Retrieval of Forest Parameters Using a Fractal-Based Coherent Scattering Model and a Genetic Algorithm

Yi-Cheng Lin, *Member, IEEE,* and Kamal Sarabandi, *Senior Member, IEEE*

Abstract— In this paper, a procedure for retrieval of forest parameters is developed using the recently developed fractal-based coherent scattering model (FCSM) and a stochastic optimization algorithm. Since the fractal scattering model is computationally extensive, first a simplified empirical model with high fidelity for a desired forest stand is constructed using FCSM. Inputs to the empirical model are the influential structural and electrical parameters of the forest stand, such as the tree density, tree height, trunk diameter, branching angle, wood moisture, and soil moisture. Other finer structural features are embedded in the fractal model. The model outputs are the polarimetric and interferometric response of the forest as a function of the incidence angle. In this study, a genetic algorithm (GA) is employed as a global search routine to characterize the input parameters of a forest stand from a set of measured polarimetric/interferometric backscatter responses of the stand. The success of the inversion algorithm is demonstrated using a set of measured single-polarized interferometric synthetic aperture radar (SAR) data and several FCSM simulation results.

Index Terms— Ecology, interferometric SAR, inversion algorithm, polarimetric SAR, radar remote sensing.

I. INTRODUCTION

RETRIEVAL of gross biophysical parameters of forest stands, such as basal area, tree height, and leaf area index (LAI), is of great importance in many environmental research programs. Radar remote sensing at lower microwave frequencies has been proposed as a sensitive instrument for such applications [1], [2]. In support of programs pertaining to radar remote sensing of vegetation, many advanced polarimetric (SIR-C, AIRSAR) [3] and interferometric (TOPSAR) [4] radar instruments have been developed.

The study of the inversion problems in geophysical science and engineering has been of great importance from the onset of the remote-sensing science [5], [6]. For example, in microwave remote sensing of vegetation, the inverse problem is defined as the application of the measured quantities, such as the polarimetric backscattering coefficients [from synthetic aperture radar (SAR)] [7] and/or the scattering phase center heights (from an interferometric SAR) [8], [9] in an algorithm to retrieve forest parameters, such as tree type, tree density and height, and moisture content of vegetation and soil.

Over the past two decades, significant effort has been devoted toward the development of scattering models for vegetation canopies [10]–[15] as well as inversion models to retrieve forest parameters from the measured data [16]–[18]. So far, the emphasis of the scattering model development has been on the construction of simplified models with as few input parameters as possible, so that the inversion problem becomes tractable. In this process, the importance of structural features of the canopy (particle arrangement), coherence effects, and multiple scattering were ignored. Even with these simplifications, the inversion process is rather complex. In [16] and [18], neural network approaches are suggested for the inversion process, in which extensive computer simulations or experimental results are used to train a neural network in reverse order (the model outputs are fed as the input to the program). This method is computationally extensive and its success depends on the fidelity and the extent of the training data. In [17], a gradient-based search routine is applied to a nested linearized model. This model is computationally efficient; however, its applicability is limited to models with small dimensionality, and its success depends on the fidelity of the forward model.

This paper describes the application of a high-fidelity scattering model in an inversion process based on a stochastic global search method. Basically, a recently developed coherent scattering model that preserves the structural features of tree canopies using fractal models is employed to generate simplified empirical models (for different tree species) that can predict the polarimetric and interferometric radar response of a forest stand efficiently and accurately. The premise for the successful development of such empirical models stems from the fact that the model outputs are averaged quantities, such as backscattering coefficients or the mean height of the scattering phase center, and therefore are very gentle functions of model inputs.

As demonstrated in [15], the fractal-based coherent scattering model (FCSM) offers two advantages over traditional scattering models; namely, FCSM is more versatile and accurate. Basically, FCSM is a first-order scattering model and capable of simulating the fully polarimetric (including the phase statistics) and polarimetric–interferometric (scattering phase centers and correlation coefficients for any polarization configuration [19]) radar responses of coniferous and deciduous forest stands. High accuracy is achieved by FCSM through incorporating the coherent effects among the individual scatterers and scattering components and by accounting for the accurate position of scatterers, which is manifested in inhomogeneous
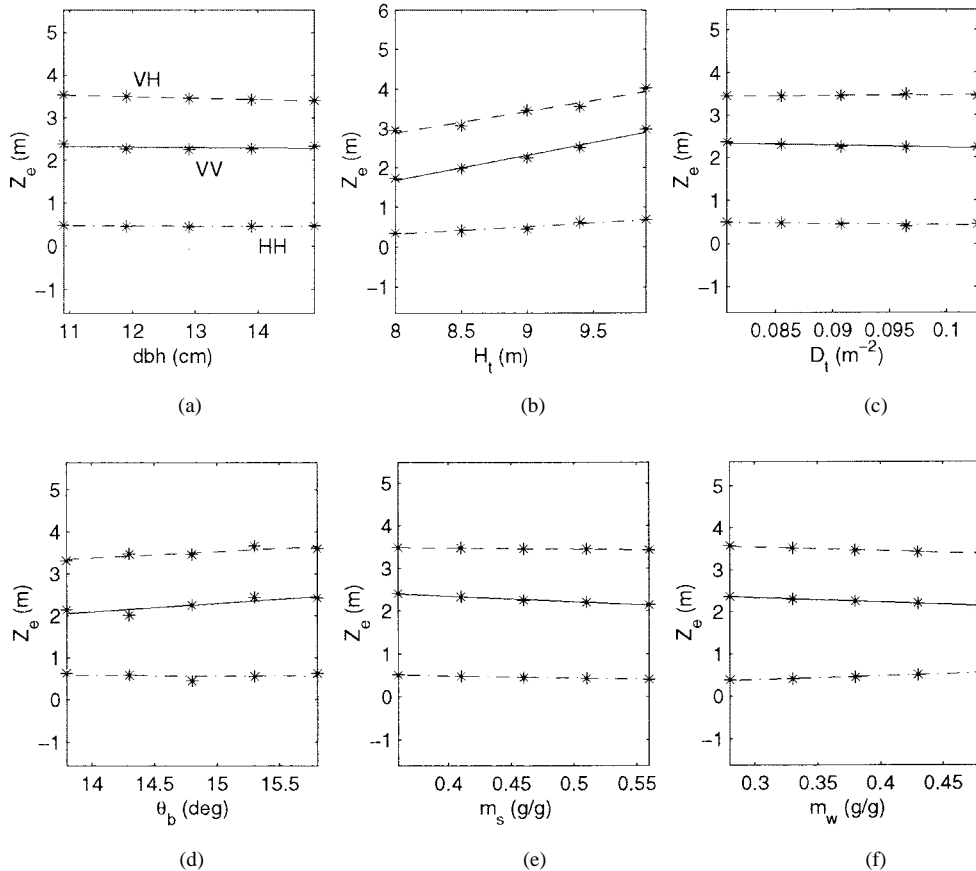
Fig. 1.   Sensitivity analysis of the C-band polarimetric scattering phase center height as a function of the physical parameters: (a) trunk diameter, (b) tree height, (c) tree density, (d) branching angle, (e) soil moisture, and (f) wood moisture, simulated at incidence angle $\theta = 25°$.

scattering and extinction profiles. This versatility and accuracy, however, has been achieved at the expense of the model complexity, which demands extensive computational power. For example, the number of input parameters needed to accurately characterize the tree structures and the environment may easily exceed 30 (it should be noted that once a tree type is chosen much fewer free parameters are needed to model the natural variabilities). On the other hand, to obtain a solution with a reasonable accuracy in the Monte Carlo simulation, a sufficiently large number (≥ 100) of realizations are required. The required computational time for each simulation limits the model's utility in inverse processes, which may demand the calculation of the forward problem many times.

To circumvent the aforementioned problem, development of empirical models based on FCSM is proposed. Construction of an empirical model can be achieved using a standard procedure, such as a curve-fitting and regression method. Unlike physical models, empirical models are simple mathematical expressions formed from a set of data acquired from measurements or a physical model prediction. Once empirical formulas are obtained, they are easy to use and require minimal computation time. It should be noted, however, that an empirical model is usually valid only for a specific case within a certain range of the parameter space over which the model is constructed.

For the development of the empirical model used in this study, first a sensitivity analysis is conducted to determine the

significant parameters, the number of which determines the dimensionality of the input vector space. A red pine stand is chosen in this paper, and six parameters are selected as the input parameters. Each selected parameter is allowed to have about 30% variation with respect to a centroid. Using the Monte Carlo simulation results obtained from FCSM, a database is constructed by varying the individual parameters over a prescribed range of the input vector space around the centroid. The parameters at the centroid are obtained from the ground truth data of a red pine test stand (Stand 22) in Raco, MI.

For the inversion process, first a least-squares estimator is used and shown to work properly when the number of measured channels is equal to or larger than the dimension of the input vector space. But since this may not be the case in general situations, a genetic algorithm (GA) [20] is developed and employed as a search routine for the nonlinear optimization problem. GA's are known to be very successful when the dimension of the input vector space is large and/or when the objective function is nonlinear.

II. EMPIRICAL MODEL DEVELOPMENT

In general, the output of the Monte Carlo coherent scattering model can be expressed as

$$\mathbf{M} = \mathcal{L}(f, p, \theta; d_a, H_t, D_t, \theta_b, m_s, m_w) \tag{1}$$
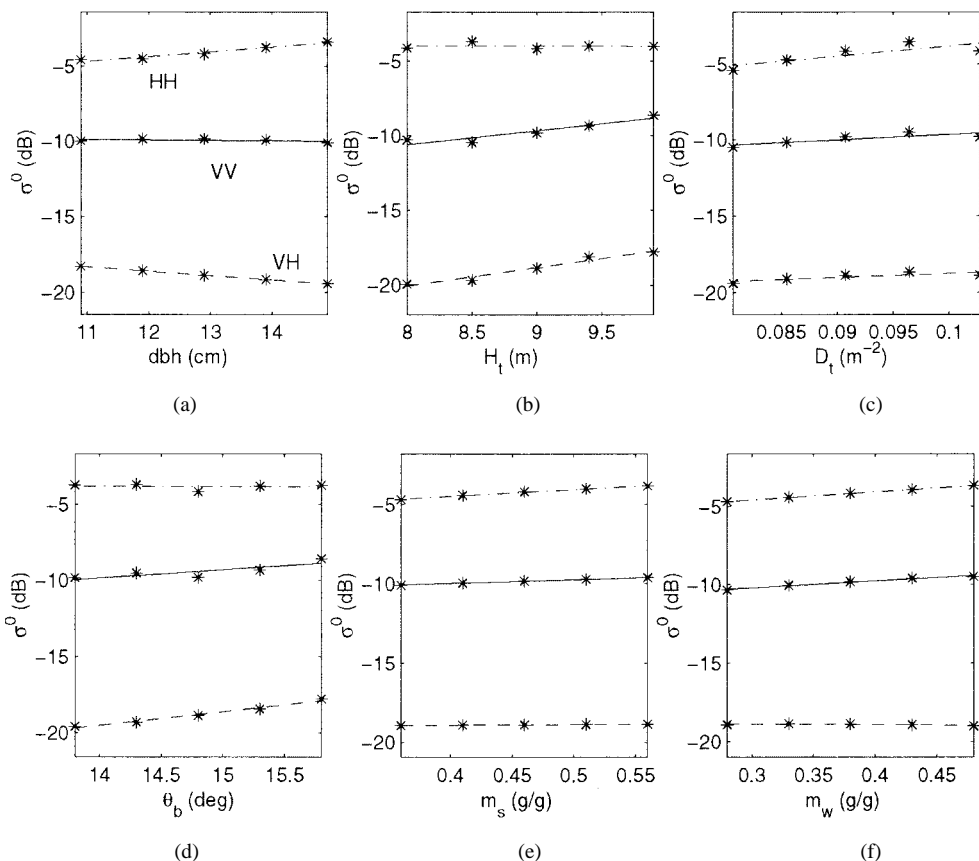
Fig. 2. Sensitivity analysis of the C-band polarimetric backscattering coefficient as a function of the physical parameters: (a) trunk diameter, (b) tree height, (c) tree density, (d) branching angle, (e) soil moisture, and (f) wood moisture, simulated at incidence angle $\theta = 25°$.

TABLE I
RANGES OF THE SELECTED GROUND TRUTH PARAMETERS AND THE
CORRESPONDING PERCENTAGE VARIATIONS TO THE CENTROID

Parameter	Range	Variation
Trunk Diameter	$10.9 \sim 14.9 (cm)$	31 %
Tree Height	$8.0 \sim 9.9 (m)$	21 %
Tree Density	$807 \sim 1027$ (trees/Hectare)	24 %
Branching Angle	$13.8 \sim 15.8$ (deg)	14 %
Soil Moisture	$0.36 \sim 0.56$ (g/g)	43 %
Wood Moisture	$0.28 \sim 0.48$ (g/g)	52 %

TABLE II
INVERSION RESULTS USING THE INTERFEROMETRIC TOPSAR DATA
AS THE MEASURED CHANNELS $\bar{\mathbf{M}}$. HERE $\mathbf{x}$ IS THE ACTUAL GROUND
TRUTH DATA AND $\mathbf{x}'$ IS THE OUTPUT OF THE INVERSION PROCESS

Measured	$Z_c^{vv}(\theta = 39°) = 3.6m$			$\sigma_{vv}^0(\theta = 39°) = -10.26$ dB		
Channels $\bar{\mathbf{M}}$	$Z_c^{vv}(\theta = 53°) = 6.1m$			$\sigma_{vv}^0(\theta = 53°) = -13.07$ dB		
Parameters	$d_a(cm)$	$H_t(m)$	$D_t(\#/Hectare)$	θ_b	m_s	m_w
Ground truth $\mathbf{x}$	12.9	9.00	907	14.8°	0.46	0.38
Inverted $\mathbf{x}'$	12.4	9.37	915	13.9°	0.44	0.37

where $\mathcal{L}$ is a complex operator relating the input and output of the model and the output $\mathbf{M}$ is a vector that may contain the backscattering coefficient (σ_{vv}^0, σ_{vh}^0, σ_{hh}^0), scattering matrix phase difference statistics, the scattering phase center height Z_e, or the interferogram correlation coefficient. The input parameters are divided into two categories: 1) radar system parameters and 2) target parameters. Radar parameters include

the radar frequency f, the polarization configuration p, and the incidence angle θ. The number of target parameters can be very large, consisting of the tree structural parameters and the dielectric properties of the constituent components. The number of these parameters is reduced drastically, however, once a tree type is chosen. In this case, only a few structural parameters are sufficient to allow for natural variabilities observed for that type of tree. The rest of the structural parameters are embedded in the fractal code of the tree. In this paper, we demonstrate development of an empirical model for a red pine tree, where only six free parameters are sufficient to describe the stand. These include the trunk diameter d_a, tree height H_t, tree density D_t, branching angle θ_b, soil moisture m_s, and wood moisture m_w. It should be noted that these parameters themselves are statistical in the coherent model with prescribed distribution functions and we are referring here to their mean value.

Multifrequency polarimetric SAR systems operate at discrete frequencies, usually at P-, L-, C-, and X-band, and the polarization configurations p are VV, VH, and HH. In this study, we demonstrate a model with three fundamental backscattering coefficients and the associated mean scattering phase center height as the model output and fix the frequency at C-band (5.3 GHz). The empirical model is developed to operate over the angular range 25–70°. Therefore, the output
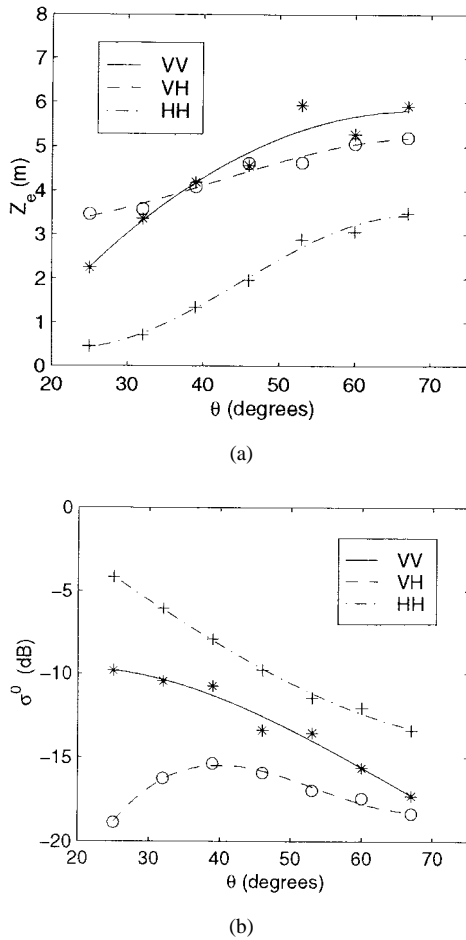
Fig. 3. Angular dependence of the polarimetric backscatter in terms of (a) the scattering phase center height Z_e and (b) the backscattering coefficient σ^0. The simulation results are fitted with polynomials of degree 3.

and input vectors $\mathbf{M}$ and $\mathbf{x}$ are defined as

$$\mathcal{L} = \begin{bmatrix} Z_e^{vv} \\ Z_e^{vh} \\ Z_e^{hh} \\ \sigma_{vv}^0 \\ \sigma_{vh}^0 \\ \sigma_{hh}^0 \end{bmatrix} \quad \text{and} \quad \mathbf{x} = \begin{bmatrix} d_a \\ H_t \\ D_t \\ \theta_b \\ m_s \\ m_w \end{bmatrix}.$$

As mentioned earlier, since no resonance behavior is expected, the output vector $\mathbf{M}$ is a gentle function of the input vector $\mathbf{x}$ and the incidence angle θ that may be related to each other via a simple empirical relationship

$$\tilde{\mathbf{M}} = \tilde{\mathcal{L}}(\theta; \mathbf{x}) \qquad (2)$$

where $\tilde{\mathcal{L}}$ is the simple empirical operator and $\tilde{\mathbf{M}}$ is the output of the empirical model. It is expected that $\tilde{\mathbf{M}}$ be as close to $\mathbf{M}$ as possible.

In general, the output parameters are nonlinear functions of the incidence angle and other input parameters. In order to establish these relationships, the coherent model must be run by varying the incidence angles and other input parameters. Through an extensive sensitivity study, it was found that over a finite domain of the input vector space, a logarithmic

relationship between the backscattering coefficient (linear in decibel scale) and a linear relation between the scattering phase center height and the input parameters exist. The dependence on the incidence angle was found to be nonlinear.

The first step in the construction of the empirical model is to choose the domain of the input vector space. In this investigation, we chose the structural parameters of a young red pine stand, a test forest stand in Raco (Stand 22), and the seasonal average of soil and vegetation moisture as the centroid of the input domain. These parameters and their range of variation used in the model development are shown in Table I. The range of parameter space is chosen so that the measured parameters of a red pine test stand (see Table II) is at the centroid of the parameter space. The Monte Carlo simulation was then carried out for specific incidence angles by varying the six free parameters within the prescribed ranges. The average scattering phase center heights (Z_e) for each polarization configuration and backscattering coefficients (σ^0 in dB) are shown in Figs. 1 and 2 as a function of each parameter, respectively. These figures clearly demonstrate the linear relationship previously described. Hence, the output vector can be readily approximated by the Taylor series expansion of the exact model to the first order, and it is given by

$$\tilde{\mathcal{L}}(\mathbf{x}) = \mathcal{L}(\mathbf{x}_0) + \mathbf{A} \cdot (\mathbf{x} - \mathbf{x}_0) \qquad (3)$$

where $\mathbf{x}_0$ denotes the input vector at the centroid and $\mathbf{A}$ is the matrix of partial derivatives whose ijth element is given by

$$a_{ij} = \left. \frac{\partial \mathcal{L}_i}{\partial x_j} \right|_{\mathbf{x}=\mathbf{x}_0}. \qquad (4)$$

a_{ij} simply represents the derivative of the ith output channel $\mathcal{L}_i$ with respect to the jth input parameter x_j, evaluated at the centroid $\mathbf{x}_0$.

In this matrix, each element was evaluated by calculating the slope of a fitting line over five sample points based on a least-squares method. In Figs. 1 and 2, the symbols ($*$) are the simulation results and the lines are the best linear estimation. It should be pointed out that each point in each figure represents an ensemble average of 200 realizations of the Monte Carlo simulation. This indicates that the initial task of generating a matrix of coefficients is very tedious and timeconsuming. Once the empirical model is obtained, however, it can provide a highly accurate solution to an arbitrary input in almost real time. This property of the empirical model is especially important in the inversion processes.

Results in Figs. 1 and 2 are for a fixed incidence angle $\theta = 25°$. The simulations at other incidence angles, however, show that the general form of (3) is valid for all incidence angles with the exception that $\tilde{\mathcal{L}}(\mathbf{x}_0)$ and $\mathbf{A}$ are functions of the incidence angle, i.e.,

$$\tilde{\mathcal{L}}(\theta; \mathbf{x}) = \mathcal{L}^0(\theta) + \mathbf{A}(\theta) \cdot (\mathbf{x} - \mathbf{x}_0). \qquad (5)$$

It is found that $\mathcal{L}^0(\theta)$ and $\mathbf{A}(\theta)$ are nonlinear, but gentle, functions of the incidence angle θ over the range of interest (25–70°). In order to obtain the functional form of $\mathcal{L}^0$ and $\mathbf{A}$ on θ, the aforementioned Monte Carlo simulation was repeated
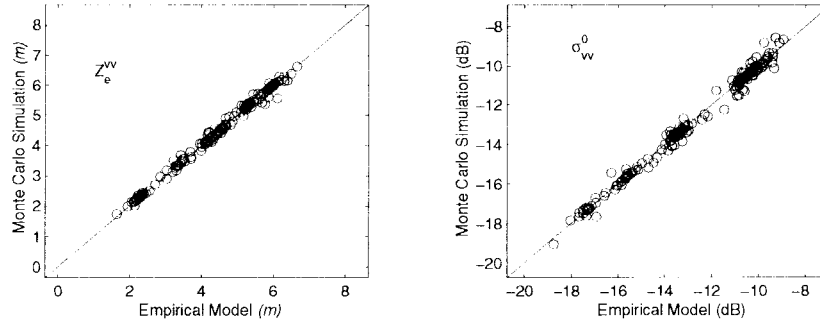
Fig. 4.   Comparison between the empirical model and the Monte Carlo coherent scattering model for a red pine stand at VV-polarization configuration.

Fig. 5.   Comparison between the empirical model and the Monte Carlo coherent scattering model for a red pine stand at VH-polarization configuration.
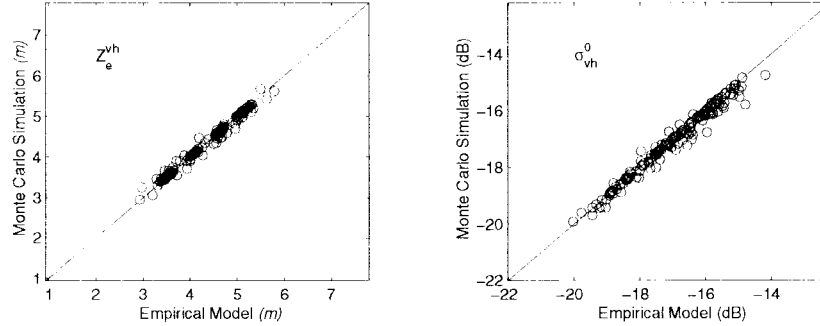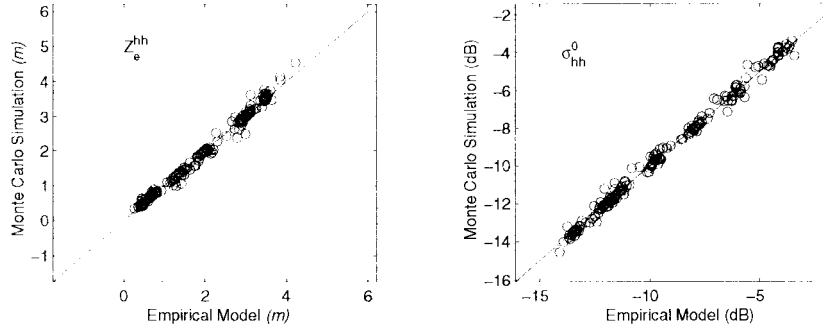

Fig. 6.   Comparison between the empirical model and the Monte Carlo coherent scattering model for a red pine stand at HH-polarization configuration.

at several different incidence angles and the corresponding values $\mathcal{L}^0$ and $\mathbf{A}$ were evaluated. Polynomial functions are used to capture the angular variations of $\mathcal{L}^0$ and $\mathbf{A}$. It was found that $\mathcal{L}^0$ and $\mathbf{A}$ can be accurately expressed by

$$\mathcal{L}^0(\theta) = \mathcal{L}_0 + \mathcal{L}_1\theta + \mathcal{L}_2\theta^2 + \mathcal{L}_3\theta^3 \tag{6}$$

and

$$\mathbf{A}(\theta) = \mathbf{A}_0 + \mathbf{A}_1\theta + \mathbf{A}_2\theta^2 + \mathbf{A}_3\theta^3 + \mathbf{A}_4\theta^4 \tag{7}$$

where $\mathcal{L}_i$ and $\mathbf{A}_i$ are 6×1 and 6×6 matrices whose values are reported in the Appendix. Fig. 3 compares the results of the empirical model given by (5) with those of the Monte Carlo simulation at the centroid ($\mathbf{x} = \mathbf{x}_0$). It should be noted that the choice of the output parameters are arbitrary and depends on the available set of input data. For example, an empirical model for a two-frequency system with three backscattering coefficients could be developed using the same procedure.

Equation (5) represents the overall empirical model whose accuracy can be evaluated through a comparison with the Monte Carlo simulations. For this comparison, a large number

of Monte Carlo simulations with independent input vectors were carried out. Figs. 4–6 show the comparison between the results of the empirical model and those of the Monte Carlo coherent model using 200 independent input data sets randomly selected within the aforementioned domain of the empirical model. The figures show excellent agreement between the empirical model and the Monte Carlo coherent model, noting that the convergence criteria for the Monte Carlo model is ± 0.5 dB. Having confidence on a fast and accurate empirical model, the inversion processes can be attempted, which is the subject of the next section.

III. INVERSION ALGORITHMS

Consider a physical system whose input–output relation is expressed by $\mathbf{M} = \tilde{\mathcal{L}}(\mathbf{x})$, where in general $\mathbf{M}$ and $\mathbf{x}$ are multidimensional vectors of arbitrary length. The inverse problem is mathematically defined as $\mathbf{x} = \tilde{\mathcal{L}}^{-1}(\mathbf{M})$, subject to certain physical constraints. Although the inverse problem may be well-defined mathematically, in practice, the inverse

solution may not exist for two reasons: 1) mathematical construction of the model may not be exact and 2) the measured vector $\mathbf{M}$ may not be exact because of measurement errors. Hence, instead of casting the problem in terms of an inverse problem, the problem of finding $\mathbf{x}$ is usually cast in terms of a constraint minimization problem.

Suppose there exists a set of measurements $\tilde{\mathbf{M}}$, the problem is defined as characterization of $\mathbf{x}$ so that the objective function (or error function), defined by

$$\mathcal{E}(\mathbf{x}) = \|\mathcal{L}(\mathbf{x}) - \tilde{\mathbf{M}}\|^{2} \tag{8}$$

is minimized over a predefined domain for $\mathbf{x}$. Here, $\|\cdot\|$ denotes the norm of the argument. As mentioned earlier, there are a number of inversion processes available in the literature; however, in this paper, by constructing a simple empirical model, a traditional least-squares minimization approach and a stochastic global minimization method are examined.

A. Least-Squares Approach

As shown in Section II, the scattering problem can be cast in terms of a linear system of equations of the form $\tilde{\mathcal{L}}(\mathbf{u}) = \mathbf{A}\mathbf{u}$, where $\mathbf{A}$ is an $m \times n$ matrix and $\mathbf{u}$ is an n-dimensional vector in D, $D \subset R^{n}$. For a given m-dimensional vector $\tilde{\mathbf{G}}$, (8) can be expanded as

$$\mathcal{E} = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij} u_{j} - \tilde{g}_{i} \right)^{2}. \tag{9}$$

A solution that minimizes $\mathcal{E}$ must satisfy

$$\frac{\partial \mathcal{E}}{\partial u_{j}} = 0, \qquad j = 1, 2, \cdots, n \tag{10}$$

and is referred to as the least-squares solution. It is shown that the solution of (10) $(\mathbf{u_m})$ can be obtained from the solution of the following matrix equation [21]:

$$(\mathbf{A}^{*}\mathbf{A}) \cdot \mathbf{u_m} = \mathbf{A}^{*} \cdot \tilde{\mathbf{G}}. \tag{11}$$

Here, $\mathbf{A}^{*}$ is the transpose of $\mathbf{A}$. It is also demonstrated that the solution $\mathbf{u_m} = (\mathbf{A}^{*}\mathbf{A})^{-1}\mathbf{A}^{*}\tilde{\mathbf{G}}$ exists if $\text{rank}(\mathbf{A}) = n$. This requirement states that the number of independent equations should exceed the number of unknowns.

To apply (11) to our empirical model using (3), it is noted that

$$\tilde{\mathcal{L}}(\mathbf{x}) - \mathcal{L}^{0} = \mathbf{A} \cdot (\mathbf{x} - \mathbf{x_0}). \tag{12}$$

Thus, we use the substitution $\mathbf{u} = \mathbf{x} - \mathbf{x_0}$ and $\tilde{\mathbf{G}} = \tilde{\mathbf{M}} - \mathcal{L}^{0}$. Here $\mathcal{L}^{0}$ and $\mathbf{A}$ are evaluated from (6) and the solution is given by

$$\mathbf{x_m} = \mathbf{x_0} + (\mathbf{A}^{*}\mathbf{A})^{-1}\mathbf{A}^{*}(\tilde{\mathbf{M}} - \mathcal{L}^{0}). \tag{13}$$

The least-squares solution may not be suitable for the inverse problem at hand for two reasons. First, the number of output channels m is usually less than that of unknown parameters n. In this case, $\text{rank}(A) < n$ and $\mathbf{A}^{*}\mathbf{A}$ is not invertible. Even when the number of channels is larger than the unknowns, the solution provided by (13) may not be accurate.
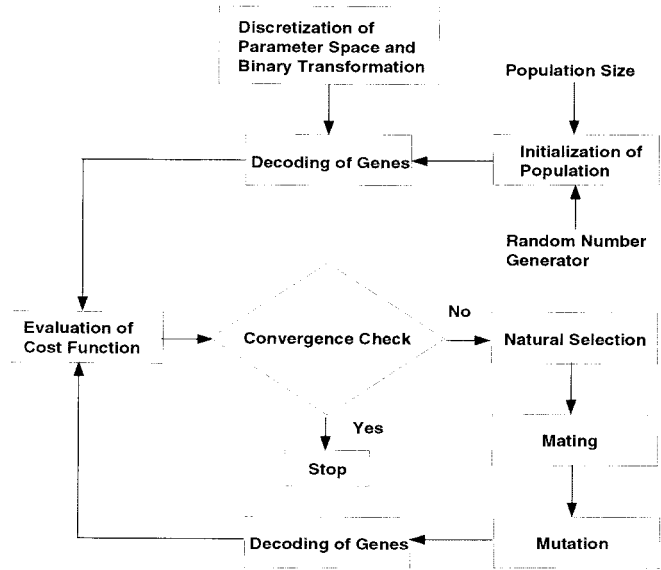

Fig. 7. Flow chart of a GA.

This happens when $\mathbf{A}^{*}\mathbf{A}$ is ill-conditioned. Basically, some elements of $(A^{*}A)^{-1}$ become very large, which amplify the errors in $\tilde{\mathbf{M}}$ [22].

B. Genetic Algorithms

In recent years, applications of GA's to a variety of optimization problems in electromagnetics have been successfully demonstrated [24], [25]. The fundamental concept of GA's is based on the concept of natural selection in the evolutionary process, which is accomplished by genetic recombination and mutation. The algorithms are based on a number of *ad hoc* steps, including the following:

1) discretization of the parameter space;
2) development of an arbitrary encoding algorithm to establish a one-to-one relationship between each code and the discrete points of the parameter space;
3) random generation of a trial set known as the initial population;
4) selection of high-performance parameters according to the objective function known as natural selection;
5) mating and mutation;
6) recursion of steps 4) and 5) until a convergence is reached.

Fig. 7 shows the flow chart of GA's. Note that the population size is provided by the user and an initial population of the given size is generated randomly.

In this study, since we have as many as six input ground truth parameters and six output channels, it is expected that the objective function is complex and highly nonlinear containing many local minima. In this case, the traditional gradient-based optimization methods usually converge to a local minimum and fail to locate the inverted data. One interesting feature of GA's is that the method would provide a list of optimal solutions instead of a solution. This is important in the sense that a solution that best meets the physical constraints (not
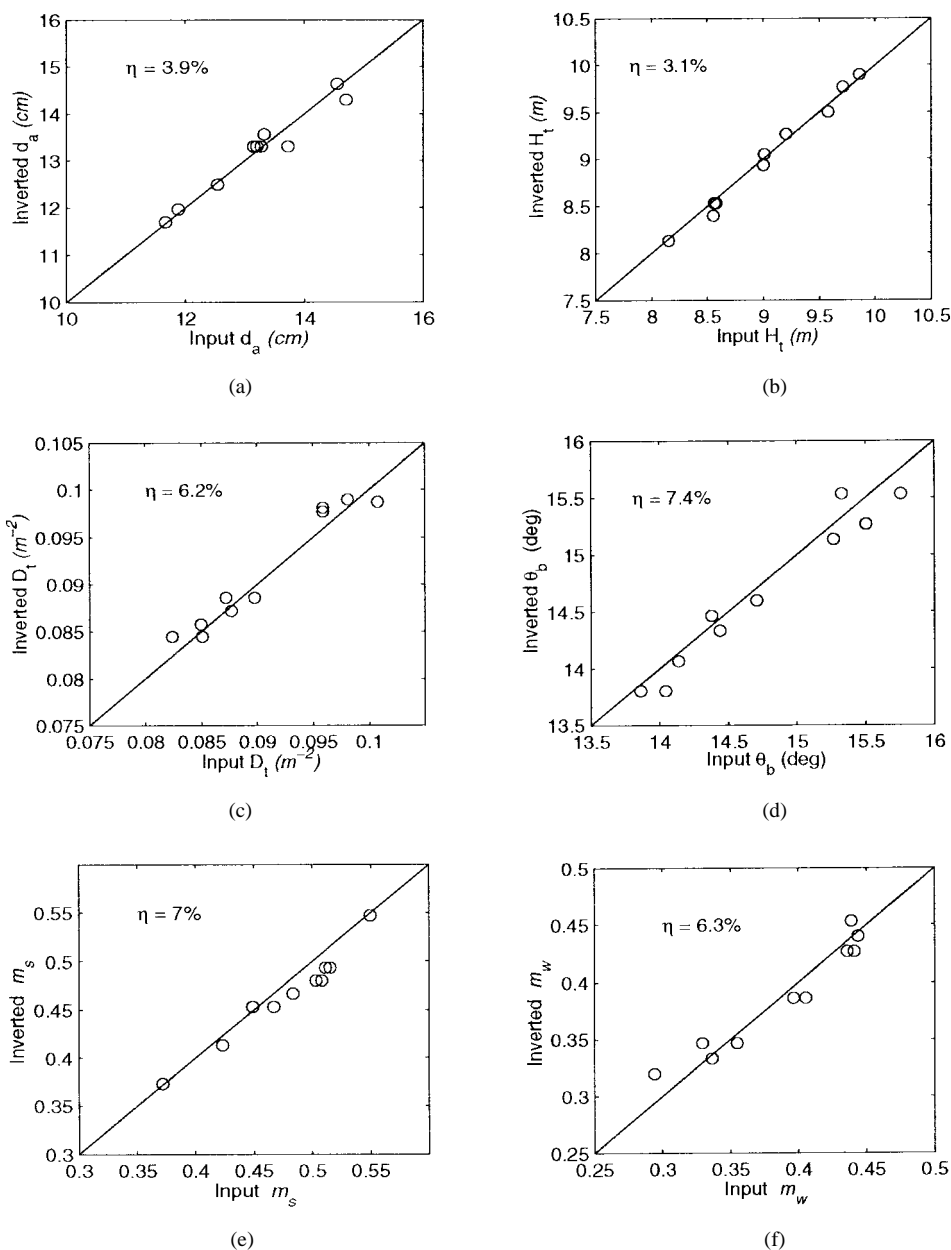
Fig. 8. Comparison of the input parameters ($\mathbf{x}$) and the output of the inversion algorithm ($\mathbf{x}'$) using the synthetic data obtained from the Monte Carlo simulation, for (a) trunk diameter d_a, (b) tree height H_t, (c) tree density D_t, (d) branching angle θ_b, (e) soil moisture m_s, and (f) vegetation moisture m_w. Here η is a measure of the average error in the inversion process defined by (15).

included in the objective function) may be selected from the list of optimal solutions.

For this problem, each of the input parameters was discretized and encoded into a 4-bit binary code, creating a discrete input vector space with 2^{24} members. A population of 240 members was used for each generation and the objective function was defined by

$$\mathcal{E}(\mathbf{x}) = \|\mathbf{w} \cdot [\tilde{\mathbf{M}} - \mathcal{L}^0 - \mathbf{A} \cdot (\mathbf{x} - \mathbf{x}_0)]\|^2 \qquad (14)$$

where $\mathbf{w}$ is a user-defined weighting function assigned to individual output channels. To examine the performance of this GA-based inversion algorithm, many arbitrary points within the domain of the input vector space were selected and then the Monte Carlo simulation was used to evaluate

the polarimetric backscattering coefficients and the scattering phase center heights at 5.3 GHz. The output of FCSM for these simulations were used as a synthetic measured data set $\tilde{\mathbf{M}}$ for the inversion algorithm. Fig. 8 shows the performance of the inversion algorithm through comparisons of the input parameters $\mathbf{x}$ and the inverted parameters $\mathbf{x}'$. Also shown in each of the figures is the calculated average error η, defined by

$$\eta = \frac{\sum_{j=1}^{N} |x_j - x'_j|}{N \Delta x} \qquad (15)$$

where N is the number of points ($N = 10$ in this case) and Δx is the range of validity of the parameter according to the empirical model. It should be noted here that the quantization
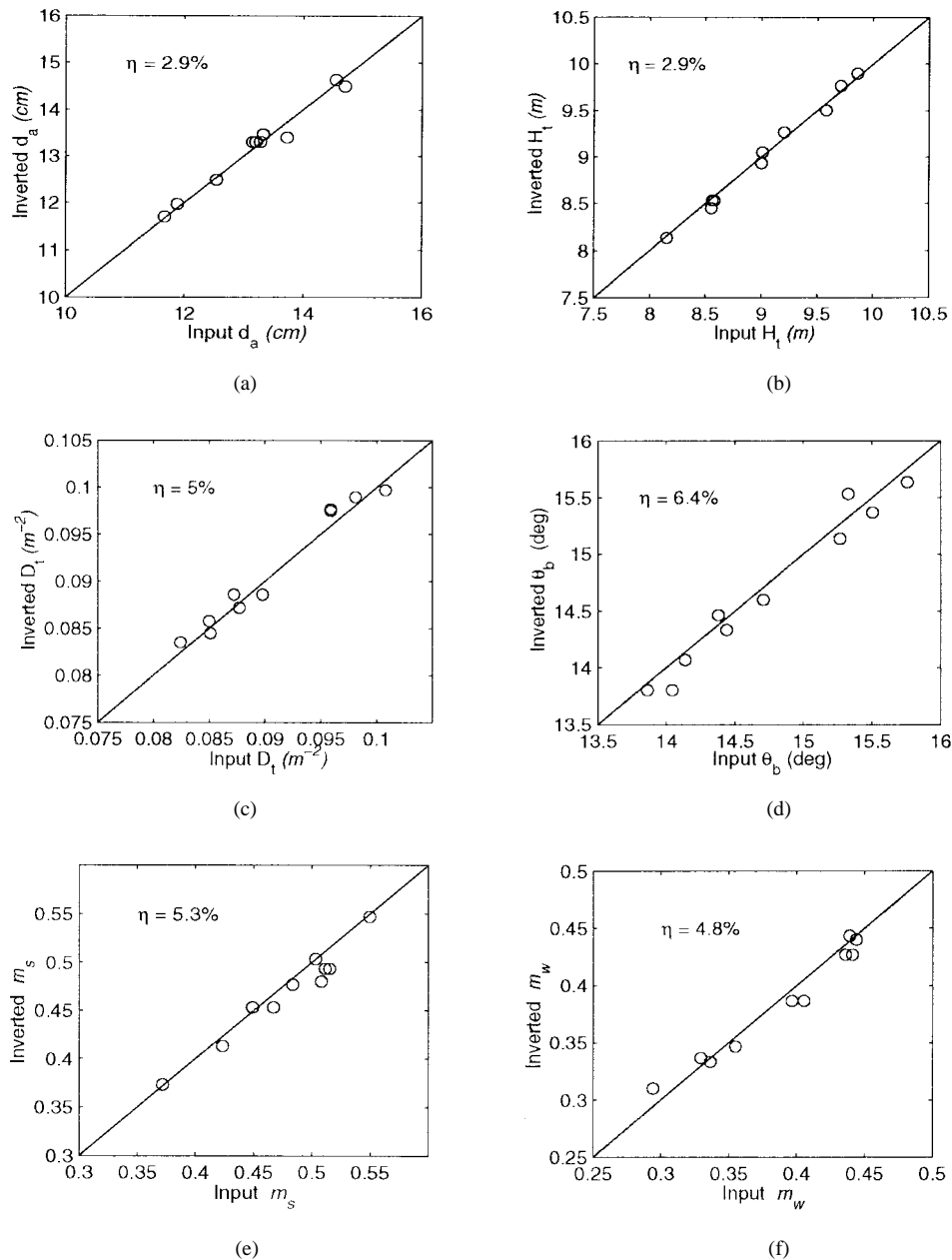
Fig. 9.   Comparison of the input parameters ($\mathbf{x}$) and the output of the inversion algorithm ($\mathbf{x}'$) using the synthetic data obtained from the empirical model, for (a) trunk diameter d_a, (b) tree height H_t, (c) tree density D_t, (d) branching angle θ_b, (e) soil moisture m_s, and (f) vegetation moisture m_w. Here η is a measure of the average error in the inversion process defined by (15).

error for 4-bit quantization ($\pm 3\%$) is also included in the results. To examine the importance of the quantization error and the stochastic nature of the solution in the inversion process, the inversion process was applied to another set of synthetic measurement data generated by the empirical model. Fig. 9 shows the comparison between the actual input $\mathbf{x}$ and the inverted solution $\mathbf{x}'$. It is noticed that the error in Fig. 9 is slightly smaller than those obtained from Fig. 8. This indicates that the quantization error and the stochastic nature of the solution are considerable factors on the overall error. Increasing the quantization level to five bits increases the members of the input vector space by factor of 2^6. This slows down the inversion process since the population in each generation must also be increased. This does not, however,

improve the overall accuracy drastically as the errors inherent in the empirical model and those caused by the stochastic nature of the GA solution are independent of quantization error.

At last, the developed inversion algorithm is tested using the real measured data acquired by the Jet Propulsion Laboratory TOPSAR over a test stand of red pine forest in Raco. Although only four data points (C-band VV-polarized backscattering coefficients and scattering phase center heights at incidence angles $\theta = 39°$ and $53°$) are available, the inversion algorithm can be easily modified via the objective function of the GA. In this case, the objective function is given by

$$\mathcal{E}(\mathbf{x}) = \mathcal{E}_1(\mathbf{x}) + \mathcal{E}_2(\mathbf{x}) \qquad (16)$$

$$[\mathcal{L}_0 \quad \mathcal{L}_1 \quad \mathcal{L}_2 \quad \mathcal{L}_3] = \begin{bmatrix} -3.1051 & 0.2642 & -0.0020 & 0.0000 \\ 3.6061 & -0.0659 & 0.0029 & 0.0000 \\ 3.6264 & -0.3188 & 0.0094 & -0.0001 \\ -12.2843 & 0.2713 & -0.0079 & 0.0000 \\ -50.1483 & 2.1623 & -0.0431 & 0.0003 \\ 3.5502 & -0.3062 & -0.0005 & 0.0000 \end{bmatrix}$$

and

$$\mathbf{A}_0 = \begin{bmatrix} 1.5294 & -0.2244 & 49.0879 & 10.8663 & 5.2289 & -59.6487 \\ -0.6424 & 0.2687 & -0.5762 & 0.7727 & -2.4209 & -7.0918 \\ 7.1314 & 0.1875 & 10.3344 & -7.0854 & -3.6540 & -2.5728 \\ -13.5348 & -1.8434 & -45.9425 & 16.2122 & 2.2269 & -108.8944 \\ -1.0359 & -1.0341 & 22.6264 & 5.2748 & 6.6677 & -73.7306 \\ -4.7396 & -2.3518 & -224.1260 & -6.9444 & -4.8994 & 114.7993 \end{bmatrix}$$

$$\mathbf{A}_1 = \begin{bmatrix} -0.1943 & 0.0320 & -4.2781 & -0.9772 & -0.7657 & 5.5036 \\ 0.0581 & -0.0244 & 0.0609 & -0.0693 & 0.1572 & 0.5872 \\ -0.7710 & -0.0071 & -0.5252 & 0.6810 & 0.3879 & 0.3659 \\ 1.3292 & 0.1763 & 3.9880 & -1.3883 & 0.3188 & 11.9278 \\ 0.0508 & 0.1137 & -2.1445 & -0.2531 & -0.5110 & 6.0220 \\ 0.6771 & 0.1919 & 19.5566 & 0.7884 & 0.9928 & -10.1014 \end{bmatrix}$$

$$\mathbf{A}_2 = \begin{bmatrix} 0.0086 & -0.0013 & 0.1302 & 0.0316 & 0.0312 & -0.1844 \\ -0.0019 & 0.0010 & -0.0025 & 0.0029 & -0.0038 & -0.0198 \\ 0.0297 & -0.0002 & 0.0005 & -0.0233 & -0.0161 & -0.0121 \\ -0.0469 & -0.0057 & -0.1266 & 0.0433 & -0.0218 & -0.4499 \\ -0.0010 & -0.0040 & 0.0673 & 0.0035 & 0.0144 & -0.1697 \\ -0.0305 & -0.0052 & -0.6039 & -0.0325 & -0.0354 & 0.3303 \end{bmatrix}$$

$$\mathbf{A}_3 = \begin{bmatrix} -0.0002 & 0.0000 & -0.0017 & -0.0004 & -0.0005 & 0.0027 \\ 0.0000 & -0.0000 & 0.0000 & -0.0001 & 0.0000 & 0.0003 \\ -0.0005 & 0.0000 & 0.0002 & 0.0003 & 0.0003 & 0.0001 \\ 0.0007 & 0.0001 & 0.0017 & -0.0006 & 0.0004 & 0.0072 \\ 0.0000 & 0.0001 & -0.0009 & -0.0000 & -0.0002 & 0.0021 \\ 0.0005 & 0.0001 & 0.0079 & 0.0006 & 0.0005 & -0.0047 \end{bmatrix}$$

$$\mathbf{A}_4 = \begin{bmatrix} 0.0097 & -0.0013 & 0.0795 & 0.0222 & 0.0306 & -0.1417 \\ -0.0014 & 0.0009 & -0.0025 & 0.0033 & -0.0016 & -0.0148 \\ 0.0278 & -0.0008 & -0.0218 & -0.0176 & -0.0142 & -0.0044 \\ -0.0390 & -0.0038 & -0.0831 & 0.0302 & -0.0284 & -0.4111 \\ -0.0003 & -0.0033 & 0.0469 & -0.0002 & 0.0079 & -0.0906 \\ -0.0339 & -0.0023 & -0.3794 & -0.0357 & -0.0218 & 0.2407 \end{bmatrix} \cdot 10^{-4}$$

where

$$\mathcal{E}_1(\mathbf{x}) = \|\mathbf{w} \cdot [\tilde{\mathbf{M}}_1 - \mathcal{L}_1^0 - \mathbf{A}_1 \cdot (\mathbf{x} - \mathbf{x}_0)]\|^2 \qquad (17)$$

$$\mathcal{E}_2(\mathbf{x}) = \|\mathbf{w} \cdot [\tilde{\mathbf{M}}_2 - \mathcal{L}_2^0 - \mathbf{A}_1 \cdot (\mathbf{x} - \mathbf{x}_0)]\|^2. \qquad (18)$$

Here the subscripts 1 and 2 denote, respectively, the case for the incidence angle $\theta = 39°$ and $53°$. Note that the weighting function and the measured vectors in this case are written as

$$\mathbf{w} = [1 \quad 0 \quad 0 \quad 1 \quad 0 \quad 0] \qquad (19)$$

$$\tilde{\mathbf{M}}_1 = [Z_e^{vv}(\theta = 39°) \quad 0 \quad 0 \quad \sigma_{vv}^0(\theta = 39°) \quad 0 \quad 0]^t \qquad (20)$$

$$\tilde{\mathbf{M}}_2 = [Z_e^{vv}(\theta = 53°) \quad 0 \quad 0 \quad \sigma_{vv}^0(\theta = 53°) \quad 0 \quad 0]^t. \qquad (21)$$

The simulation results are compared with ground truth data [2] in Table II, where a very good agreement is shown.

IV. CONCLUSIONS

In this paper, a simplified empirical model was developed using a high-fidelity Monte Carlo coherent scattering model to be incorporated in an efficient inversion algorithm. The empirical model was specifically developed for a red pine forest stand that provides simple expressions for the polarimetric backscattering coefficients and scattering phase center heights at C-band as a function of the incidence angle. The accuracy of the empirical model was examined by comparing its output with that of the Monte Carlo FCSM. The empirical model in conjunction with a stochastic search algorithm (GA) were used to construct an inversion algorithm. The accuracy of the inversion algorithm was demonstrated by first using synthetic-measured data generated from the empirical model and the Monte Carlo FCSM. It was shown that the inversion algorithm can accurately estimate the input parameters when synthetic data were used. Next, we applied the inversion algorithm to an actual data set, obtained from TOPSAR, composed of VV-polarized backscattering coefficient and scattering phase center height at C-band and at two incidence angles. Excellent

agreement was obtained between the ground truth data and the output of the inversion algorithm.

APPENDIX

In this appendix, the values of coefficient matrices $\mathcal{L}_i$ and $\mathbf{A}_i$ used in (6) and (7) are reported for the red pine stand investigated in this study, as shown at the top of the previous page.

REFERENCES

[1] M. C. Dobson, F. T. Ulaby, T. L. Toan, A. Beaudoin, and E. S. Kasischke, "Dependence of radar backscatter on conifer forest biomass," *IEEE Trans. Geosci. Remote Sensing,* vol. 30, pp. 402–415, Mar. 1992.

[2] M. C. Dobson, F. T. Ulaby, L. E. Pierce, T. L. Sharik, K. M. Bergen, J. Kellndorfer, J. R. Kendra, Y. C. Lin, A. Nashashibi, K. Sarabandi, and P. Siqueira, "Estimation of forest biophysical characteristics in Northern Michigan with SIR-C/X-SAR," *IEEE Trans. Geosci. Remote Sensing,* vol. 33, pp. 877–895, July 1995.

[3] R. L. Jordan, B. L. Huneycutt, and M. Werner, "The SIR-C/X-SAR synthetic aperture radar system," *IEEE Trans. Geosci. Remote Sensing,* vol. 33, pp. 829–839, July 1996.

[4] H. A. Zebker, S. N. Madsen, J. Martin, K. B. Wheeler, T. Miller, Y. Lou, G. Alberti, S. Vetrella, and A. Cucci, "The TOPSAR interferometric radar topographic mapping instrument," *IEEE Trans. Geosci. Remote Sensing,* vol. 30, pp. 933–940, Sept. 1992.

[5] V. Dimri, *Deconvolution and Inverse Theory: Application to Geophysical Problems.* Amsterdam, The Netherlands: Elsevier, 1992.

[6] M. K. Sen, *Global Optimization Methods in Geophysical Inversion.* Amsterdam, The Netherlands: Elsevier, 1995.

[7] F. T. Ulaby and C. Elachi, *Radar Polarimetry for Geoscience Applications.* Norwell, MA: Artech House, 1990.

[8] K. Sarabandi, "Δk-radar equivalent of interferometric SARs: A theoretical study for determination of vegetation height," *IEEE Trans. Geosci. Remote Sensing,* vol. 35, pp. 1267–1276, Sept. 1997.

[9] R. N. Treuhaft, S. N. Madsen, M. Moghaddam, and J. J. van Zyl, "Vegetation characteristics and underlying topography from interferometric radar," *Radio Sci.,* vol. 31, pp. 1449–1485, 1996.

[10] K. Sarabandi, "Electromagnetic scattering from vegetation canopies," Ph.D. dissertation, Univ. Michigan, Ann Arbor, 1989.

[11] F. T. Ulaby, K. Sarabandi, K. MacDonald, M. Whitt, and M. C. Dobson, "Michigan microwave canopy scattering model," *Int. J. Remote Sensing,* vol. 11, pp. 1223–1253, 1990.

[12] L. Tsang, C. H. Chan, J. A. Kong, and J. Joseph, "Polarimetric signature of a canopy of dielectric cylinders based on first and second order vector radiative transfer theory," *J. Electromagn. Waves Applicat.,* vol. 6, pp. 19–51, 1992.

[13] M. A. Karam, A. K. Fung, R. H. Lang, and N. H. Chauhan, "A microwave scattering model for layered vegetation," *IEEE Trans. Geosci. Remote Sensing,* vol. 30, pp. 767–784, July 1992.

[14] N. S. Chauhan, R. H. Lang, and K. J. Ranson, "Radar modeling of a boreal forest," *IEEE Trans. Geosci. Remote Sensing,* vol. 29, pp. 627–638, July 1991.

[15] Y.-C. Lin and K. Sarabandi, "A Monte Carlo coherent scattering model for forest canopies using fractal-generated trees," *IEEE Trans. Geosci. Remote Sensing,* vol. 37, pp. 440–451, Jan. 1999.

[16] L. Pierce, K. Sarabandi, and F. Ulaby, "Application of an artificial neural network in canopy scattering inversion," *Int. J. Remote Sensing,* vol. 15, pp. 3263–3270, 1994.

[17] P. F. Polatin, K. Sarabandi, and F. T. Ulaby, "An iterative inversion algorithm with application to the polarimetric radar response of vegetation canopies," *IEEE Trans. Geosci. Remote Sensing,* vol. 32, pp. 62–71, Jan. 1994.

[18] F. Amar, M. S. Dawson, and A. K. Fung, "Inversion of the relevant forest and vegetation parameters using neural networks," in *Proc. Progress Electromagn. Res. Symp. (PIERS),* 1993.

[19] K. Sarabandi and Y. C. Lin, "Simulation of interferometric SAR response for characterizing the scattering phase center statistics of forest canopies," *IEEE Trans. Geosci. Remote Sensing,* to be published.

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

[21] S. H. Friedberg, A. J. Insel, and L. E. Spence, *Linear Algebra.* Englewood Cliffs, NJ: Prentice-Hall, 1979.

[22] A. Ishimaru, *Wave Propagation and Scattering in Random Media,* vol. II. New York: Academic, 1978.

[23] J. Bosworth, "Comparison of genetic algorithms with conjugate gradient methods: Technical report," Tech. Rep. UMR0554, Comput. Commun. Sci. Dept., Univ. Michigan, Ann Arbor, 1972.

[24] K. Sarabandi and E. S. Li, "Characterization of optimum polarization for multiple target discrimination using genetic algorithms," *IEEE Trans. Antennas Propagat.,* to be published.

[25] R. L. Haupt, "An introduction to genetic algorithms for electromagnetics," *IEEE Antennas Propagat. Mag.,* vol. 37, pp. 7–15, 1995.

Yi-Cheng Lin (S'92–M'98), for a photograph and biography, see p. 451 of the January 1999 issue of this TRANSACTIONS.

Kamal Sarabandi (S'87–M'90–SM'92), for a photograph and biography, see p. 35 of the January 1999 issue of this TRANSACTIONS.