Statistical estimation of transmission loss from geoacoustic inversion using a towed array

Yong Han Goh,[a] Peter Gerstoft, and William S. Hodgkiss
Marine Physical Laboratory, Scripps Institution of Oceanography, La Jolla, California 92093-0238, USA

Chen-Fen Huang
Department of Marine Environmental Informatics, National Taiwan Ocean University, Keelung, Taiwan

Geoacoustic inversion estimates environmental parameters from measured acoustic fields (e.g., received on a towed array). The inversion results have some uncertainty due to noise in the data and modeling errors. Based on the posterior probability density of environmental parameters obtained from inversion, a statistical estimation of transmission loss (TL) can be performed and a credibility level envelope or uncertainty band for the TL generated. This uncertainty band accounts for the inherent variability of the environment not usually contained in sonar performance prediction model inputs. The approach follows [Gerstoft *et al.* IEEE J. Ocean. Eng. **31**, 299–307 (2006)] and is demonstrated with data obtained from the MAPEX2000 experiment conducted by the NATO Undersea Research Center using a towed array and a moored source in the Mediterranean Sea in November 2000. Based on the geoacoustic inversion results, the TL and its variability are estimated and compared with the measured TL. © *2007 Acoustical Society of America.*
[DOI: 10.1121/1.2782915]

I. INTRODUCTION

Geoacoustic inversion using matched-field processing is a model-based technique that has been applied successfully to derive environmental and seabed parameters for propagation prediction.[1-4] 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. For example, if the results of the inversion are only required for sonar performance prediction, 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 usually is sufficient to describe a simpler "effective" seabed model having a similar acoustic effect on the underwater sound field within the range-frequency domain of interest.[5]

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 use of horizontal line arrays (HLAs) has been gaining in popularity as it offers several advantages over a VLA. This includes 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,[2,6,7] or using the ship self-noise for inversion.[8,9] Others have used bottom-moored horizontal line arrays and a towed source.[3,10] In this paper, a towed array with a fixed source was used for geoacoustic inversion, which to the authors' knowledge, is a setup that has not been presented in current literature on geoacoustic inversion.

There are uncertainties associated with the seabed parameters obtained from geoacoustic inversion, and in this paper, the mapping of these uncertainties to the transmission loss domain is also presented. This has been performed using results from a VLA with a moving source.[11,12] In this paper, we present results obtained using a HLA with a stationary source. Sonar performance prediction using a probability density function based on environmental variability has also been discussed in Refs. [13-15].

II. PARAMETER ESTIMATION USING GENETIC ALGORITHMS AND INVERSION PROCEDURE

A. Base line environment model

The base line environmental model established for the North Elba site from Ref. [16] was used. Figure 1 illustrates the base line environmental model, comprising the water column, a constant thickness sediment layer with depth dependent sound speed increasing with depth, and a bottom half-space layer.

B. Inversion procedure and cost function

The inversion is performed as follows:
(1) Record the acoustic field at the site of interest.

---

[a]Author to whom correspondence should be addressed. Now with DSO National Laboratories, 20 Science Park Drive, Singapore 118230. Electronic mail: gyonghan@dso.org.sg.
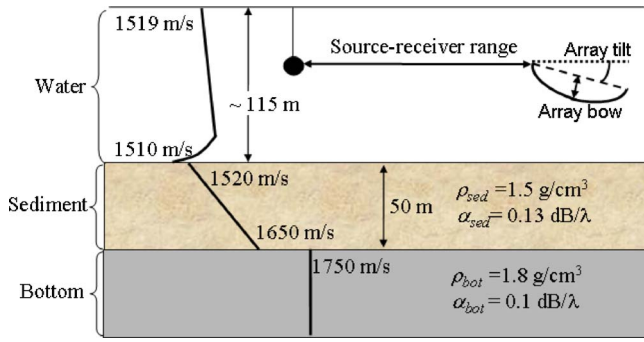
FIG. 1. (Color online) The base line environmental model for the North Elba experiment site (see Ref. 16).

FIG. 2. An observation $\mathbf{q}$ ($\in \mathcal{Q}$) is mapped into a distribution of environmental parameters $\mathbf{m}$ ($\in \mathcal{M}$) that potentially could have generated it. These environmental parameters are then mapped into the usage domain $\mathcal{U}$.

(2) Choose a suitable propagation model. In this paper, the NURC SNAP normal-mode propagation model,[17] which is robust, fast, and suitable for low frequencies, was used.

(3) Choose a suitable cost function to minimize. The phone-coherent method traditionally has been used. However, recent analysis[2,7,18] has shown that frequency-coherent matched-field inversion works well for multi-frequency HLA data when the source spectrum is known, where matched-field correlations between data and replica frequency vectors are performed coherently over frequency but incoherently over range (hydrophones). The frequency-coherent cost function is derived using the maximum likelihood method in Sec. III A and is given by

$$C_F = \left[\prod_{i=1}^{N_H} (1 - \phi_{F_i}) \right]^{-N_H}, \tag{1}$$

where

$$\phi_{F_i} = \frac{\left| \sum_{j=1}^{N_F} [s(\omega_j) a_{ij}]^* q_{ij} \right|^2}{\sum_{j=1}^{N_F} |s(\omega_j) a_{ij}|^2 \sum_{j=1}^{N_F} |q_{ij}|^2}. \tag{2}$$

In Eqs. (1) and (2), N_F is the number of frequency components, N_H is the number of hydrophones, a_{ij} corresponds to the modeled Green's function (* denotes the complex conjugate), q_{ij} is the measured complex pressure, and $s(\omega_j)$ is the complex source term at the jth frequency. In Eq. (2), the correlation is between the measured and modeled acoustic pressure vectors. The cost function takes on a value of 0 for two identical signals and 1 for completely uncorrelated signals.

Equation (2) may be expressed equivalently as

$$\phi_{F_i} = \frac{\left| \sum_{j=1}^{N_F} a_{ij}^* \frac{q_{ij}}{s(\omega_j)} \right|^2}{\sum_{j=1}^{N_F} |a_{ij}|^2 \sum_{j=1}^{N_F} \left| \frac{q_{ij}}{s(\omega_j)} \right|^2}. \tag{3}$$

In Eq. (3), the correlation is between the measured and modeled Green's function, where the measured Green's function is obtained by dividing the measured acoustic data q_{ij} by the known source spectrum over the frequency band. Equation (2) is preferred as Eq. (3) may suffer from numerical instabilities when $s(\omega_j)$ is close to zero. Note that the source term
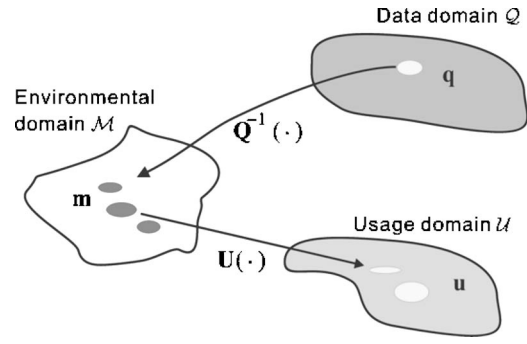
needs to be explicitly included in the frequency-coherent cost function in both Eqs. (2) and (3).

The frequency-coherent cost function was used in this paper, as we found the phone-coherent method yields worse results, as in Refs. 2, 7, and 18.

(4) An efficient algorithm is needed to navigate the enormous search space and find the global minimum of the cost function. In this paper, a genetic algorithm search is used with the propagation model SNAP as implemented in the inversion package, SAGA,[19] 40,000 forward models were used in the inversion searches. Reference 20 provides a detailed description of GA and their application to geoacoustic parameter estimation.

III. MAPPING GEOACOUSTIC PARAMETER UNCERTAINTIES TO TRANSMISSION LOSS DOMAIN

The mapping of geoacoustic parameter uncertainties to the transmission loss (TL) domain has been described in Refs. 11 and 12 and is summarized in this section for completeness.

Figure 2 summarizes the estimation of TL (usage domain $\mathcal{U}$) from ocean acoustic data observed on a vertical or horizontal array (data domain $\mathcal{Q}$).[11] Based on the ocean acoustic data, we statistically characterize TL, the usage domain $\mathcal{U}$. The vector $\mathbf{q}_i$ represents the acoustic data observed at the ith hydrophone and the vector $\mathbf{u}$ represents TL at several ranges and depths. This is mapped via a set of environmental parameters $\mathbf{m}$ in the environmental domain $\mathcal{M}$. Both the experimental data $\mathbf{q}_i$ and the usage domain quantity $\mathbf{u}$ are related to $\mathbf{m}$ via forward models $\mathbf{Q}(\mathbf{m})$ and $\mathbf{U}(\mathbf{m})$. The geoacoustic inverse problem is solved as an intermediate step to obtain the posterior distribution of environmental parameters $p(\mathbf{m}|\mathbf{q}_i)$. We are not just interested in the environment itself, but also a statistical estimation of the TL field. Based on the posterior distribution $p(\mathbf{m}|\mathbf{q}_i)$, the probability distribution of the TL $p(\mathbf{u}|\mathbf{q}_i)$ is obtained via Monte Carlo integration. From this TL probability distribution, all relevant statistics of TL such as the median, percentiles and correlation coefficients can be obtained. The vector $\mathbf{u}$ is used to denote the transmission loss as a K-dimensional vector at discrete (r_k, z_k) positions, where $u_k = u(r_k, z_k)$.

Goh *et al.*: Statistical estimation of transmission loss

A. Frequency-coherent likelihood and cost functions

The likelihood function for the frequency-coherent model [Eq. (4)] is derived following a similar approach as in Refs. 21 and 22 which focused on the phone-coherent model.

In the case of the coherent Bartlett processor, the source strength is known from frequency to frequency in a relative sense, but the absolute amplitude and phase are unknown. This is often the case, for example, when a Linear Frequency Modulation (LFM) signal is transmitted, where the absolute amplitude and phase are unknown. At a single phone, the relation between the observed complex-valued data vector $\mathbf{q}_i = [q_i(\omega_1) \ldots q_i(\omega_{N_F})]^T$ and the modeled data may be described by the model[23]

$$\mathbf{q}_i = \alpha \mathbf{S} \mathbf{a}_i(\mathbf{m}) + \mathbf{e}. \tag{4}$$

Here $\mathbf{S}$ is the diagonal matrix with its diagonal being the source vector at zero phase,

$$\mathbf{S} = \begin{bmatrix} s(\omega_1) & 0 & \ldots & 0 \\ 0 & s(\omega_2) & \ldots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \ldots & s(\omega_{N_F}) \end{bmatrix},$$

and $\alpha = Ae^{i\theta}$ represents the unknown amplitude scaling and phase shift which are independent of frequency. The Green's function $\mathbf{a}_i(\mathbf{m})$ at the corresponding N_F frequencies is obtained using an acoustic propagation model and an environmental model $\mathbf{m}$. The error term is represented by $\mathbf{e}$.

In the Bayesian inference framework, the solution to the inverse problem is given by

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

where $\mathcal{L}(\mathbf{m})$ is used to denote the likelihood function $p(\mathbf{q}_i|\mathbf{m})$.

Following the derivation in Ref. 11 where the error vector $\mathbf{e}$ is assumed to be Gaussian distributed with zero mean and covariance $\mathbf{C}_e = \nu \mathbf{I}$, the likelihood function is given by

$$\mathcal{L}(\mathbf{m}, \nu, \mathbf{S}, \alpha) = \frac{1}{\pi^{N_F} \nu^{N_F}} \exp\left(-\frac{\|\mathbf{q}_i - \alpha \mathbf{S} \mathbf{a}_i(\mathbf{m})\|^2}{\nu}\right). \tag{6}$$

Here, ν is assumed to be constant over the frequency band.

The maximum-likelihood estimate of α is obtained by solving $\frac{\partial \log \mathcal{L}}{\partial \alpha} = 0$, giving

$$\alpha_{\mathrm{ML}} = \frac{[\mathbf{S} \mathbf{a}_i(\mathbf{m})]^\dagger \mathbf{q}_i}{\|\mathbf{S} \mathbf{a}_i(\mathbf{m})\|^2}. \tag{7}$$

Substituting α_{ML} back into Eq. (6), the likelihood function is then

$$\mathcal{L}(\mathbf{m}, \nu) = \frac{1}{\pi^N \nu^N} \exp\left(-\frac{\phi_{F_i}(\mathbf{m})}{\nu}\right), \tag{8}$$

where

$$\phi_{F_i}(\mathbf{m}) = \|\mathbf{q}_i\|^2 - \frac{|[\mathbf{S} \mathbf{a}_i(\mathbf{m})]^\dagger \mathbf{q}_i|^2}{\|[\mathbf{S} \mathbf{a}_i(\mathbf{m})]\|^2} \tag{9}$$

is the cost function for the ith hydrophone. For a single receiver, the cost function in Eq. (1) can be obtained by normalizing Eq. (9) with $\|\mathbf{q}_i\|^2$. For N_H hydrophones, the likelihood function may be written as

$$\mathcal{L}(\mathbf{m}, \boldsymbol{\nu}) = \prod_{i=1}^{N_H} (\pi \nu_i)^{-N_F} \exp\left(-\frac{\phi_{F_i}(\mathbf{m})}{\nu_i}\right), \tag{10}$$

where the noise variance ν_i is assumed to be different at different hydrophones.

For an HLA, the noise variance at the first hydrophone could be different from the last hydrophone due to distance from the towing vessel and different flow noise. The error variance $\boldsymbol{\nu}$ may be treated as a nuisance parameter and eliminated via integrating Eq. (10) weighted by a noninformative prior of $\boldsymbol{\nu}[p(\nu_i) \propto 1/\nu_i]$ over its entire range[24]

$$\mathcal{L}(\mathbf{m}) = \int_0^\infty \cdots \int_0^\infty \mathcal{L}(\mathbf{m}, \boldsymbol{\nu}) p(\nu_1) \ldots p(\nu_{N_H}) d\nu_1 \ldots d\nu_{N_H}. \tag{11}$$

Integrating out $\boldsymbol{\nu}$, the likelihood function for the frequency-coherent cost function case is then

$$\mathcal{L}(\mathbf{m}) = \frac{1}{\pi^{N_F}} \prod_{i=1}^{N_H} \frac{(N_F - 1)!}{\phi_{F_i}(\mathbf{m})^{N_F}} \tag{12}$$

or,

$$\mathcal{L}(\mathbf{m}) \propto \left[\frac{1}{\phi_{F_i}^g(\mathbf{m})}\right]^{N_F N_H} = \left[\prod_{i=1}^{N_H} \phi_{F_i}(\mathbf{m})\right]^{-N_F}, \tag{13}$$

where $\bar{\phi}_{F_i}^g(\mathbf{m}) = \sqrt[N_H]{\prod_{i=1}^{N_H} \phi_{F_i}(\mathbf{m})}$ is the geometric mean of the cost function over N_H hydrophones. The cost function defined in Eq. (2) is thus derived from the likelihood function $\mathcal{L}(\mathbf{m})$.

The above derivation assumes that the errors are independent across both spatial samples of the acoustic field and frequencies. In practice, these errors may be strongly correlated, for example, when the dominant source of errors is due to frequency-dependent modeling mismatch, the modeling errors may not be independent across the frequencies used. In this case, the full data error covariance matrix $\mathbf{C}_e$, which would be nondiagonal in general, should be used.[25]

B. Prediction in the TL domain

Probability density functions that describe yet unobserved events are referred to as predictive distributions. The posterior predictive distribution of $\mathbf{u}$ for a set of discrete ranges and depths given the observed acoustic data $\mathbf{q}_i$ may be obtained by integrating the values of the TL with respect to the posterior distribution of the model parameters[11]

J. Acoust. Soc. Am., Vol. 122, No. 5, November 2007

Goh *et al.*: Statistical estimation of transmission loss   2573

$$p(\mathbf{u}|\mathbf{q}_i) = \int_{\mathcal{M}} \delta(\mathbf{U}(\mathbf{m}) - \mathbf{u})p(\mathbf{m}|\mathbf{q}_i)d\mathbf{m}. \qquad (14)$$

It is possible to implement Eq. (14) directly using Markov chain Monte Carlo (MCMC) methods described in the next subsection. The posterior distribution $p(\mathbf{u}|\mathbf{q}_i)$ 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 5th and 95th percentiles instead of means and standard deviations. The βth percentile of the TL distribution at a given position, denoted by u^β, is computed by finding the TL value that satisfies

$$\int_{-\infty}^{u^\beta} p(u|\mathbf{q}_i)du = \beta/100. \qquad (15)$$

The present formulation in the usage domain, Eq. (14), has the advantage that it is easy to incorporate additional *independent* information about the environment $\mathbf{m}$, as carried out in Eqs. (16)–(18). If the additional parameters $\mathbf{m}_a$ with probability density $p(\mathbf{m}_a)$ are independent of $\mathbf{m}$, then

$$p(\mathbf{u}|\mathbf{q}_i) = \int_{\mathcal{M}_a} \delta(\mathbf{U}(\mathbf{m},\mathbf{m}_a) - \mathbf{u})p(\mathbf{m}|\mathbf{q}_i)p(\mathbf{m}_a)d\mathbf{m}d\mathbf{m}_a, \qquad (16)$$

where $\mathcal{M}_a$ is the environmental domain spanned by $\mathbf{m}$ and $\mathbf{m}_a$. Equation (16) is used in the computation of the influence of sound speed variability on TL prediction in Sec. VI.

In addition, suppose the posterior distribution $p(\mathbf{m}_i|\mathbf{q}_i)$ of a parameter m_i from $\mathbf{m}$ is not the correct distribution to be used in mapping to the usage domain (e.g., assumption of constant source depth throughout the trial). Under the assumption that this parameter is independent of the other parameters and the observed data vector $\mathbf{q}_i$, $p(m_i|\mathbf{q}_i)$ can be replaced with an alternative distribution $p(m_i)$. Then

$$p(\mathbf{u}|\mathbf{q}_i) = \int_{\mathcal{M}} \delta(\mathbf{U}(m_i,\mathbf{m}_{-i}) - \mathbf{u})p(\mathbf{m}_{-i}|\mathbf{q}_i)p(m_i)dm_id\mathbf{m}_{-i}, \qquad (17)$$

where $\mathbf{m}_{-i}$ is the parameter vector $\mathbf{m}$ with its ith component removed. This is used for including uncertainty in the source depth using statistics from the matched-field estimated source depths in the TL prediction in Sec. VI.

Finally, not all parameters in $\mathbf{m}$ are required for input in the forward mapping $\mathbf{u} = \mathbf{U}(\mathbf{m})$. Only a subset $\mathbf{m}'$ is required, e.g., receiver geometric parameters such as receiver range, receiver depth, array bow, and array tilt are not used in the forward mapping. Then $\mathbf{u} = \mathbf{U}(\mathbf{m}) = \mathbf{U}(\mathbf{m}')$ and

$$p(\mathbf{u}|\mathbf{q}_i) = \int_{\mathcal{M}'} \delta(\mathbf{U}(\mathbf{m}') - \mathbf{u})p(\mathbf{m}'|\mathbf{q}_i)d\mathbf{m}', \qquad (18)$$

where $\mathcal{M}'$ is the environmental domain spanned by $\mathbf{m}'$ and $p(\mathbf{m}'|\mathbf{q}_i)$ can be obtained from $p(\mathbf{m}|\mathbf{q}_i)$ by integrating out the parameters which are not required.

C. 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[26,27] 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.[9,28] The commonly used MCMC methods are the Metropolis-Hastings algorithm,[29,30] and Gibbs sampling.[31]

The integral in Eq. (14) is the expectation of function $\delta(\mathbf{U}(\mathbf{m}) - \mathbf{u})$ with respect to the posterior distribution of the model parameters. It can be approximated by using the MCMC samples $\{\mathbf{m}^{(t)}\}$ drawn from the posterior distribution of model parameters $p(\mathbf{m}|\mathbf{q}_i)$,

$$p(\mathbf{u}|\mathbf{q}_i) = \frac{1}{T}\sum_{t=1}^{T} \delta(\mathbf{U}(\mathbf{m}^{(t)}) - \mathbf{u}), \qquad (19)$$

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. Equation (19) is implemented using a numerical approximation by binning the calculated TL values. The bin width should be selected small enough to have negligible effect on the distribution. In this paper, a 1 dB bin width is used.

Using all samples from MCMC runs would consume a large amount of computation time to compute $p(\mathbf{u}|\mathbf{q}_i)$ and memory storage to save all the samples $\{\mathbf{m}^{(t)}\}$.[32] Statistical literature has suggested that inferences could be based on a subsampling of each sequence, with a subsampling factor high enough that successive draws of $\mathbf{m}$ are approximately independent.[27,32,33] The subsampling reduces the number of samples needed to calculate $p(\mathbf{m}|\mathbf{q}_i)$, translating into a large saving in computer time for calculating $p(\mathbf{u}|\mathbf{q}_i)$. The subsampled model parameter vectors then are used to compute $p(\mathbf{u}|\mathbf{q}_i)$.[11]

All results presented in this paper are generated by SAGA,[19] which implements the MCMC method using the Metropolis-Hastings algorithm described in Ref. 34.

IV. EXPERIMENTAL SETUP

The experiment was conducted as part of the MAPEX2000 experiment by the NATO Undersea Research Center (NURC) 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 and 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
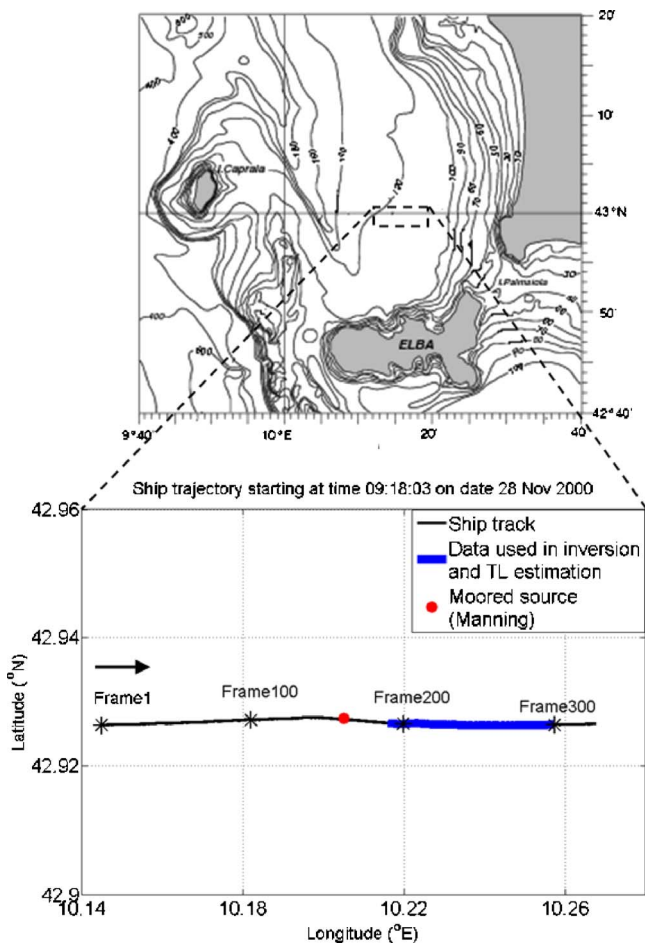
FIG. 3. (Color online) Bathymetry of experimental location, and track of NRV Alliance during the MAPEX2000 experiment. All times are Coordinated Universal Time (UTC). Each frame represents a 15 s increment.

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 to 500 Hz was transmitted every 15 s. The received time series was converted to the frequency domain using a fast Fourier transform 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.

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 speed.

V. 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 NRV Alliance.

Before inversion, the received acoustic data were deconvolved by dividing by the ideal 2 s LFM source spectrum over the 150–500 Hz frequency band.[7] The frequency bins used for inversion correspond to the frequencies Doppler
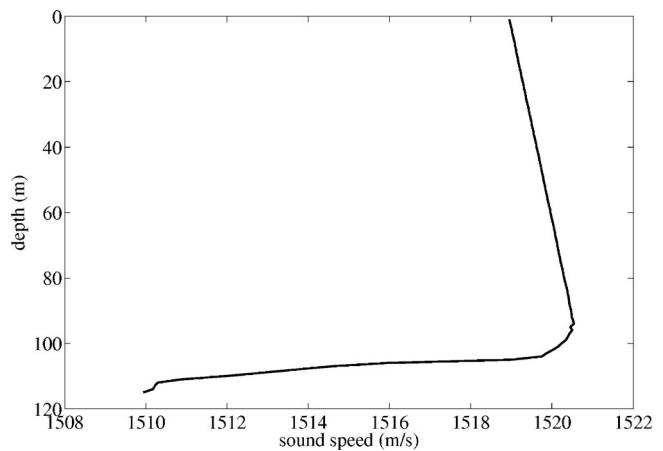

FIG. 4. Sound speed profile taken from a Conductivity Temperature Depth (CTD) cast at position (42.943° N 10.127° E) on 28 November 2000.

shifted according to the Doppler factor of $(1+v/c)$, where c is the sound speed, and v is the ship velocity (positive when moving towards the source, and negative when moving away). With the ship speed at 2 m/s, this corresponds to a maximum shift of about 8 bins at 500 Hz.

Figure 5 shows the one-dimensional posterior probability distribution plots of the geometric and sediment parameters for an inversion performed at frame 195 where the range between the source and the first hydrophone is 750 m. The plots indicate the uncertainty existing in the parameters obtained from the geoacoustic inversion results.

VI. TRANSMISSION LOSS (TL) ESTIMATION FROM INVERSION RESULTS

In this section, the mapping of the uncertainties in the estimation of the seabed parameters resulting from the geoacoustic inversion to uncertainties in the transmission loss domain is discussed.

A. Predictive distribution of transmission loss

The MCMC method using the Metropolis-Hastings algorithm is applied to the geoacoustic inversion results to compute the TL uncertainty as a result of uncertainty in the environmental parameters. The results were obtained based on the posterior probability distributions of the parameters obtained for frame 195 (Fig. 5). The posterior predictive distribution of TL for the position (r_i, z_i) is obtained by integrating the predictions of TL with respect to the posterior distribution of the model parameters, using Eq. (14). Figure 6(a) shows the TL uncertainty plot at a receiver depth of 60 m over range at a frequency of 300 Hz. Figures 6(b) and 6(c) show the uncertainty spread in TL at a region of destructive and constructive interference at 1030 and 1510 m, respectively. The TL uncertainty band is about 5 dB around the region of constructive interference and widens to approximately 15 dB around the region of destructive interference, i.e., regions of destructive interference are predicted with much more uncertainty than regions of constructive interference. Figure 6(d) summarizes the predictive distribution by
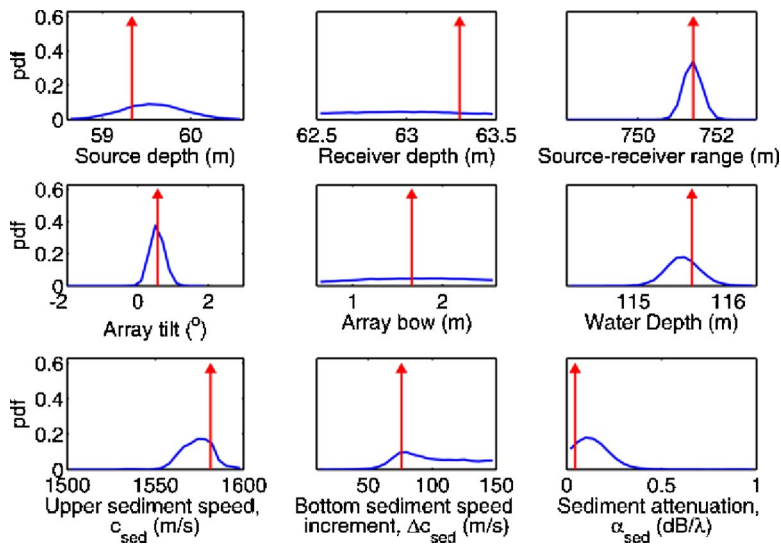
FIG. 5. (Color online) Posterior probability distributions for geometric parameters and sediment parameters, obtained from inversion for frame 195. The arrow indicates the MAP value.

the median (heavy line) and the 90% credibility interval (gray area). This is a practical way to convey the uncertainty in TL.

B. Experimental comparisons

The predictive TL distributions obtained in the preceding subsection are now compared with the actual TL observations. Bayesian inference provides the posterior distribution of the full parameter vector. To estimate the statistical properties of TL, only the posterior probability distribution of the input parameters to the forward model are required. These parameters include the geoacoustic parameters and water depth (water depth is included as it affects the number of propagating modes in the ocean waveguide). Source depth is also an important parameter for predicting TL fields accurately. In this data set, the matched-field estimated source depths varied with a mean of 52.3 m and standard deviation of 3.2 m. Other geometric parameters such as range, receiver depth, array bow and tilt are not used as input parameters to the forward model, and are removed from the parameter vector $\mathbf{m}$ [see Eq. (18)]. The initial prediction was performed at three frequencies, 300, 400, and 500 Hz, with the source depth fixed at the value obtained at the time of the inversion of frame 195. Figure 7(a) compares the observed TL (crosses) with the initial TL prediction statistics (solid line with gray area) for the frequencies 300, 400, and 500 Hz and for an array depth of 55 m. The initial prediction was not satisfactory as uncertainties in parameters such as the source depth were not modeled, resulting in 30–50% of the observed TL values falling outside the 90% credibility interval (CI). As frequency increases, larger spreads in TL predictions also are observed. This is most pronounced near regions of destructive interference.

Next, the uncertainty in the source depth was modeled using a Gaussian model with a mean of 52.3 m and standard deviation of 3.2 using the statistics from the matched field estimated source depths, and introduced into the parameter vector $\mathbf{m}$ (see also Sec. III B). The TL prediction process was then repeated with the inclusion of the source depth uncertainties. There is now a marked improvement with >80% of the observed TL values falling within the predicted 90% CI. The TL prediction plot is shown in Fig. 7(b).
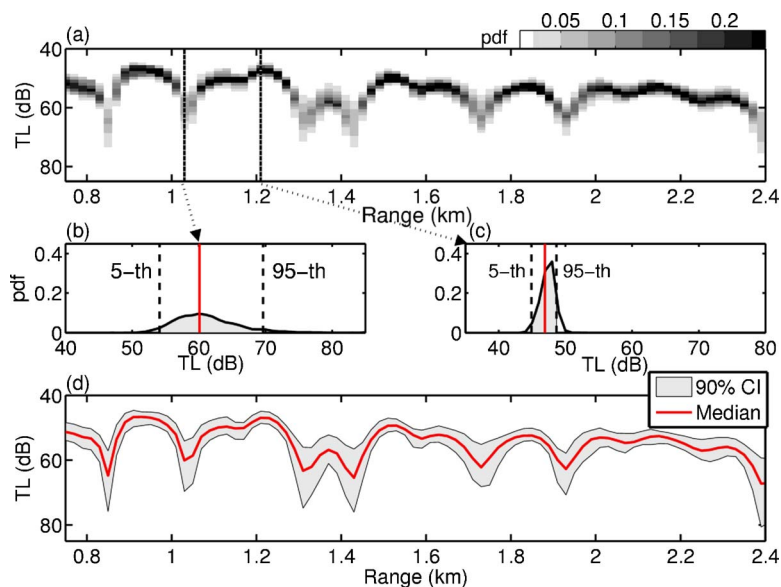

FIG. 6. (Color online) Posterior distribution of TL versus range for 300 Hz at 60 m array depth. (a) Contour of posterior distribution for TL versus range. (b) Region of destructive interference. (c) Region of constructive interference. (d) Statistics of the predicted TL versus range. The heavy line and the surrounding gray area represent the median and the 90% credibility interval, respectively.
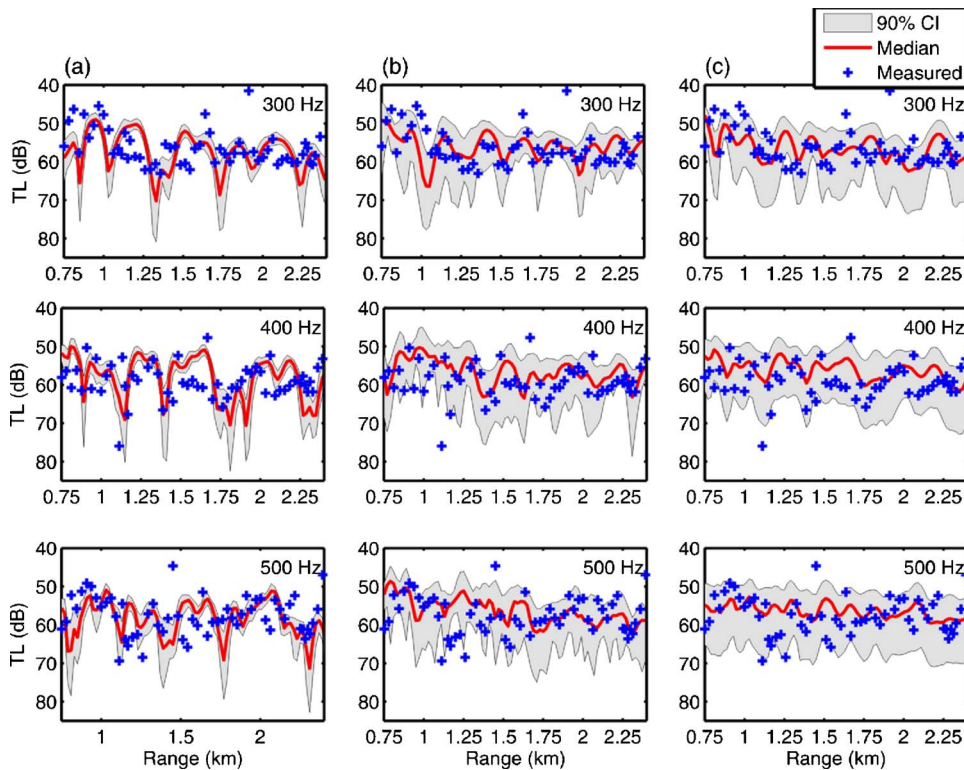
FIG. 7. (Color online) Predicted and measured TL (crosses) for frequencies 300, 400, and 500 Hz. The median of the predicted TL (heavy line) is shown together with the 90% CI (gray area). (a) Initial prediction. (b) Prediction with uncertainty introduced in source depth. (c) Prediction with uncertainty introduced in sound speed profile.

The ocean sound speed profile has been treated as known, but it also could be treated as an uncertain parameter [see Eq. (16)]. To observe the effects of sound speed profile variability on TL prediction, the variability was modeled with a Gaussian model with the measured values at each depth as the mean, and a uniform standard deviation along the whole water column (from 0 to 110 m with a 10 m spacing). Figure 7(c) shows the predicted results using a standard deviation of 1.7 m/s for the whole water column. This value gives >80% of the observed TL values falling within the predicted 90% CI.

Figure 8 plots on the right the variation of the percentage of the observed TL values falling within the predicted 90% CI versus the standard deviation of the modeled sound speed profile at 300 Hz. To the left, the median over range of
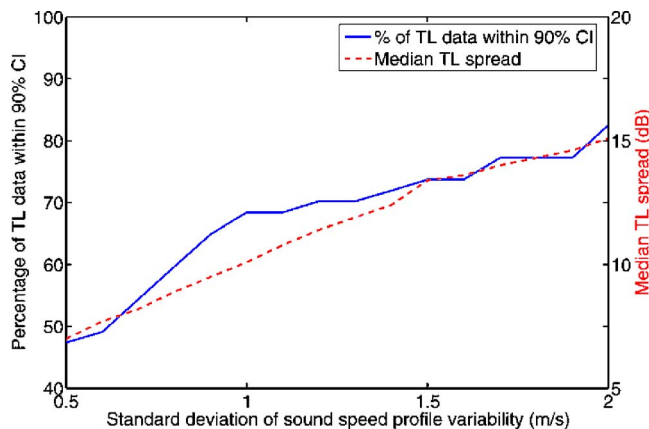

FIG. 8. (Color online) TL variation at 300 Hz versus standard deviation of the modeled sound speed profile. Right axis (solid): Variation of the percentage of the observed TL falling within the predicted 90% CI. Left axis (dashed): The median over range of the predicted TL spread at each range.

the TL spread at a single range is plotted. As expected, the median of the spread of the predicted TL and the percentage of the observed TL values falling within the predicted 90% CI increases as the standard deviation increases. Table I summarizes the comparison of the measured and predicted TL from Fig. 7.

Complicated environments, such as spatial and temporal fluctuations in the water column, sediment, sea surface and water-sediment interface, are not modeled and this will increase the error. Further, not all noise sources have been taken into account. Therefore, the percentage of observed data points within the computed 90% CI is less than predicted.

The TL measurement was carried out with a different source depth at each range, due to the motion of the ship (the matched-field inversion indicated a standard deviation of 3 m). It is the uncertainty in source depth that is reflected in the large band of TL uncertainty in Fig. 7(b). Thus, it is expected that if the true source depth at each range was included in the statistical TL prediction, then the predicted TL would lie in a much narrower credibility band (about 8 dB as in Fig. 7(a)) and contain the same fraction of the measured TL. The statistical prediction of TL also could be used in a sensitivity study to explore the importance of each parameter. In such an approach, the reduction in TL uncertainty if one parameter was fixed could be explored systematically.

VII. CONCLUSION

In this paper, the use of geoacoustic inversion results to estimate statistically transmission loss are demonstrated for data received on a towed horizontal array.

A frequency-coherent likelihood function based on a noninformative prior probability distribution was derived for

J. Acoust. Soc. Am., Vol. 122, No. 5, November 2007

Goh *et al.*: Statistical estimation of transmission loss   2577

TABLE I. Summary of transmission loss prediction performance at depth of 55 m. Numbers in dB indicated the median value of the predicted TL spread over range, while numbers in % represent the percentage of measured TL points that lie within the 99% CI.

Frequency (Hz)	Initial prediction	With uncertainty introduced to source depth	With uncertainty introduced to sound speed profile
300	49.1% / 5.2 dB	87.7% / 13.2 dB	84.2% / 15.3 dB
400	50.9% / 5.2 dB	87.7% / 15.5 dB	91.2% / 15.5 dB
500	59.7% / 6.0 dB	84.2% / 14.2 dB	89.5% / 16.4 dB

the horizontal array. A Markov chain Monte Carlo sampling was used to sample the posteriori distribution. Then, the posterior probability density of environmental parameters is utilized for the statistical estimation of TL. This is done by mapping of the probabilities from the geoacoustic domain to the transmission loss domain. It also is shown how additional uncertainty can be incorporated into the TL uncertainty.

The approach is demonstrated on towed array data from a sea trial in the North Elba area. Parameter uncertainties obtained from geoacoustic inversion are mapped into the transmission loss domain, where the probability distribution of transmission loss over different ranges and frequencies is obtained. The characteristic features such as the median and lower/upper percentiles from the distribution also are extracted. The predicted TL statistics are compared with TL measurements from the MAPEX2000 experiment. It is observed that more than 80% of the measured TL data fall within 90% of the range-varying predicted TL probability distribution, demonstrating that the statistical estimation approach presented is reasonable.

[1]D. F. Gingras and P. Gerstoft, "Inversion for geometric and geoacoustic parameters in shallow water: Experimental results," J. Acoust. Soc. Am. **97**, 3589–3598 (1995).

[2]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**, 1523–1535 (2002).

[3]D. P. Knobles, R. A. Koch, L. A. Thompson, K. C. Focke, and P. E. Eisman, "Broadband sound propagation in shallow water and geoacoustic inversion," J. Acoust. Soc. Am. **113**, 205–222 (2003).

[4]L. Jaschke and N. R. Chapman, "Matched field inversion of broadband data using the freeze bath method," J. Acoust. Soc. Am. **106**, 1838–1851 (1999).

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

[6]A. Caiti, S. M. Jesus, and A. Kristensen, "Geoacoustic seafloor exploration with a towed array in a shallow water area of the Strait of Sicily," IEEE J. Ocean. Eng. **21**, 355–366 (1996).

[7]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. Ocean. Eng. **30**, 198–206 (2005).

[8]D. Battle, P. Gerstoft, W. A. Kuperman, W. S. Hodgkiss, and M. Siderius, "Geoacoustic inversion of tow-ship noise via near-field-matched-field pro-

cessing," IEEE J. Ocean. Eng. **28**, 454–467 (2003).

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

[10]D. Tollefsen, S. E. Dosso, and M. J. Wilmut, "Matched-field geoacoustic inversion with a horizontal array and low-level source," J. Acoust. Soc. Am. **120**, 221–230 (2006).

[11]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**, 1932–1941 (2006).

[12]P. Gerstoft, C.-F. Huang, and W. S. Hodgkiss, "Estimation of transmission loss in the presence of geoacoustic inversion uncertainty," IEEE J. Ocean. Eng. **31**, 299–307 (2006).

[13]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 (Kluwer, Dordrecht, 2002), pp. 611–618.

[14]K. R. James and D. R. Dowling, "A probability density function method for acoustic field uncertainty analysis," J. Acoust. Soc. Am. **118**, 2802–2810 (2005).

[15]S. Finette, "Embedding uncertainty into ocean acoustic propagation models," J. Acoust. Soc. Am. **117**, 997–1000 (2005).

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

[17]F. B. Jensen and M. C. Ferla, *SNAP: the SACLANTCEN normal-mode acoustic propagation model*, SACLANT Undersea Research Center, SM–121, La Spezia, Italy (1979).

[18]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**, 231–246 (2006).

[19]P. Gerstoft, *SAGA Users guide 5.0, an inversion software package*, An updated version of "SAGA 2.0," SACLANT Undersea Research Centre, SM–333, La Spezia, Italy (1997).

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

[21]C. F. Mecklenbräuker and P. Gerstoft, "Objective functions for ocean acoustic inversion derived by likelihood methods," J. Comput. Acoust. **8**, 259–270 (2000).

[22]P. Gerstoft and C. F. Mecklenbräuker, "Ocean acoustic inversion with estimation of *a posteriori* probability distributions," J. Acoust. Soc. Am. **104**, 808–819 (1998).

[23]C. Soares and S. M. Jesus, "Broadband matched-field processing: Coherent and incoherent approaches," J. Acoust. Soc. Am. **113**, 2587–2598 (2003).

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

[25]C.-F. Huang, P. Gerstoft, and W. S. Hodgkiss, "On the effect of error correlation on matched-field geoacoustic inversion," J. Acoust. Soc. Am. **121**, EL64–69 (2007).

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

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

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

[29]N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E.

Teller, "Equation of state calculations by fast computing machines," J. Chem. Phys. **21**, 1087–1092 (1953).

[30] W. K. Hastings, "Monte Carlo sampling methods using Markov chains and their applications," Biometrika **57**, 97–109 (1970).

[31] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," IEEE Trans. Pattern Anal. Mach. Intell. **6**, 721–741 (1984).

[32] C. J. Geyer, "Practical Markov chain Monte Carlo (with discussion)," Stat. Sci. **7**, 473–483 (1992).

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

[34] C. Yardim, P. Gerstoft, and W. S. Hodgkiss, "Estimation of radio refractivity from radar clutter using Bayesian Monte Carlo analysis," IEEE Trans. Antennas Propag. **54**, 1318–1327 (2006).