Estimation of Radio Refractivity From Radar Clutter Using Bayesian Monte Carlo Analysis

Caglar Yardim, *Student Member, IEEE*, Peter Gerstoft, and William S. Hodgkiss, *Member, IEEE*

***Abstract*—This paper describes a Markov chain Monte Carlo (MCMC) sampling approach for the estimation of not only the radio refractivity profiles from radar clutter but also the uncertainties in these estimates. This is done by treating the refractivity from clutter (RFC) problem in a Bayesian framework. It uses unbiased MCMC sampling techniques, such as Metropolis and Gibbs sampling algorithms, to gather more accurate information about the uncertainties. Application of these sampling techniques using an electromagnetic split-step fast Fourier transform parabolic equation propagation model within a Bayesian inversion framework can provide accurate posterior probability distributions of the estimated refractivity parameters. Then these distributions can be used to estimate the uncertainties in the parameters of interest. Two different MCMC samplers (Metropolis and Gibbs) are analyzed and the results compared not only with the exhaustive search results but also with the genetic algorithm results and helicopter refractivity profile measurements. Although it is slower than global optimizers, the probability densities obtained by this method are closer to the true distributions.**

***Index Terms*—Atmospheric ducts, genetic algorithms, Markov chain Monte Carlo (MCMC) techniques, radar clutter, refractivity estimation.**

I. Introduction

An accurate knowledge of radio refractivity is essential in many radar and propagation applications. Especially at low altitudes, radio refractivity can vary considerably with both height and range, heavily affecting the propagation characteristics. One important example is the formation of an electromagnetic duct. A signal sent from a surface or low altitude source, such as a ship or low-flying object, can be totally trapped inside the duct. This will result in multiple reflections from the surface and they will appear as clutter rings in the radar plan position indicator (PPI) screen (Fig. 1). In such cases, a standard atmospheric assumption with a slope of modified refractivity of 0.118 M-units/m may not give reliable predictions for a radar system operating in such an environment.

Ducting is a phenomenon that is encountered mostly in sea-borne applications due to the abrupt changes in the vertical temperature and humidity profiles just above large water masses, which may result in an sharp decrease in the modified refractivity (M-profile) with increasing altitude. This will, in turn, cause the electromagnetic signal to bend downward, effectively trapping the signal within the duct. It is frequently encountered in many regions of the world such as the Persian Gulf, the Mediterranean, and California. In many cases, a simple trilinear M-profile is used to describe this variation. The coverage diagram of a trapped signal in such an environment is given in Fig. 2.

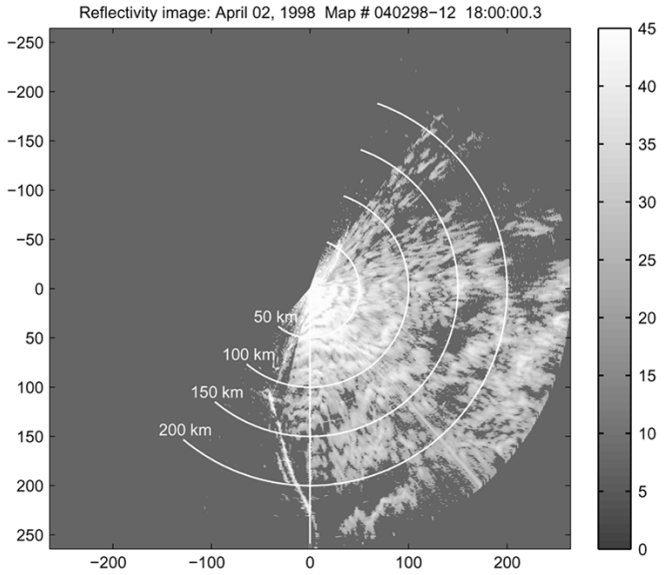

Fig. 1. Clutter map from Space Range Radar (SPANDAR) at Wallops Island, VA.

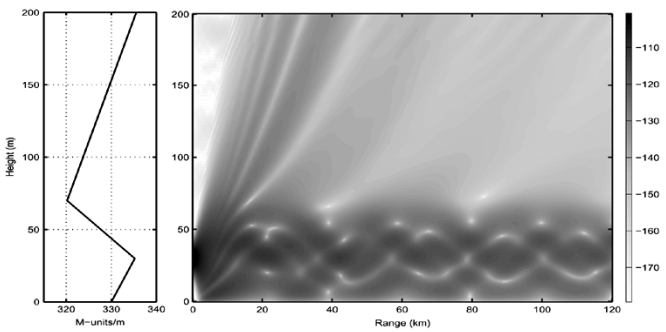

Fig. 2. Trilinear M-profile and its corresponding coverage diagram.

The first attempt in estimating the M-profile from radar clutter returns using a maximum likelihood (ML) approach was made in [1] and was followed by similar studies, which used either a marching-algorithm approach [2] or the global-parametrization approach [3], [4]. The latter is adopted in this paper. The main

Manuscript received May 18, 2005; revised October 10, 2005. This work was supported by the Office of Naval Research Code 32 under Grant N00014-05-1-0369.

C. Yardim and W. S. Hodgkiss are with the Electrical and Computer Engineering Department, University of California, San Diego, La Jolla, CA 92093 USA and also with the Marine Physical Laboratory, Scripps Institution of Oceanography, University of California, San Diego, La Jolla, CA 92093 USA (e-mail: cyardim@ucsd.edu, whodgkiss@ucsd.edu).

P. Gerstoft is with the Marine Physical Laboratory, Scripps Institution of Oceanography, University of California, San Diego, La Jolla, CA 92093 USA (e-mail: gerstoft@ucsd.edu)

Digital Object Identifier 10.1109/TAP.2006.872673

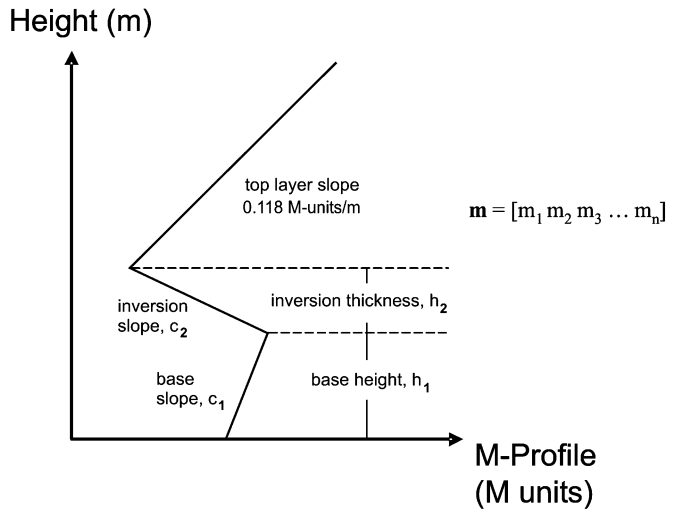

Fig. 3. The four-parameter trilinear M-profile model used in this work.

purpose of these studies is to estimate the M-profile using only the radar clutter return, which can readily be obtained during the normal radar operation, without requiring any additional measurements or hardware. A near-real-time estimation can be achieved with a sufficiently fast optimizer. Moreover, information obtained from other sources can be easily incorporated into the Bayesian formulation, e.g., the statistics of M-profiles in the region, meteorological model simulation results, helicopter, radiosonde, or some other ship-launched in situ instrument measurements.

To address the uncertainties in the M-profile parameter estimates, determination of the basic quantities such as the mean, variance, and marginal posterior probability distribution of each estimated parameter is necessary. They can be computed by taking multidimensional integrals of the posterior probability density (PPD), which can be accomplished by a Markov chain Monte Carlo (MCMC) sampling method within a Bayesian inversion structure. MCMC is selected because it provides unbiased sampling of the PPD, unlike global optimizers such as the genetic algorithm, which usually oversample the peaks of the PPD and introduce a bias [5].

Bayesian inversion is a likelihood-based technique which, when combined with a powerful sampling algorithm such as MCMC, can be an effective tool in the estimation of uncertainty in nonlinear inversion problems such as the electromagnetic refractivity from clutter (RFC) inversion. An alternative approach, which does not make use of likelihood, is given in [6]. The likelihood formulations are based on those used in [7]. The MCMC sampler employs a split-step fast Fourier transform (FFT) parabolic equation code as its forward propagation model.

II. Theory

The M-profile is assumed range-independent, and a simple trilinear profile is used to model the vertical M-profile (Fig. 3). An M-profile with n parameters is represented by the vector $\mathbf{m}$, with the element m_i being the value of the ith parameter. Each of these environmental parameters is then treated as an unknown random variable. Therefore, an n-dimensional joint posterior PPD can be defined using all of the parameters. All of the desired quantities such as the means, variances, and marginal posterior distributions can be found using the PPD. A summary of the notation used is given in Table I.

TABLE I
NOTATION

$\mathbf{d}$	Measured radar clutter data vector
$\mathbf{m}$	Refractivity (M-profile) model parameter vector
$f(\cdot)$	Forward model (split-step FFT parabolic equation)
$f(\mathbf{m})$	Replica radar clutter vector that would be observed in an environment characterized by $\mathbf{m}$
$\phi(\mathbf{m})$	Error function
n	Number of model parameters
N_{R}	Number of data points collected at different ranges
N	Number of Metropolis/Gibbs samples
N_{eff}	Number of effective data points
$\mathbf{C}_{\mathrm{d}}$	Data covariance matrix
ν	Error variance
$\hat{\nu}$	Error variance estimate
$\widehat{\mathbf{m}}$	Model vector GA estimate
D	Kolmogorov-Smirnov (K–S) statistic
$\mathrm{p}(\mathbf{m}\|\mathbf{d})$	Posterior probability density, PPD
$\mathcal{L}(\mathbf{m})$	Likelihood function
$\mathrm{p}(\mathbf{m})$	Prior function
$\mathrm{p}(\mathbf{d})$	Evidence function
$\mathrm{P}(m_i\|\mathbf{d})$	Cumulative marginal distribution function
$\mathbf{C}_{\mathrm{m}}$	Model covariance matrix
$\mathbf{\Lambda}$	Diagonal matrix containing eigenvalues, λ_i
$\mathbf{R}$	Rotation matrix
$\widetilde{\mathbf{m}}$	Rotated model parameter vector
T	Simulated annealing temperature
r	Range
L	One-way propagation loss
F	Propagation factor

The n-parameter refractivity model $\mathbf{m}$ is given to a forward model, an electromagnetic split-step FFT parabolic equation, along with the other necessary input parameters such as frequency, transmitter height, beamwidth, and antenna beam pattern [8], [9]. The forward model propagates the field in a medium characterized by $\mathbf{m}$ and outputs the radar clutter $f(\mathbf{m})$. This is then compared with the measured clutter data $\mathbf{d}$, and an error function $\phi(\mathbf{m})$ is derived for the likelihood function. In previous global-parametrization approaches, this error function was used by a global optimization algorithm such as the genetic algorithm (GA), which would minimize $\phi(\mathbf{m})$ and reach the ML solution. Instead of GA, a likelihood function based on $\phi(\mathbf{m})$ can be used in a Metropolis or Gibbs sampler. This makes it possible to get not only the ML solution but also better estimates for the uncertainties in terms of variances, marginal and multidimensional PPDs.

A. Bayesian Inversion

RFC can be solved using a Bayesian framework, where the unknown environmental parameters are taken as random variables with corresponding one-dimensional (1-D) probability density functions (pdfs) and an n-dimensional joint pdf. This probability function can be defined as the probability of the model vector $\mathbf{m}$ given the experimental data vector $\mathbf{d}$, $\mathrm{p}(\mathbf{m}|\mathbf{d})$, and it is called the posterior pdf (PPD). $\mathbf{m}$ with the highest probability is referred to as the maximum a posteriori (MAP) solution. For complex probabilities, global optimizers such as the genetic algorithm or simulated annealing can be used to find the MAP solution. An alternative to this is minimizing the

mean square error between clutter data $\mathbf{d}$ and the reconstructed clutter $f(\mathbf{m})$. It is referred to as the Bayesian minimum mean square error estimator and can easily be shown to be equal to the vector mean of $\mathrm{p}(\mathbf{m}|\mathbf{d})$ [10]. Both estimates are calculated in this paper. Due to the fact that a noninformative prior is used, MAP and ML solutions are the same and will be referred to simply as ML from now on. The posterior means, variances, and marginal probability distributions then can be found by taking n- or $(n-1)$-dimensional integrals of this PPD

$$\mu_i = \int \cdots \int_{\mathbf{m}'} m_i' \mathrm{p}(\mathbf{m}' \,|\, \mathbf{d}) d\mathbf{m}' \tag{1}$$

$$\sigma_i^2 = \int \cdots \int_{\mathbf{m}'} (m_i' - \mu_i)^2 \mathrm{p}(\mathbf{m}' \,|\, \mathbf{d}) d\mathbf{m}' \tag{2}$$

$$\mathrm{p}(m_i \,|\, \mathbf{d}) = \int \cdots \int_{\mathbf{m}'} \delta(m_i' - m_i) \mathrm{p}(\mathbf{m}' \,|\, \mathbf{d}) \mathrm{d}\mathbf{m}'. \tag{3}$$

Posterior density of any specific environmental parameter such as the M-deficit or the duct height can be obtained by marginalizing the n-dimensional PPD as given in (3). Joint probability distributions can be obtained using similar integrations. For further details about Bayesian inverse theory, see [10] and [11].

The posterior probability can be found using Bayes' formula

$$\mathrm{p}(\mathbf{m} \,|\, \mathbf{d}) = \frac{\mathcal{L}(\mathbf{m}) \mathrm{p}(\mathbf{m})}{\mathrm{p}(\mathbf{d})} \tag{4}$$

with

$$\mathrm{p}(\mathbf{d}) = \int_{\mathbf{m}} \mathrm{p}(\mathbf{d} \,|\, \mathbf{m}) \mathrm{p}(\mathbf{m}) \mathrm{d}\mathbf{m}. \tag{5}$$

The likelihood function $\mathcal{L}(\mathbf{m})$ will be defined in the next section. The prior $\mathrm{p}(\mathbf{m})$ represents the a priori knowledge about the environmental parameters $\mathbf{m}$ before the experiment. Therefore, it is independent of the experimental results and, hence, $\mathbf{d}$. The evidence appears in the denominator of the Bayes' formula as $\mathrm{p}(\mathbf{d})$. It is the normalizing factor for $\mathrm{p}(\mathbf{m}|\mathbf{d})$ and is independent of the parameter vector $\mathbf{m}$. A noninformative or flat prior assumption will reduce (4) to

$$\mathrm{p}(\mathbf{m} \,|\, \mathbf{d}) \propto \mathcal{L}(\mathbf{m}). \tag{6}$$

B. Likelihood Function

Assuming a zero-mean Gaussian-distributed error, the likelihood function can be written as

$$\mathcal{L}(\mathbf{m}) = (2\pi)^{-N_\mathrm{R}/2} \,|\, \mathbf{C}_\mathrm{d} \,|\, ^{-1/2} \times \exp\left[-\frac{(\mathbf{d} - f(\mathbf{m}))^\mathrm{T} \mathbf{C}_\mathrm{d}^{-1} (\mathbf{d} - f(\mathbf{m}))}{2} \right] \tag{7}$$

where $\mathbf{C}_\mathrm{d}$ is the data covariance matrix, $(\,\cdot\,)^\mathrm{T}$ is the transpose, and N_R is the number of range bins used (length of the data vector $\mathbf{d}$). Further simplification can be achieved by assuming that the errors are spatially uncorrelated with identical distribution for each data point forming the vector $\mathbf{d}$. For this case, $\mathbf{C}_\mathrm{d} = \nu \mathbf{I}$, where ν is the variance and $\mathbf{I}$ the identity matrix. Defining an error function $\phi(\mathbf{m})$ by

$$\phi(\mathbf{m}) = |\mathbf{d} - f(\mathbf{m})|^2 = \sum_{i=1}^{N_\mathrm{R}} |d_i - f_i(\mathbf{m})|^2 \tag{8}$$

the likelihood function can be written as

$$\mathcal{L}(\mathbf{m}) = (2\pi\nu)^{-N_\mathrm{R}/2} \exp\left[-\frac{\phi(\mathbf{m})}{2\nu} \right]. \tag{9}$$

Therefore, recalling (6), the posterior probability density can be expressed as

$$\mathrm{p}(\mathbf{m} \,|\, \mathbf{d}) \propto \exp\left[-\frac{\phi(\mathbf{m})}{2\nu} \right]. \tag{10}$$

The calculation of probabilities of all possible combinations along a predetermined grid is known as the exhaustive search and practically can be used for up to four to five parameters, depending on the forward modeling CPU speed. However, as n increases further, it becomes impractical. Therefore, there is a need for an effective technique that can more efficiently estimate not only the posterior probability distribution but also the multidimensional integrals given in (1)–(3).

C. Markov Chain Monte Carlo Sampling

MCMC algorithms are mathematically proven to sample the n-dimensional state space in such a way that the PPD obtained using these samples will asymptotically be convergent to the true probability distribution. There are various implementations of MCMC such as the famous Metropolis–Hastings (or simplified Metropolis version) algorithm, which was first introduced in [12], and Gibbs sampling, which was made popular among the image processing community by [13]. They are extensively used in many other fields such as geophysics [11] and ocean acoustics [5], [14], [15].

To have asymptotic convergence in the PPD, a Markov chain sampler must satisfy the detailed balance [16]. Markov chains with this property are also called reversible Markov chains, and it guarantees the chain to have a stationary distribution. MCMC samplers satisfy this detailed balance. Moreover, we can set up the MCMC such that the desired distribution (PPD) is the stationary distribution. Hence, the sample distribution will converge to the PPD as more and more samples are collected. Global optimizers do not satisfy detailed balance, and they usually oversample the high probability regions since they are designed as point estimators, trying to get to the optimal solution point fast, not to wander around the state space to sample the distribution. On the contrary, an unbiased sampler following MCMC rules will spend just enough time on lower probability regions and the histogram formed from the samples will converge to the true distribution.

Metropolis and Gibbs samplers are selected as the MCMC algorithms to be used here. The working principle for both methods is similar. Assume the ith sample is $\mathbf{m}^i = [m_1^i m_2^i \ldots m_n^i]$. In the n-dimensional parameter space, new MCMC samples are obtained by drawing from

n successive 1-D densities. Selection of the next MCMC sample $\mathbf{m}^{i+1} = [m_1^{i+1} m_2^{i+1} \dots m_n^{i+1}]$ starts with fixing all coordinates except the first one. Then the intermediate point $[m_1^{i+1} m_2^{i} \dots m_n^{i}]$ is selected by drawing a random value from a one-dimensional distribution around m_1^i. The new point $[m_1^{i+1} m_2^{i} \dots m_n^{i}]$ is then used as the starting point to update the second parameter by drawing a value for m_2. The next sample $\mathbf{m}^{i+1}$ is formed when all the parameters are updated once successively. The difference between the two methods lies in the selection of the 1-D distributions.

1) Metropolis Algorithm: In the more general Metropolis–Hastings algorithm, the 1-D distribution can be any distribution. However, in the simplified version, called the Metropolis algorithm, the 1-D distribution has to be symmetric. The most common ones used in practice are the uniform and Gaussian distributions (a variance of $0.2 \times$ search interval is used) centered around the m_j. Likelihood is assumed to be zero outside the search interval. After each 1-D movement, the Metropolis algorithm acceptance/rejection criterion is applied. The criterion to update the jth parameter can be given as

$$a = \frac{\mathrm{p}(m_j^{\text{proposed}})}{\mathrm{p}(m_j^i)}$$
$$m_j^{i+1} = \begin{cases} m_j^{\text{proposed}} & \text{(Accept)}, \quad \text{if } a > \mathrm{rand}[0,1] \\ m_j^i & \text{(Reject)}, \quad \text{else.} \end{cases} \tag{11}$$

The probability $\mathrm{p}(\mathbf{m} \,|\, \mathbf{d})$ for any model vector $\mathbf{m}$ can be calculated using (10).

2) Gibbs Algorithm: For Gibbs sampling, the 1-D distribution is not any random distribution but the conditional pdf at that point itself with all other parameters fixed. Therefore, the jth parameter m_j^{i+1} is drawn from the conditional pdf $\mathrm{p}(m_j \,|\, m_1^{i+1} m_2^{i+1} \dots m_{j-1}^{i+1} m_{j+1}^{i} \dots m_n^{i})$. For example, the first intermediate point $[m_1^{i+1} m_2^{i} \dots m_n^{i}]$ is selected by drawing from the conditional 1-D pdf obtained by fixing all m except $m_1, \mathrm{p}(m_1 \,|\, m_2^i m_3^i \dots m_n^i)$.

There are two possible advantages in this method. First of all, this method is especially powerful in applications, where 1-D conditional pdfs are known. Secondly, since the samples are selected from the conditional pdfs instead of some random distribution, Metropolis–Hastings acceptance/rejection criterion will always accept any selected point. Unfortunately, these pdfs are not known here and are found using exhaustive search, which makes this method less attractive in RFC applications. For further details, see [17] and [18].

Once the algorithm converges (see Section III-C), we will have a set $\{\mathbf{m}_1, \mathbf{m}_2, \mathbf{m}_3, \dots, \mathbf{m}_N\}$ of N samples that will be used to form the PPD and any other desired quantity as explained in the next section.

D. Monte Carlo Integration

Monte Carlo integration method can be used to evaluate (1)–(3). Notice that all of those equations are of the form

$$I = \int g(x) \mathrm{p}(x) \mathrm{d}x \tag{12}$$

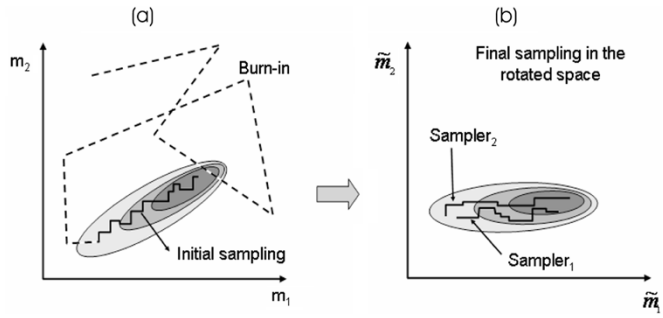

Fig. 4. Full implementation of the MCMC algorithm: (a) burn-in and initial sampling phases in the original parameter space and (b) two parallel-running samplers operating on the new parameter landscape after coordinate rotation.

where x is a random variable with a pdf of $\mathrm{p}(x)$ and $g(x)$ is some function of x. Monte Carlo integration [16] states that if a large enough set of random x values are drawn from its own pdf, $\{x_1, x_2, x_3, \dots, x_N\}$, the integral can be estimated as

$$I \approx \frac{1}{N} \sum_{i=1}^{N} g(x_i). \tag{13}$$

One of the biggest advantages of using an MCMC sampler comes from the ease at which the MCMC and the MC integration can be combined. Remembering that MCMC actually samples the n-dimensional posterior distribution $\mathrm{p}(\mathbf{m} \,|\, \mathbf{d})$, it will be clear that once MCMC converges, the set of MCMC samples can be directly used to calculate these multidimensional integrals.

III. Implementation

The most attractive property of the MCMC algorithm is that the result is guaranteed to converge to the true distribution when a large enough set of samples is collected. However, classical Metropolis/Gibbs sampling can be slow computationally, and some modifications are needed to make it faster without affecting the end results. There are two main drawbacks of MCMCs that decrease their speed, and they are discussed in the following two sections.

A. Burn-In Phase

The first drawback is related to the distance between the starting point and the high probability regions. Without any prior information, the algorithm would start from a random $\mathbf{m}$, which may be far from the high probability regions. This is a concern since, unlike a global optimizer, it may take a considerable number of iterations for the MCMC to reach the high probability regions. Hence, it must be correctly guided or initialized before the sampling is started. The classical MCMC can be modified to include an initialization phase, which is commonly referred as the "burn-in" phase [Fig. 4(a)]. In most cases, the burn-in section itself is actually a global optimizer. A GA or a fast cooling simulated annealing (SA) algorithm can be good candidates. Both were used here and gave good results. If a fast cooling SA is to be used, the temperature T should be

lowered until it reaches $T = 1$ so that the Boltzmann distribution used for SA actually becomes the likelihood function itself at that temperature [16]

$$\mathrm{p_{SA}}(\mathbf{m}) = \exp\left[-\frac{\phi(\mathbf{m})}{2\nu T}\right] \Longrightarrow \mathcal{L}(\mathbf{m}). \tag{14}$$

B. Initial Sampling and Coordinate Rotation

The second drawback is based on the effects of interparameter correlation. Recalling that there is freedom of movement only in the directions parallel to parameter axes, it is hard to sample highly correlated PPDs [Fig. 4(a)]. With only vertical and horizontal movements allowed, it will take many samples to move from one end to the other of the slanted high probability areas, whereas it can be done with a much smaller number of samples in an uncorrelated case.

The remedy lies in a rotation of coordinate in the n-dimensional parameter space [19]. Instead of applying Metropolis sampling in the correlated parameter space, a new set of uncorrelated parameters are defined by applying an orthogonal coordinate transformation that diagonalizes the model covariance matrix $\mathbf{C_m}$. The rotation matrix $\mathbf{R}$ is found by eigenvalue-eigenvector decomposition of $\mathbf{C_m}$

$$\mathbf{C_m} = \mathbf{R}\mathbf{\Lambda}\mathbf{R}^{\mathrm{T}}, \qquad \tilde{\mathbf{m}} = \mathbf{R}^{\mathrm{T}}\mathbf{m} \tag{15}$$

where $\mathbf{\Lambda}$ is a diagonal matrix containing the eigenvalues of the covariance matrix, $\mathbf{R}$ is the orthonormal rotation matrix, whose columns contain the eigenvectors of $\mathbf{C}_{\mathrm{m}}$, and $\tilde{\mathbf{m}}$ is the rotated model vector. The model covariance matrix is found using samples drawn from the PPD after the burn-in phase. Only a small number of MCMC samples (about 1000) are enough to find $\mathbf{C_m}$ due to its fast converging nature [5], [14]. This is referred to as the initial sampling phase. After this phase, the parameter space is rotated before the final sampling phase starts.

C. Final Sampling Phase and Convergence

The final sampling phase simply is composed of two independent, parallel-running Metropolis (or Gibbs) samplers in the rotated space, sampling the same PPD [Fig. 4(b)]. The algorithm uses a convergence criterion based on the Kolmogorov–Smirnov (K–S) statistic function D of the marginal posterior distributions [20]. $\mathrm{p}(m_j \,|\, \mathbf{d})$s are calculated using (3) as each new sample is drawn. A pair of $\mathrm{p}(m_j \,|\, \mathbf{d})$, one from each sampler, is calculated for each parameter m_j. Then the K–S statistic is calculated for each parameter as

$$D_j = \max_{m_j} \left|\mathrm{P}_2(m_j|\mathbf{d}) - \mathrm{P}_1(m_j|\mathbf{d})\right| \tag{16}$$

where $\mathrm{P}_1(m_j \,|\, \mathbf{d})$ and $\mathrm{P}_2(m_j \,|\, \mathbf{d})$ represent the cumulative marginal distribution functions for the two parallel-running samplers. The simulation is assumed to have converged when the largest D_j term is less than ϵ ($\epsilon = 0.05$ is used here). After the convergence criterion is met, these two independent sets are combined to form a final set twice as large so that the difference between the true distribution and the estimated one is expected to be even less.

To use the likelihood (9), the error variance ν must be estimated. In this paper, an estimate for ν is calculated using an ML approach. To find the ML estimate of the variance, $\partial\mathcal{L}/\partial\nu = 0$ is solved and evaluated at the ML model estimate $\hat{\mathbf{m}}$. This results in

$$\hat{\nu} = |\mathbf{d} - f(\hat{\mathbf{m}})|^2 / N_{\mathrm{R}}. \tag{17}$$

$\hat{\mathbf{m}}$ itself is estimated using a global optimizer, usually the value obtained at the end of the burn-in phase.

Equation (17) assumes the measurements at different ranges are uncorrelated, which may not be the case. Gerstoft and Mecklenbräuker [7] suggested replacing N_{R} with N_{eff}, number of effectively uncorrelated data points. More detailed discussion can be found in [14] and [15].

Received clutter power can be calculated as given in [3], in terms of the one-way loss term L in decibels as

$$f(\mathbf{m}) = -2L + 10\log(r) + C \tag{18}$$

where C is a constant that includes wavelength, radar cross-section (RCS), transmitter power, antenna gain, etc. If these values are known, they can be included into the equation. However, one or more may be unavailable such as the mean RCS or transmitter specifications. An alternative, which is used here, is discarding these constant terms, which is done by subtracting both the mean of the replica clutter $f(\mathbf{m})$ from itself and the mean of the measured clutter $\mathbf{d}$ from itself before inserting them into the likelihood function.

D. Postprocessing

After the MCMC sampler converges with a set of refractivity parameters $\{\mathbf{m}_1, \mathbf{m}_2, \mathbf{m}_3, \ldots, \mathbf{m}_N\}$, a postprocessing section is needed to convert the uncertainty in the M-profile into other parameters of interest that could be used by the radar operator, such as the propagation factor and the one-way propagation loss.

Assume an end-user parameter u, which can be expressed as a function of model parameters $u = \mathrm{g}(\mathbf{m})$. The posterior probability distribution of u, PPD_u can be simply calculated by drawing a large amount of samples of $\mathbf{m}$ from its own PPD and calculating $\mathrm{g}(\cdot)$ for each one. Then this set of $\mathrm{g}(\mathbf{m})$ can be used to obtain the PPD_u or perform any other uncertainty analysis of u using (1)–(3) and MC integration.

Since the usage of an MCMC sampler guarantees that $\{\mathbf{m}_1, \mathbf{m}_2, \mathbf{m}_3, \ldots, \mathbf{m}_N\}$ is a large enough set drawn from its own PPD, this set can readily be used to obtain the statistics of u.

IV. Examples

This section is composed of both synthetic and experimental examples. Four different algorithms, two of which are MCMC, are first tested using synthetic data generated by the terrain parabolic equation program (TPEM) [8]. Then the data gathered during the Wallops'98 Space Range Radar (SPANDAR) measurement are analyzed using the Metropolis sampler and GA.

TABLE II
SYNTHETIC DATA CASE: GA ESTIMATES AND METROPOLIS ALGORITHM RESULTS

Parameter	Units	Search Limits Lower	Search Limits Upper	True Value	GA Estimate	Metropolis ML Estimate	Mean	STD	90% Credibility Intervals Lower	90% Credibility Intervals Upper
c_1	M-units/m	0	0.25	0.13	0.132	0.128	0.1363	0.019	0.111	0.171
c_2	M-units/m	−3.5	−1	−2.5	−2.504	−2.503	−2.47	0.077	−2.59	−2.34
h_1	m	25	50	40	39.93	40.05	39.80	0.147	39.938	40.488
h_2	m	0	50	20	19.86	19.96	29.65	9.184	18.375	46.750

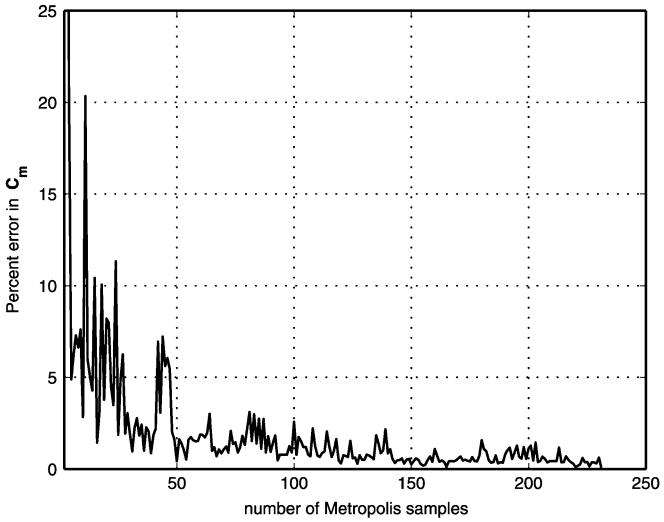

Fig. 5. Initial sampling phase—convergence of the model covariance matrix in terms of percent error. $\mathbf{C}_{\mathrm{m}}$ later will be used for coordinate rotation.

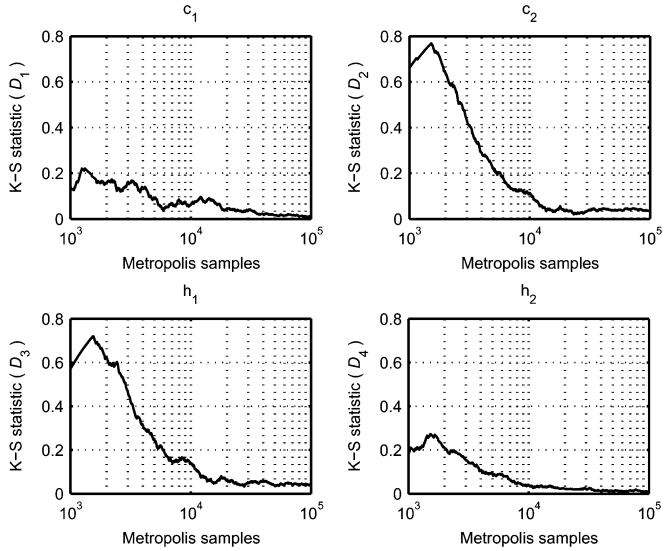

Fig. 6. Final sampling phase: Kolmogorov–Smirnov statistic D for each parameter. Used for the convergence of the posterior probability density.

A. Algorithm Validation

To validate the MCMC algorithms, a comparison with the true distribution is necessary. The true distribution can be obtained by using exhaustive search; however, it is extremely inefficient and demands a large number of forward model runs. Even if only 25 discrete possible values are assumed for each of the four parameters used for the model in Fig. 2, the state space consists of $25^4 = 3.9 \times 10^5$ (390 k) possible states. A simple range-independent trilinear model with only four parameters is used since an exhaustive search would need around 10 000-k forward model runs for five parameters. The number of forward model runs needed for MCMC is proportional to the dimension size, so as dimension size increases, it requires much fewer samples than the exhaustive search. The selected parameters are the slope and height of the base layer (c_1 and h_1) and the slope and thickness of the inversion layer (c_2 and h_2), as shown in Fig. 3. Their test case values and the search bounds are given in Table II. A standard atmosphere with a vertical refractivity gradient (top layer slope) of 0.118 M-units/m is assumed above the inversion layer. Parameters are selected in terms of the heights and slopes instead of the classical heights and widths (such as the frequently used inversion thickness and M-deficit) due to their relatively smaller interparameter correlation.

The synthetic data are generated by TPEM at a frequency of 2.84 GHz, antenna 3-dB beamwidth of 0.4°, source height of 30.78 m, and radar clutter standard deviation of 10 dB, a typical value reported also in [21]. Inversion is done using four different methods for a range of 10–60 km.

The convergence characteristics of both MCMC algorithms are similar. The results of the initial sampling phase are given in Fig. 5 in terms of a convergence plot for the model covariance matrix $\mathbf{C}_{\mathrm{m}}$. The percent error is calculated as the average absolute percent change in all matrix entries of $\mathbf{C}_{\mathrm{m}}$ as the matrix is recalculated while new samples are taken. The matrix converges quickly and for this case, only 250 samples were sufficient to have a good estimate of $\mathbf{C}_{\mathrm{m}}$.

After the rotation matrix is obtained, two independent Metropolis/Gibbs algorithms run during the final sampling phase until the convergence criterion given in Section III-C are met (Fig. 6). Dosso [5] reported convergence in 7×10^5 (700 k) forward models with a Gibbs sampler and around 100 k with a fast Gibbs sampler (FGS). Similarly, Battle *et al.* [15] reported convergence in 63-k models using a less strict convergence criterion, again with an FGS. Due to the modifications in the Metropolis/Gibbs algorithms used here, they also can be classified as "fast" samplers. Therefore, a typical value of 100-k models is in agreement with the previous applications of MCMC algorithms. During the simulations, values as low as 20-k models and higher than 150-k models were encountered.

The marginal distributions of the four parameters are given Fig. 7. Except for exhaustive search, the results for all others are calculated using the MC integration. The exhaustive search result [Fig. 7(a)] is obtained with 25 discrete values per parameter and 390-k samples, whereas both Metropolis [Fig. 7(b)] and Gibbs [Fig. 7(c)] samplers use approximately 70-k samples and the genetic algorithm [Fig. 7(d)] uses less than 10-k samples.

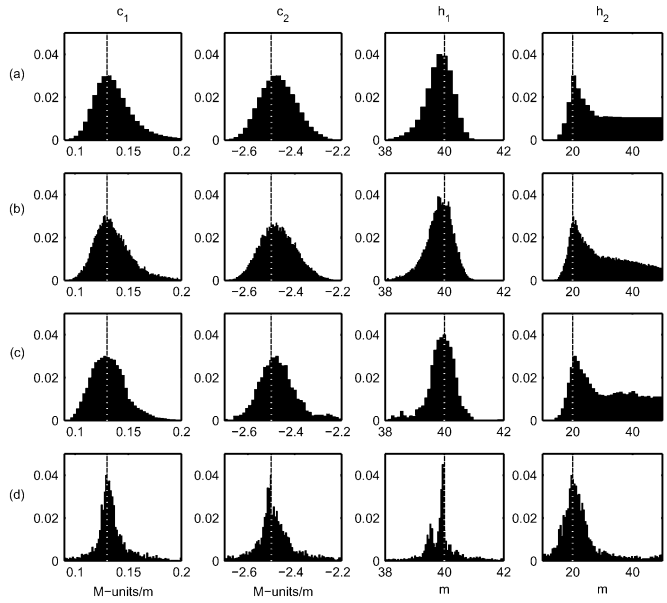

Fig. 7. Marginal posterior probability distributions for the synthetic test case. Vertical lines show the true values of the parameters. (a) Exhaustive search, (b) Metropolis algorithm, (c) Gibbs algorithm, and (d) genetic algorithm.

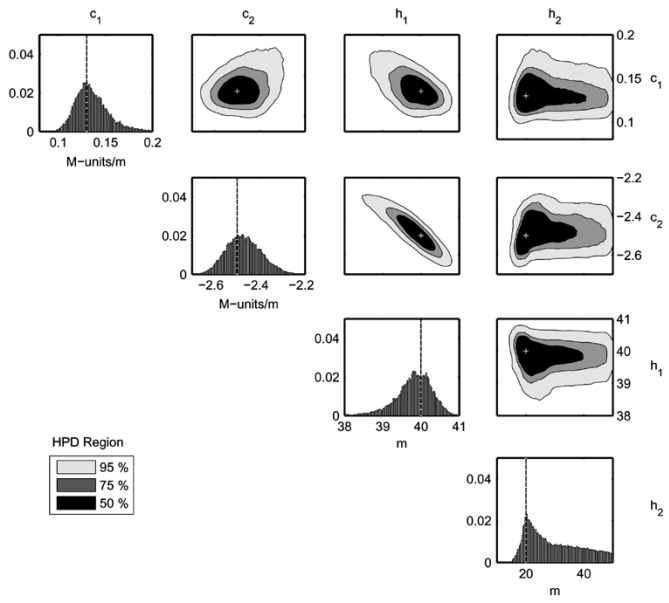

Fig. 8. Both 1-D marginal (diagonal) and 2-D marginal (upper diagonal) PPDs for the synthetic test case obtained by the Metropolis algorithm. Vertical lines (in 1-D plots) and crosses (in 2-D plots) show the true values of the parameters.

The results of the Metropolis algorithm and GA are also summarized in Table II. The MATLAB GA toolbox [22] is used for GA results. It uses three isolated populations of size 40 each with a migration rate of 0.025 per ten generations, a mutation rate of 0.10 and a crossover fraction of 0.8.

All four algorithms have almost identical ML estimates for the parameters. However, it should be noted that MCMC samplers converged after collecting nearly seven times more samples than GA. On the other hand, the distributions obtained from the MCMC samplers are closer to the true distributions given by exhaustive search.

Marginal and two-dimensional (2-D) posterior distributions obtained by the Metropolis sampler are given in Fig. 8. The diagonal pdfs are the 1-D marginal pdfs and the off-diagonal plots are the 2-D marginal pdfs, where the 50, 75, and 95% highest posterior density (HPD) regions are plotted, with the ML solution points (white crosses). In Bayesian statistics, credibility intervals and HPD regions are used to analyze the posterior distributions. They are very similar to the confidence interval and their definitions can be found in [23].

B. Wallops Island Experiment

The Metropolis sampler is used to analyze the data collected from the Wallops Island 1998 experiment conducted by the Naval Surface Warfare Center, Dahlgren Division (Fig. 9). The radar clutter data gathered by SPANDAR are inverted using the Metropolis algorithm and the results compared with helicopter measurements. Data used in the inversion were taken during a surface-based ducting event on April 2, 1998 [24], [3], at a frequency of 2.84 GHz, power of 91.4 dBm, 3 dB beamwidth of 0.4°, antenna gain of 52.8 dB, 30.78 m above the mean sea level (MSL), VV polarization, and 600 m range bins. Only data between 10–60 km are used in the inversion. The results are summarized in Table III.

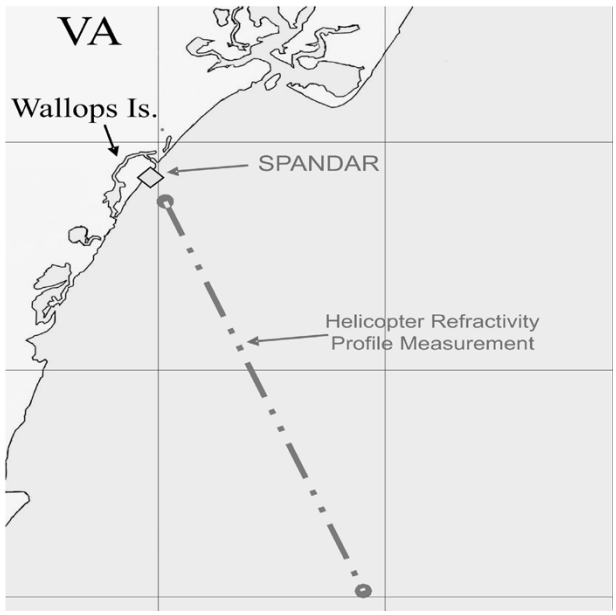

Fig. 9. Wallops '98 Experiment: SPANDAR radar and the helicopter measurements (37.83° N, 75.48° W)

The same four-parameter model used in the synthetic test case is selected. The marginal distributions obtained by Metropolis and GA in Fig. 10 use 80- and 10-k samples, respectively. Fig. 11 shows the 1-D and 2-D marginal PPD plots obtained by the Metropolis algorithm.

The helicopter measurements at different ranges can be seen in Fig. 12(a). Profiles measured at 10, 20, 30, 40, and 50 km are all plotted in Fig. 12(b) together with HPD regions obtained from the postprocessing of the 80-k Metropolis samples. These HPD regions show the regions where the values of the trilinear M-profile at various altitudes are expected to be. Therefore, it

TABLE III
WALLOPS ISLAND EXPERIMENT (CLUTTER MAP 17): GA ESTIMATES AND METROPOLIS ALGORITHM RESULTS

Parameter	Units	Search Limits Lower	Search Limits Upper	GA Estimate	Metropolis ML Estimate	Mean	STD	90% Credibility Intervals Lower	90% Credibility Intervals Upper
c_1	M-units/m	-1	0	-0.603	-0.604	-0.444	0.186	-0.79	-0.11
c_2	M-units/m	-1	1	-0.014	-0.010	-0.180	0.475	-0.61	0.89
h_1	m	10	75	31.00	30.98	33.76	11.95	16.79	59.63
h_2	m	0	75	24.63	22.93	34.41	20.50	4.69	68.66

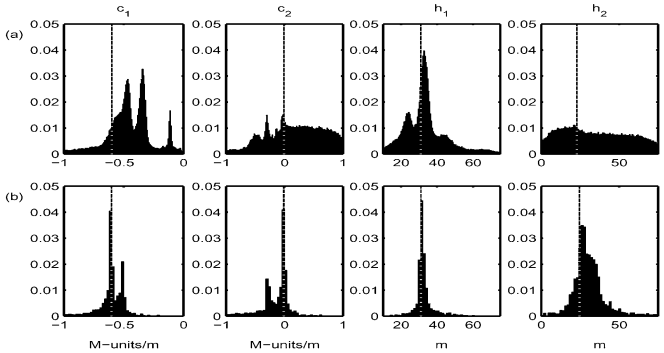

Fig. 10. Marginal posterior probability distributions obtained using SPANDAR data. (a) Metropolis algorithm and (b) genetic algorithm. Vertical lines show the estimated optimum values of the parameters.

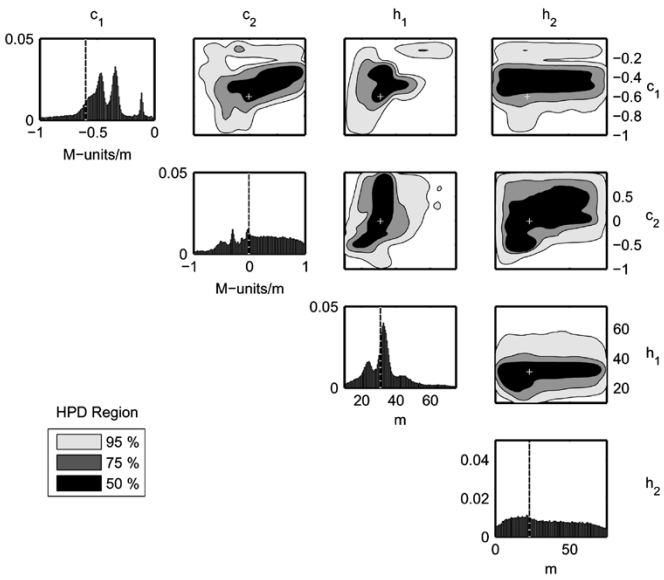

Fig. 11. Both 1-D marginal (diagonal) and 2-D marginal (upper diagonal) PPDs obtained from the SPANDAR data obtained by the Metropolis algorithm. Vertical lines (in 1-D plots) and crosses (in 2-D plots) show the optimum values of the parameters.

roughly can be compared to the mean of the helicopter M-profiles measured at different ranges. This mean helicopter profile is plotted together with the ML solution obtained from the Metropolis sampler in Fig. 12(c).

The improvement obtained by the inversion is analyzed in Fig. 13. The coverage diagrams (dB) in Fig. 13(a)–(c) are obtained using a standard atmospheric assumption, helicopter refractivity profile measurements, and the Metropolis inversion results. Fig. 13(d)–(e) are difference plots and are calculated by subtracting the dB coverage diagrams obtained by two different

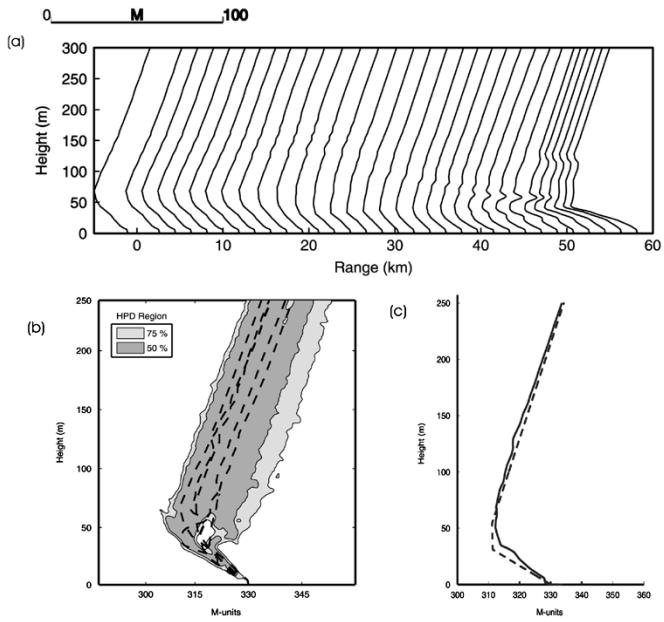

Fig. 12. M-profiles 0–60 km: (a) helicopter measurements at different ranges, (b) helicopter profiles (dashed) at 10, 20, 30, 40, 50 km together with 50% and 75% HPD regions for the range independent model, and (c) range-independent maximum likelihood solution (dashed line) and mean of the profiles measured at different ranges (solid line).

methods. Fig. 13(d) is the difference between the standard atmosphere and the helicopter results, whereas Fig. 13(e) is the difference between the Metropolis inversion and the helicopter results. The improvement inside the duct ($h < 60$ m) easily can be noticed. However, the results outside the duct are not as good. This is an expected result since the inversion is done using the radar measured sea surface reflected clutter caused by the electromagnetic duct. No signal outside the duct is used and hence the inversion algorithm has poor accuracy outside the duct.

In Fig. 14, the clutter power versus range plots are given. The relative clutter return measured by SPANDAR is plotted together with the clutter found using the Metropolis ML estimate $\hat{\mathbf{m}}$ and the clutter obtained using the range-varying helicopter profile. Surface layer evaporation ducting was appended to the bottom of the helicopter refractivity profiles, with the evaporation duct heights being less than 5 m. Then, this profile [Fig. 12(a)] is simulated using the parabolic equation model to estimate the helicopter clutter. Misfits can be explained by the range independent assumption of the simple trilinear M-profile, which cannot exactly duplicate the real radar clutter. Details of errors associated with the range independent assumption can be found in [25].

The PPD of the environmental parameters can be used to get PPDs of parameters of interest. This easily can be done by

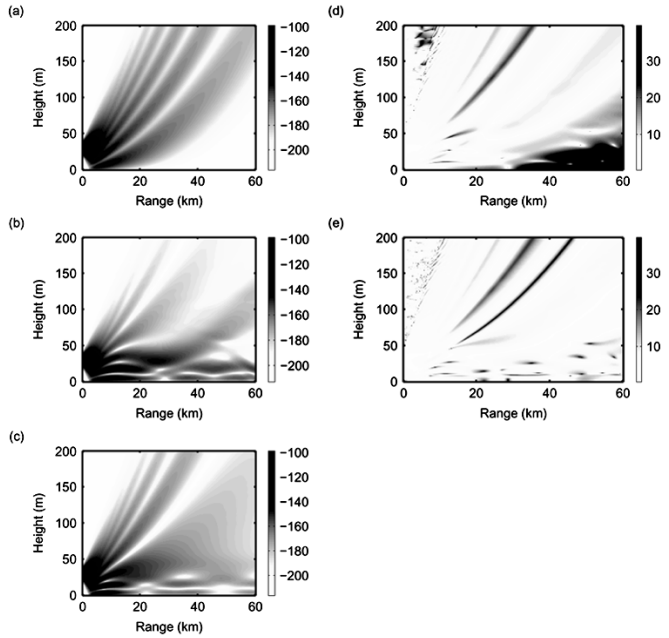

Fig. 13. Coverage diagrams (dB). One-way propagation loss for (a) standard atmosphere (0.118 M-units/m), (b) helicopter-measured refractivity profile, and (c) Metropolis inversion result. The difference (dB) between (d) helicopter and standard atmosphere results and (e) helicopter and Metropolis inversion results.

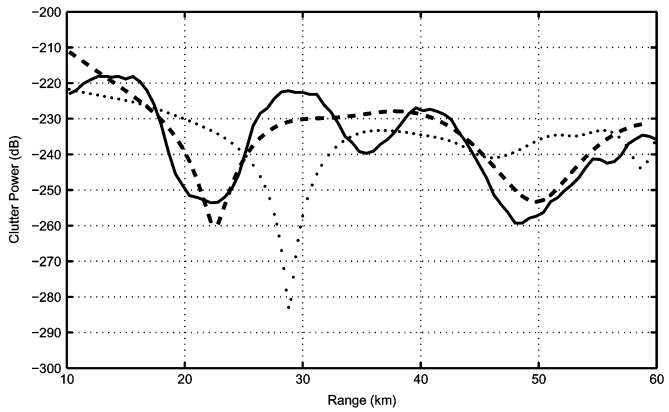

Fig. 14. Clutter power versus range plots. Clutter measured by SPANDAR (solid), the predicted clutter obtained using the Metropolis ML solution $\hat{\mathbf{m}}$ (dashed), and the predicted clutter obtained using the helicopter-measured refractivity profile (dotted).

postprocessing the 80-k Metropolis samples of the refractivity model parameters. Posterior densities for propagation factor (F) at a range of 60 km with height values above MSL of 28 and 180 m are given in Fig. 15(a) and (b), respectively. These two height values specifically are selected to compare the quality of estimates inside and outside of the duct. As expected, F in case (a), which is inside the duct, has a much smaller variance, whereas the uncertainty in case (b) is much larger.

Finally, Fig. 15(c) shows the effects of uncertainty in the environmental parameters on establishing a successful communication link. Assume that the transmitter–receiver pair requires a propagation loss less than 135 dB to attain the necessary signal-to-noise ratio to operate in this environment. The diagram shows the probability of establishing a successful

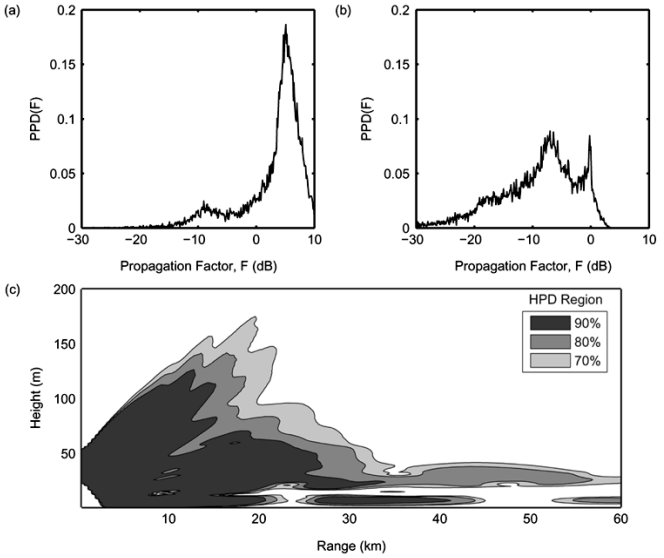

Fig. 15. PPD for propagation factor F at range 60 km with altitudes of (a) 28 m and (b) 180 m above MSL. (c) PPD of the coverage area for the given communication link.

link in an environment known to an accuracy of $\mathrm{p}(\mathbf{m}\,|\,\mathbf{d})$ as a coverage area HPD plot. The results are given as 70, 80, and 90% HPD regions. As expected, the link can be maintained over an extended range inside the duct.

V. Conclusion

A method for estimation of the radio refractivity from radar clutter using Markov chain Monte Carlo samplers with a likelihood-based Bayesian inversion formulation has been introduced. This approach enables us to obtain full n-dimensional posterior probability distributions for the unknown parameters as well as the maximum likelihood solution itself.

Comparisons with exhaustive search and genetic algorithm results show that MCMC samplers require more samples than a classical global optimizer but are better in estimating probability distributions. The need for a relatively large number of forward model runs limits its usage as a near-real-time M-profile estimator. However, it can be used together with a fast global optimizer, which will do the near-real-time inversion. The MCMC sampler will then provide the credibility intervals and the uncertainties, which may not be needed frequently.

One immediate benefit of the method is the ability to assess the quality of the inversion and obtain highest posterior density plots for other parameters that could be of interest to an end-user, such as the one-way propagation loss, propagation factor for different heights and ranges, or variability in the coverage diagrams. They easily can be obtained by postprocessing the Metropolis samples of the refractivity parameters.

Acknowledgment

The authors would like to thank T. Rogers, SPAWAR, San Diego, CA, for providing the SPANDAR'98 clutter maps and D. Dockery, Applied Physics Laboratory, Johns Hopkins University, Baltimore, MD, for providing the helicopter refractivity measurements.

REFERENCES

[1] S. Vasudevan and J. Krolik, "Refractivity estimation from radar clutter by sequential importance sampling with a Markov model for microwave propagation," in *Proc. ICASSP*, 2001, pp. 2905–2908.

[2] R. Anderson, S. Vasudevan, J. Krolik, and L. T. Rogers, "Maximum *a posteriori* refractivity estimation from radar clutter using a Markov model for microwave propagation," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, vol. 2, Sydney, Australia, 2001, pp. 906–909.

[3] P. Gerstoft, L. T. Rogers, J. Krolik, and W. S. Hodgkiss, "Inversion for refractivity parameters from radar sea clutter," *Radio Sci.*, vol. 38, no. 18, pp. 1–22, 2003.

[4] P. Gerstoft, W. S. Hodgkiss, L. T. Rogers, and M. Jablecki, "Probability distribution of low altitude propagation loss from radar sea-clutter data," *Radio Sci.*, vol. 39, pp. 1–9, 2004.

[5] S. E. Dosso, "Quantifying uncertainty in geoacoustic inversion I: A fast Gibbs sampler approach," *J. Acoust. Soc. Amer.*, vol. 111, no. 1, pp. 129–142, 2002.

[6] L. T. Rogers, M. Jablecki, and P. Gerstoft, "Posterior distributions of a statistic of propagation loss inferred from radar sea clutter," *Radio Sci.*, 2005.

[7] P. Gerstoft and C. F. Mecklenbräuker, "Ocean acoustic inversion with estimation of *a posteriori* probability distributions," *J. Acoust. Soc. Amer.*, vol. 104, no. 2, pp. 808–819, 1998.

[8] A. E. Barrios, "A terrain parabolic equation model for propagation in the troposphere," *IEEE Trans. Antennas Propag.*, vol. 42, no. 1, pp. 90–98, 1994.

[9] M. Levy, *Parabolic Equation Methods for Electromagnetic Wave Propagation*, London, U.K.: Institution of Electrical Engineers, 2000.

[10] S. M. Kay, *Fundamentals of Statistical Signal Processing—Volume I: Estimation Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1993.

[11] M. K. Sen and P. L. Stoffa, "Bayesian inference, Gibbs' sampler and uncertainty estimation in geophysical inversion," *Geophys. Prosp.*, vol. 44, pp. 313–350, 1996.

[12] 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.*, vol. 21, pp. 1087–1092, 1953.

[13] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 6, pp. 721–741, 1984.

[14] S. E. Dosso and P. L. Nielsen, "Quantifying uncertainty in geoacoustic inversion II. Application to broadband, shallow-water data," *J. Acoust. Soc. Amer.*, vol. 111, no. 1, pp. 143–159, 2002.

[15] D. J. Battle, P. Gerstoft, W. S. Hodgkiss, W. A. Kuperman, and P. L. Nielsen, "Bayesian model selection applied to self-noise geoacoustic inversion," *J. Acoust. Soc. Amer.*, vol. 116, no. 4, pp. 2043–2056, Oct. 2004.

[16] J. J. K. O. Ruanaidh and W. J. Fitzgerald, *Numerical Bayesian Methods Applied to Signal Processing*, ser. Statistics and Computing Series. New York: Springer-Verlag, 1996.

[17] K. Mosegaard and M. Sambridge, "Monte Carlo analysis of inverse problems," *Inv. Prob.*, vol. 18, no. 3, pp. 29–54, 2002.

[18] D. J. C. MacKay, *Information Theory, Inference and Learning Algorithms*. Cambridge, U.K.: Cambridge Univ. Press, 2003.

[19] M. D. Collins and L. Fishman, "Efficient navigation of parameter landscapes," *J. Acoust. Soc. Amer.*, vol. 98, pp. 1637–1644, 1995.

[20] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. Flannery, *Numerical Recipes in C*, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 1995.

[21] K. D. Anderson, "Radar detection of low-altitude targets in a maritime environment," *IEEE Trans. Antennas Propag.*, vol. 43, pp. 609–613, Jun. 1995.

[22] D. E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*. Reading, MA: Addison-Wesley, 1989.

[23] G. E. P. Box and G. C. Tiao, *Bayesian Inference in Statistical Analysis*. New York: Addison-Wesley, 1992.

[24] L. T. Rogers, C. P. Hattan, and J. K. Stapleton, "Estimating evaporation duct heights from radar sea echo," *Radio Sci.*, vol. 35, no. 4, pp. 955–966, 2000.

[25] J. Goldhirsh and D. Dockery, "Propagation factor errors due to the assumption of lateral homogenity," *Radio Sci.*, vol. 33, no. 2, pp. 239–249, 1998.

Caglar Yardim (S'98) was born in Ankara, Turkey, on December 5, 1977. He received the B.S. and M.S. degrees in electrical engineering from the Middle East Technical University (METU), Ankara, in 1998 and 2001, respectively. He is currently pursuing the Ph.D. degree at the Electrical and Computer Engineering Department, University of California, San Diego.

From 1998 to 2000, he was a Research Assistant with METU. From 2000 to 2002, he was a Research Engineer with the Microwave and System Technologies Division, ASELSAN Electronic Industries Inc. Since 2002, he has been a Graduate Research Associate with the Marine Physical Laboratory, University of California, San Diego. His research interests include propagation, optimization, modeling, and inversion of electromagnetic signals.

Peter Gerstoft received the M.Sc. and Ph.D. degrees from the Technical University of Denmark, Lyngby, in 1983 and 1986, respectively, and the M.Sc. degree from the University of Western Ontario, London, ON, Canada, in 1984.

From 1987 to 1992, he was with Ødegaard and Danneskiold-Samsøe, Copenhagen, Denmark, working on forward modeling and inversion for seismic exploration, with a year as a Visiting Scientist at the Massachusetts Institute of Technology, Cambridge. From 1992 to 1997, he was Senior Scientist at SACLANT Undersea Research Centre, La Spezia, Italy, where he developed the SAGA inversion code, which is used for ocean acoustic and electromagnetic signals. Since 1997, he has been with the Marine Physical Laboratory, University of California, San Diego. His research interests include global optimization, modeling, and inversion of acoustic, elastic, and electromagnetic signals.

Dr. Gerstoft is a Fellow of the Acoustical Society of America and an elected member of the International Union of Radio Science, Commission F.

William S. Hodgkiss (S'68–M'75) was born in Bellefonte, PA, on August 20, 1950. He received the B.S.E.E. degree from Bucknell University, Lewisburg, PA, in 1972 and the M.S. and Ph.D. degrees in electrical engineering from Duke University, Durham, NC, in 1973 and 1975, respectively.

From 1975 to 1977, he was with the Naval Ocean Systems Center, San Diego, CA. From 1977 to 1978, he was a Faculty Member in the Electrical Engineering Department, Bucknell University. Since 1978, he has been a Member of the Faculty, Scripps Institution of Oceanography, University of California, San Diego, and on the staff of the Marine Physical Laboratory, where he is currently Deputy Director. His present research interests are in the areas of adaptive array processing, propagation modeling, and environmental inversions with applications of those to underwater acoustics and electromagnetic wave propagation.

Dr. Hodgkiss is a Fellow of the Acoustical Society of America.