Seabed Characterization by Inversion of Acoustic Data from a Towed Array

Yong Han Goh, Chen-Fen Huang, Peter Gerstoft, and William S. Hodgkiss

***Abstract*—The goal of geoacoustic inversion is to estimate environmental parameters from measured acoustic fields from e.g. a towed array. Though the inversion results have some uncertainty, inversion is an efficient technique to estimate environment parameters. Based on the *a posteriori* probability density of environmental parameters obtained from inversion, we perform statistical estimation of transmission loss (TL) and generate a 90% credibility level envelope or uncertainty band for the TL. This uncertainty band accounts for the inherent variability of the environment not usually contained in sonar performance prediction model inputs, and presents a useful probabilistic description of the environment's variability. Towed arrays are advantageous to fixed systems as they are easy to deploy and the moving ship enables estimation of spatially variable seabed properties. The approach is demonstrated with data obtained from the MAPEX2000 experiment conducted by NATO Undersea Research Centre (NURC) in the Mediterranean Sea in November 2000.**

***Index Terms*— geoacoustic inversion, seabed characterization, towed array, transmission loss estimation.**

I. Introduction

GEO-ACOUSTIC inversion using matched-field processing is a model-based technique that has been applied successfully to derive environment and seabed parameters for propagation prediction [1],[7],[12]. Computer simulations are used to model the acoustic response to different sea-bed types (forward models), and efficient search algorithms used to find the environment giving an optimal match between the modeled and measured data. It should be noted, however, that inverse problems are usually under-determined, and solutions may not be unique, i.e. there could be several solutions giving rise to the same objective function. If the results of the inversion are only required for sonar performance prediction (for example), it is only the resulting acoustic field in the water that matters, often at long range and within a restricted range of frequencies. In this context, a precise description of the seabed is not necessary, and it is usually sufficient to describe a simpler "effective" seabed model having the same acoustic effect on the underwater sound field within the range-frequency domain of interest [2].

Yong Han Goh is with the Defence Science and Technology Agency, Singapore, and Electrical and Computer Engineering Department, University of California, San Diego (phone: 1-858-610-6154; fax: 1-858-534-7641; e-mail: yhgoh@ ucsd.edu).

Chen-Fen Huang, Peter Gerstoft and William S. Hodgkiss are with the Marine Physical Laboratory, Scripps Institute of Oceanography, University of California, San Diego.

There are uncertainties associated with the sea-bed parameters obtained from geoacoustic inversion, and in this paper, the translation of these uncertainties to the transmission loss domain is also presented. Sonar performance prediction using a probability density function (pdf) based on environmental variability has been discussed in [3].

By far, a sound source in combination with a receiving vertical line array (VLA) is the most common configuration for collection of acoustic data for geoacoustic inversion. It is a sensible choice as the propagating acoustic field is received at almost all angles if the VLA spans a large portion of the water column. The disadvantage is that for reliable seabed estimates, a geometry with a fixed VLA receiver array is usually limited to 2 km (or less) separation between the two due to ocean sound-speed variability [7]. In addition, in strongly range-dependent regions where distinct bottom types impact acoustic propagation in significantly different ways, the behaviour of the acoustic field will not be modeled accurately using averaged seabed properties as input to the propagation model.

The use of horizontal line arrays (HLA) has been gaining in popularity as it offers several advantages over the VLA. This include the ease of deployment from a ship, and the ability to cover large areas of interest as the ship travels, together with either a separate towed source [7], or using the ship noise as the acoustic source [8], though in this paper, a fixed source was used.

In Section II of this paper, we describe the inversion procedure for geoacoustic parameter estimation. Section III describes an algorithm for the estimation of transmission loss (TL) from ocean acoustic data recorded from a receiver array which can be either a VLA or HLA. Section IV presents the experimental setup. In Section V, a simulation study is presented to study the sensitivity of different seabed parameters to matched-field geoacoustic inversion. Section VI presents the geoacoustic inversion results using the MAPEX2000 experimental data recorded on the HLA. Section VII presents the TL estimation results using the posterior probability distributions of the parameters (or uncertainties of the parameters) derived from geoacoustic inversion in Section VI.

II. Parameter estimation using genetic algorithms

Genetic algorithms (GA) are based on an analogy with biological evolution. The basic principle of GA is simple. From all possible model vectors, an initial population of n members is selected. The fitness (or objective function) of

each member is computed. Then, through a set of evolutionary steps, the initial population evolves in order to become fitter. An evolutionary step consists of selecting a parental distribution from the initial population based on the individual's fitness. The parents are then combined in pairs, and operators are applied to them to form a set of children. The operators are the crossover and mutation operators. Finally, the children replace part of the initial distribution to get a fitter population. For a detailed description of GA and their application to geoacoustic parameter estimations, refer to [12].

A. Baseline environmental model

We use the baseline environmental model established for the North Elba site from [4]. Fig. 1 illustrates the baseline environmental model, comprising the water column, a constant thickness sediment layer with depth dependent sound speed (increasing with depth), and subbottom half-space layer.

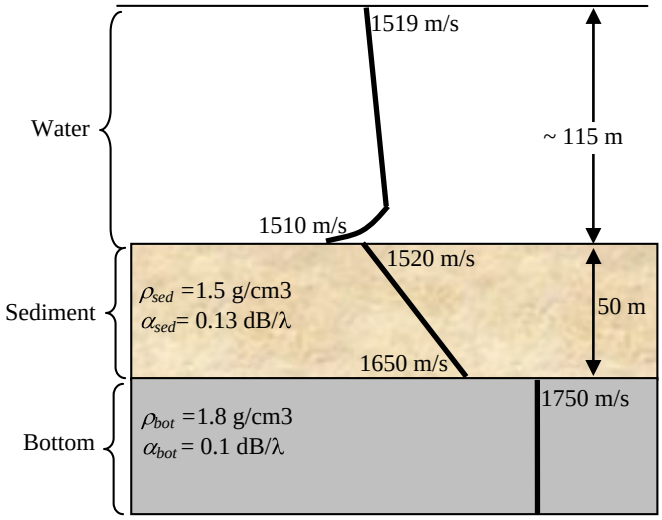


Fig. 1. The baseline environmental model based on [4] for the North Elba experiment site.

B. Inversion procedure and objective function

The inversion is performed as follows:

(1) Record the acoustic field at the site of interest. A signal transmission covering a broad band of frequencies contains more information than a single tone and will generally produce better (more stable) inversion results. In this experiment, data from the 300-500 Hz band are used for the inversion.

(2) Choose a suitable propagation model. In this paper, we use the SACLANTCEN SNAP normal-mode propagation model [10], which is robust, fast, and suitable for low frequencies.

(3) Choose a suitable cost function to minimize. The phone-coherent method has traditionally been used. However, recent analyses [7] show that frequency-coherent matched-field inversions (MFI) work well for multi-frequency HLA data, where matched-field correlations between data and replica frequency vectors are summed incoherently over range (hydrophones). The frequency-coherent cost function based on the Bartlett correlator is defined by

$$C_F = \frac{1}{N_H}\sum_{i=1}^{N_H}(1-B_{F_i}), \tag{1}$$

where

$$B_{F_i} = \frac{\left|\sum_{j=1}^{N_F} p_{ij}^{*} q_{ij}\right|^2}{\sum_{j=1}^{N_F}\left|p_{ij}\right|^2 \sum_{j=1}^{N_F}\left|q_{ij}\right|^2}, \tag{2}$$

and the phone-coherent cost function is defined by

$$C_H = \frac{1}{N_F}\sum_{j=1}^{N_F}(1-B_{H_j}), \tag{3}$$

where

$$B_{H_j} = \frac{\left|\sum_{i=1}^{N_H} p_{ij}^{*} q_{ij}\right|^2}{\sum_{i=1}^{N_H}\left|p_{ij}\right|^2 \sum_{i=1}^{N_H}\left|q_{ij}\right|^2}. \tag{4}$$

In Eqs. (1), (2), (3) and (4), N_F is the number of frequency components, N_H is the number of hydrophones, and p_{ij} and q_{ij} are the modeled and measured complex pressure vectors respectively (* denotes the complex conjugate). Both cost functions take on a value of 0 for two identical signals and 1 for completely uncorrelated signals.

The frequency-coherent MFI requires knowledge of the source spectrum (as is the case for this experiment), as frequency correlation is the same as correlation of the time series by the convolution theorem [5]. We use the frequency-coherent cost function in this paper, as we found the phone-coherent method yields poor results, as in [13]. The failure of the phone-coherent method here may be due to large errors in positions, as the phone-coherent method requires accurate information on the array shape and position [5],[13]. However, if using ships of opportunity or ship self-noise, since the source spectrum is unknown, it will be more convenient to use the phone-coherent cost function, where matched-field correlations between data and replica phone vectors are summed incoherently over frequency.

(4) An efficient algorithm is needed to navigate the enormous search space and find the global minimum to the cost function. In this paper, a genetic algorithm search is used with the propagation model SNAP as implemented in the inversion package SAGA version 5.3 [11]. 40 000 forward models were used in the inversion searches.

III. Algorithm to Map Geoacoustic Parameters' Uncertainties to Transmission Loss Domain

The mapping of geoacoustic parameters' uncertainties to the transmission loss (TL) domain has been discussed in [9] and [15] and is summarized in this section for completeness.

Fig. 2 summarizes the estimation of TL (usage domain *U*)

from ocean acoustic data observed on a vertical or horizontal array (data domain Q). Based on the ocean acoustic data **q**, we statistically characterize TL, the usage domain U. The vector **q** represents the acoustic data observed at N_H hydrophones and the vector **u** represents TL at several ranges and depths. This is mapped via a set of environmental parameters **m** in the environmental domain M. The geoacoustic inverse problem is solved as an intermediate step to obtain the posterior distribution of environmental parameters p(**m**|**q**) (environmental domain M). We are not just interested in the environment itself but also a statistical estimation of the TL field (usage domain U). Based on the posterior distribution p(**m**|**q**), the probability distribution of the TL p(**u**|**q**) is obtained via Monte Carlo integration. From this TL probability distribution, all relevant statistics of TL can be obtained, such as the median, percentiles and correlation coefficients. The vector **u** is used to denote the transmission loss as an I-dimensional vector at discrete (r_i, z_i) positions,

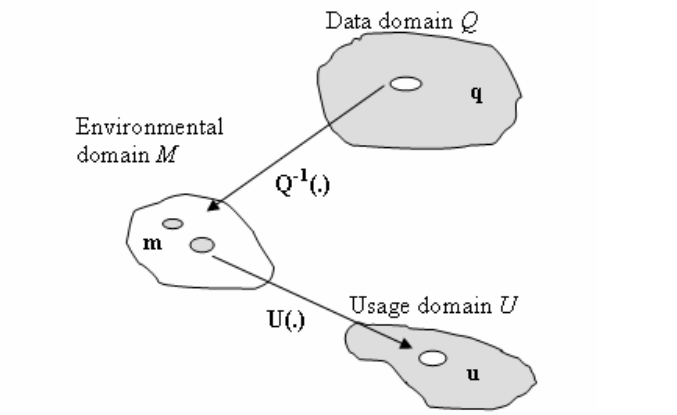


Fig. 2. An observation **q** ∈ Q is mapped into a distribution of environmental parameters **m** ∈ M that potentially could have generated it. These environmental parameters are then mapped into the usage domain U, which is the transmission loss domain in this paper.

where $u_i = u(r_i, z_i)$.

Both the experimental data **q** and the usage domain model **u** are related to **m** via forward models **Q**(**m**) and thus formally, if the data were error free and the mappings were unique, we would have $\mathbf{u} = \mathbf{U}(\mathbf{Q}^{-1}(\mathbf{q}))$. It is assumed that the mappings **Q**(**m**) and **U**(**m**) are deterministic and all uncertainties (including noise and modeling errors) are in the data. Due to the uncertainties in the data, the inverse mapping from **q** to **m** is formulated in a probabilistic framework where one also can include prior information. The forward mapping could also be probabilistic as in [16] and [17].

A. Bayesian inference

We use the Bayesian inference framework to determine the parameters of interest **m** given an observation **q**. The solution is characterized by the posterior probability p(**m**|**q**). First, the prior information about the model parameter vector is quantified by the probability density function p(**m**). Then, this information is combined with the likelihood function p(**q**|**m**) provided by the combination of data and the physical model to give the posterior information of the model parameters p(**m**|**q**). A complete discussion of inverse theory from a probabilistic point of view may be found in the recent textbook by Tarantola [16]. The solution to the inverse problem is then given by

$$\mathrm{p}(\mathbf{m}|\mathbf{q}) = \frac{\mathrm{p}(\mathbf{q}|\mathbf{m})\,\mathrm{p}(\mathbf{m})}{\mathrm{p}(\mathbf{q})} \propto L(\mathbf{m})\,\mathrm{p}(\mathbf{m}), \tag{5}$$

where p(**q**) is a normalizing factor that makes the posterior probability density p(**m**|**q**) integrate to one. Since p(**q**) does not depend on environmental model **m**, it is typically ignored in parameter estimation. Hence, in the second representation, the normalization constant p(**q**) is omitted and a brief notation L(**m**) is used to denote the likelihood function p(**q**|**m**). The posterior distribution p(**m**|**q**) carries all information available on models originating from the data and from data-independent prior information. From this distribution, all relevant features of the environmental model can be estimated using e.g., the maximum *a posteriori* (MAP) estimator.

The posterior probability distribution p(**m**|**q**) is M-dimensional, where M is the dimension of **m**. The method used to compute the posterior distribution p(**m**|**q**) depends on the dimension M. For small scale problems where $M < 8$, evaluating the likelihood function over a grid of parameter values seems most efficient. For medium scale problems, utilizing sequences of random numbers sampling from the posterior distribution using Markov chain Monte Carlo (MCMC) methods is most efficient [9]. A brief description on MCMC algorithms are given in section III-D.

B. Likelihood and objective functions

This section derives a likelihood function to be used in the probabilistic inversion following the same approach as described in [18],[19]. At a single frequency, the relation between the observed complex-valued data vector **q** sampled at a N-element array and the modeled data **D**(**m**) is described by the model

$$\mathbf{q} = \mathbf{D}(\mathbf{m}) + \mathbf{e}\,, \tag{6}$$

where **e** represents the error term. The modeled data are given by D(**m**) = s**p**(**m**), where the complex deterministic source term s is unknown. The transfer function **p**(**m**) is obtained using an acoustic propagation model and an environmental model **m**. For simplicity in the development below, data from only one frequency is first assumed. The theory for multifrequency is also described in [18],[19].

Assume the error vector **e** to be Gaussian distributed with zero mean and covariance $\mathbf{C}_e$. The error vector represents all features that are not modeled in the data such as noise, theoretical errors, and modeling errors. Hence, the likelihood function is

$$L(\mathbf{m},\mathbf{C_e},s)=\frac{1}{\pi^N\left|\mathbf{C_e}\right|\exp([\mathbf{q}-s\mathbf{p}(\mathbf{m})]^{\mathrm{H}}\mathbf{C_e}^{-1}[\mathbf{q}-s\mathbf{p}(\mathbf{m})])}, \quad (7)$$

where N is the number of data points and superscript H denotes the complex conjugate transpose. Although in general not true, we assume $\mathbf{C_e} = \nu\mathbf{I}$ to make the derivation more tractable mathematically. The source term s can be estimated in closed form by setting $\frac{\partial \log L}{\partial s}=0$, whereby

$$s_{\mathrm{ML}}=\frac{\mathbf{q}^{\mathrm{H}}\mathbf{p}(\mathbf{m})}{\left\|\mathbf{p}(\mathbf{m})\right\|^2}. \quad (8)$$

It is seen that s depends on $\mathbf{m}$ but not on ν. After substituting s_{ML} back into Eq. (7), the likelihood function is then

$$L(\mathbf{m},\nu)=\frac{1}{\pi^N\nu^N}\exp\left(-\frac{\phi(\mathbf{m})}{\nu}\right), \quad (9)$$

where

$$\phi(\mathbf{m})=\left\|\mathbf{q}\right\|^2\left(1-\frac{\left|\mathbf{q}^{\mathrm{H}}\mathbf{p}(\mathbf{m})\right|^2}{\left\|\mathbf{q}\right\|^2\left\|\mathbf{p}(\mathbf{m})\right\|^2}\right) \quad (10)$$

is the objective function. The maximum likelihood (ML) estimate of the noise ν_{ML} can be estimated in closed form by solving $\frac{\partial \log L}{\partial \nu}=0$, giving

$$\nu_{\mathrm{ML}}=\frac{\phi(\mathbf{m})}{N}. \quad (11)$$

It is straightforward to extend the above to the multi-frequency phone-coherent cost function in Eqs. (3) and (4) and the frequency-coherent cost function in Eqs. (1) and (2).

With the phone-coherent cost function case with constant weighting over frequency, we have, for a particular frequency j,

$$\phi_{H_j}(\mathbf{m})=1-B_{H_j}, \quad (12)$$

where we have dropped the term $\left\|\mathbf{q}_j\right\|^2$ by normalizing $\phi_{H_j}(\mathbf{m})$ as we are using uniform weighting over each frequency, so that for N_F frequencies, the likelihood function is given by

$$L(\mathbf{m},\mathbf{v})=\prod_{j=1}^{N_F}(\pi\nu_j)^{-N_H}\exp\left(-\frac{\phi_{H_j}(\mathbf{m})}{\nu_j}\right). \quad (13)$$

Similarly, with the frequency-coherent cost function case, for a particular hydrophone i, we have

$$\phi_{F_i}(\mathbf{m})=1-B_{F_i}, \quad (14)$$

so that for N_H hydrophones, the likelihood function is given by

$$L(\mathbf{m},\mathbf{v})=\prod_{i=1}^{N_H}(\pi\nu_i)^{-N_F}\exp\left(-\frac{\phi_{F_i}(\mathbf{m})}{\nu_i}\right). \quad (15)$$

The ML solution of the model parameter vector $\mathbf{m}_{\mathrm{ML}}$ is obtained by maximizing the objective function over all $\mathbf{m}$. Finally, an overall estimate for the error power ν is obtained from Eq. (11) at the environmental ML solution: $\nu_{\mathrm{ML}}(\mathbf{m}_{\mathrm{ML}})$ and can be re-inserted into the likelihood function. For simplicity, we consider the error as known and only keep the free argument $\mathbf{m}$ of the objective function ϕ. This approach leads to [19]

$$L(\mathbf{m})=\left(\frac{N}{\pi\phi(\mathbf{m}_{\mathrm{ML}})}\right)^N\exp\left(-N\frac{\phi(\mathbf{m})}{\phi(\mathbf{m}_{\mathrm{ML}})}\right). \quad (16)$$

The above derivation assumes that the error in each sample is uncorrelated with the next sample. In practice, these are strongly correlated because the independent information is limited by the number of propagating modes. Therefore the number of samples N in the above equations must be replaced with the effective number of samples, N_{eff}.

C. Prediction in the TL domain

Probability density functions that describe yet unobserved events are referred to as predictive distributions. Based on the posterior distribution $\mathrm{p}(\mathbf{m}|\mathbf{q})$, the posterior predictive distribution $\mathrm{p}(\mathbf{u}|\mathbf{q})$ is obtained from the joint posterior pdf of $\mathbf{u}$ and $\mathbf{m}$ given $\mathbf{q}$,

$$\begin{aligned}\mathrm{p}(\mathbf{u}|\mathbf{q}) &= \int_M \mathrm{p}(\mathbf{u},\mathbf{m}|\mathbf{q})\,\mathrm{d}\,\mathbf{m}\\ &= \int_M \mathrm{p}(\mathbf{u}|\mathbf{m},\mathbf{q})\,\mathrm{p}(\mathbf{m}|\mathbf{q})\,\mathrm{d}\,\mathbf{m}\end{aligned} \quad (17)$$

where the second equation follows from the definition of conditional probability [15]. Since all uncertainties are assumed to be in the data $\mathbf{q}$ and all information in $\mathbf{q}$ has been mapped into $\mathbf{m}$ (see Fig. 2), conditioning on $\mathbf{q}$ adds no information in our prediction of $\mathbf{u}$. Therefore,

$$\mathrm{p}(\mathbf{u}|\mathbf{m},\mathbf{q}) = \mathrm{p}(\mathbf{u}|\mathbf{m}). \quad (18)$$

The conditional probability density $\mathrm{p}(\mathbf{u}|\mathbf{m})$ is used to describe uncertainties in the forward mapping due to imperfect knowledge of the environment (e.g., parameterization) [16],[17]. Here, the forward mapping is assumed exact: a functional relationship $\mathbf{u} = \mathbf{U}(\mathbf{m})$ gives the transmission loss $\mathbf{u}$ exactly for each value of $\mathbf{m}$. Note that $\mathbf{u} = \mathbf{U}(\mathbf{m})$ is a short notation for the set of equations $u_i = U_i(\mathbf{m})$, i =

1, . . . , *I*. Therefore, the probability density is

$$p(\mathbf{u}|\mathbf{m}) = \delta(\mathbf{U}(\mathbf{m})-\mathbf{u}), \tag{19}$$

where the vector delta function is defined as the product of the delta functions for the elements [16] as in

$$\delta(\mathbf{U}(\mathbf{m}) - \mathbf{u}) = \prod_{i=1}^{I} \delta(U_i(\mathbf{m}) - u(r_i, z_i)). \tag{20}$$

The posterior predictive distribution of **u** for a set of discrete ranges and depths given the observed acoustic data **q** is obtained by integrating the values of the TL with respect to the posterior distribution of the model parameters

$$p(\mathbf{u}|\mathbf{q}) = \int_M \delta(\mathbf{U}(\mathbf{m}) - \mathbf{u})\, p(\mathbf{m}|\mathbf{q})\, d\mathbf{m}, \tag{21}$$

where *M* represents the environmental model domain [15]. It is possible to implement Eq. (21) directly using MCMC methods described in the next subsection. The posterior distribution p(**u**|**q**) carries all the information about TL in the presence of the geoacoustic inversion uncertainties. As the predictive distributions are not necessarily Gaussian, it is preferable to characterize the distributions with medians and distance between the 5-th and 95-th percentiles instead of means and standard deviations. Note that the median corresponds to the 50th percentile of the distribution. The β-th percentile of the TL distribution at a given position, denoted by u^{β}, is computed by finding the TL value that satisfies

$$\int_{-\infty}^{u^{\beta}} p(u|\mathbf{q})\, du = \beta/100. \tag{22}$$

D. Markov chain Monte Carlo methods

Markov chain Monte Carlo (MCMC) is essentially Monte Carlo integration using Markov chains. MCMC methods are a class of algorithms for sampling from probability distributions based on constructing a Markov chain that has the desired distribution as its stationary distribution. In the Bayesian framework, there is often a need to integrate over high-dimensional probability distributions to make inference about model parameters or to make predictions. MCMC methods are able to evaluate integrals in high dimensional space efficiently [20] and have been found to be well suited for problems of Bayesian inference. They are extensively used in various fields of inverse problems, including ocean acoustics [23], [24]. The commonly used MCMC methods are the Metropolis-Hastings algorithm, which was introduced first in [21], and Gibbs sampling, which was developed originally in [22]. MCMC algorithms consist of a random walk in the parameter space where the next parameter value depends only on the current value and not on past values (Markov chain). After an initial "burn-in" period in which the random walker moves toward the high posterior probability region, the set of parameter vectors sampled by the chain will be dependent samples approximately distributed as in the posterior pdf, i.e. the distribution of the samples converges to the posterior pdf after a sufficiently long time.

In the MCMC, samples are generated from the posterior distribution p(**m**|**q**). The difficulty is to create a Markov chain which converges rapidly. As noted in [15], parameter coupling is frequently encountered in ocean acoustics. High correlation between parameters can slow down the convergence of a MCMC sampler considerably. Thus, a parameter covariance matrix estimated from the sampled models during the initial "burn in" period [23] is used for determining appropriate coordinate rotations. MCMC convergence is established by implementing two independent runs in parallel and periodically comparing the marginal distributions of the parameters estimated from each run. The procedure is terminated when the maximum difference between two cumulative marginal distributions for all parameters is less than 0.05 in this paper. A good introduction to MCMC methods is in [25], which also contains many applications in statistical data analysis.

The integral in Eq. (21) is the expectation of the function $\delta(\mathbf{U}(\mathbf{m}) - \mathbf{u})$ with respect to the posterior distribution of the model parameters. This and other expectations can be approximated by using the MCMC samples $\{\mathbf{m}^{(t)}\}$ drawn from the posterior distribution of model parameters p(**m**|**q**), with the probability distribution of the TL

$$p(\mathbf{u}|\mathbf{q}) = \frac{1}{T}\sum_{t=1}^{T} \delta\left(\mathbf{U}\left(\mathbf{m}^{(t)}\right) - \mathbf{u}\right), \tag{23}$$

where the superscript *t* is used to label the sequence of states in a Markov chain and *T* denotes the total length of the sequence. To implement Eq. (23), a numerical approximation is made by binning the calculated TL values. The bin width is selected small enough to have negligible effect on the distribution [15]. In this paper, a 1-dB bin width is used.

Using all samples from MCMC runs can consume a large amount computation time to compute p(**u**|**q**) and storage to save all $\mathbf{m}^{(t)}$. It has been suggested in the statistical literature [25],[26],[27] that inferences should be based on a subsampling of each sequence, with a subsampling factor high enough that successive draws of **m** are approximately independent. This can save a large amount of storage and computation time. This subsampling reduces the number of samples needed to calculate p(**m**|**q**) and thus translates into a large saving in computer time for calculating p(**u**|**q**). Practically, we use a Monte Carlo (random) subsampling of the MCMC samples $\{\mathbf{m}^{(t)}\}$ and monitor the convergence of the maximum difference between the marginal cumulative distributions estimated from subsamples and from all MCMC samples. We define the maximum difference to be less than 0.05 for all parameters in this paper. These subsampled model parameter vectors are then used to compute p(**u**|**q**).

All results presented in this paper are generated by SAGA version 5.3 [11], which implements the MCMC method using the Metropolis-Hastings algorithm described in [28].

IV. Experimental Setup

The NATO Undersea Research Centre (NURC) conducted the MAPEX2000 experiment on 28 November 2000 in a shallow water area north of Elba island off the Italian west coast (see Fig. 3). This area is characterized by a flat bottom covered with clay and sand-clay sediments. The bathymetry was measured to be between 110-120 m along the track (42.928°N, 10.145°E to 42.928°N, 10.260°E). The HLA was towed by NRV Alliance at approximately 4 knots with the first hydrophone approximately 350 m behind the ship's stern. The acoustic source deployed from NRV Manning which was moored at 42.926°N, 10.206°E. The tow depth of the HLA was 55-65 m during the trial. The HLA is 254 m in length, and data recorded along the entire length was used (128 hydrophones spaced at 2 m). A sequence of 2-s LFM sweeps from 150-500 Hz was transmitted every 15 seconds. The received time-series was converted to the frequency domain using a Fast Fourier Transform (FFT) with a frequency bin width of 0.09 Hz. Frequency bins corresponding to 300-500 Hz in 10 Hz increments were used in the inversion for comparison with modeled results.

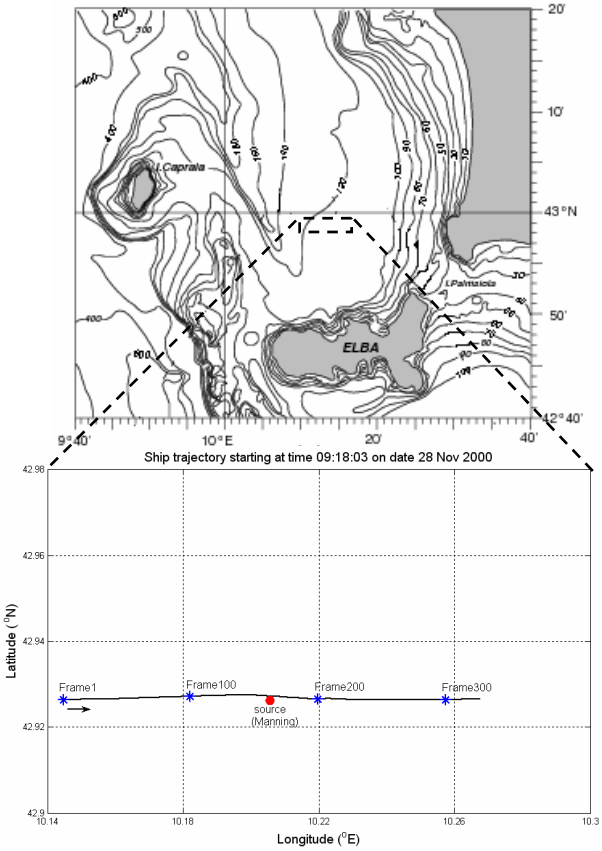


Fig. 3. Bathymetry of experimental location, and track of NRV Alliance during the MAPEX2000 experiment. All times are UTC. Each frame represents a 15-s increment.

The sound speed profile was measured before the experiment and shown in Fig. 4. The profile exhibits a slight positive gradient for most of the water column, except near the bottom, where there is a sharp decrease in sound velocity.

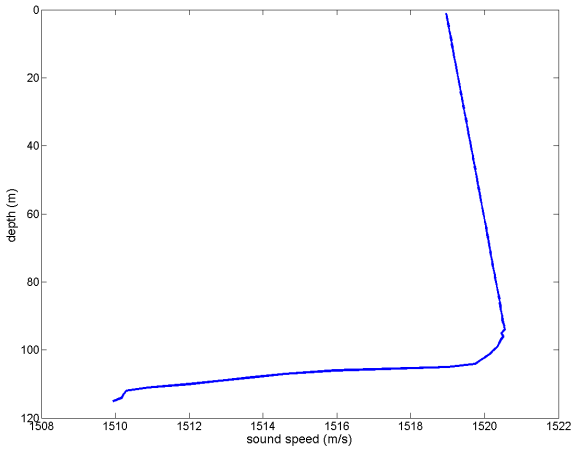


Fig. 4. Sound velocity profile taken from a CTD cast at position 42.943°N 10.127°E on 28 November 2000.

V. HLA Seabed Characterization : A Simulation Study

In this section, the sensitivity of different seabed and geometric parameters using a HLA inversion method is determined via simulations. The sensitivity test is performed by looking at each seabed parameter separately. For a selected parameter, the acoustic fields are computed for all values in its search space. All other parameters are held fixed at their respective reference values. The geometric and seabed parameters considered and their reference values are provided in table I.

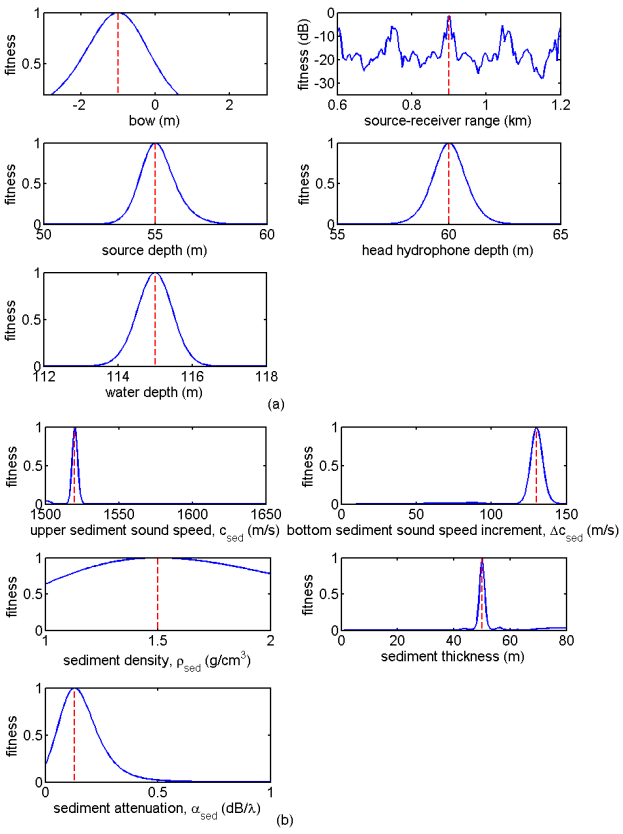
TABLE I
GEOMETRIC AND SEABED PARAMETERS USED IN SENSITIVITY STUDY FOR SOURCE-1ST HYDROPHONE RANGE OF 900 M

Model parameter	Value	Lower search interval	Upper Search Interval
Geometric			
Source range (m)	900	600	1200
Source depth (m)	55	50	60
1st hydrophone depth (m)	60	55	65
Array tilt (°)	0	-3	2
Array bow (m)	1	-3	3
Water depth (m)	115	112	118
Sediment			
Sediment thickness (m)	50	1	80
Sediment density (g/cm^3)	1.5	1.0	2.0
Sediment top comp. speed (m/s)	1520	1520	1650
Comp. speed increment at sediment bottom (m/s)	130	10	150
Attenuation in sediment (dB/λ)	0.13	0.0	1.0
Bottom			
Bottom density (g/cm^3)	1.8	1.0	2.0
Bottom comp. speed as increment of sediment bottom comp. speed (m/s)	100	10	150
Attenuation in bottom (dB/λ)	0.1	0.0	1.0

For the experimental geometry considered, the HLA was positioned with the first hydrophone 900 m from the source. The water depth was taken as 115 m and the source depth at 55 m. The HLA was at 60 m depth and comprises 128 elements with 2-m spacing, spanning 900-1154 m from the

source. The fitness for each parameter, given by 1-C_F, where the cost function C_F is given by Eq. (1), is plotted in Fig. 5.

Referring to Fig. 5(a) for geometric parameters, it can be seen that the HLA has the greatest sensitivity to the source-receiver range. The fitness graph for the sensitivity of the source-receiver range is plotted in dB scale as it falls off rapidly from the actual source-receiver range. We also observe that there are other cyclical peaks due to sidelobes from linear matched-field processing. This is to be expected as range determination of a source to a HLA at endfire is reported to be more ambiguous than for a vertical array [14]. For seabed parameters, the HLA shows good sensitivity to the sediment sound speed and sediment thickness. A caveat to the sensitivity test is the interdependency of each parameter on the others. For example, the sensitivity to sediment thickness and sub-bottom properties will also depend on the sediment properties, as the ability to sense the sub-bottom will depend on the amount of penetration through the sediment. Given the sensitivity curves, and the interdependency of sediment properties, we refine the geoacoustic model by holding the sediment thickness constant, and varying only the top sediment sound speed, and the sediment sound speed gradient. We also hold the bottom density and bottom attenuation fixed as they are insensitive to the inversion. This is because including insensitive parameters increases the computational effort of the inversion without providing more meaningful results, and it is preferable to hold insensitive parameters at physically meaningful values [13]. A full inversion was then performed for both the geometric and seabed parameters.



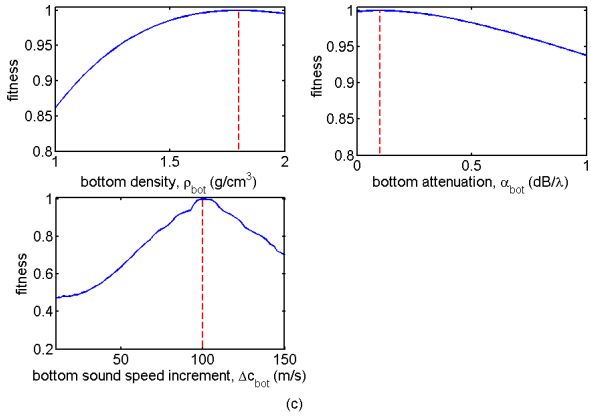


Fig. 5. Sensitivity study of parameters based on simulation. (a) Geometric parameters. (b) Sediment parameters. (c) Bottom parameters.

VI. MAPEX2000 HLA SEABED CHARACTERIZATION : INVERSION RESULTS

During the experiment, a 2-s LFM signal (150-500 Hz) was transmitted from the fixed source at 15-s intervals. This was recorded by the HLA towed by Alliance. Fig. 6 shows the signal received on the last hydrophone of the HLA at frame 185. Before inversion, the received acoustic data were calibrated by dividing by the ideal 2-s LFM source spectrum over the 150-500 Hz frequency band [13]. The frequency bins used for inversion are corresponding to the frequencies Doppler-shifted according to the Doppler factor of (1+v/c), where c is the sound velocity, and v is the ship velocity (positive when moving towards the source, and negative when moving away) [6]. We follow a similar approach as in [7] in performing the inversions. First, we perform a general inversion for all the parameters. We then constrain the geometric parameters by allowing them to vary only within narrow intervals, and perform a full inversion of the seabed properties. This prevents the highly sensitive geometric parameters from dominating the inversion.

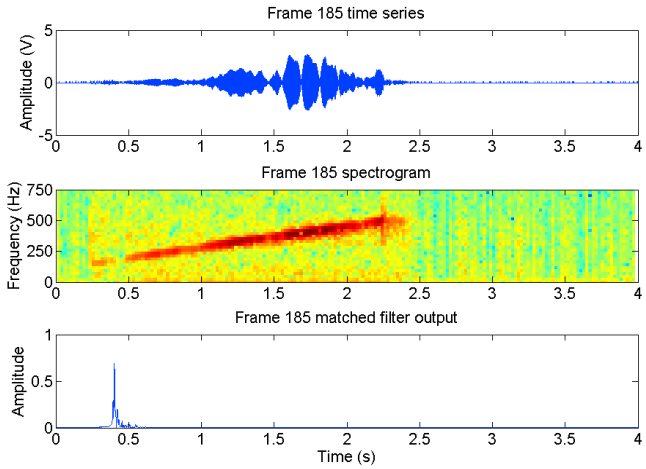


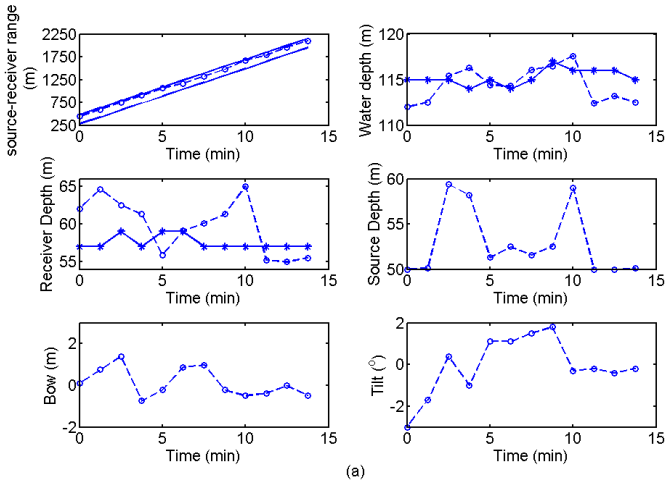
Fig. 6. Data received on the last hydrophone on the HLA at frame 185. (a) Time series. (b) Spectrogram of the received LFM sequence. (c) Matched filter output showing the arrival time of the LFM pulse.

Table II shows the geometric and geoacoustic parameters and their search bounds used in the final inversion, where the geometric parameters are only allowed to vary in a narrow interval. We invert for selected seabed parameters which are dominant in propagation modeling (sediment top speed, increase of sediment speed at bottom, and sediment attenuation) while holding the rest of the parameters, which are insensitive (sediment density and all other bottom parameters) or coupled to other parameters (e.g. sediment thickness, which is coupled to sediment speed), constant.

TABLE II
GEOMETRIC AND SEABED PARAMETERS USED IN FINAL INVERSIONS

Model parameter	Value	Lower search interval	Upper Search Interval
Geometric			
Source range (m)	From initial inversion results	-5	+5
Source depth (m)		-2.5	+2.5
1st hydrophone depth (m)		-2.5	+2.5
Array tilt (°)		-1	+2
Array bow (m)		-1	+2
Water depth (m)	115	112	118
Sediment			
Sediment top comp. speed (m/s)	1520	1500	1600
Comp. speed increment at Sediment bottom (m/s)	130	10	150
Attenuation in sediment (dB/λ)	0.13	0.0	1.0

Fig. 7 shows the variation of the inverted parameters with time, as Alliance traveled along the track from frame 185 (10:04:00) to frame 240 (10:17:45). The geometric parameters such as source-receiver range, water depth, receiver depth and source depth have been inverted within the bounds of errors in measurement. It should also be noted that the inversion results for other parameters represent "averaged" values over range, rather than results at a particular location.



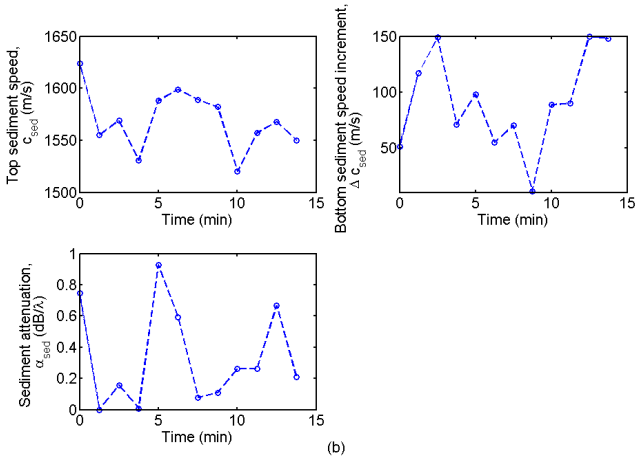


Fig. 7. Variation of measured and inverted geometric and sediment properties with time along the track. (a) Geometric parameters. (b) Sediment parameters. The 'o' indicates the values obtained after inversion, and '*' indicates the measured values where available. The source-receiver ranges obtained from inversion are well-determined along the track, and plotted with the upper and lower intervals of the measured range (solid lines) using GPS with an error of ±100 m.

Fig. 8 shows the 1-D posterior probability distribution plots of the geometric and sediment parameters for an inversion performed at frame 195. The plot indicates that uncertainty exists in the parameters obtained from geoacoustic inversion results. In the next section, we discuss how the uncertainties in the estimation of the sea-bed parameters resulting from the geoacoustic inversion are used to map to uncertainties in the transmission loss domain.

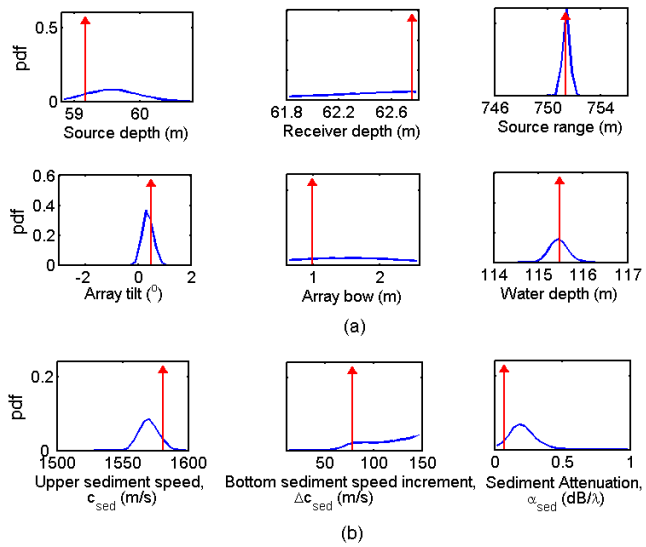


Fig. 8. *A posteriori* probability distributions for (a) geometric parameters and (b) sediment parameters, obtained from inversion for frame 195. The red arrow shows the MAP value.

VII. TRANSMISSION LOSS (TL) ESTIMATION FROM EXPERIMENTAL INVERSION RESULTS

We apply the MCMC method presented in section III to the geoacoustic inversion results obtained from the preceding section to obtain the TL uncertainty plot as a result of

uncertainty in the environmental parameters. Fig. 9(a) shows the TL uncertainty plot at a receiver depth of 60 m over range at a frequency of 300 Hz. The plot shows peaks and troughs in the TL due to the constructive and destructive interference of the modes propagating in the water column. Fig. 9 (b) and (c) shows the uncertainty spread in TL at a region of constructive and destructive interference at 1530 and 1640 m respectively. We see that the TL uncertainty band is about 5 dB around the regions of constructive interference and widens to approximately 18 dB around the regions of destructive interference, i.e. regions of destructive interference are predicted with much more uncertainty than at regions of constructive interference.

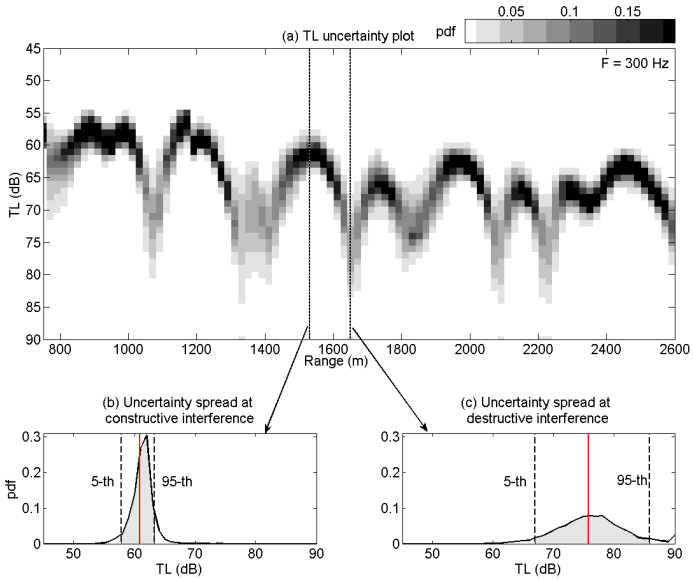


Fig. 9. (a) TL uncertainty plot at a receiver depth of 60 m at 300 Hz. (b) Uncertainty spread of TL at a range of 1530 m where there is constructive interference. (c) Uncertainty spread of TL at a range of 1640 m where there is destructive interference.

In Fig. 10, we plot the predicted TL in terms of the median (solid red line) and bounded by the 90% credibility interval (CI) (gray region) for the 3 frequencies 300 Hz, 400 Hz and 500 Hz. We also insert the measured TL values in the same plot to see if the TL values match the predicted TL within the 90% CI. We observe that approximately 70% of the measured TL values fall within the predicted TL's 90% CI uncertainty envelope for 300 Hz and 400 Hz, while for 500 Hz, the prediction is less accurate, with only 52% of the measured TL values falling within the predicted TL's 90% CI uncertainty envelope. This is likely due to the increased difficulty in prediction as frequency increases, which leads to a corresponding increase in the number of propagating modes and their interaction with each other.

VIII. Conclusions

In this paper, we have demonstrated the inversion of seabed parameters using data received on a towed array with a frequency-coherent cost function. The Markov chain Monte Carlo method was then applied to sample the posterior probability density of the geoacoustic parameters. Then, these parameter uncertainties are translated to the transmission loss domain, where the probability distribution of transmission loss over different range and frequencies are obtained. We also extracted the characteristic features such as the median and lower/upper percentiles from the distribution.

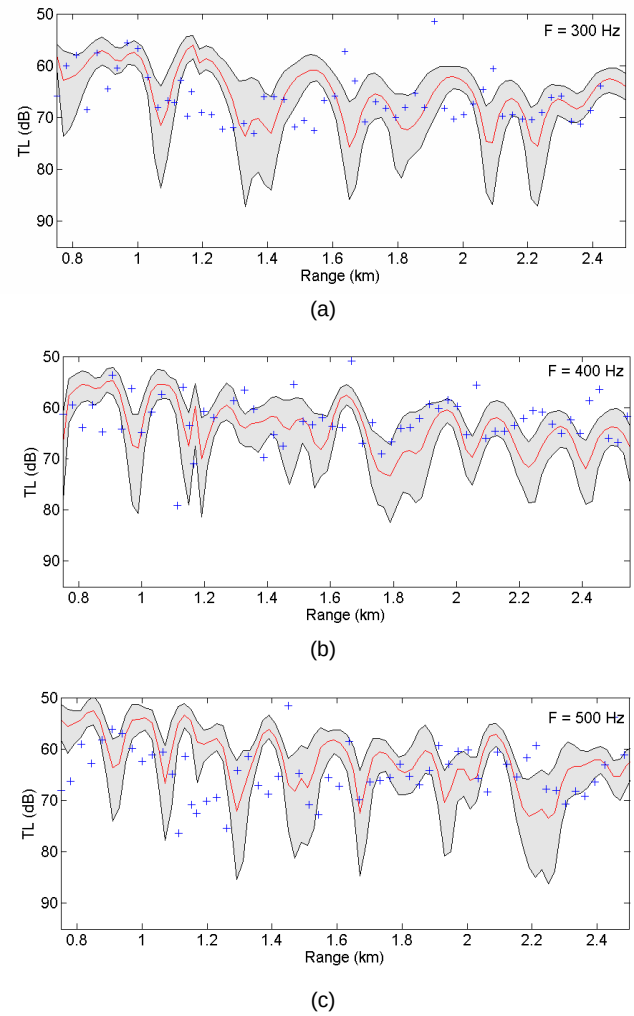


Fig. 10. Predicted and measured TL at depth 60 m and for frequencies (a) 300 Hz, (b) 400 Hz and (c) 500 Hz. The median of the predicted TL (red line) is shown together with the 90% CI (gray area). The measured TL values are marked with crosses.

Acknowledgment

We would like to thank the NATO Undersea Research Centre for the provision of the experimental data.

References

[1] D. F. Gingras and P. Gerstoft, "Inversion of geometric and geoacoustic parameters in shallow water: Experimental resuls," *J. Acoust. Soc. Am.* 95, pp. 770-782, 1994.

[2] D. M. F. Chapman, "What are we inverting for," in *Inverse Problems in Underwater Acoustics*, edited by M. I. Taroudakis and G. N. Makrakis, pp. 1-14. Springer-Verlag, New York, 2001

[3] P. Abbot and I. Dyer, "Sonar performance predictions based on environmental variability," in *Impact of Littoral environmental variability on Acoustic Predictions and Sonar Performance*, edited by N.

G. Pace and F. B. Jensen, pp. 611-618. Kluwer Academic Publishers, The Netherlands, 2002.

[4] P. Gerstoft and D. F. Gingras, "Parameter estimation using multifrequency range-dependent acoustic data in shallow water," *J. Acoust. Soc. Am.* 99, pp. 2839-2850, 1996.

[5] L. T. Fialkowski, T. C. Yang, K. Yoo, E. Kim, and D. K. Dacol, "Consistency and reliability of geoacoustic inversions with a horizontal line array," *J. Acoust. Soc. Am.* 120, pp. 231-246, 2006.

[6] B. R. Mahafza, *Radar Systems Analysis and Design Using Matlab*, pp. 12-19. Chapman & Hall/CRC, Taylor and Francis Group, 2005.

[7] M. Siderius, P. L. Nielsen and P. Gerstoft, "Range-dependent seabed characterization by inversion of acoustic data from a towed receiver array," *J. Acoust. Soc. Am*. 112, pp. 1523-1535, 2002.

[8] D. J. Battle, P. Gerstoft, W. A. Kuperman, W. S. Hodgkiss, and M. Siderius, "Geoacoustic inversion of tow-ship noise via near-field-matched field processing," *IEEE J. Oceanic Eng*., vol. 28, pp. 454-467, 2003.

[9] P. Gerstoft, C. F. Huang and W. S. Hodgkiss, "Estimation of transmission loss in the presence of geoacoustic inversion uncertainty," *IEEE J. Oceanic Eng*., vol. 31, pp.299-307, 2006.

[10] F. B. Jensen and M. C. Ferla, "SNAP-The SACLANTCEN normal mode acoustic propagation model," SACLANT Undersea Research Centre, SM-121, La Spezia, Italy, 1979.

[11] P. Gerstoft, SAGA Users Guide 5.0, an inversion software package. An updated version of "SAGA Users guide 2.0, an inversion software package," SACLANT Undersea Research Centre, SM-333, La Spezia, Italy, 1997.

[12] P. Gerstoft, "Inversion of seismoacoustic data using genetic algorithms and *a posteriori* probability distributions," *J. Acoust. Soc. Am.* 95, pp. 770-782, 1994.

[13] M. R. Fallat, P. L. Nielsen, S. E. Dosso, and M. Siderius, "Geoacoustic Characterization of a Range-Dependent Ocean Environment Using Towed Array Data," *IEEE J. Oceanic Eng*., vol. 30, pp. 198-206, 2005.

[14] A. Tolstoy, *Matched Field Processing for Underwater Acoustics*, pp. 134-145. World Scientific Publishing, Singapore, 1993.

[15] C. F. Huang, P. Gerstoft, and W. S. Hodgkiss, "Validation of statistical estimation of transmission loss in the presence of geoacoustic inversion uncertainty," *J. Acoust. Soc. Am.* 120, pp. 1931-1941, 2006.

[16] A. Tarantola, *Inverse Problem Theory and Methods for Model Parameter Estimation*. SIAM, Philadelphia, 2005.

[17] K. Mosegaard and A. Tarantola, "Probabilistic approach to inverse problems," in *International Handbook of Earthquake & Engineering Seismology, Part A*, pp. 237–265. Academic Press, London, 2002.

[18] C. F. Mecklenbrauker and P. Gerstoft, "Objective functions for ocean acoustic inversion derived by likelihood methods," *J. Comp. Acoust*, vol. 8, pp. 259-270, 2000.

[19] P. Gerstoft and C. F. Mecklenbrauker, "Ocean acoustic inversion with estimation of a posteriori probability distributions," *J. Acoust. Soc. Am.* 104, pp. 808-819, 1998.

[20] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes in Fortran 77*, 2nd edition. Cambridge University Press, London, 1992.

[21] N. Metropolis, A.W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, "Equation of state calculations by fast computing machines," *Journal of Chemistry and Physics*, vol. 21, pp. 1087-1092, 1953.

[22] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 6, pp. 721-741, 1984.

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

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

[25] W. R. Gilks, S. Richardson, and D. J. Spiegelhalter, *Markov Chain Monte Carlo in Practice*. Chapman and Hall, London, 1996.

[26] C. J. Geyer, "Practical Markov chain Monte Carlo (with discussion)," *Statistical Science* 7, 473-483, 1992.

[27] C. P. Robert and G. Casella, *Monte Carlo Statistical Methods*. Springer-Verlag, New York, 1999.

[28] C. Yardim, P. Gerstoft, and W. S. Hodgkiss, "Estimation of radio refractivity from radar clutter using Bayesian Monte Carlo analysis," *IEEE Trans. on Antennas and Propagation*, vol. 54, pp. 1318-1327, 2006.