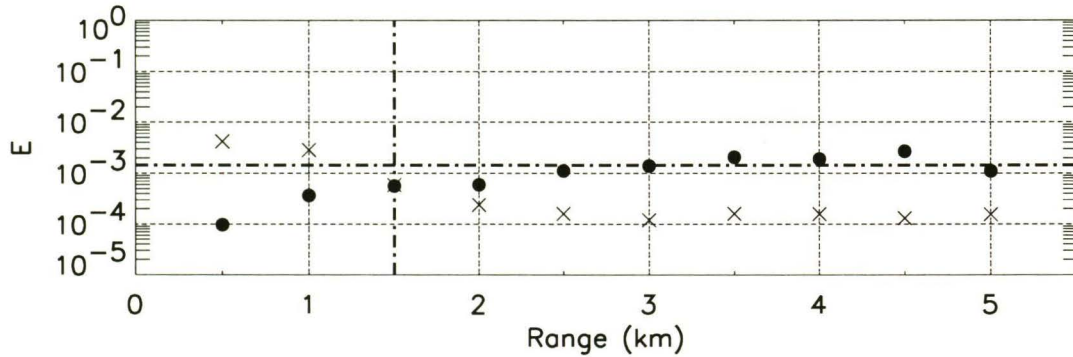


Figure 4.10 Bartlett mismatches calculated for 50 (•) and 200 (×) Hz data for the 10 ranges, using the TC1  $L = 3$  solution. The horizontal dashed-dotted line represents the mismatch found for the model in the ASSA inversion. The vertical dashed-dotted line is at the limiting data range used in the inversion.

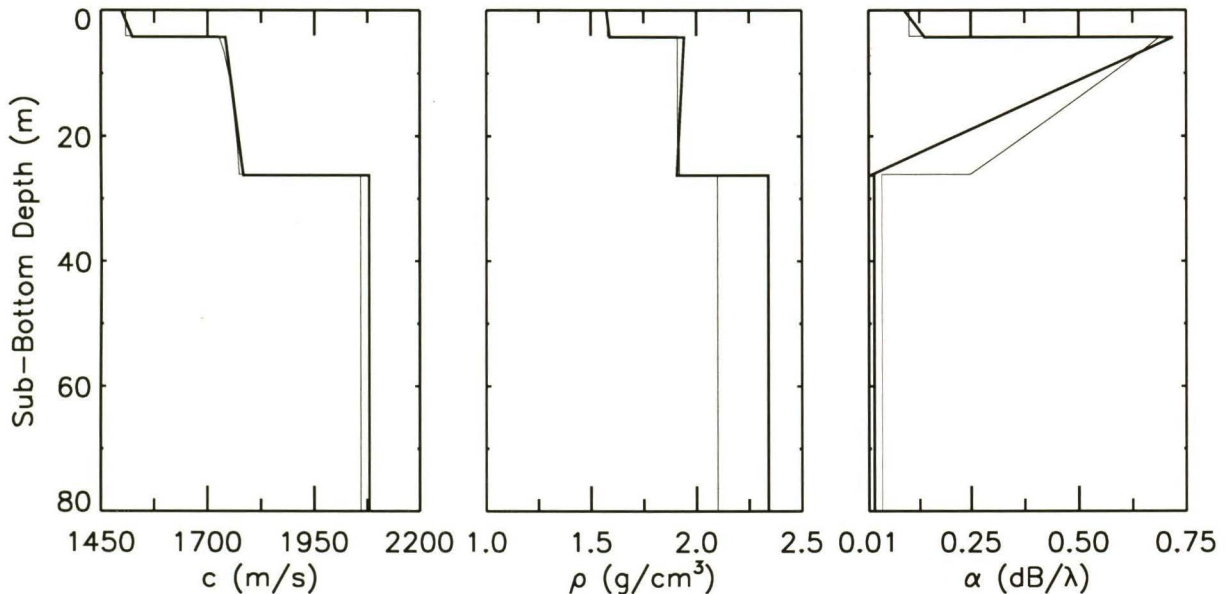


Figure 4.11 Three frequency profiles of compressional speed  $c$ , density  $\rho$ , and attenuation  $\alpha$  for TC1. The thin line represents the true model and the thicker line represents the ASSA solution for  $L = 3$  with 50, 200 and 400 Hz data.

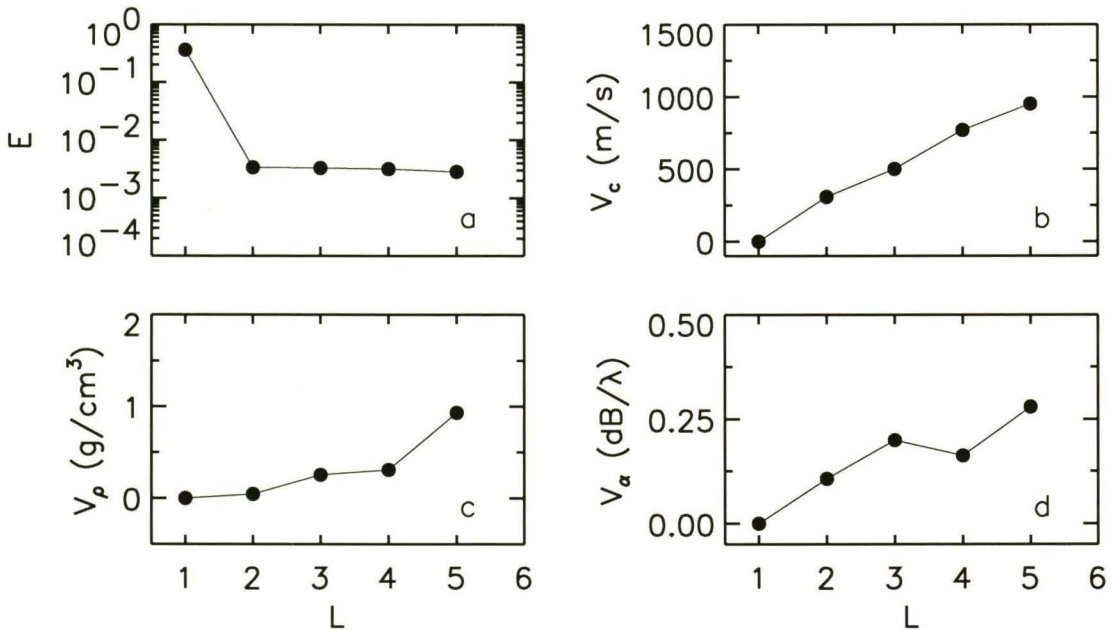


Figure 4.12 Under-parameterized approach to determining the appropriate number of sediment layers,  $L$ , for TC2: (a) the mismatch  $E$  and (b) the  $l_1$  norm variations of the compressional speed,  $V_c$ , density  $V_\rho$  and attenuation  $V_\alpha$ .

attenuation profiles also approximate the true profiles well.

The 1-D sensitivity plots are shown in Fig. 4.14. The most sensitive parameters are  $h_1$ ,  $c_{11}$  and  $c_{12}$  and therefore these should be determined the best of all the parameters. Those parameters are determined well in Fig. 4.13, however  $c_b$  is also very well determined. Sensitivity decreases from  $c$  to  $\rho$  to  $\alpha$ . Again,  $\rho_{12}$  was required to be less than  $\rho_{11}$ , which influenced the ASSA parameter distribution for  $\rho_{12}$ . The difference in sensitivity between  $\rho$  and  $\alpha$  is not observable in Fig. 4.13.

To determine how well the  $L = 2$  solution can fit data not used in the inversions, Fig. 4.15 shows the mismatches calculated for 50 and 200 Hz source frequency data at ten ranges. The 200 and 50 Hz mismatches are typically above and below the mismatch of the  $L = 2$  solution, respectively. Relative to the ASSA solution mismatch, the model predicts the 50 Hz data well, however the 200 Hz data is not predicted as well.

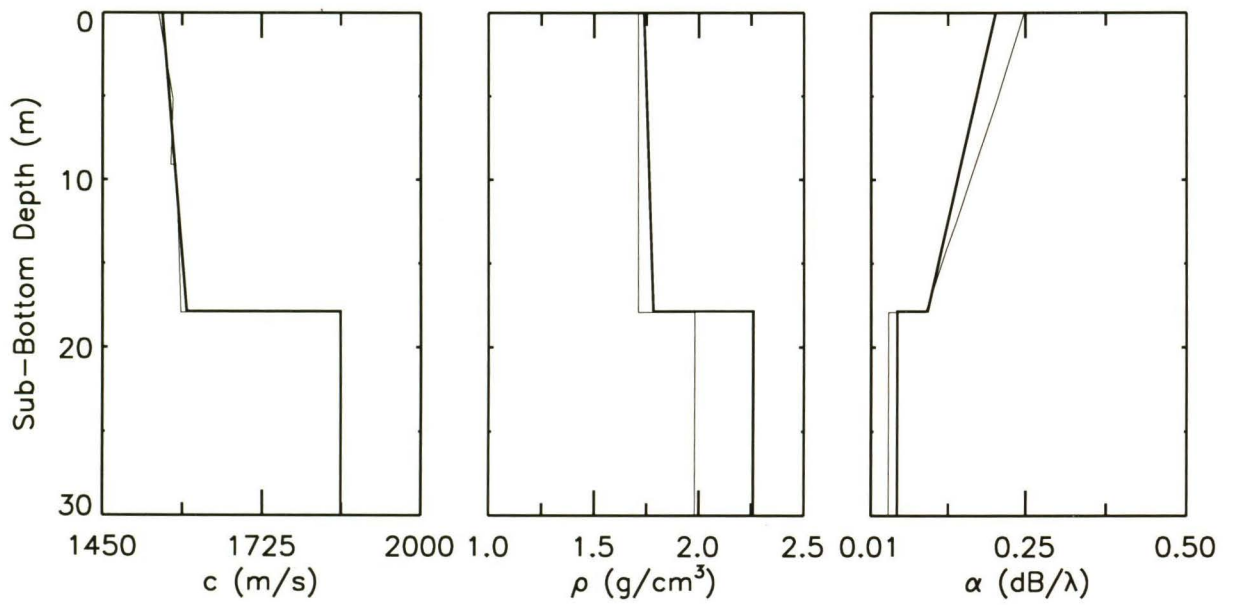


Figure 4.13 Profiles of compressional speed  $c$ , density  $\rho$ , and attenuation  $\alpha$  for TC2. The thin line represents the true model and the thicker line represents the ASSA solution for  $L = 2$  sediment layers.

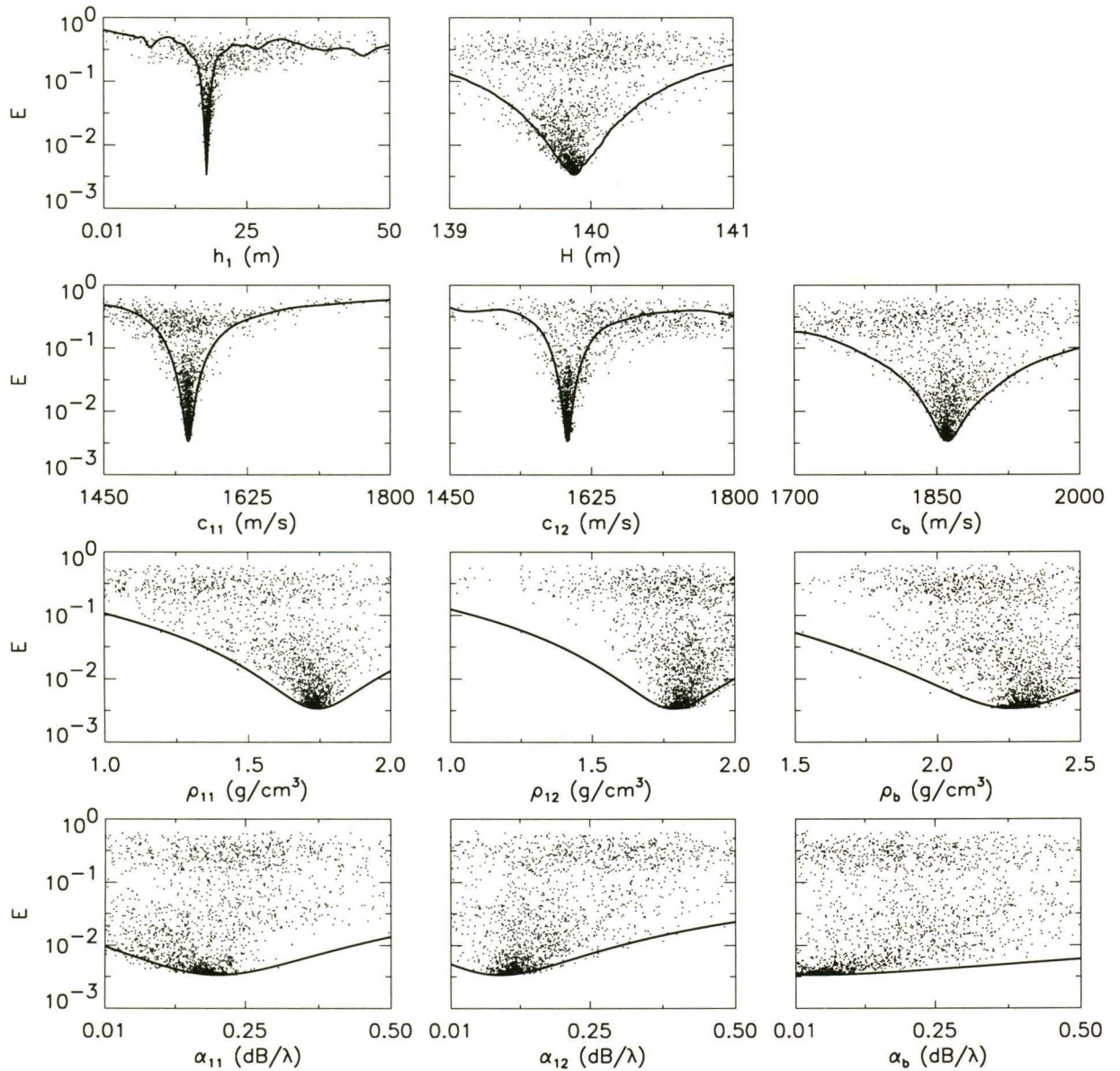


Figure 4.14 Sensitivity analysis for TC2 and  $L = 2$ . Solid line represents 1-D cross-sections of the objective function. Small dots represent the mismatch  $E$  for every accepted model. The  $x$ -axis bounds represent the parameter bounds used in the inversion.

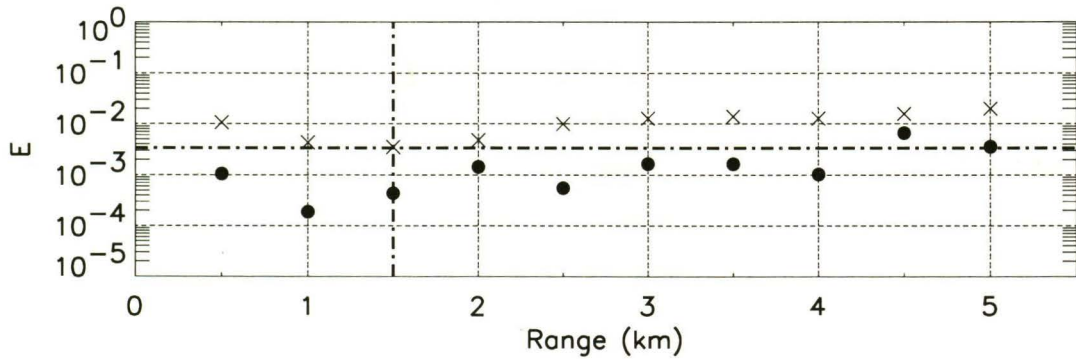


Figure 4.15 Bartlett mismatches calculated for 50 (•) and 200 (×) Hz data for the ten ranges, using the TC2  $L = 2$  solution. The horizontal dashed-dotted line represents the mismatch found for the model in the ASSA inversion. The vertical dashed-dotted line is at the limiting data range used in the inversion.

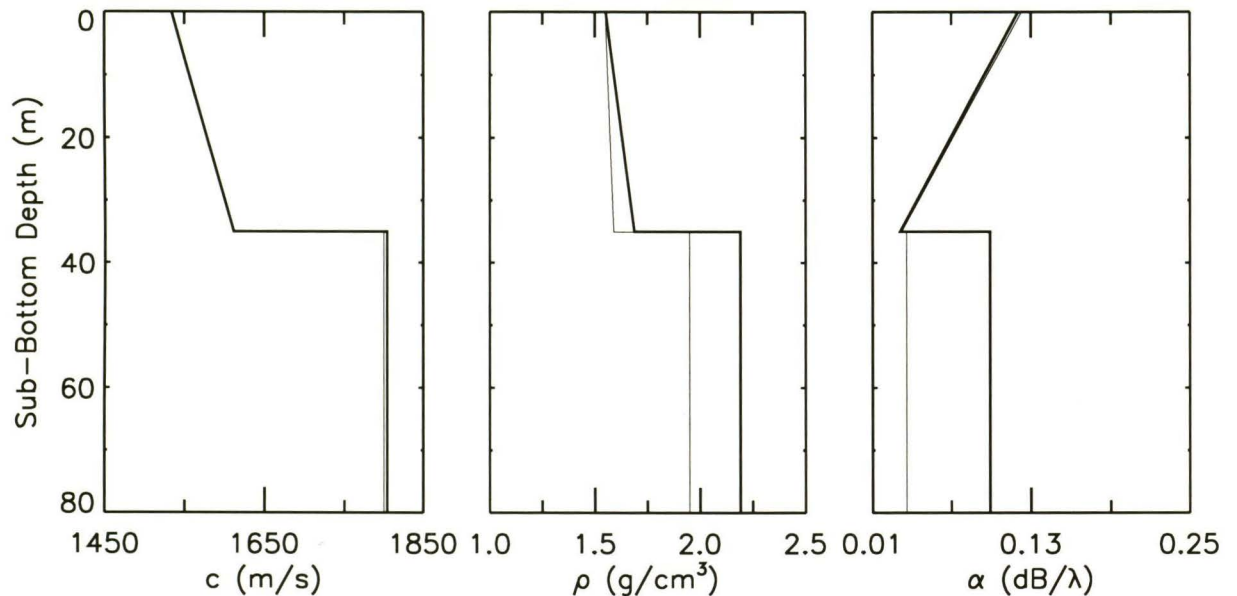


Figure 4.16 Profiles of compressional speed  $c$ , density  $\rho$ , and attenuation  $\alpha$  for the TC0 solution given  $L = 2$  sediment layers. The thin line represents the true model and the thicker line represents the ASSA solution.

ASSA was also applied to the TC0 data. The parameterization of the TC0 ocean bottom environment was known prior to the workshop and included two sediment layers, so a model parameterization analysis was not performed. The ASSA control parameters included  $T_o = 0.3$  and  $\beta = 0.99$ . Four perturbations were required to be accepted at each of the 800 temperature steps. It took  $\sim 10$  hours of computing time to complete the algorithm. Fig 4.16 shows the parameter profiles for the solution found through ASSA. To a depth of about 30 m, the profiles predict the true profiles very well. The entire compressional speed profile is virtually identical to the true profile. The model shown has a mismatch of  $E = 9.2 \times 10^{-4}$  which is lower than the true model mismatch of  $E = 1.8 \times 10^{-3}$ .

### 4.3 FGS Results

In this section, the results of applying the FGS to the same data used in Sec. 4.2 are presented. The FGS (Sec. 2.2.4) is a powerful but computationally intensive inversion technique that provides a wealth of statistical information. The FGS does not lend itself to being applied to determine the optimal number of layers for the test cases. Therefore, the FGS inversion technique was applied to the workshop data using the optimal parameterizations found in Sec. 4.2. In addition, since a good inversion result was available from the ASSA inversion, the FGS cooling stage (Sec. 2.2.4) was omitted and the algorithm was initiated at  $T = 1$  with the ASSA solution. This ensures that the algorithm starts in a region of likely models and it reduces the computation time. In essence, ASSA is applied as the cooling stage. The initial maximum perturbation size was set to 0.002 times the width of the parameter bounds. As mentioned in Sec. 2.2.4, the maximum perturbation is allowed to grow by continually replacing the maximum perturbation size with  $k$  times the most recently accepted perturbation, if the latter is larger than the former.

The incoherent processor used as the error function in the FGS is different than the one used in ASSA. To get a proper PPD estimate from the FGS analysis it could no longer be assumed that the data variances are approximately equal. Also, based on early results (not shown), the assumption that there are independent errors at each VLA sensor is not appropriate here. Therefore, Eq. (2.26) was used to calculate the variance estimates  $\hat{\nu}_{fa}$  and Eq. (2.27) was used as the FGS error function. To evaluate Eq. (2.26),  $\hat{\mathbf{m}}$  should be the solution found by minimizing the objective function given in Eq. (2.25) for unknown data variances. However, for efficiency, the ASSA solutions found earlier were used instead, which should have negligible effect. Finally, it should be noted that the Workshop test cases are somewhat artificial in that no noise was added. Hence, the data errors that limit the accuracy of inversion are essentially the numerical errors of the propagation models. The actual distribution of these errors is unknown, and may affect the accuracy of the Bayesian approach.

#### 4.3.1 Test Case 1

To evaluate Eq. (2.26) for the TC1 data, the effective mode number,  $N_e$ , had to be determined with respect to each frequency through a principal component analysis. The analysis was performed on a data covariance matrix computed using the acoustic data from all ten ranges. The data covariance matrix is composed of the covariances between data sets with respect to range. The maximum number of principal components is ten ( $N_{pc} = 10$ ) because ten data sets were used per frequency. The principal components of the matrix and their affiliated energies are eigenvectors and eigenvalues, respectively. The number of significant eigenvalues indicates  $N_e$  because the set of principal components and the set of propagating modes are both orthogonal bases for the acoustic data (Neilson *et al.* 1997). The eigenvalues of the data covariance matrix are obtained through an eigenvector decomposition of the matrix.

The number of important modes can be determined by analyzing the eigenvalues.

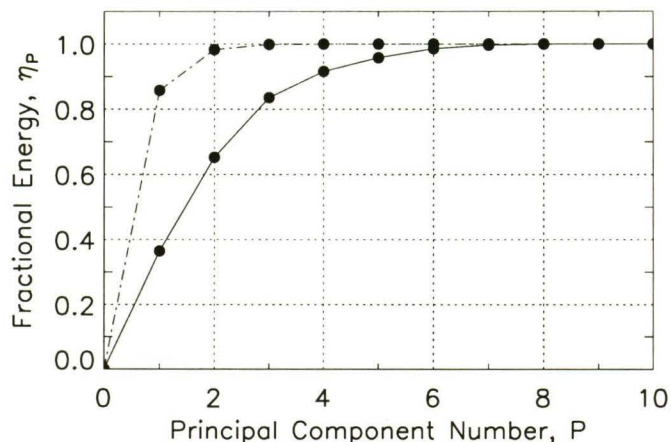


Figure 4.17 Cumulative fractional energy  $\eta_P$  (4.3) for the eigenvector decomposition of the TC1 data covariance matrix as a function of principal component number,  $P$ . The dashed-dotted line represents  $\eta_P$  for the 50 Hz data and the solid line represents  $\eta_P$  for the 200 Hz data.

For the  $P^{\text{th}}$  principal component, the cumulative fractional energy,  $\eta_P$ , is defined

$$\eta_P = \frac{\sum_{k=1}^P \mu_k}{\sum_{k=1}^{N_{pc}} \mu_k}, \quad (4.3)$$

where  $\mu_k$  is the  $k^{\text{th}}$  largest eigenvalue and  $\eta_P$  is the fraction of acoustic energy contained in the first  $P$  principal components. Fig. 4.17 shows  $\eta_P$  plotted for the TC1 environment. Based on this figure,  $N_e = 3$  and  $N_e = 7$  were chosen for the 50 and 200 Hz data, respectively, because these account for a high percentage of the acoustic energy. Appropriate choices also include  $N_e = 2$  and  $N_e = 6$ , respectively.

The FGS was applied to the TC1 data, assuming an  $L = 3$  model. The FGS algorithm generated two independent samples. During the run, the parameter space axes were rotated twice: after  $2 \cdot 18 \cdot 195 \approx 7 \times 10^3$  (i.e., (the number of independent samples)  $\times$  (the number of model parameters)  $\times$  (the number of cycles through the parameters)) and after  $2 \cdot 18 \cdot 670 \approx 2 \times 10^4$  models had been accepted. It took  $\sim 12$  days

of computation to accept  $\sim 5 \times 10^4$  models and for the threshold of  $\epsilon_2 = 0.1$  to be achieved. It was found that the run time of the algorithm could vary substantially due to the random nature of the independent samples collected. For the difficult problem of TC1, rotating the model space more than once did seem to improve efficiency.

The PPD represents the general Bayesian solution to the inverse problem, however for an inversion problem with more than two parameters, the PPD cannot be visually interpreted. One method that can be used to interpret the estimated PPD is to generate the marginal probability distributions for each parameter (Eq. 2.13). Analysis of the PPD can result in quantitative predictions of parameter uncertainties which are evident in the marginal distributions. Fig. 4.18 shows the estimated marginal probability distributions for TC1. To avoid dependence on the starting model, all models collected prior to the first rotation were removed from the sample. The structure of the marginal distribution for the  $h_1$  distribution implies that two values of  $h_1$  are preferred. In addition, the lower layer compressional speeds have smaller distribution widths than  $c_{11}$  and  $c_{12}$ , likely because the acoustic data used lacks information about the thin top layer. Including data at a higher frequency, such as 400 Hz, would help better resolve the top layer parameters, but is too computationally intensive.

The distribution widths quantify the uncertainty of the parameters. The smallest interval of each marginal distribution which contains 95% of the samples defines the 95% highest posterior density (HPD) interval (e.g., Gamerman 1997). Fig. 4.19 shows the parameter profiles of the MAP solution and the schematic representations of the 95% HPD intervals. Because the layer thickness and water depth at the source are parameters, the plot of the 95% HPD intervals can only be approximated. The compressional speed and layer thicknesses are well determined parameters. For the compressional speed, the HPD interval bounds below the first layer are quite narrow. The top layer has HPD interval bounds that are slightly wider because these parameters are not as well determined. The compressional speed profile is also a good representation of the true profile. The HPD interval bounds for the density

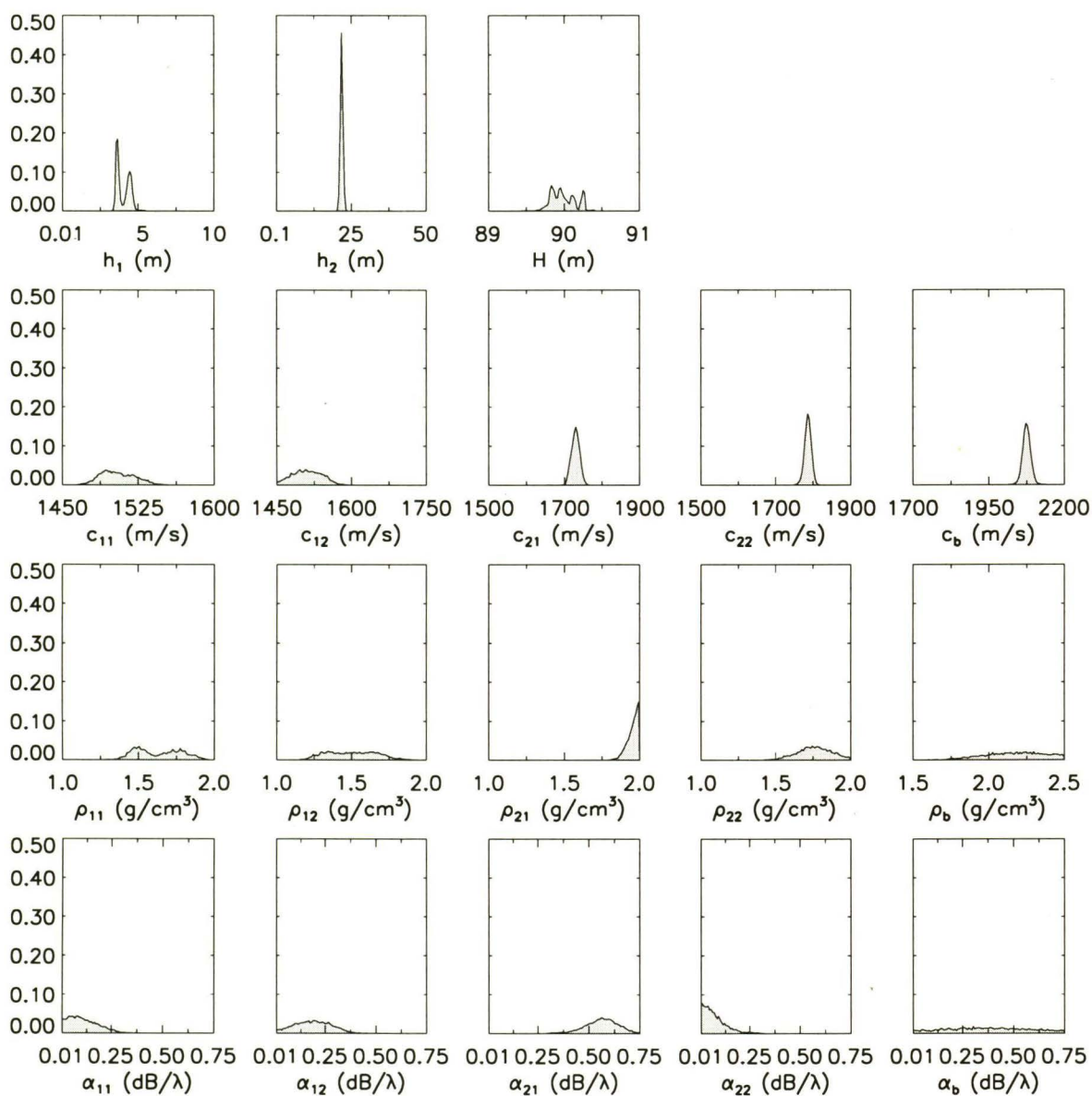


Figure 4.18 Estimated marginal probability distributions for TC1 and  $L = 3$ . The  $x$ -axis limits represent the bounds used in the inversion.

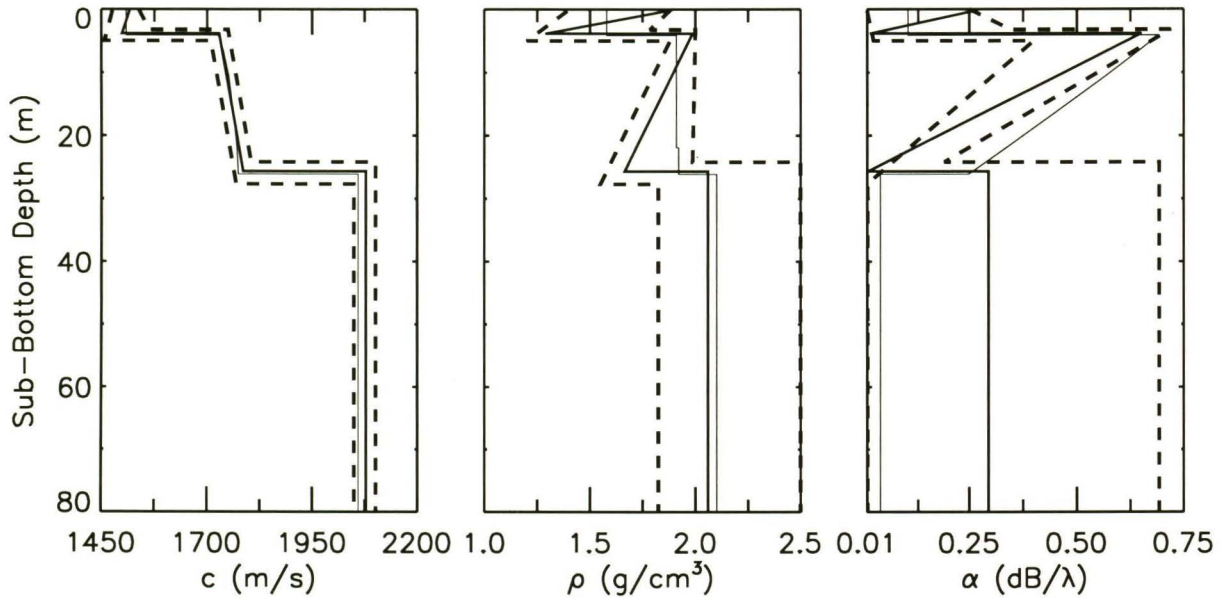


Figure 4.19 FGS MAP profiles of compressional speed  $c$ , density  $\rho$ , and attenuation  $\alpha$  for TC1. The thin line represents the true model and the thicker line represents the MAP. The dashed lines represent 95% HPD interval bounds.

and attenuation indicate that the parameters cannot be as well determined as for the compressional speed, especially for the basement attenuation.

For ASSA, the sensitivity plots provide substantially less information than the marginal distributions. However, to assess the ASSA relative sensitivity analyses carried out earlier, a relative sensitivity analysis can be performed using the marginal distributions. For TC1, it is found that the relative sensitivities predicted from Fig. 4.8 only reflect some of the features of the relative sensitivities that can be estimated using Fig. 4.18. One notable difference is that Fig. 4.8 does not indicate that the compressional speed parameters for the top layer are less resolvable than the compressional speed parameters of the lower layers.

The ASSA solution and the FGS MAP estimate are generally comparable and represent the true model well. The FGS MAP determines the top layer thickness

better than the two-frequency ASSA solution shown in Fig. 4.7. The layer thickness is determined as well as in the three-frequency ASSA solution (Fig. 4.11), for which 400 Hz data was added. The density profile is quite different than the ASSA profile in Fig. 4.7. There was no stipulation in the FGS algorithm, unlike in the ASSA algorithm, that the density at the top of a layer had to be lower than the density at the bottom of a layer. To quantitatively compare the FGS MAP with the ASSA solution and the true model, the mismatch of the FGS MAP was calculated using the ASSA error function (Eq. 4.2). Using the ASSA error function to calculate the mismatch means that the assumption of constant variance with range and frequency is being made. Since the FGS did not assume a constant variance, calculating the ASSA mismatch for the FGS MAP is only an approximate way to compare the two models. The mismatch for the FGS MAP was calculated to be  $E = 8.7 \times 10^{-4}$ , which is better than the two-frequency ASSA solution ( $E = 1.3 \times 10^{-3}$ ).

Marginal probability distributions and confidence intervals provide uncertainty estimates for particular parameters with the effects of all other parameters integrated out. To investigate inter-parameter relationships, the covariance or correlation matrix can be examined. The correlation between parameters  $m_i$  and  $m_j$  is defined as  $R_{ij}$  where

$$R_{ij} = \frac{C_{ij}}{\sqrt{C_{ii} C_{jj}}}, \quad (4.4)$$

and  $C_{ij}$  is element  $(i, j)$  in the model covariance matrix  $\mathbf{C}_M$ . A correlation of 1 implies perfect correlation, as found between a parameter and itself. A correlation of  $-1$  implies perfect anti-correlation and a correlation of 0 implies no correlation. Fig. 4.20 shows the rows of the model correlation matrix,  $\mathbf{R}_M$ . The parameter  $h_1$  is strongly correlated with  $c_{12}$ ,  $c_{21}$  and  $\rho_{12}$ , and strongly anti-correlated with  $h_2$ ,  $H$ ,  $c_{11}$ , and  $\rho_{11}$ . The parameter  $h_2$  is strongly correlated with  $c_{22}$  and strongly anti-correlated with  $c_{12}$  and  $c_{21}$ . The parameter  $H$  is strongly correlated with  $c_{11}$  and  $\rho_{11}$ . In addition,  $H$  and  $c_{12}$ ,  $c_{11}$  and  $c_{12}$ ,  $c_{12}$  and  $\rho_{11}$ ,  $c_{21}$  and  $c_{22}$ ,  $\rho_{11}$  and  $\rho_{12}$ ,  $\alpha_{11}$

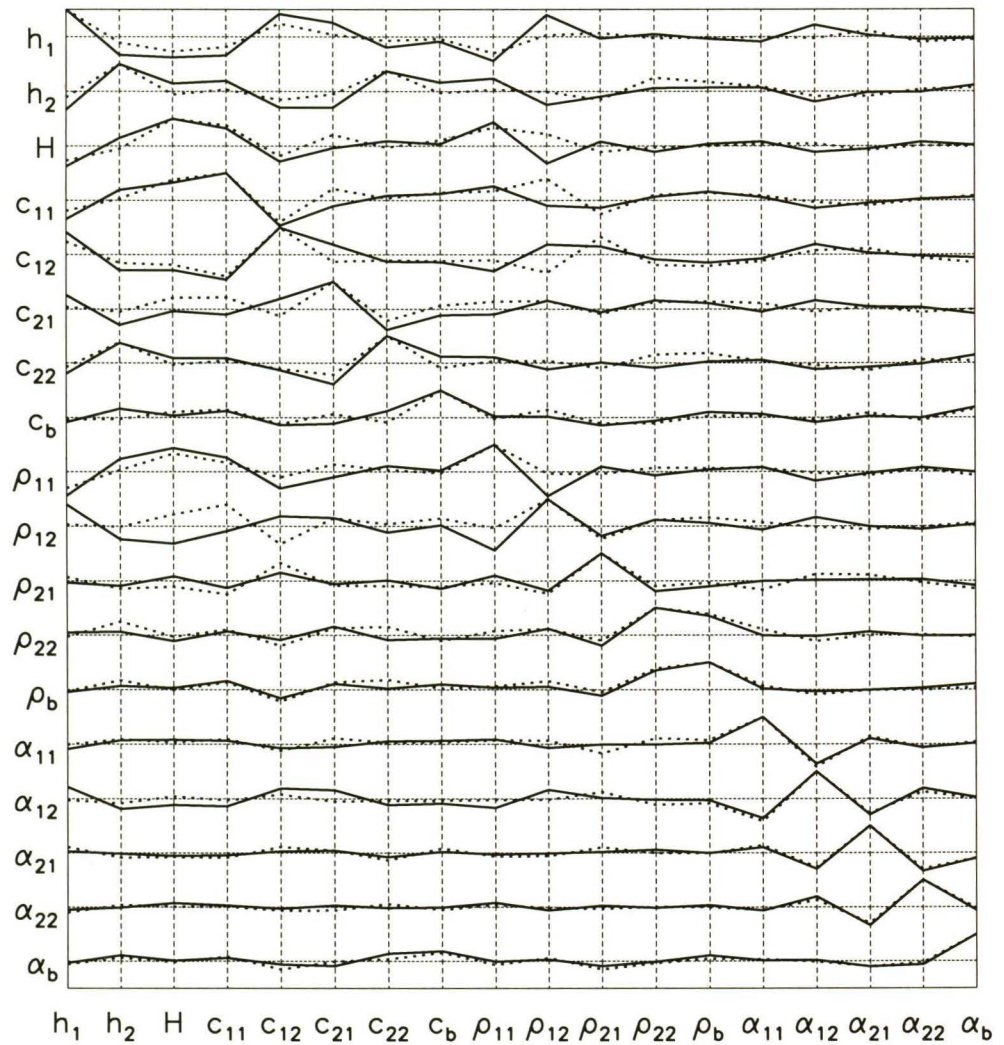


Figure 4.20 Rows of the TC1 model correlation matrix with respect to the parameters. The solid line represents the correlation matrix at the conclusion of the algorithm and the dotted line represents the correlation matrix used to rotate the parameter space axes for the first time.

and  $\alpha_{12}$ ,  $\alpha_{12}$  and  $\alpha_{21}$ , and  $\alpha_{21}$  and  $\alpha_{22}$  are all strongly anti-correlated, and  $\rho_{22}$  and  $\rho_b$  are strongly correlated. Some of these correlations can be understood with simple physical explanations. For instance, the compressional speed at the top and bottom of each layer are anti-correlated with each other because the acoustic travel time through the layer must remain constant. To conserve the amount of attenuation in each layer, the attenuation at the top of each layer is anti-correlated with the attenuation at the bottom of each layer. In addition, the first layer thickness is strongly anti-correlated with second layer thickness and the water depth, suggesting that the depth at which the basement begins is being conserved. Even though the parameter correlations are weighted averages according to Eq. (2.13), Fig. 4.20 still implies that the objective function for TC1 is a complicated function with many parameter correlations. The model covariance matrix used for the first rotation, the dotted line in Fig. 4.20, looks similar to the matrix at the conclusion of the algorithm, with some significant differences, particularly for parameter  $\rho_{12}$ . Hence, it was wise to re-rotate the model space.

#### 4.3.2 Test Case 2 and Calibration Case

The results of applying the FGS to TC2 data and then to TC0 data will now be presented. Before applying the FGS to TC2 data,  $N_e$  had to be estimated. Fig. 4.21 shows the cumulative fractional energies resulting from the TC2 principal component analysis. Based on this figure,  $N_e = 2$  for 50 Hz data and  $N_e = 6$  for 200 Hz data were picked. The FGS was applied to the TC2 data twice assuming an  $L = 2$  model. Each application reached the threshold of  $\epsilon_2 = 0.1$ . The first application took  $\sim 5$  days of computing time with  $2 \cdot 11 \cdot 300 \approx 7 \times 10^3$  models accepted prior to the rotation of the model space and  $\sim 2 \times 10^4$  models accepted at the completion of the algorithm. The second application took  $\sim 10$  days of computing time, with  $2 \cdot 11 \cdot 200 \approx 4 \times 10^3$  models accepted prior to rotating the model space and  $\sim 4 \times 10^4$  models accepted in

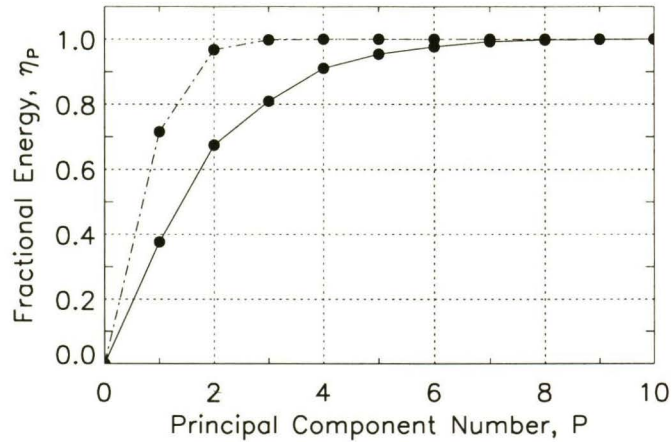


Figure 4.21 Cumulative fractional energy  $\eta_P$  (4.3) for the eigenvector decomposition of the TC2 data covariance matrix as a function of principal component number,  $P$ . The dashed-dotted line represents  $\eta_P$  for the 50 Hz data and the solid line represents  $\eta_P$  for the 200 Hz data.

total. The factor of 2 difference between the accepted models of the first and second run could be due to rotating the model space too early in the second run or it could be coincidental.

The estimated marginal probability distributions of the two independent applications are very similar (not shown), but for the best estimate of the distributions the samples collected from both applications were combined. Fig. 4.22 shows the marginal probability distributions generated from the accepted models of both runs. The layer thickness and compressional speed parameters have very narrow distributions and are very well determined parameters. The density parameters are also well determined, though not as well as compressional speed and layer thickness. The uncertainty in density also increases with depth. The water depth distribution is approximately half the width of its TC1 counterpart. It has two distinct peaks: one at the value the FGS samples were initiated with (i.e., the ASSA solution value) and one at the true

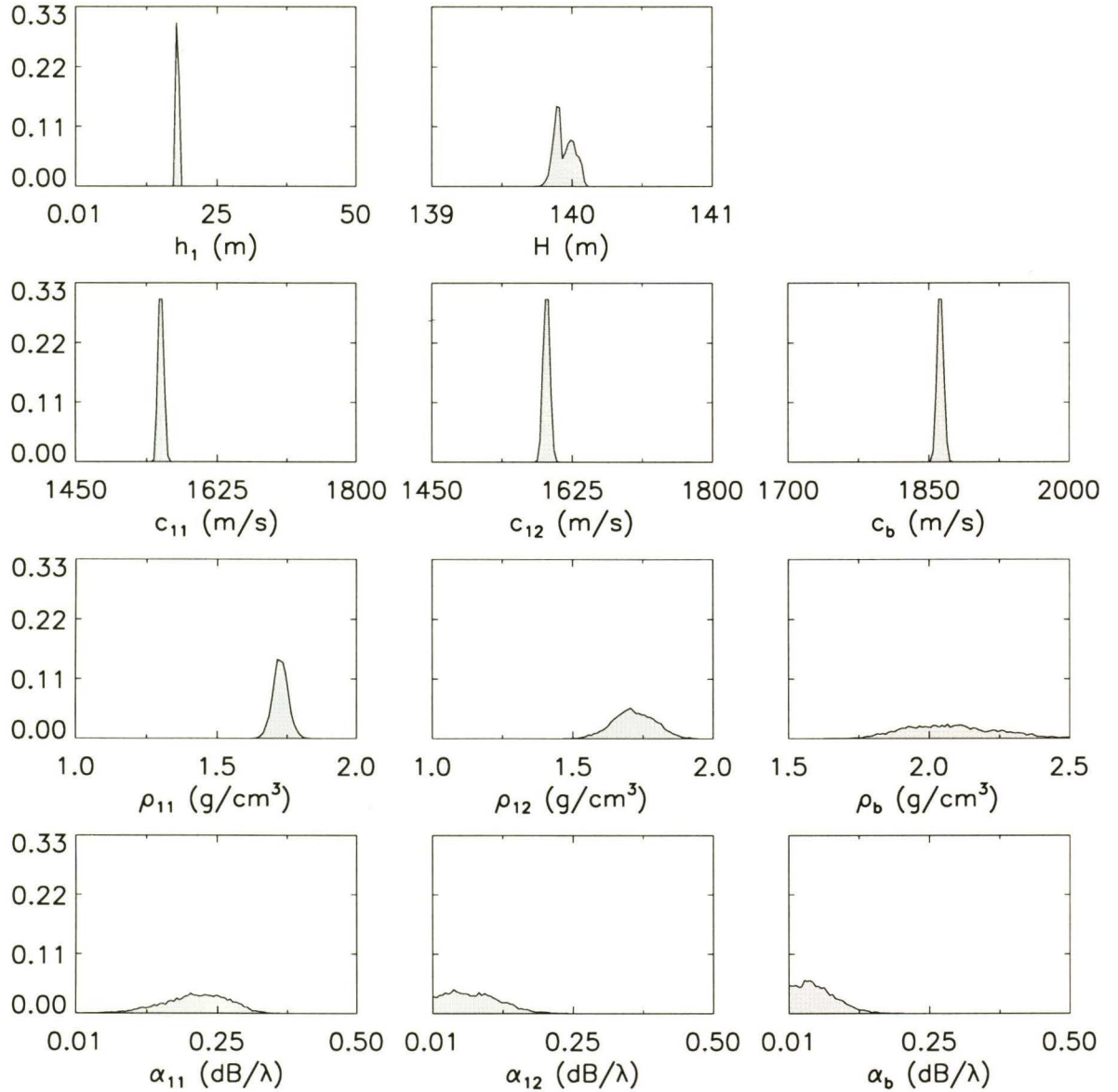


Figure 4.22 Estimated marginal probability distributions for TC2 and  $L = 2$ . The  $x$ -axis limits represent the bounds used in the inversion.

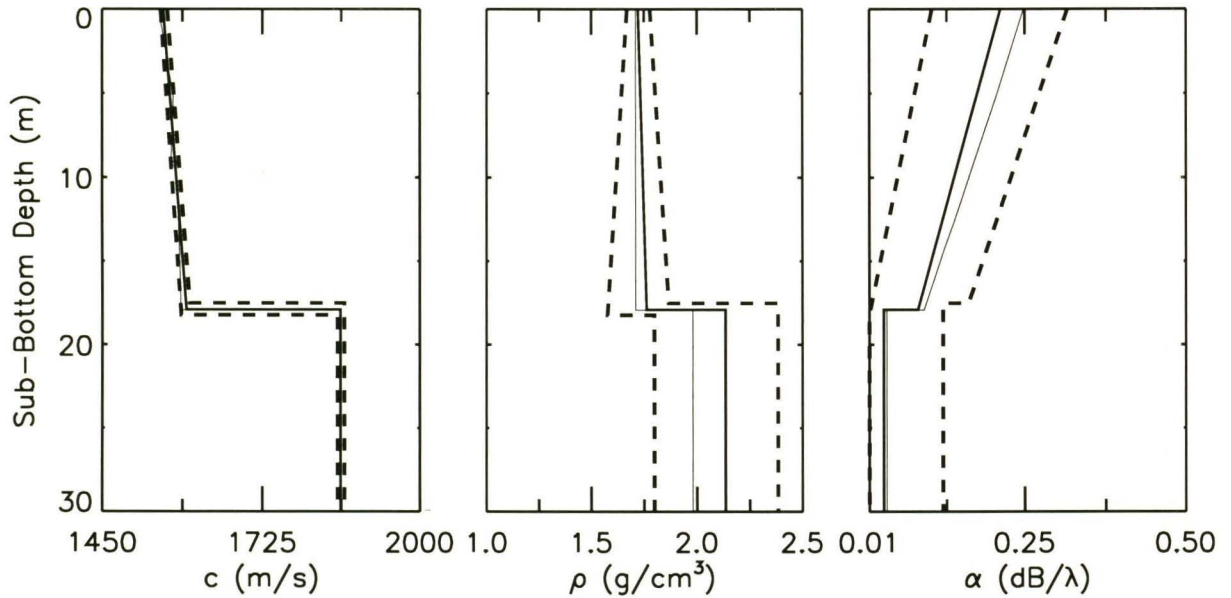


Figure 4.23 FGS MAP profiles of compressional speed  $c$ , density  $\rho$ , and attenuation  $\alpha$  for TC2. The thin line represents the true model and the thicker line represents the MAP. The dashed lines represent the 95% HPD interval bounds.

value (i.e.,  $H = 140$  m). To assess the ASSA relative sensitivity analysis (Fig. 4.14), a similar analysis can be done using Fig. 4.22. The relative sensitivity trends seen in Fig. 4.22 are comparable to the trends seen in Fig. 4.14, with the exception of  $\alpha_b$ .

To quantify the parameter uncertainties, Fig. 4.23 shows the schematic representation of the 95% HPD intervals that are determined from the marginal distributions. The 95% HPD intervals for the layer thickness and the compressional speed parameters are much narrower for TC2 than for TC1 (Fig. 4.19), implying that those properties can be determined better for TC2. The density and attenuation parameters generally have 95% HPD intervals comparable to the narrower intervals found for those properties in TC1. However, attenuation is better constrained for TC2 than for TC1.

The parameter profiles of the FGS MAP estimate are also shown in Fig. 4.23. The

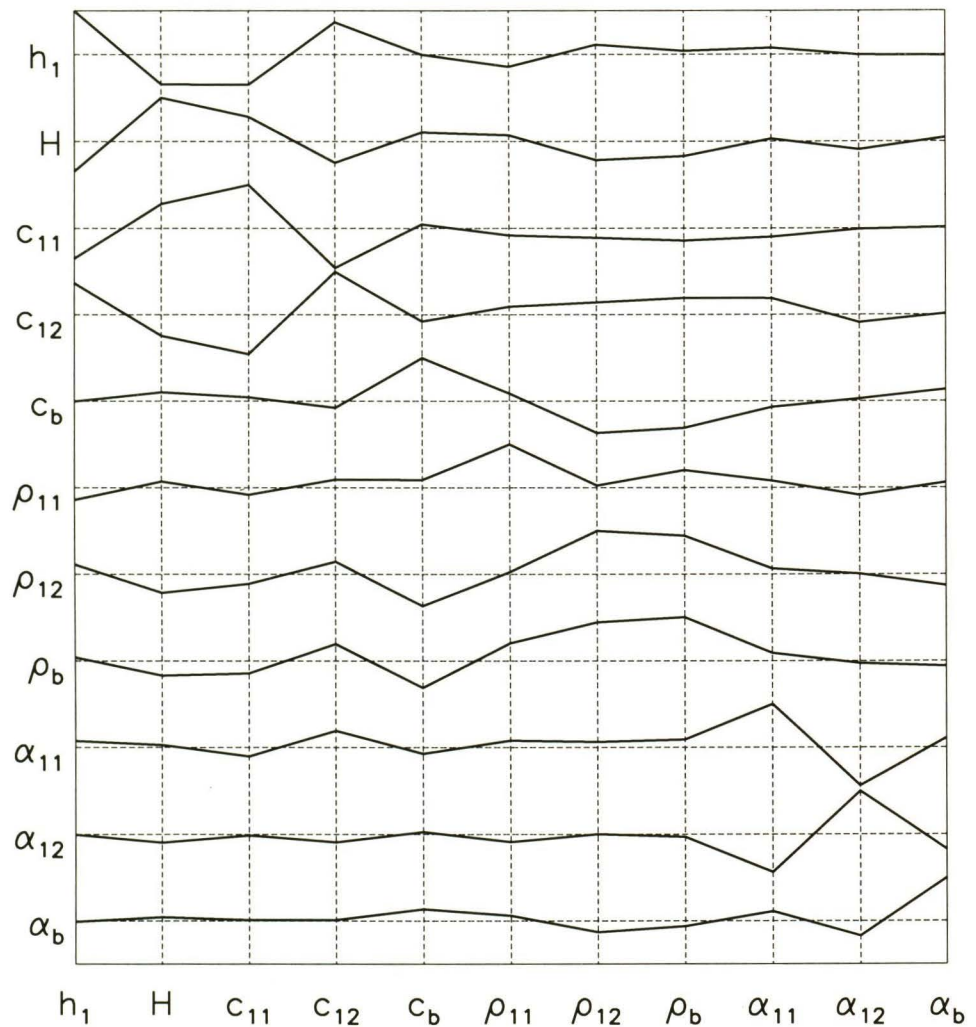


Figure 4.24 Rows of the TC2 model correlation matrix.

FGS MAP solution is a very good estimate of the true solution. The compressional speed profile is virtually identical to the true profile. The MAP estimate is also similar to the ASSA solution (Fig. 4.13). The MAP estimate has a mismatch of  $E = 3.5 \times 10^{-3}$ , with respect to the ASSA error function.

The correlations for TC2 are less complicated than for TC1, partly due to the

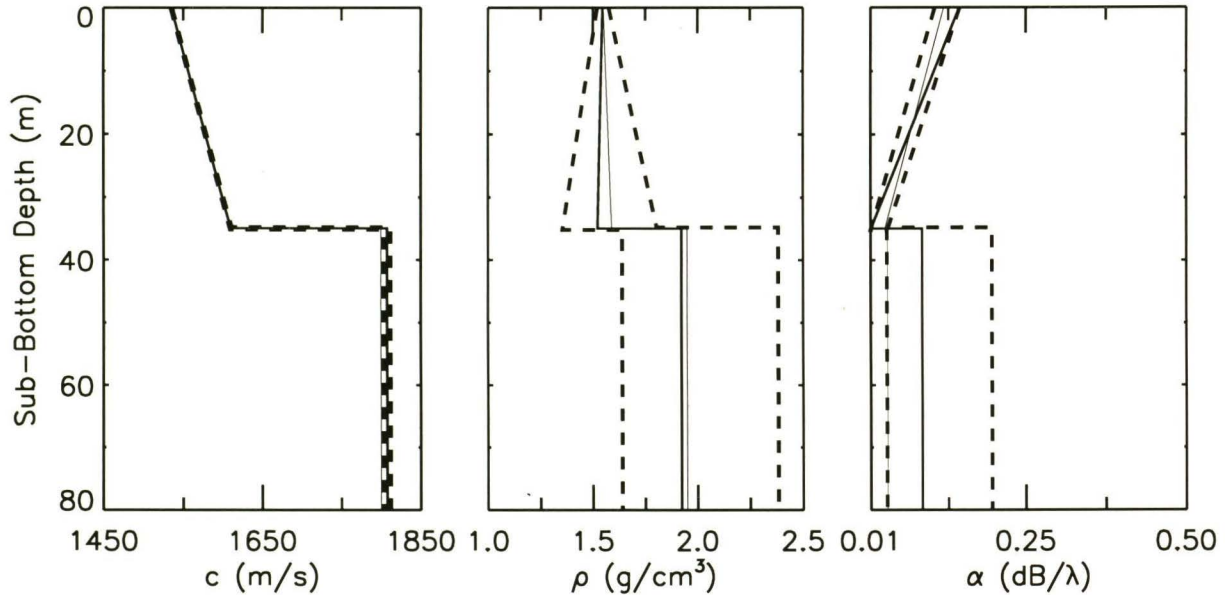


Figure 4.25 FGS MAP profiles of compressional speed  $c$ , density  $\rho$ , and attenuation  $\alpha$  for TC0. The thin line represents the true model and the thicker line represents the MAP. The dashed lines represent the 95% HPD interval bounds.

reduced number of parameters. Fig. 4.24 shows the rows of the model correlation matrix. The combinations of correlations and anti-correlations between compressional speeds, layer thicknesses, water depth and attenuations that were explained earlier for TC1 are also evident in Fig. 4.24. In addition,  $c_b$  is strongly anti-correlated with  $\rho_{12}$  and  $\rho_b$ , and  $\rho_{12}$  is very strongly correlated with  $\rho_b$ .

FGS was also applied to TC0. For the 50 and 200 Hz source frequency data,  $N_e = 2$  and  $N_e = 6$  were chosen respectively (figure not shown). The application of the FGS took  $\sim 11$  days of computation time, during which  $2 \cdot 11 \cdot 4035 \approx 9 \times 10^4$  models were accepted. Rotation of the model space occurred after  $\sim 4 \times 10^3$  models had been accepted. Fig. 4.25 shows the parameter profiles for the FGS MAP and the schematic representation of the 95% HPD intervals. Based on the HPD intervals, the model is well determined, with the possible exception of the basement density and basement

---

attenuation. The MAP estimate is a very good approximation of the true model. The compressional speed profile is almost identical to the true profile. However, the true basement compressional speed falls slightly outside the HPD interval bounds for  $c_b$ , possibly because the data variance was not estimated properly. The mismatch for the FGS MAP estimate with respect to the ASSA objective function, is  $E = 7.8 \times 10^{-4}$ .

---

## Chapter 5

### Range Dependent Inversion for Noisy Synthetic Data

The noise that exists on synthetic data provided for Workshop '01 is due to the numerical errors of the propagation models used to generate the data and the replica fields. The signal-to-noise ratio (SNR) is therefore much higher than it would be for real measured data and the distribution of errors is unknown. To investigate how well ASSA and the FGS work when applying the techniques to more realistic range dependent data, spatially-white complex Gaussian noise with zero mean was added to the synthetic data of TC1 and TC2 to achieve an SNR of 10 dB at each frequency. In this chapter, the results of applying the two inversion techniques to the noisy data are presented. The data variance was known; therefore, the FGS used the error function given in Eq. (2.23).

#### 5.1 Test Case 1

The optimal model parameterization for noise-free data is not necessarily the same as for noisy data. Hence, another parameterization analysis was carried out for TC1. The ASSA control parameters previously used in TC1 were applied, with the exception that  $\beta = 0.975$ . The computing time increased with the number of layers from  $\sim 1$  hour for  $L = 1$  to  $\sim 14.5$  hours for  $L = 5$ . The number of temperature steps also increased with the number of layers from  $\sim 200$  for  $L = 1$  to 800 for  $L = 5$ . Fig. 5.1 shows the results of the under-parameterized approach for finding the optimal number of layers. Based on the figure, it was decided that the optimal number of layers needed to represent the ocean bottom is  $L = 2$  ( $L = 3$  was used for noise-free

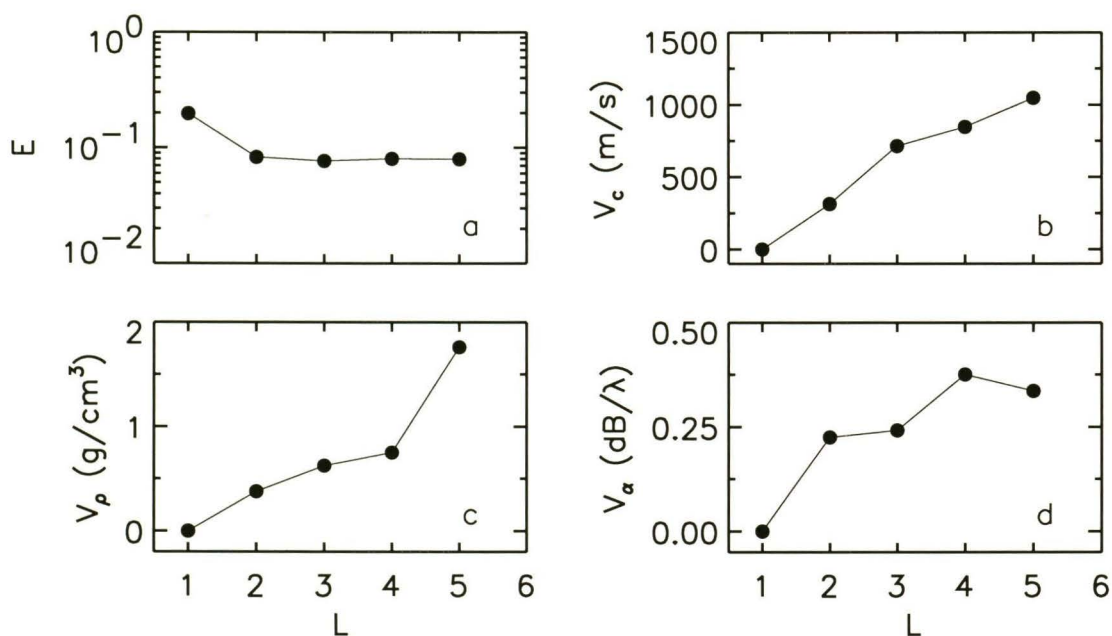


Figure 5.1 Under-parameterized approach to determining the appropriate number of sediment layers,  $L$ , for noisy TC1 data: (a) the mismatch  $E$  and (b) the  $l_1$  norm variations of the compressional speed,  $V_c$ , density  $V_\rho$  and attenuation  $V_\alpha$ .

data in Chapter 4).

The parameter profiles for the ASSA solution are shown in Fig. 5.2. The model solution has a mismatch of  $E = 0.083$ , which is only slightly larger than the true model mismatch for the noisy data,  $E = 0.080$ . The profiles have little structure, although it is clear from the model structure analysis that the data does not have enough information to adequately resolve more features. The compressional speed profile is a good simplified representation of the true profile until  $\sim 30$  m depth. The density profile is a good approximation of the entire profile. Attenuation is adequately approximated for the first 10–15 m depth. Shallow structure, therefore, is better determined than the deeper structure for the model estimate. Fig. 5.3 shows the 1-D sensitivity plots for the noisy data. According to the figure, the most sensitive parameters are  $h$ ,  $H$  and  $c$  while  $\rho$  and  $\alpha$  are the least sensitive.

The FGS was also applied to the noisy TC1 data, using the same  $L = 2$  model

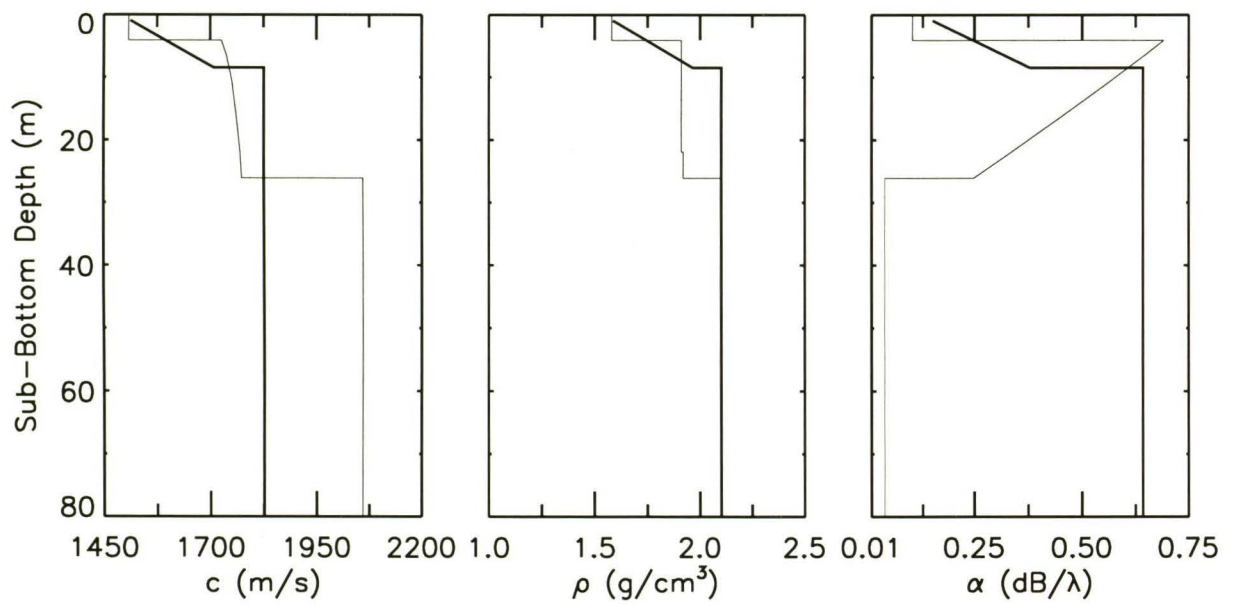


Figure 5.2 Profiles of compressional speed  $c$ , density  $\rho$ , and attenuation  $\alpha$  for noisy TC1 data. The thin line represents the true model and the thicker line represents the ASSA solution for  $L = 2$  sediment layers.

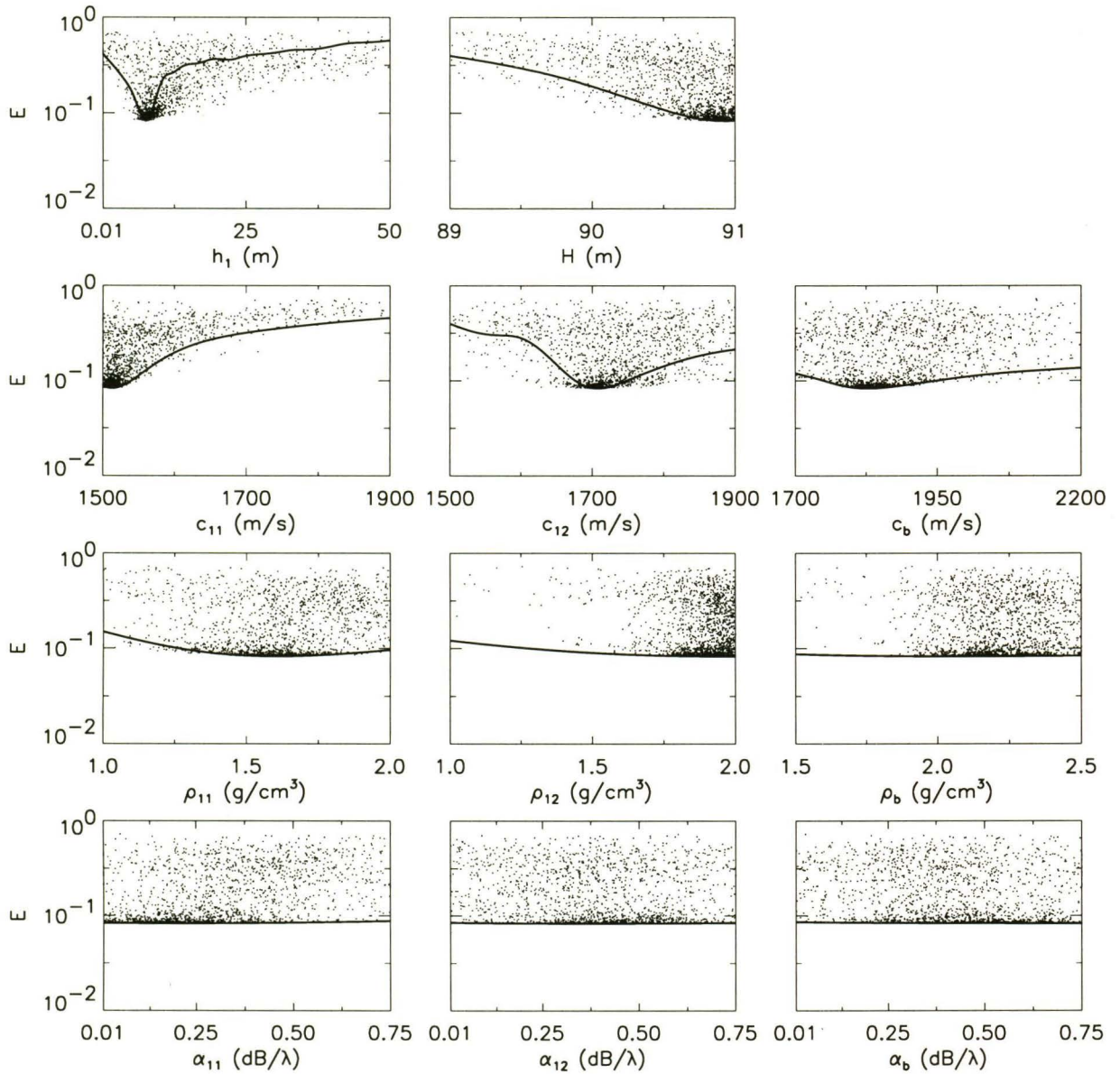


Figure 5.3 Sensitivity analysis for noisy TC1 data. Solid line represents 1-D cross-sections of the objective function. Small dots represent the mismatch  $E$  for every accepted model. The  $x$ -axis limits represent the parameter bounds used in the inversion.

parameterization. The parameter space axes were rotated twice, once after  $2 \cdot 11 \cdot 200 \approx 4 \times 10^3$  models had been accepted, and again after  $\sim 10^4$  models had been accepted. The threshold of  $\epsilon_2 = 0.1$  was reached after  $\sim 4$  days of computing time and after  $\sim 4 \times 10^4$  models had been accepted.

Fig. 5.4 shows the estimated marginal probability distributions of the parameters and Fig. 5.5 shows the schematic representation of the 95% HPD intervals. The uncertainty of the layer thickness has increased considerably compared to the noise-free cases examined earlier. Of the compressional speed parameters,  $c_{11}$  is determined well, while the remainder of the parameters are not. The density parameters of the top layer have distributions comparable to the noise-free case (Fig. 4.18). The basement density and the attenuation are not well determined. In addition, the relative sensitivities determined from Fig. 5.3 generally agree with the relative sensitivities that could be determined using Fig. 5.4.

Fig. 5.5 also shows the FGS MAP estimate profiles. The compressional speed profile is an acceptable two-layer approximation of the true profile. The 95% HPD interval bounds for  $c_b$  do not encompass the true solution because the data are not sensitive to deep structure. The MAP estimate and 95% HPD interval bounds for  $c_b$  appear to be influenced most by the first 30–40 m depth of the environment. The density profile approximates the true profile well, despite the large uncertainty of the basement density. The attenuation profile does not reflect the true profile below 15 m depth. The mismatch of the FGS MAP, with respect to the ASSA objective function, is  $E = 0.083$ , which is equivalent to the ASSA solution mismatch.

## 5.2 Test Case 2

ASSA was applied to the noisy TC2 data using the same  $L = 2$  model parameterization that was found to be optimal for the noise-free data (Sec. 4.2.2). Using

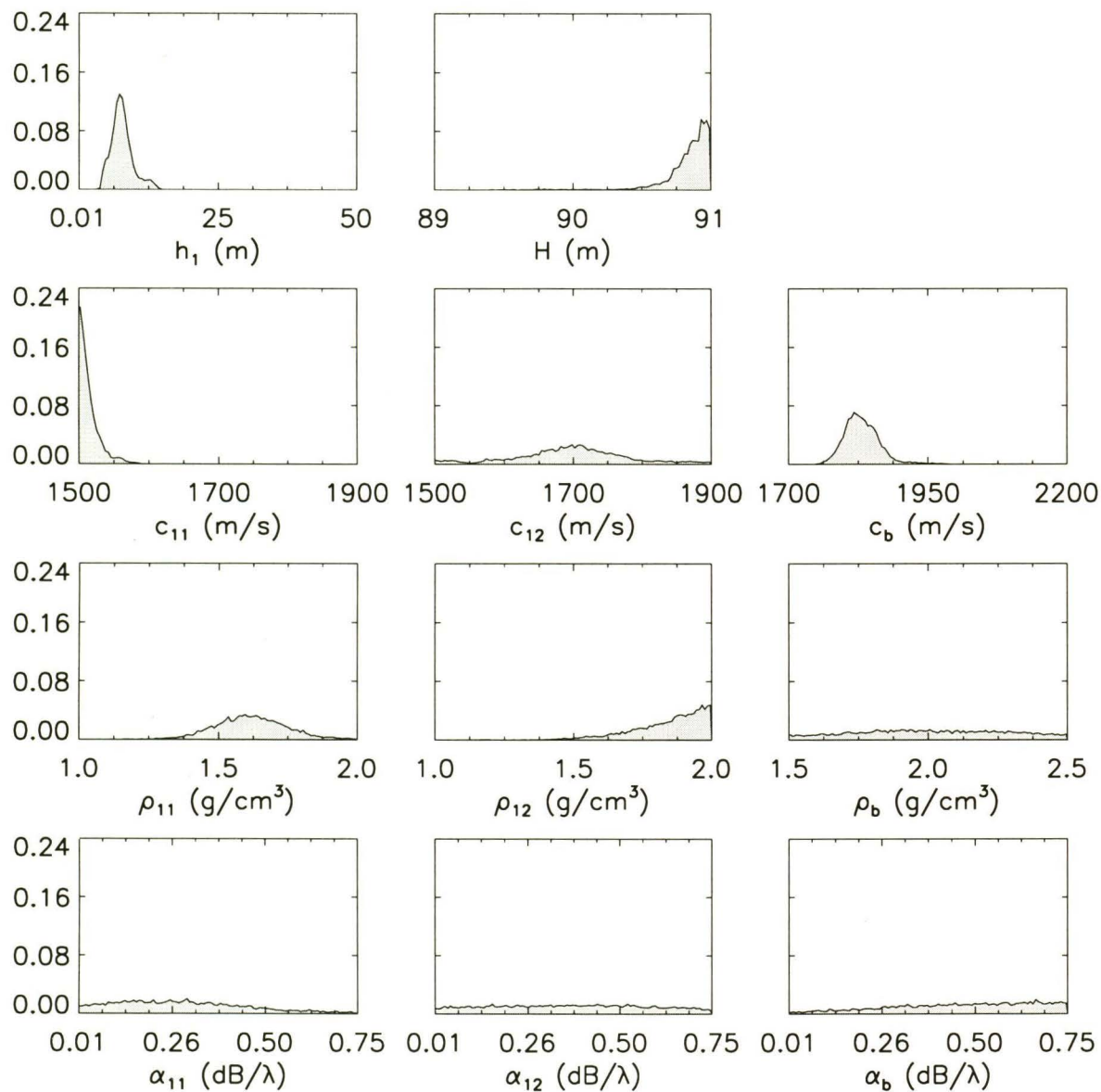


Figure 5.4 Estimated marginal probability distributions for noisy TC1 data and  $L = 2$ . The  $x$ -axis limits represent the parameter bounds used in the inversion.

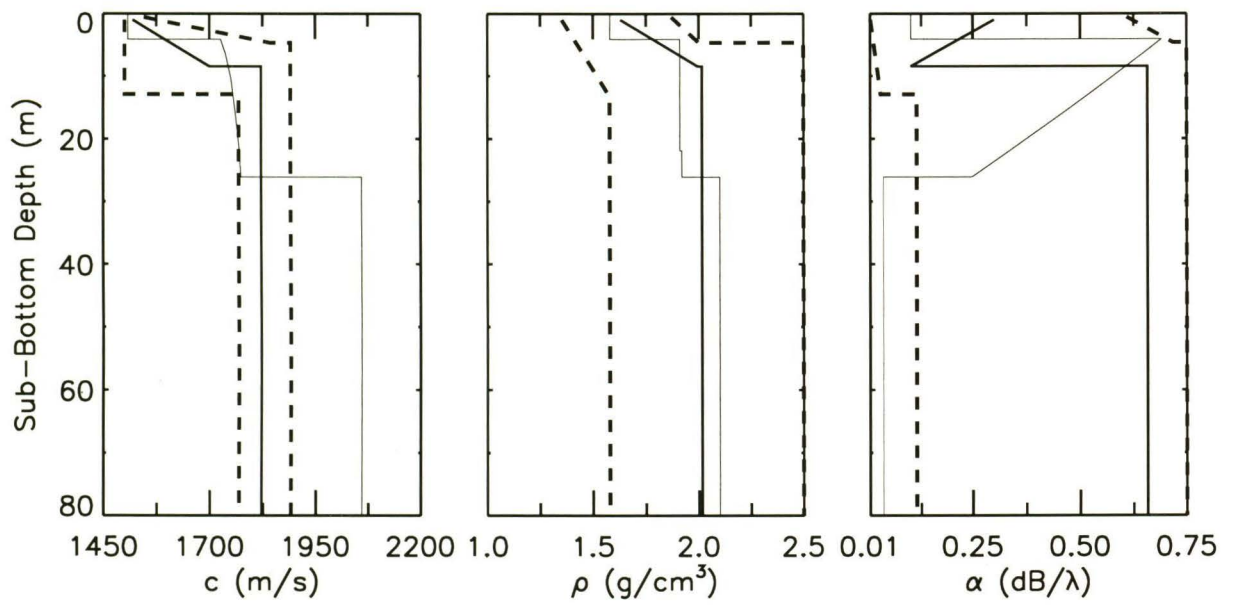


Figure 5.5 FGS MAP profiles of compressional speed  $c$ , density  $\rho$ , and attenuation  $\alpha$  for noisy TC1 data. The thin line represents the true model and the thicker line represents the MAP. The dashed lines represent the 95% HPD interval bounds.

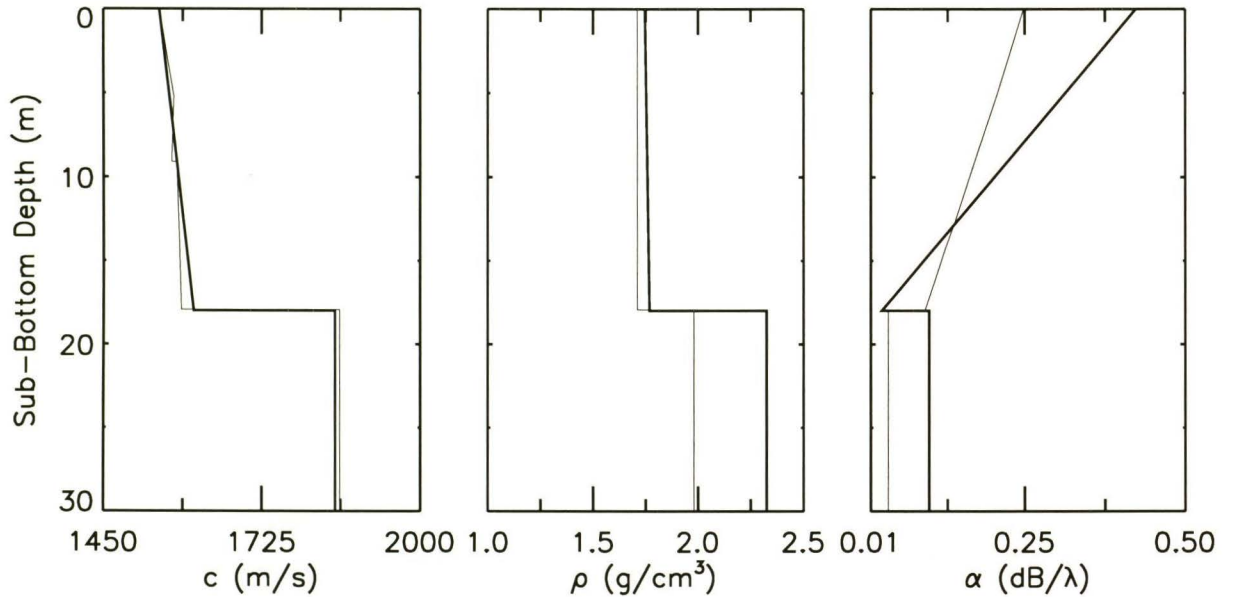


Figure 5.6 Profiles of compressional speed  $c$ , density  $\rho$ , and attenuation  $\alpha$  for noisy TC2 data. The thin line represents the true model and the thicker line represents the ASSA solution for  $L = 2$  sediment layers.

the same ASSA control parameters as for the noise-free application, the algorithm required  $\sim 250$  temperature steps and  $\sim 3.5$  hours of computing time to complete. The parameter profiles of the ASSA solution are shown in Fig. 5.6. The compressional speed profile and layer thickness are well determined; as is the density profile, though to a lesser degree. The attenuation profile is also reasonably predicted. The best model has a mismatch of  $E = 0.076$ , which is lower than the true model mismatch for the noisy data,  $E = 0.082$ . Fig. 5.7 shows the 1-D sensitivity plots. The relative sensitivities are unchanged compared to the noise-free analysis (Fig. 4.14): sensitivity decreases from  $c$  to  $\rho$  to  $\alpha$  and  $h_1$ ,  $c_{11}$  and  $c_{12}$  are the most sensitive parameters. The relative sensitivities are generally reflected by the ASSA solution in Fig. 5.6.

The FGS was also applied to the noisy TC2 data using the  $L = 2$  model parameterization. The threshold of  $\epsilon_2 = 0.1$  was achieved after  $\sim 6$  days of computing time

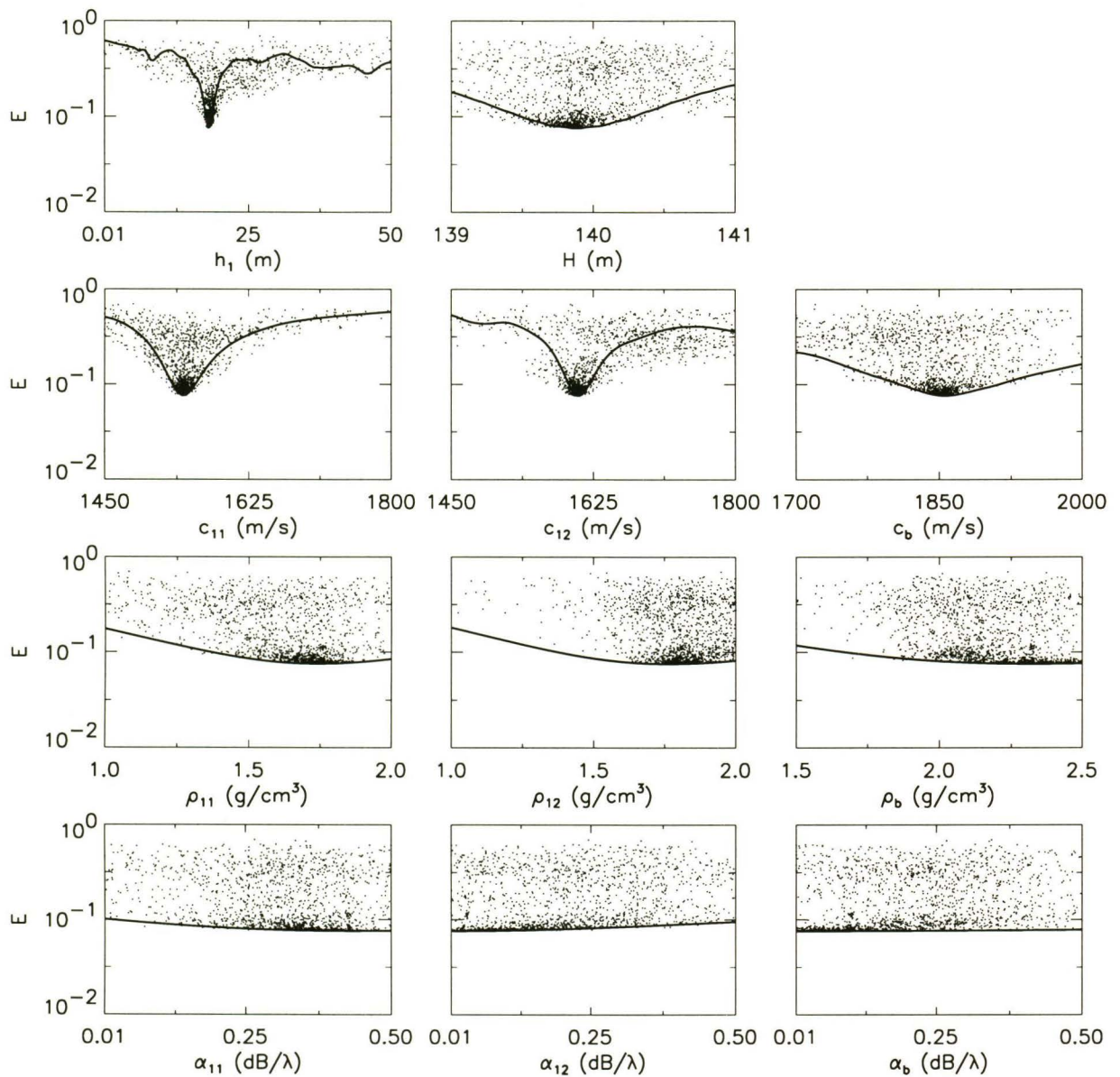


Figure 5.7 Sensitivity analysis for noisy TC2 data. Solid line represents 1-D cross-sections of the objective function. Small dots represent the mismatch  $E$  for every accepted model. The  $x$ -axis limits represent the parameter bounds used in the inversion.

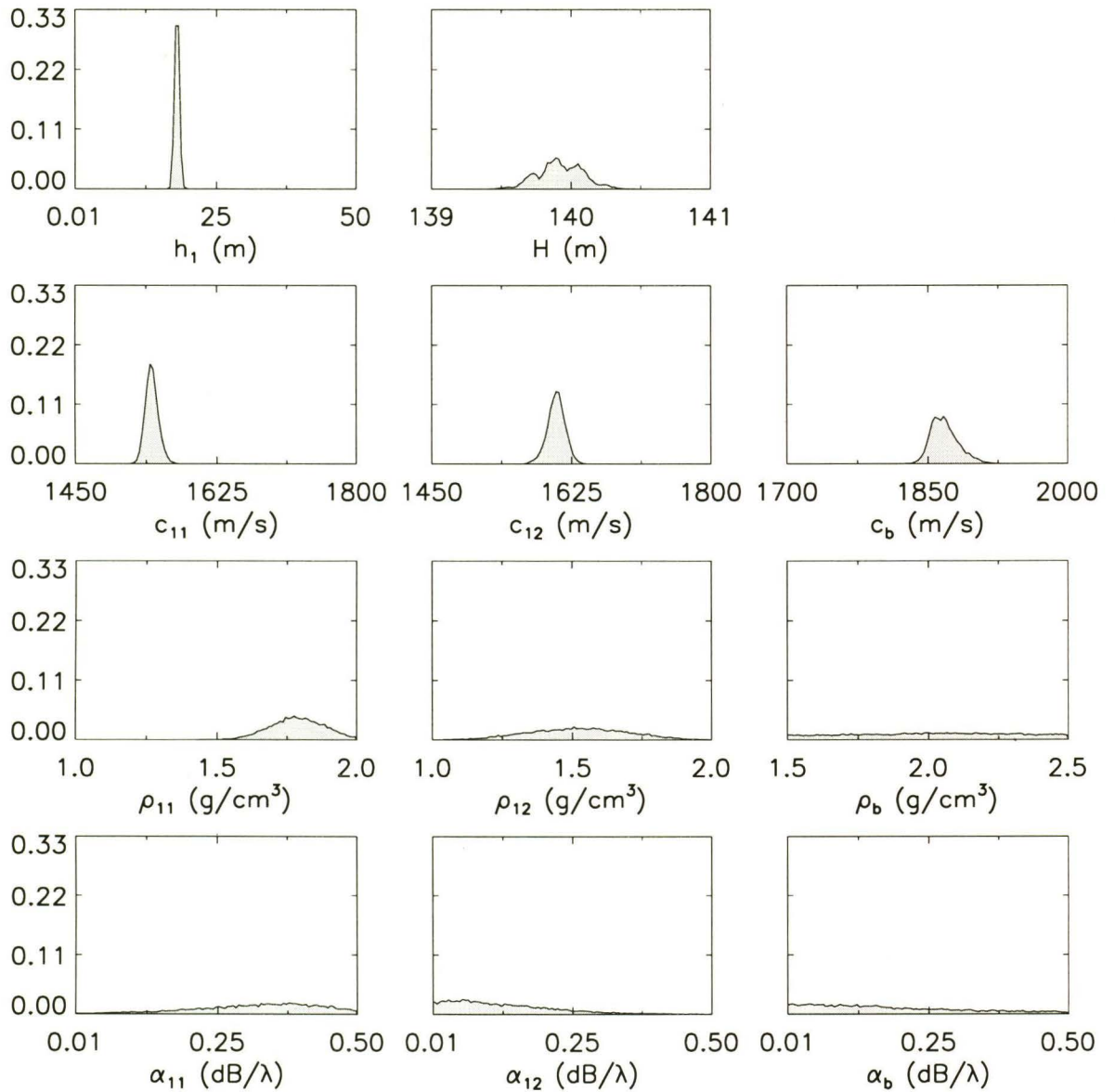


Figure 5.8 Estimated marginal probability distributions for noisy TC2 data and  $L = 2$ .

The  $x$ -axis limits represent the parameter bounds used in the inversion.

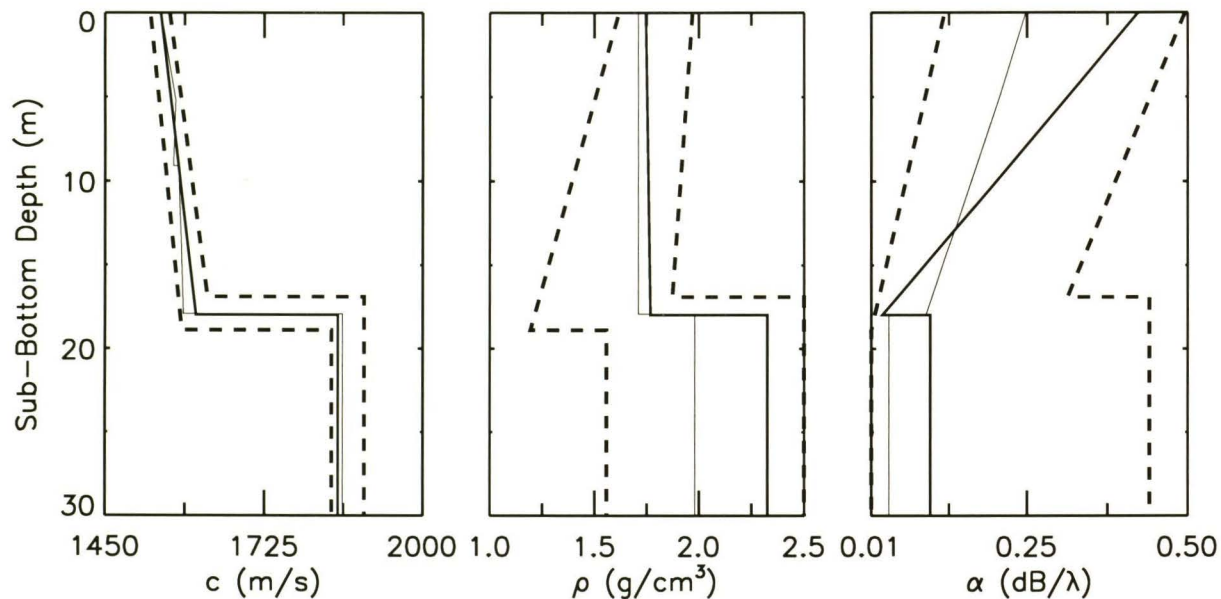


Figure 5.9 FGS MAP profiles of compressional speed  $c$ , density  $\rho$ , and attenuation  $\alpha$  for noisy TC2 data. The thin line represents the true model and the thicker line represents the MAP. The dashed lines represent the 95% HPD interval bounds.

with  $2 \cdot 11 \cdot 2340 \approx 5 \times 10^4$  models accepted. Rotation was applied after  $\sim 6 \times 10^3$  models had been accepted. Fig. 5.8 shows the FGS estimated marginal probability distributions and Fig. 5.9 shows the schematic representation of the 95% HPD intervals. The uncertainties of all the parameters have increased compared to their noise-free counterparts (Figs. 4.22 and 4.23). The layer thickness HPD interval is approximately double that found in the noise-free TC2 results; however the uncertainty is still small. The compressional speed uncertainties are also still reasonable. The density, with the exception of  $\rho_{11}$ , and attenuation parameters have HPD intervals that nearly encompass their bounds. The relative sensitivities determined in Fig. 5.3 are the same as those determined from the FGS results.

Fig. 5.9 also shows the profiles for the estimated FGS MAP. The FGS MAP has a mismatch of  $E = 0.076$ , with respect the ASSA error function, which is equivalent

---

to the ASSA solution mismatch. The model approximates the true model well and is virtually identical to the ASSA solution.

## Chapter 6

### Conclusions

This thesis considered applications of optimization (ASSA) and Bayesian (FGS) inversion methods for several types of geoacoustic inverse problems. There were two primary objectives: to assess an under-parameterized approach to determining appropriate model parameterizations for both range independent and range dependent problems and to assess the ability of ASSA and the FGS to generate solutions for range dependent inversion problems. To facilitate this investigation, noise-free synthetic benchmark data from two recent workshops were used. The investigation also included noisy synthetic data, which provided a more realistic test.

The model parameterization is often unknown in geoacoustic inversion; therefore, determining parameterization can be an important aspect of an inversion. The under-parameterized approach in conjunction with ASSA successfully predicted appropriate model structures for range independent and range dependent problems which included both noise-free and noisy data. The results of the least structure analysis did not require any prior knowledge about the true model structure. Through this procedure, ASSA was also shown to be able to invert for a large number of parameters (up to 39 for TC1).

Once parameterization is established, the solution to the inversion problem can also be established. For example, the PPD represents the Bayesian solution to an inverse problem. Assessing the ability of the FGS to estimate the PPDs for range dependent problems was also an aim of this work. For each range dependent problem, the FGS successfully estimated the 1-D marginal probability distributions, the MAP and the correlation matrix by sampling the PPD. The uncertainties were quantified by calculating the 95% HPD intervals using the marginal probability distributions.

Table 6.1 The mismatches (Eq. (4.2)) for test case NL, TC0, TC1, TC2, noisy TC1, and noisy TC2 parameter estimates are presented with respect to ASSA and the FGS. The mismatches calculated using the true models are also included.

Test case	True Model	ASSA	FGS
NLa	$9.5 \times 10^{-5}$	$1.8 \times 10^{-5}$	—
NLb	$9.8 \times 10^{-6}$	$5.6 \times 10^{-5}$	—
NLc	$3.2 \times 10^{-6}$	$8.3 \times 10^{-6}$	—
TC0	$1.8 \times 10^{-3}$	$9.2 \times 10^{-4}$	$7.8 \times 10^{-4}$
TC1	$2.3 \times 10^{-3}$	$1.4 \times 10^{-3}$	$8.7 \times 10^{-4}$
TC2	$6.5 \times 10^{-3}$	$3.4 \times 10^{-3}$	$3.5 \times 10^{-3}$
TC1 (noise)	$8.0 \times 10^{-2}$	$8.3 \times 10^{-2}$	$8.3 \times 10^{-2}$
TC2 (noise)	$8.2 \times 10^{-2}$	$7.6 \times 10^{-2}$	$7.6 \times 10^{-2}$

This quantified how well the parameters could be determined as well as allowed comparisons with respect to test cases and noise on the data. The MAP estimates approximated the true models to varying degrees, but there was always a subset of parameters that were well determined. This subset typically included layer thickness and compressional speed parameters. The MAP estimate mismatches, as shown in Table 6.1, are typically lower than their true model mismatches. The FGS was able to invert for a large number of parameters (18 for TC1).

The FGS is a very computationally intensive inversion method. When parameter estimates are of primary importance, optimization methods, such as ASSA, can be more appropriate. Another objective of this work was to assess the ability of ASSA to converge to an appropriate model solution for range dependent problems. Like the MAP estimates, the ASSA solutions approximated the true models to varying degrees. Table 6.1 contains the mismatches for all the model solutions. The mismatches were

usually lower than the true model mismatches. Determining relative sensitivities using 1-D cross-sections and the models available from the ASSA inversion provided an approximate idea of how well the parameters are determined. The ASSA relative sensitivity analysis was generally reliable when compared to the same analysis applied to the FGS marginal probability distributions.

One inefficiency of the procedure used in this thesis was that to apply the FGS an additional optimization had to be performed in order to estimate the data errors. A method to bypass that inefficiency could be to perform the FGS inversion while including the data variance as an unknown parameter. In this work, it was not resolved whether the error estimates for the noise-free data were appropriate. Gaussian-distributed random errors were assumed for the data. This may not have been appropriate.

In conclusion, the FGS and ASSA algorithms were successfully applied to the problems examined here. The ASSA is an efficient method for determining appropriate model estimates and relative sensitivities. For unknown model parameterization, it can also be effectively used with the under-parameterized approach to determine the optimal parameterization. In addition, the FGS is a useful approach for generating the PPD for a given model parameterization. Future work might involve applying ASSA and the FGS to inversion problems that have range dependence in the parameters not just the bathymetry. Applying the methods to measured range dependent data would also be appropriate. It would also be interesting to apply ASSA to the Workshop '01 test cases using an adiabatic normal mode propagation model to test if equivalent results can be obtained despite the adiabatic approximation. Adiabatic normal mode models are not as computationally intensive as PE models. Developing the FGS with unknown variances as parameters would also be a worthwhile advance.

---

## References

- Ainslie, M.A., R.M. Hamson, G.D. Horsley, A.R. James, R.A. Laker, M.A. Lee, D.A. Miles and S.D. Richards, Deductive Multi-Tone Inversion of Seabed parameters, *J. Comp. Acoust.*, **8** (2), 271–284, 2000.
- Brooke, G.H. and D.J. Thomson, Non-local boundary conditions for high-order parabolic equation algorithms, *Wave Motion*, **31**, 117–129, 2000.
- Brooke, G.H., D.J. Thomson and G.R. Ebbeson, PECan: a Canadian parabolic equation model for underwater sound propagation, *J. Comp. Acoust.*, **9** (1), 69–100, 2001.
- Collins, M.D., W.A. Kuperman, and H. Schmidt, Nonlinear inversion for ocean-bottom properties, *J. Acoust. Soc. Am.*, **92** (5), 2270–2283, 1992.
- Collins, M.D., The split-step Padé solution for the parabolic equation method, *J. Acoust. Soc. Am.*, **93**, 1736–1742, 1993.
- Chapman, N.R., S. Chin-Bing, D. King, and R. Evans, Benchmark workshop for geoacoustic inversion techniques in range dependent waveguides, *J. Acoust. Soc. Am.*, **110**, 2660, 2001.
- Dizaji, R.M., N.R. Chapman, and R.L. Kirilin, A Phase Regulated Back Wave Propagation Technique for Geoacoustic Inversion, *J. Acoust. Soc. Am.*, **111** (2), 800–808, 2002.
- Dosso, S.E., M.J. Wilmut, and A.S. Lapinski, An Adaptive-Hybrid Algorithm for Geoacoustic Inversion, *IEEE J. Ocean. Eng.*, **26** (3), 324–336, 2001.
- Dosso, S.E., Quantifying uncertainty in geoacoustic inversion. I. A fast Gibbs sampler approach, *J. Acoust. Soc. Am.*, **111** (1), 129–142, 2002.
- Dosso, S.E., and P.L. Nielsen, Quantifying uncertainty in geoacoustic inversion. II. Application to broadband, shallow-water data, *J. Acoust. Soc. Am.*, **111** (1), 143–159, 2002.
- Dosso, S.E., and M.J. Wilmut, Effects of Incoherent and Coherent Source Spectral Information in Geoacoustic Inversion, *J. Acoust. Soc. Am.*, in press 2002.
- Evans, R.B., A coupled mode solution for acoustic propagation in a waveguide with stepwise depth variations of a penetrable bottom, *J. Acoust. Soc. Am.*, **74** (1), 188–195, 1983.

- 
- Fallat, M.R., and S.E. Dosso, Geoacoustic inversion for the Workshop '97 benchmark test cases using simulated annealing, *J. Comput. Acoust.*, **6** (1/2), 29–44, 1998.
- Fallat, M.R., and S.E. Dosso, Geoacoustic inversion via local, global and hybrid algorithms, *J. Acoust. Soc. Amer.*, **105** (1/2), 3219–3230, 1999.
- Fallat, M.R., P.L. Nielsen, and S.E. Dosso, Hybrid geoacoustic inversion of broadband Mediterranean Sea data, *J. Acoust. Soc. Amer.*, **107** (4), 1967–1977, 2000.
- Gamerman, D., *Markov Chain Monte Carlo: Stochastic simulation for Bayesian inference*, Chapman & Hall, New York, 1997.
- Gerstoft, P., Inversion of seismoacoustic data using genetic algorithms and *a posteriori* probability, *J. Acoust. Soc. Am.*, **95**, 770–782, 1994.
- Gerstoft, P., Inversion of acoustic data using a combination of genetic algorithms and the Gauss-Newton approach, *J. Acoust. Soc. Am.*, **97** (4), 2181–2190, 1995.
- Gerstoft, P. and C.F. Mecklenbräuker, Ocean acoustic inversion with estimation of *a posteriori* probability distributions, *J. Acoust. Soc. Am.*, **104** (2), 808–819, 1998.
- Gingras, D.F. and P. Gerstoft, Inversion for geometric and geoacoustic parameters in shallow water: Experimental results, *J. Acoust. Soc. Am.*, **97** (6), 3589–3598, 1995
- Hammersley, J.M. and D.C. Handscomb, *Monte Carlo Methods*, Wiley, New York, 1964.
- Heard, G.J., D. Hannay and S. Carr, Genetic algorithm inversion of the 1997 Geoacoustic Inversion Workshop test case data, *J. Comput. Acoust.*, **6** (1/2), 61–71, 1998.
- Jaschkr, L., and N.R. Chapman, Matched field inversion of broadband data using the freeze bath method, *J. Acoust. Soc. Am.*, **106** (4), 1838–1851, 1999.
- Jensen, F.B., W.A. Kuperman, M.B. Porter, and H. Schmidt, *Computational Ocean Acoustics*, American Institute of Physics, New York, 1994.
- Kirkpatrick, S., C.D. Gelatt Jr. and M.P. Vecchi, Optimization by simulated annealing, *Science*, **220**, 671–680, 1983.
- Knobles, D.P., R.A. Koch, E.K. Westwood, and T. Udagawa, The inversion of ocean waveguide parameters using a nonlinear least squares approach, *J. Comp. Acoust.*, **6** (1/2), 83–97, 1998.
- Lane, F., Estimation of the kinematic rupture parameters from historical seismograms: An application of simulated annealing to a nonlinear optimization problem, Ph.D. dissertation, Colorado School of Mines, Golden, CO, 1992.

- 
- Mecklenbräuker, C.F., and P. Gerstoft, Objective functions for ocean acoustic inversion derived by likelihood methods, *J. Comp. Acoust.*, **8** (2), 259–270, 2000.
- Metropolis, N., A. Rosenbluth, M. Rosenbluth, A. Teller, and E. Teller, Equation of state calculation by fast computing machines, *J. Chem. Phys.*, **21**, 1087–1092, 1953.
- Musil, M., M.J. Wilmut, and N.R. Chapman, A hybrid simplex genetic algorithm for estimating geoacoustic parameters using matched-field inversion, *IEEE J. Oceanic Eng.*, **24** (3), 358–369, 1999.
- Nedler, J.A. and R. Mead, A simplex method for function minimization, *Comput. J.*, **7**, 308–313, 1965.
- Neilsen, T.B., E.K. Westwood, and T. Udagawa, Mode function extraction from a VLA using singular value decomposition, *J. Acoust. Soc. Am.*, **101** (5), 3025, 1997.
- Olson, C.L., *Statistics: Making Sense of Data*, WCB/McGraw-Hill, 1987.
- Pan, L., and L. Wu, A hybrid global optimization method for inverse estimation of hydraulic parameters: Annealing-simplex method, *Water Resources Research*, **34** (9), 2261–2269, 1998.
- Press, W.H., S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery, *Numerical Recipes in Fortran 77 – The Art of Scientific Computing, 2nd ed.*, Cambridge University Press, New York, 1992.
- Rajan, S.D., J.F. Lynch, and G.V. Frisk, Perturbative inversion methods for obtaining bottom geoacoustic parameters in shallow water, *J. Acoust. Soc. Am.*, **82** (3), 998–1017, 1987.
- Ratilal, P., P. Gerstoft, and J.T. Goh, Subspace approach to inversion by genetic algorithms involving multiple frequencies, *J. Comp. Acoust.*, **6** (1/2), 99–115, 1998.
- Reidel, M., S.E. Dosso, and L. Beran, Uncertainty estimation for AVO inversion, *Geophysics*, submitted 2002.
- Schmidt, H., SAFARI: Seismo-acoustic fast field algorithm for range-independent environments, SACLANT ASW Research Centre, La Spezia, Italy, Rep. SR-112, 1988.
- Sen, M.K., and P.L. Stoffa, *Global Optimization methods in Geophysical Inversion*, Elsevier, New York, 1995.

- 
- Sen, M.K., and P.L. Stoffa, Bayesian inference, Gibbs' sampler and uncertainty estimation in geophysical inversion, *Geophysical Prospecting*, **44**, 313–350, 1996.
- Siderius, M., M. Snellen, D.G. Simons, and R. Onken, An environmental assessment in the strait of Sicily: measurement and analysis techniques for determining bottom and oceanographic properties, *IEEE J. Oceanic Eng.*, **25** (3), 2000.
- Szu, H., and R. Hartley, Fast Simulated Annealing, *Phys. Letters A*, **122**, 157–162, 1987.
- Taroudakis, M.I., and M.G. Markaki, Bottom geoacoustic inversion by matched field processing – a sensitivity study, *Inverse Problems*, **16**, 1679–1692, 2000.
- Tolstoy, A., Using matched-field processing to estimate shallow-water bottom properties from shot data taken in the Mediterranean Sea, *IEEE J. Oceanic Eng.*, **21** (4), 1996.
- Tolstoy, A., N.R. Chapman, and G. Brooke, Workshop '97: benchmarking for geoacoustic inversion in shallow water, *J. Comput. Acoust.*, **6** (1/2), 1–28, 1998.
- Tolstoy, A., MFP benchmark inversions via the RIGS method, *J. Comput. Acoust.*, **6** (1/2), 185–203, 1998.
- Westwood, E.K., C.T. Tindle, and N.R. Chapman, A normal mode model for acousto-elastic ocean environments, *J. Acoust. Soc. Am.*, **100** (6), 3631–3645, 1996.
- Zala, C.A. and J.M. Ozard, Estimation of geoacoustic parameters from narrow band data using a search-optimization technique, *J. Comp. Acoust.*, **6** (1/2), 223–244, 1998.

## Vita

**Surname:** Lapinski  
**Given Names:** Anna-Liesa Salome  
**Place of Birth:** St. Catharines, Ontario, Canada

### **Educational Institutions Attended:**

University of Victoria, Victoria, BC 1999 to 2002  
Brock University, St. Catharines, ON 1993 to 1994, 1995 to 1999

### **Degrees Awarded:**

B.Sc. (Honours) in Physics, Brock University 1999

### **Honours and Awards:**

CAA Fessenden Student Prize in Underwater  
Acoustics, Canadian Acoustical Association 2001  
NSERC Postgraduate Scholarship 1999 to 2001  
The Captain Wilfred T. Gagnon and  
Mrs. Emma Gagnon Memorial Scholarship 2001, 2000 & 1999  
President's Scholarship 2000 & declined 1999  
Howard E. Petch Research Scholarship 1999

### **Publications:**

S. E. Dosso, M. J. Wilmut, and A. S. Lapinski, "An Adaptive Hybrid Algorithm for Geoacoustic Inversion," *IEEE Journal of Oceanic Engineering*, **26**, 324–336 (2001).

## Partial Copyright License

I hereby grant the right to lend my thesis to users of the University of Victoria Library, and to make single copies only for such users or in response to a request from the Library of any other university, or similar institution, on its behalf or for one of its users. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by me or a member of the University designated by me. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Title of thesis:

Optimization and Bayesian Approaches to Geoacoustic Inversion.

Author:



Anna-Liesa Salome Lapinski

July 1, 2002