An Iterative Inversion Algorithm with Application to the Polarimetric Radar Response of Vegetation Canopies

Paul F. Polatin, Kamal Sarabandi, *Senior Member, IEEE,* and Fawwaz T. Ulaby, *Fellow, IEEE*

Abstract—**The retrieval of scene parameters from polarimetric radar data using an iterative inversion approach is considered in this paper. The theoretical development of a general, model-based iterative algorithm for inversion of polarimetric radar data is presented. Factors relevant to its implementation, such as sensor configuration, algorithm optimization and computational structure are discussed. The algorithm is applied to the specific problem of inverting the vector radiative transfer model for a simplified, representative vegetation canopy consisting of vertical trunks, leaves, and a rough ground surface. The results of this inversion are in excellent agreement with simulated data generated using the radiative transfer model. The convergence properties of the algorithm are evaluated, and it is found that successful convergence is achieved in about 90% to 95% of the cases tested for the implementation used in this work. An error analysis is presented which considers the effect of both systematic and measurement derived errors. Typical error bounds for the current application are approximately $\pm 3\%$, allowing for ± 0.5 dB accuracy in the measured radar data.**

I. Introduction

In RECENT years a great deal of emphasis has been placed on the retrieval of information from synthetic aperture radars and radar polarimeters [1]–[3]. Some of the applications have been in remote sensing of soil moisture for bare soil [4], [5] and vegetation-covered soil [6], [7]. Other applications include the determination of vegetation canopy parameters [8]–[11], sea surface characteristics [12], [13], and snow parameters [14], [15]. Within the general problem of classification of remotely sensed data, there exists the subproblem of inversion of radar data to obtain parameters of interest for the scene under observation. The vast majority of the literature in this area has been concerned with two major approaches to the problem of inversion of radar data.

The first approach involves the construction of an empirical scattering model specific to the type of problem being studied [4], [16], [17]. In this technique the scattering characteristics of a particular type of terrain or vegetation canopy are determined experimentally, usually at several frequencies using a polarimetric radar scatterometer. The results are then fit to a fairly simple equation or a set of equations that describe the scattering behavior as a function of polarization and frequency over a specified range of parameters (region of validity) for the type of terrain or canopy being characterized. The empirical model obtained in this way is designed to be invertible over its region of validity. This type of technique can give accurate results for the case it has been designed to treat but is completely specific to that case and provides little physical insight.

The second approach, which has received a large amount of attention lately, is based on the use of artificial neural networks [18]–[21]. The transfer function of the network is determined by using training sequences of known input and output data. In the present case, this would consist of polarimetric radar data as inputs and scene parameters as outputs. The network characteristics relating the radar data to the scene parameters are determined using a back propagation algorithm to successively refine the transfer function based on the set of training sequences. After the network has been trained, it can be used to estimate unknown scene parameters given an input set of radar data. The neural network has great flexibility and, depeding on the set of training data, the results may be excellent [18], [22]. However, the neural network is, in most cases, used essentially as a black box. There is no way currently known to discern the underlying physical processes that give rise to the network behavior, which means that there is no systematic way of selecting the optimum set of data channels for use in a neural network-based inversion given the network response alone. To do this, one must ultimately rely on information provided by theoretical models and/or measured data. There is also no way of determining if the decision path taken by the neural network in arriving at its result is a reasonable one.

In this paper, we present an iterative algorithm for the inversion of polarimetric radar data. The algorithm is completely general and may be applied, in principle, to any type of radar scattering problem for which a model exists. The behavior of this algorithm is derived entirely from the physical scattering mechanisms existing within the system being studied and represents the summation of knowledge gained in applying the scattering model across a representative range of states of the system. The amount of computation time required in applying the algorithm to any given case is small since all the information necessary for inversion of a set of input data has been precomputed. It is also possible to monitor the decision path taken by the algorithm in arriving at a result and therefore some measure of control over the reliability of such results is achievable.

Manuscript received August 12, 1992; revised April 15, 1993.

The authors are with the Radiation Laboratory, Department of Electrical Engineering and Computer Science, The University of Michigan, Ann Arbor, MI 48109-2122.

IEEE Log Number 9213887.

The second and third sections of this paper present the theoretical development of the iterative inversion algorithm and some important considerations involved in its implementation. Sections IV and V discuss the application of the algorithm to the particular case of retrieval of vegetation canopy parameters from radar data and give the results of an application based on inversion of the radiative transfer model for a simplified representative canopy. The final section presents the results of an error analysis for the algorithm in terms of both systematic and measured quantities and discusses how sensor configuration, algorithm implementation, and data uncertainty influence the inversion accuracy.

II. Iterative Algorithm

Let us assume that backscatter data is provided for a particular target at a given frequency and for a range of incidence angles $\theta \in [a, b]$. This data may be represented in terms of a Fourier series in the restricted angular range. The Fourier coefficients obtained in this way provide an equivalent representation of the system response for this type of target. The behavior of the Fourier coefficients as a function of the target parameters provides a convenient measure of the system response that is independent of incidence angle. If a subset of the Fourier coefficients can be found that represents the angular response of the system to sufficient accuracy and their behavior is known for all values of the target parameters, then, in effect, we have an empirical model that describes the target behavior in detail. In addition, if the functional forms of these Fourier coefficients with respect to the target parameters are known, it should be possible to construct subsets of the parameter space over which piecewise linear representations of the coefficients are obtainable. The resulting sets of linear equations should be directly invertible and, by means of an iterative process that successively converges on smaller domains, a solution set is found.

Because it is impossible to determine experimentally the behavior of a complicated physical system under all conditions, a model that represents its behavior as a function of all its important parameters is used in the construction of the empirical Fourier representation. This leaves the determination of the optimum sensor configuration to the discretion of the system designer. Let us consider then a model $\mathcal{M}$ that operates on parameters $[\alpha_1, \cdots, \alpha_m]$ and the angle θ to produce an estimate of σ^o for a particular scene

$$\hat{\sigma}^o(\alpha_1, \cdots, \alpha_m, \theta) = \mathcal{M}(\alpha_1, \cdots, \alpha_m, \theta). \tag{1}$$

We now restrict the range of validity of the model to a subset of the range $[0, \pi/2]$. In other words let $\theta \in [a, b]$ such that $0 < a \leq \theta \leq b < \pi/2$. A linear transformation $\mathcal{L}$ exists that maps the subrange into the full range

$$\theta \in [a, b] \xrightarrow{\mathcal{L}} \theta' \in \left[0, \frac{\pi}{2}\right] \tag{2}$$

with $\theta' = \beta(\theta - a)$ such that $\theta' = 0$ when $\theta = a$ and $\theta' = \pi/2$ when $\theta = b$. Thus, we find that $\beta = \pi/2(b - a)$ and $\theta = a + \theta'/\beta$. We now construct the function $g(\theta') = \hat{\sigma}^o(a + \theta'/\beta)$. Since the domain of g is $[0, \pi/2]$, we can expand g as a Fourier cosine series in θ' as follows:

$$g(\theta') = \tfrac{1}{2}C_o + \sum_{n=1}^{\infty} C_n \cos(2n\theta') \tag{3}$$

where

$$C_n = \frac{4}{\pi}\int_0^{\pi/2} g(\theta')\cos(2n\theta')\, d\theta'. \tag{4}$$

We can then

$$\hat{\sigma}^o(\theta) + \tfrac{1}{2}C_o + \sum_{n=1}^{\infty} C_n \cos[2n\beta(\theta - a)] \tag{5}$$

with

$$C_n = \frac{4\beta}{\pi}\int_a^b \hat{\sigma}^o(\theta)\cos[2n\beta(\theta - a)]\, d\theta. \tag{6}$$

These are the coefficients for the restricted range Fourier series description of $\hat{\sigma}^o(\theta)$.

We describe the general polarimetric model response as a function of angle in the angular range $[a, b]$ by its restricted range Fourier series

$$\hat{\sigma}^o_{pq}(\alpha_1, \cdots, \alpha_m; f, \theta) = \tfrac{1}{2}C^{pq}_o + \sum_{n=1}^{N} C^{pq}_n \cos[2n\beta(\theta - a)] \tag{7}$$

with

$$C^{pq}_n = C^{pq}_n(\alpha_1, \cdots, \alpha_m; f)$$

where $\alpha_1, \cdots, \alpha_m$ are the model parameters, f is the frequency and p, q is each either v or h for vertical or horizontal polarization. The series has been truncated at the $(N + 1)$th term which is assumed to give a satisfactory approximation to the system response. This representation generally consists of six or seven terms in the case of vegetation since the angular behavior of the radar cross-section of such canopies away from normal incidence is a gentle function of θ. Thus, the unique set of Fourier coefficients has been determined for the model response as a function of angle within the restricted angular range for any given set of input parameters. By way of an example, the behavior of the first four Fourier coefficients as a function of the volumetric soil moisture beneath a vegetation canopy simulated using vector radiative transfer theory at 1.5 GHz for VV and VH polarizations is shown in Figs. 1 and 2. It is seen that the functional behavior of the Fourier coefficients is fairly linear over the entire operational range of volumetric soil moisture and that the convergence of the Fourier series is rapid. This has been found to be true in most vegetation canopy applications. Figs. 3 and 4 show the magnitude angular response of the radiative transfer model as compared with a Fourier series representation utilizing many terms and synthetic data created by using the four coefficient series of Figs. 1 and 2. In this case, the four-term approximation agrees with the model to better than 0.25 dB over the entire angular range.

The behavior of the $C^{pq}_n(\alpha_1, \cdots, \alpha_m; f)'s$ is now approximated as linear functions of the $\alpha'_k s$ over the initial range of these parameters. That is, it is known initially, from experience, that each α_k falls into a range $\alpha_k^{\min} \leq \alpha_k \leq \alpha_k^{\max}$;

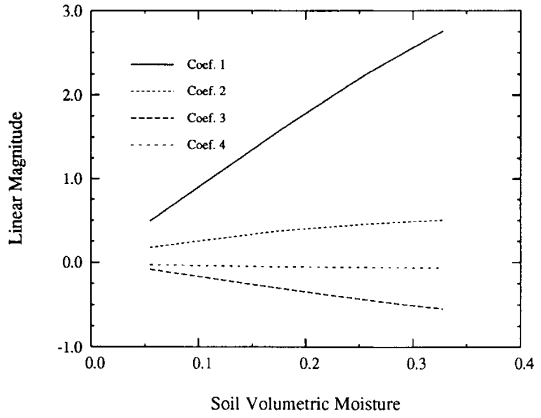


Fig. 1. The first four VV polarized Fourier coefficients of a simulated vegetation canopy as a function of volumetric soil moisture at 1.5 GHz.

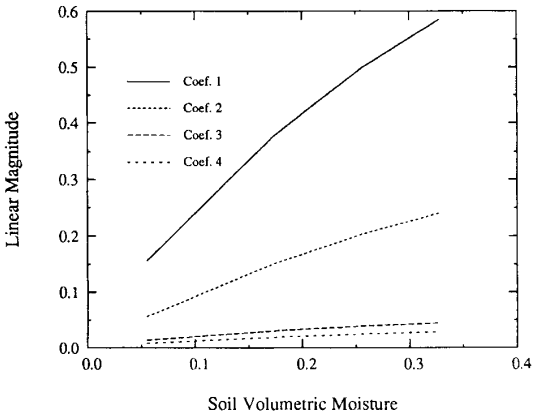


Fig. 2. The first four VH polarized Fourier coefficients of a simulated vegetation canopy as a function of volumetric soil moisture at 1.5 GHz.

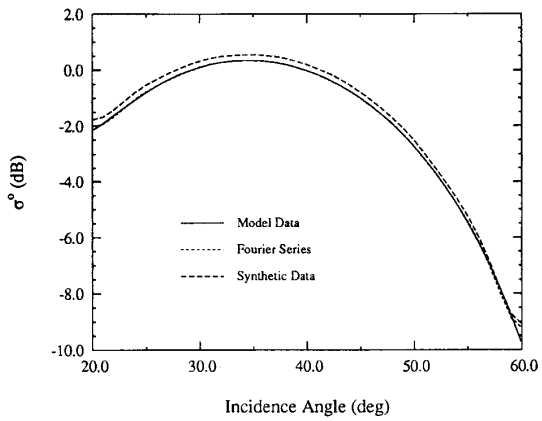


Fig. 3. Comparison of radiative transfer model with Fourier representations for VV polarization at 1.5 GHz.

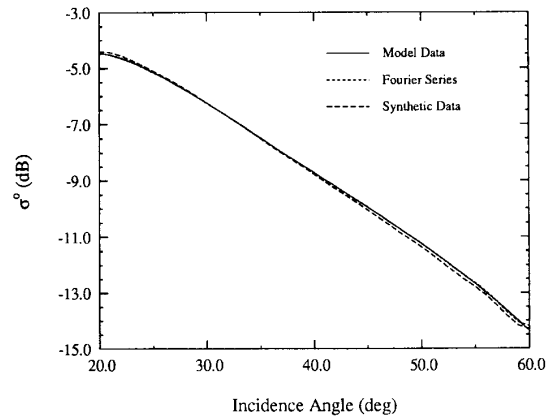


Fig. 4. Comparison of radiative transfer model with Fourier representations for VH polarization at 1.5 GHz.

which is not unreasonable in, for example, typical vegetation canopies of this kind at this time within the growing season. Then we say that each initial range has a centroid or central value of that parameter range. The centroid for that parameter range is denoted by α_k^o and is defined by

$$\alpha_k^o = (\alpha_k^{\max} + \alpha_k^{\min})/2. \tag{8}$$

To the first order of approximation, the Taylor series expansion of the Fourier coefficients about the range centroids is given by

$$C_n^{pq}(\boldsymbol{\alpha}) = \langle C_n^{pq}\rangle_o + \sum_{k=1}^{m}\left(\frac{\partial C_n^{pq}}{\partial \alpha_k}\right)_{\boldsymbol{\alpha}=\boldsymbol{\alpha}^o}(\alpha_k - \alpha_k^o) \tag{9}$$

where $\boldsymbol{\alpha} = (\alpha_1, \cdots, \alpha_m), \boldsymbol{\alpha}^o = (\alpha_1^o, \cdots, \alpha_m^o)$, and $\langle C_n^{pq}\rangle_o = C_n^{pq}(\boldsymbol{\alpha}^o)$.

The coefficients, $(\partial C_n^{pq}/\partial \alpha_k)_{\boldsymbol{\alpha}=\boldsymbol{\alpha}^o}$ are determined by varying each α_k over its range while holding all other $\alpha's$ fixed at their centroid values. At each value of α_k in its range, the model is evaluated as a function of θ and the Fourier coefficients are extracted. In this way, all the partial derivatives may be computed from a linear least squares fit of the coefficients for each parameter to be estimated.

From this discussion it is seen that $\hat{\sigma}_{pq}^o$ has, in effect, been linearized as a function of the model parameters for any arbitrary value of incidence angle in the range $[a, b]$. Thus, we write

$$\hat{\sigma}_{pq}^o(\theta) \simeq \langle \hat{\sigma}_{pq}^o(\theta)\rangle + \sum_{k=1}^{m}\langle r_k^{pq}(\theta)\rangle(\alpha_k - \alpha_k^o) \tag{10}$$

where

$$\langle \hat{\sigma}_{pq}^o(\theta)\rangle \equiv \frac{1}{2}\langle C_o^{pq}\rangle_o + \sum_{n=1}^{N}\langle C_n^{pq}\rangle_o \cos[2n\beta(\theta - a)] \tag{11}$$

and

$$\langle r_k^{pq}(\theta)\rangle \equiv \frac{1}{2}\left(\frac{\partial C_o^{pq}}{\partial \alpha_k}\right)_{\boldsymbol{\alpha}=\boldsymbol{\alpha}^o} + \sum_{n=1}^{N}\left(\frac{\partial C_n^{pq}}{\partial \alpha_k}\right)_{\boldsymbol{\alpha}=\boldsymbol{\alpha}^o} \cdot \cos[2n\beta(\theta - \alpha)]. \tag{12}$$

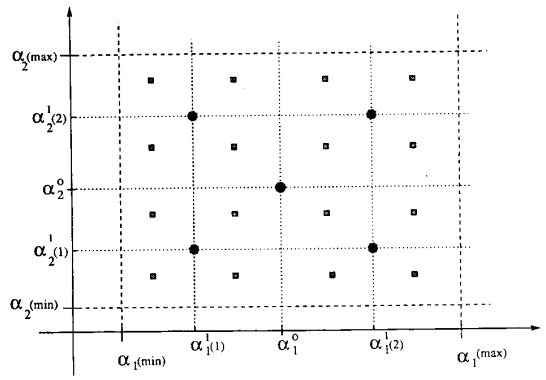


Fig. 5. 2-D parameter space discretization showing subrange centroids as solid marks.

It is possible to divide all the initial parameter ranges $\alpha_k \in [\alpha_k^{min}, \alpha_k^{max}]$ into two equal subranges for each parameter. These subranges are denoted by $\{\alpha_k(1,1)\} = [\alpha_k^{min}, \alpha_k^o]$ and $\{\alpha_k(1,2)\} = [\alpha_k^o, \alpha_k^{max}]$. Each of these new subranges has associated with it a centroid. These centroids are denoted by $\alpha_k^{(1)}(1)$ and $\alpha_k^{(1)}(2)$, respectively. It is now possible to construct 2^m new centroid vectors consisting of m tuples of centroids of the parameter subranges. According to the procedure described previously, the Fourier coefficients and gradients about the 2^m new centroid vectors are now computed. This may be repeated successively to obtain any desired degree of discretization of the parameter space. The pth-order set of subranges is denoted by $\{\alpha_k(p,1)\}, \cdots, \{\alpha_k(p,2p)\}$ and their respective subrange centroids are $\alpha_k^{(p)}(1), \cdots, \alpha_k^{(p)}(2p)$. The resulting $(2p)^m$ centroid vectors may be numbered consecutively starting with $\boldsymbol{\alpha}_p^o(1) = (\alpha_1^{(p)}(1), \cdots, \alpha_m^{(p)}(1))$ and ending with $\boldsymbol{\alpha}_p^o((2p)^m) = (\alpha_1^{(p)}(2p), \cdots, \alpha_m^{(p)}(2p))$. The approach is illustrated in Fig. 5 for a two parameter space.

Now, let us assume we have m measurements of $\sigma^o; \sigma_1^o = \sigma^o_{p_1,q_1}(\theta_1, f_1), \cdots, \sigma_m^o = \sigma^o_{p_m,q_m}(\theta_m, f_m)$. Then the vector $\boldsymbol{\sigma}_s^o = (\sigma_1^o, \cdots, \sigma_m^o)$ can be formed. A system of linear equations is obtained from (10) and may be represented by:

$$\boldsymbol{\sigma}_s^o = \langle \hat{\boldsymbol{\sigma}}_s^o \rangle + \langle \bar{\boldsymbol{r}} \rangle (\hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}^o) \tag{13}$$

where

$$\langle \hat{\boldsymbol{\sigma}}_s^o \rangle \equiv (\langle \sigma_1^o \rangle, \cdots, \langle \sigma_m^o \rangle)^T \tag{14}$$

from (11), and $\langle \bar{\boldsymbol{r}} \rangle_{ij}$ is the value of $\langle r_i^{p_j,q_j}(\theta_j, f_j) \rangle$ in the matrix of Fourier expansions of the coefficient slopes as in (12). $\hat{\boldsymbol{\alpha}} = (\hat{\alpha}_1, \cdots, \hat{\alpha}_m)$ is the estimate vector for the parameter set and $\boldsymbol{\alpha}^o$ is the centroid vector for the set of parameter ranges. Then the zeroth-order estimate for the parameter set may be found from

$$\hat{\boldsymbol{\alpha}}(0) = \langle \bar{\boldsymbol{r}} \rangle_o^{-1} (\boldsymbol{\sigma}_s^o - \langle \hat{\boldsymbol{\sigma}}_s^o \rangle_o) + \boldsymbol{\alpha}^o(0) \tag{15}$$

where $\boldsymbol{\alpha}^o(0)$ is the centroid vector for the initial set of parameter ranges, $\langle \hat{\boldsymbol{\sigma}}_s^o \rangle_o$ is the vector (14) for these initial ranges, and $\langle \bar{\boldsymbol{r}} \rangle_o$ is the matrix whose elements are the original values of (12) formed as described previously.

This estimate will have errors due to the fact that the Fourier coefficients computed using the initial parameter ranges are not actually linear functions of the parameters. However, it may be expected that for parameter ranges sufficiently small, this first-order Taylor expansion approximation will become increasingly accurate. This forms the basis for an iterative algorithm. The parameter space is discretized to produce successively higher-order centroid vectors about subrange spaces of successively smaller volume as described previously. Within these successively smaller parameter subspaces, the first-order Taylor representation of the Fourier coefficients becomes an increasingly better approximation. Then let us suppose that we have a performance measure $P(|\boldsymbol{\alpha} - \boldsymbol{\alpha}'|)$ that decreases monotonically with argument and that satisfies $P(0) = P_{min}$. The first-order estimate of the parameter set is then computed as follows

$$\hat{\boldsymbol{\alpha}}(1) = \langle \bar{\boldsymbol{r}} \rangle_1^{-1} (\boldsymbol{\sigma}_s^o - \langle \hat{\boldsymbol{\sigma}}_s^o \rangle_1) + \boldsymbol{\alpha}_1^o(\text{opt}) \tag{16}$$

where $\boldsymbol{\alpha}_1^o(\text{opt})$ is a first-order centroid vector that satisfies

$$P(|\hat{\boldsymbol{\alpha}}(0) - \boldsymbol{\alpha}_1^o(i)|) \geq P(|\hat{\boldsymbol{\alpha}}(0) - \boldsymbol{\alpha}_1^o(\text{opt})|) \geq P_{\min}, \quad \forall i \in \{1, \cdots, 2^m\} \tag{17}$$

and the other quantities with subscript "1" are the same type as in (15) with values computed at $\boldsymbol{\alpha}_1^o(\text{opt})$.

This process may be repeated as many times as desired up to the highest order of discretization of the model parameter space, or until the rms difference between successive solution vectors stabilizes to within some arbitrary percentage.

III. Implementation of the Algorithm

In practice, it may not be necessary to compute the Fourier representation of the model behavior. In many cases, one is constrained to the use of a particular sensor configuration. In such cases, there is a limited set of channels available to users of the data. For example, given a polarimetric SAR with one look angle and two frequencies of operation, and ruling out the use of phase information, there would be a maximum of six channels (four copolarized and two cross-polarized amplitudes) available for use in inversion. If there are more data channels available than there are parameters of interest for the system under observation, then it would be useful to perform a sensitivity analysis on the model to determine which of the available channels will provide an optimum set. In any event, if the set of data channels available for use is pre-determined then the piecewise linear representation of the model must be based on information provided by those channels and (13) may be inverted directly. This reduces the number of values that need to be pre-computed by a factor equal to the number of coefficients required in the Fourier representation.

Whether the Fourier representation is used or not, the matrix of slopes and vector of intercepts as indicated in (13) must be generated from the pre-computed model outputs for each centroid subrange in the parameter space, the number of such subranges being determined by the maximum level of discretization of the space. The intercept vector for a particular centroid is fixed quantity which depends on the set of model

outputs obtained for the inputs determined by the position of that centroid in parameter space. The slope matrix, on the other hand, is dependent on the set of data points used in sampling the parameter subranges connected with a particular centroid. The functional behavior of the model within a subrange of one of the parameters is considered to be linear, and the partial derivative with respect to that parameter is determined by a weighted linear least squares fit of the model outputs holding all other parameters fixed at the centroid value and varying the parameter in question over its subrange. The point that represents the centroid value is strongly weighted relative to the other points because that is the point about which the derivative must be computed.

There are then at least two basic types of optimization that may be utilized in implementing the iterative inversion scheme. The first type will be referred to as intra-centroid optimization. In intra-centroid optimization, the data points used in computing the linear fit (derivatives) about a particular centroid are initially equally weighted except that the centroid itself is strongly weighted. In successive iterations, points that are found to be most distant from the refined estimate are assigned increasingly lower weights until the estimated value of the parameter vector no longer changes by more than a specified amount. The second type of optimization, referred to here as intercentroid optimization, is utilized after applying the intracentroid scheme. If the refined estimate obtained using intracentroid optimization represents a state of the system closer in parameter space to another centroid on either the same or a different level as was shown in Fig. 5, the algorithm uses this new centroid as the basis for further refinement. It is evident that the algorithm may "wander" from centroid to centroid until it finds a point of local stability on which to converge. The convergence behavior of the algorithm using these optimization schemes will be discussed in a later section.

Convergence of the algorithm itself is fairly rapid since most of the CPU intensive computations are pre-evaluated using the scattering model for the system under investigation and are stored on disk as data arrays to be loaded into memory at runtime. In this way, systems of functions that are fairly nonlinear may be inverted in a direct way, although the more the model behavior deviates from linearity in a given parameter, the less likely it will be to obtain a unique solution for that parameter.

IV. Application to a Vegetation Canopy

Modeling of vegetation canopies using the radiative transfer approach has become increasingly popular in recent years [23]–[27]. The vector radiative transfer equations take into account the individual scattering properties of the vegetation canopy components through a phase matrix which relates the incident to scattered intensities as a function of the wave directions, scatterer geometries, and the material composition of the scatterers. This formulation has the advantage of being general, mathematically tractable, and computationally fairly nonintensive which makes it convenient from the standpoint of inversion.

In this section, the iterative algorithm developed above is used to invert the radiative transfer model of a simplified vegetation canopy. The canopy consists of leaves and vertical trunks above a rough ground layer. The layered structure of the canopy model is illustrated in Fig. 6. The leaves are considered to be uniformly distributed in orientation and are modeled as thin dielectric sheets using physical optics to generate the scattering matrix [28], [29]. The scattered field from the rough ground surface is generated using the Kirchoff and scalar approximations [30]. Trunks are treated as finite length circular dielectric cylinders for which it has been shown that an approximate solution for the scattered field can be obtained based on the solution for the infinite length case provided the diameters of the cylinders are much smaller than their lengths [31]–[33]. The dielectric function of all canopy components is calculated using the dual-dispersion model of El-Rayes and Ulaby [34], [35].

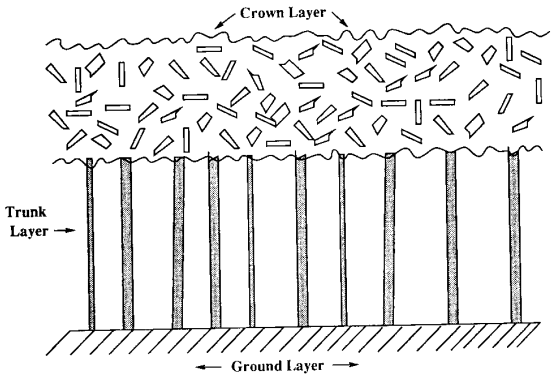


Fig. 6. Layered structure of the radiative transfer model for a vegetation canopy.

TABLE I
Variable Parameter Set for the Vegetation Canopy Used to Test the Inversion Algorithm

Variable Parameter	Parameter Range
Volumetric Soil Moisture	0.05–0.33
Canopy Density: Trunks	0.05–0.25 trunks/m^2
Trunk Height	2.40–3.68 m
Trunk Diameter	11.0–14.2 cm
Canopy Density: Leaves	100–300 leaves/m^3

In this simplified vegetation canopy, five parameters are considered as variables: volumetric soil moisture, trunk canopy density, trunk height, trunk diameter, and leaf canopy density. Actually, there are only four independent variables since the trunk height is usually related to the trunk diameter through a function. The ranges of these variable parameters and the values of the parameters considered to be fixed are listed in Tables I and II. While the ranges of some of these parameters may not be exceptionally representative of all typical real canopies, they do model the important general features of electromagnetic scattering from vegetation layers and serve to demonstrate the invertibility of the radiative transfer equations using the iterative technique developed in this work.

The range of volumetric soil moistures given is representative of the entire wetness scale for typical soils, from that just following a rainstorm to drought conditions [36]. The trunk canopy density range is for conditions from fairly dense woods to sparsely wooded areas [37]. The number density of leaves in

TABLE II
FIXED PARAMETER SET FOR THE VEGETATION CANOPY USED TO TEST THE INVERSION ALGORITHM

Fixed Parameter	Parameter Value
Trunk Gravimetric Moisture	0.7
Leaf Gravimetric Moisture	0.7
Dry Density of Plant Materials	0.33 gm/cc
Leaf Diameter	10.0 cm
Leaf Thickness	0.025 cm
Crown Height	2.5 m
Surface RMS Roughness	1.4 cm
Surface Correlation Length	20.0 cm
Soil Composition	sand: 23.6%
	clay: 33.7%

the canopy, considered together with the size of the leaves and the height of the crown layer combine to give a range of leaf area index (LAI) from 2.0–5.9 which covers almost the entire range of this parameter [38]. The trunk height and diameter ranges are not extremely representative of an average forest canopy but would be more in accord with what one would expect to find in an orchard or grove [25]. Values of 0.7 for the gravimetric moisture of leaves and trunks would give an L-band dielectric constant of about 28-j9 and a C-band dielectric constant of around 25-j9 for both canopy constituents using the model of El-Rayes and Ulaby. This is consistent with typical measured data for trees [8]. Soil surface characteristics are within the validity region for the physical optics model [30].

To test the inversion algorithm, it was first necessary to select four radar channels as data sources to be used as model outputs and inversion algorithm inputs. These four channels provide the correct number of equations to invert the linear system (13) for the four variable model parameters. The assumption was made that a predetermined system configuration consisting of an L-band radar operating at 1.5 GHz, and a C-band radar operating at 5.0 GHz were available for use. Further, an analysis taking into account the overall backscatter signal levels and sensitivity to variation of the inversion parameters indicated that an incidence angle of 45 degrees would be useful in this application. It was decided to use data from the co-polarized channels because they produce the highest signal levels and they can, in practice, be calibrated with the greatest accuracy.

The four-parameter space was discretized up to the second level producing sixteen secondary range centroids. The number of centroids at each level is given by $n = 2^{m(l-1)}$, where m is the number of parameters and l is the level of discretization. The model was then run in the forward direction five times within each full parameter range for the primary centroid and three times within each subrange for every secondary centroid to provide the precomputed data used in determining the slope matrices and intercept vectors. After the necessary information was produced to provide a mapping between the parameter space and the model output space, the model was then used as a data simulator for use in testing the inversion.

V. INVERSION RESULTS

The region in parameter space for which the least inversion accuracy is to be expected is the region near the primary range centroid. In this region, the estimate of the parameter vector is based on only a single level of refinement and intracentroid optimization. Because the primary centroid is the nearest centroid to the estimated parameter vector in this case, the algorithm does not take advantage of the higher degree of discretization available on the second level, and intercentroid optimization is not used. When the estimated parameter vector is closer in parameter space to any higher level centroid the intercentroid optimization may then be used to obtain further refinement of the estimate.

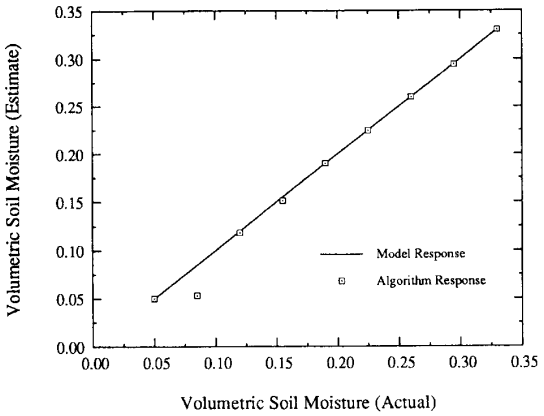


Fig. 7. The determination of soil moisture over the entire parameter range near the primary centroid.

Typical results for inversions obtained near the primary range centroid are shown in Figs. 7 through 10. It is seen that primary level optimization alone provides excellent inversion accuracy for most of the parameters over almost the entire parameter range. However, there is a significant degree of instability in the trunk height determination for the region below about 2.7 meters (Fig. 10). This is due to the small range of heights considered in this study and the fact that the radar response is not a sensitive function of either trunk height or diameter within the selected ranges of these two parameters for the data channels used in the inversion. This problem could be remedied by using a finer division of the parameter space on the second level so as to provide a convenient secondary range centroid corresponding to the position of the primary range centroid but with a restricted parameter subrange. A larger range of trunk heights and diameters such as one might find in an actual forest, or a set of data channels more sensitive to these parameters, would also improve the relative accuracy of the trunk height estimate.

The excellent results obtained for regions close in parameter space to the primary centroid are actually representative of the worst-case estimates produced by the algorithm. Typically, when the algorithm is able to take advantage of a higher level of optimization by using a convenient second-level centroid, excellent convergence stability is achieved. This is illustrated in Figs. 11 and 12 which demonstrate that the stability problems encountered in using the entire parameter range of the primary centroid are greatly reduced when the parameter subranges of the second level are utilized.

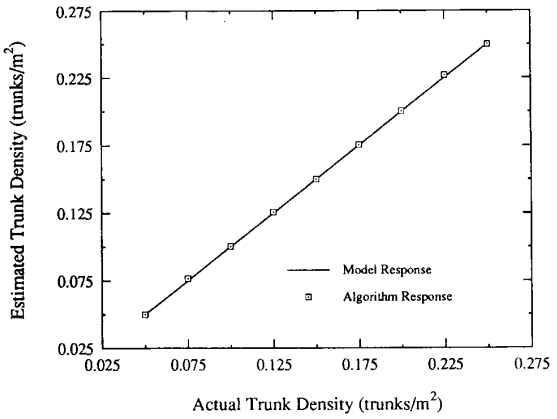


Fig. 8. The determination of trunk canopy density over the entire parameter range near the primary centroid.

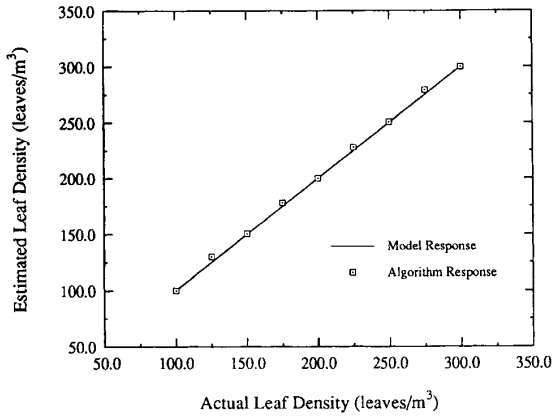


Fig. 9. The determination of leaf canopy density over the entire parameter range near the primary centroid.

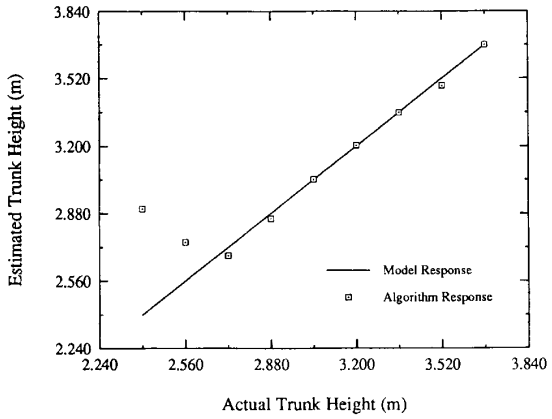


Fig. 10. The determination of trunk height over the entire parameter range near the primary centroid.

The iterative inversion algorithm exhibits the four general types of convergence behavior shown in Figs. 13 through 16. In about 90% to 95% of the cases studied, the algorithm converges to the correct value as shown in Figs. 13 and 14. In the remaining cases, the algorithm either does not converge at all as shown in Fig. 16 or converges to the wrong value as in Fig. 15. Of these last two situations it is definitely preferable to have nonconvergence of the algorithm. With nonconvergence it is clear that the inversion must be rejected and thus an erroneous result is easily avoided.

Anomalous convergence represents less than 5% of all convergence behavior exhibited in this implementation of the iterative inversion algorithm. With anomalous convergence, it is not always possible to distinguish an incorrect result, therefore this type of convergence behavior represents the most difficult problem to correct in applying the algorithm. Because the convergence behavior of the algorithm is to some degree implementation-specific, there are also some steps that can be taken to reduce the risk of obtaining incorrect results. It should be possible to utilize other types of algorithms to monitor the decision processes that the inversion algorithm uses in achieving convergence. In this way, it would be possible to minimize the probability of taking an incorrect decision path. However, the investigation of this approach is outside the scope of this work.

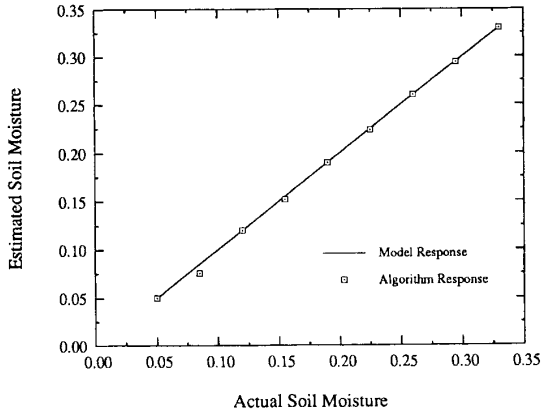


Fig. 11. The determination of leaf canopy density over a second level parameter subrange.

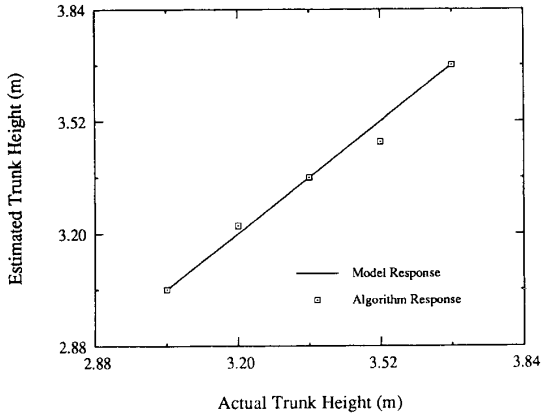


Fig. 12. The determination of trunk height over a second level parameter subrange.

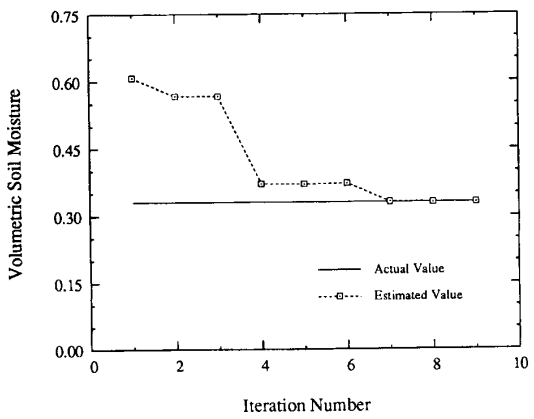


Fig. 13. Uniform convergence of the iterative inversion algorithm.

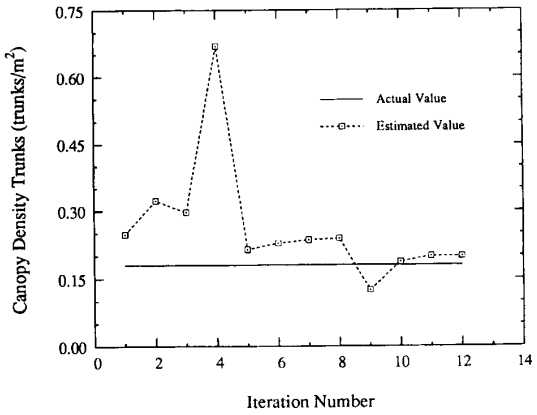


Fig. 14. Nonuniform convergence of the iterative inversion algorithm.

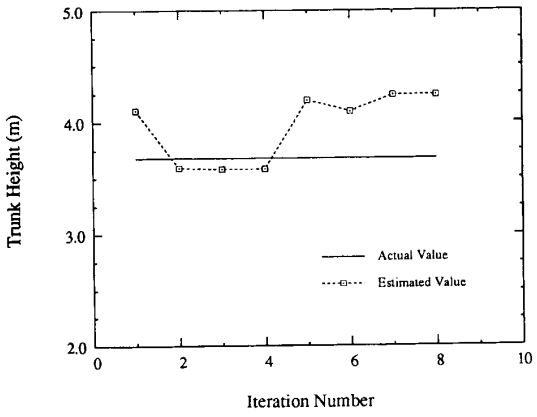


Fig. 15. Anomalous convergence of the iterative inversion algorithm.

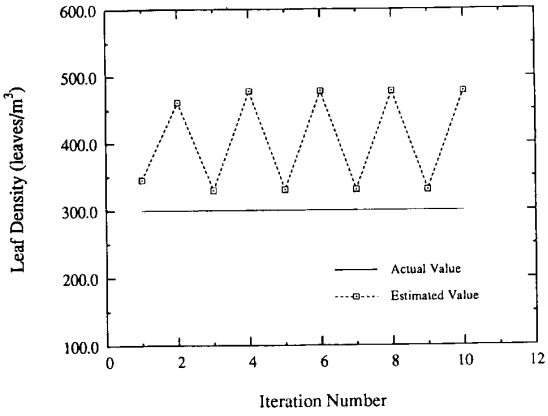


Fig. 16. Oscillatory nonconvergence of the iterative inversion algorithm.

VI. Error Analysis

A complete analysis of errors for the inversion algorithm is lengthy and the derivation will not be given here. The rms error in the i th-parameter α_i will be denoted $\epsilon(\alpha_i)$. This error originates from two major sources. The first source is the overall systematic error which depends on the combined data channel characteristics, the accuracy of the scattering model, and the degree of discretization utilized in the algorithm implementation. The systematic error in parameter α_i is directly proportional to the offsets of the individual data channels from their respective range centroids for a given level of discretization. This offset is defined as $\Delta\sigma_j^o = (\sigma_j^o - \langle\sigma_j^o\rangle)$ for the jth data channel as in (15). The second source of error in the inversion algorithm derives from the measurement uncertainties in the data channels. The measurement uncertainty in σ^o for the jth data channel, $\epsilon_m(\sigma_j^o)$ produces an error in the inversion dependent on the gain characteristics of that channel. Both sources of error are inversely proportional to the determinant Δ of the slope matrix $\bar{\boldsymbol{r}}$ as given in (13).

The expression for the rms error in the parameter α_i for an inversion using m data channels is given by

$$\epsilon(\alpha_i) = \pm\frac{1}{|\Delta|}\left\{\sum_{j=1}^{m}[\mu_{ij}^2(\Delta\sigma_j^o)^2 + cof_{ji}^2(\bar{\boldsymbol{r}})\epsilon_m^2(\sigma_j^o)]\right\}^{1/2}. \tag{18}$$

Where $cof_{ji}(\bar{\boldsymbol{r}})$ is the ji cofactor of $\bar{\boldsymbol{r}}$, and the factor μ_{ij} represents the overall uncertainty in parameter α_i produced by an error in estimating the response of data channel j to that parameter. It is defined by

$$\mu_{ij} = [\epsilon^2(cof_{ji}(\bar{\boldsymbol{r}})) + cof_{ji}^2(\bar{\boldsymbol{r}})(\epsilon(\Delta)/\Delta)^2]^{1/2} \tag{19}$$

where $\epsilon(cof_{ji}(\bar{\boldsymbol{r}}))$ is the error in $cof_{ji}(\bar{\boldsymbol{r}})$ and $\epsilon(\Delta)$ is the uncertainty in the determinant of the slope matrix.

Examination of (18) and (19) reveal some important factors that affect the accuracy of the inversion. The overall rms system error will be minimized when Δ is maximized, the $\epsilon(cof_{ji}(\bar{\boldsymbol{r}}))$'s are minimized and the individual slope matrix element errors that contribute to $\epsilon(\Delta)$ are minimized simultaneously. The magnitude of Δ is a function of the data channel gain characteristics and the structure of the linear system of equations provided by the combination of channels used in the inversion. The slope matrix determinant is maximized when

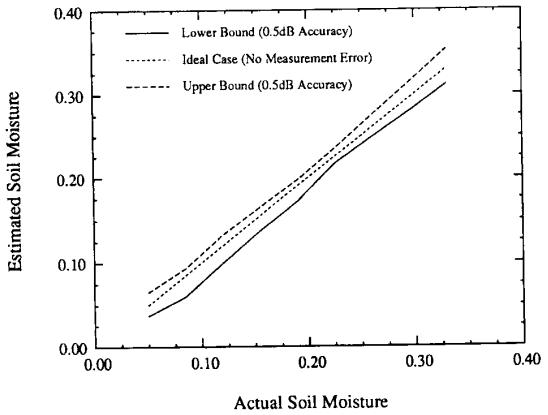


Fig. 17. Typical inversion error bounds for variable soil moisture with other parameters held fixed.

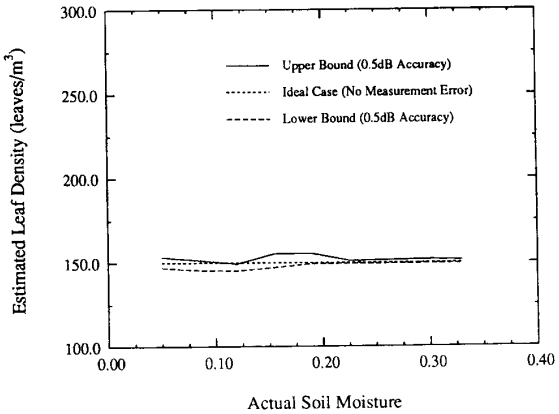


Fig. 18. Typical inversion error bounds for fixed leaf density with variation of the soil moisture parameter.

the overall channel gains are as large as possible and the system of equations is as close to being linearly independent as is practical for the set of sensors available. The slope matrix errors are minimized as the linear regression slopes approach the local derivatives in the proximity of the parameter estimate vector. This occurs as the level of optimization increases. The upper limit on the reduction of systematic error as a result of increasing the level of discretization will ultimately be determined by the contribution to the $\epsilon(cof_{ji}(\bar{r}))$'s originating from the model itself. If the model is inaccurate, no amount of optimization will permit accurate inversion of the measured data. However, if the data channels are chosen appropriately and the model is sufficiently accurate, the systematic error can be minimized by providing a suitable degree of discretization of the parameter space. In the case of systematic errors, it is to be expected that the best error performance will occur when the actual parameter vector lies close to a range centroid.

The second contribution to (18) is dependent on the measurement uncertainty in the data channels. If the data channels have been properly selected and the algorithm has been implemented so as to provide sufficient discretization of the parameter space, then measurement errors can be expected to give the largest contribution to the overall inversion error. These errors have a slight dependence on algorithm implementation but are largely influenced by the ratio of the measurement uncertainty to the sensitivity of the data channels. This type of error response is illustrated in Figs. 17 and 18 which show typical error bounds for the inversion if there is a uniform ± 0.5 dB measurement uncertainty in all of the data channels. If soil moisture were the only variable parameter in the inversion, measurement accuracy would control the error response to yield a worst case average uncertainty of about $\pm 3.5\%$. The inversion accuracy for the leaf canopy density, with the density fixed at mid-range, is within about $\pm 2.0\%$. In this case, the systematic error and the measurement uncertainty error are of the same order of magnitude.

VII. Conclusions

In this work, a general model-based iterative algorithm has been presented for use in the inversion of polarimetric radar data. The algorithm was implemented using two levels of discretization and two types of optimization and was applied to the case of a general representative vegetation canopy. This simplified test canopy was modeled using vector radiative transfer theory and consisted of vertical trunks, leaves, and a rough ground surface. Four canopy parameters, soil moisture, trunk canopy density, trunk height, and leaf canopy density were estimated over their ranges utilizing four radar data channels, which were simulated with the radiative transfer model, as inputs to the algorithm. Excellent inversion accuracy was obtained over almost the entire range of all four parameters. Successful convergence was achieved in 90% to 95% of the cases tested for the algorithm as implemented in this work, and it is concluded that further improvement in the convergence characteristics should be obtainable in future implementations. The results of the error analysis show that if the radar data channels are appropriately chosen so as to maximize the sensitivity of each channel to a separate inversion parameter, and the degree of discretization of the parameter space is sufficient, the overall inversion error will be minimized for a given level of uncertainty in the measured input data.

References

[1] F. T. Ulaby, R. K. Moore, and A. K. Fung, *Microwave Remote Sensing: Active and Passive, Vol. III—From Theory to Applications.* Dedham, MA: Artech, 1986.

[2] F. T. Ulaby, D. Held, M. C. Dobson, K. McDonald, and T. B. A. Senior, "Relating polarization phase difference of SAR signals to scene properties," *IEEE Trans. Geosci. Remote Sens.*, vol. 25, pp. 83–92, 1987.

[3] J. J. van Zyl, H. A. Zebker, and C. Elachi, "Imaging radar polarization signatures: Theory and observations," *Radio Sci.*, vol. 22, pp. 529–543, 1987.

[4] Y. Oh, K. Sarabandi and F. T. Ulaby, "An empirical model and an inversion technique for radar scattering from bare soil surfaces," *IEEE Trans. Geosci. Remote Sens.*, vol. 30, pp. 370–381, 1992.

[5] J. Shi, J. V. Soares, L. Hess, E. T. Engman, and J. van Zyl, "SAR-derived soil moisture measurements for bare fields," *Proc. IGARSS '91, Espoo, Finland*, vol. I, pp. 393–396, 1991.

[6] J. R. Wang, E. T. Engman, T. J. Schmugge, T. Mo and J. C. Shiue, "The effects of soil moisture, surface roughness, and vegetation on L-band emission and backscatter," *IEEE Trans. Geosci. Remote Sens.*, vol. 25, pp. 825–833, 1987.

[7] S. Gogineni, A. K. Fung, K. S. Chen, and J. Wang, "Comparison of measurements and theory for backscatter from vegetation-covered soil on the konzaprairie," *Proc. IGARSS '92, Houston, Texas,* vol. II, pp. 920–922, 1992.
[8] K. C. McDonald, M. C. Dobson, and F. T. Ulaby, "Modeling multi-frequency diurnal backscatter from a walnut orchard," *IEEE Trans. Geosci. Remote Sens.*, vol. 29, pp. 852–863, 1991.
[9] T. LeToan, H. Laur, E. Mougin, and A. Lopes, "Multitemporal and dualpolarization observations of agricultural vegetation covers by X-band SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 27, pp. 709–718, 1989.
[10] N. S. Chauhan, R. H. Lang, and K. J. Ranson, "Radar modeling of a boreal forest," *IEEE Trans. Goesci. Remote Sens.*, vol. 29, pp. 627–638, 1991.
[11] Y. Hussin, R. Reich, and R. Hoffer, "Estimating slash pine biomass using radar backscatter," *IEEE Trans. Geosci. Remote Sens.*, vol. 29, pp. 427–431, 1991.
[12] D. Holliday, G. St-Cyr, and N. E. Woods, "Comparison of a new radar ocean imaging model with sarsex internal wave data," *Int. J. Remote Sens.*, vol. 8, pp. 1423–1430, 1987.
[13] K. S. Chen and A. K. Fung, "An empirical bispectrum model for sea surface scattering," *IGARSS Symp.*, vol. II, pp. 1454–1455, 1992.
[14] C. Matzler and E. Schanda, "Snow mapping with active microwave sensors," *Int. J. Remote Sens.*, vol. 5, pp. 409–422, 1984.
[15] F. T. Ulaby and W. H. Stiles, "The active and pasive response to snow parameters—Part II: Water equivalent of dry snow," *J. Geophys. Research,* vol. 85, pp. 1045–1049, 1980.
[16] E. P. W. Attema and F. T. Ulaby, "Vegetation modeled as a water cloud," *Radio Sci.*, vol. 13, pp. 357–364, 1978.
[17] H. Shinohara, T. Homma, H. Nohmi, H. Hirosawa, and T. Tagawa, "Relation between L-band microwave penetration/backscattering characteristics and state of trees," *Proc. Igarss '92, Houston, Texas,* vol. I, pp. 59–541, 1992.
[18] L. Pierce, K. Sarabandi, and F. T. Ulaby, "Application of an artificial neural network in canopy scattering inversion," *Proc. IGARSS '92, Houston, Texas,* vol. II, pp. 1067–1069, 1992.
[19] S. E. Decatur, "Application of neural networks to terrain classification," *Proc. IJCNN,* vol. 1, pp. 283–285, 1989.
[20] D. E. Rumelhart, J. L. McClelland, and the PDP Research Group, *Parallel Distributed Processing, Vol. I—Foundation.* Cambridge, MA: M.I.T. Press, 1986.
[21] G. E. McClelland, R. N. DeWitt, T. H. Hemmer, M. L. Matheson, and G. O. Moe, "Multispectral image processing with a three-layer backpropagation network," *Proc. IJCNN,* vol. 1, pp. 151–153, 1989.
[22] S. Hong, K. Fukue, H. Shimoda, and T. Sakata, "Nonparametric texture extraction using neural network," *Proc. IGARSS '92, Houston, Texas,* vol. II, pp. 1084–1086, 1992.
[23] M. A. Karam and A. K. Fung, "Electromagnetic scattering from a layer of finite length, randomly oriented, dielectric circular cylinders over a rough interface with application to vegetation," *Internat. J. Remote Sens.*, vol. 9, pp. 1109–1134, 1988.
[24] F. T. Ulaby, K. Sarabandi, K. McDonald, M. Whitt, and M. C. Dobson, "Michigan microwave canopy scattering model," *Int. J. Remote Sens.*, vol. 11, pp. 1223–1253, 1990.
[25] K. C. McDonald, M. C. Dobson, and F. T. Ulaby, "Using MIMICS to model L-band multiangle and multitemporal backscatter from a walnut orchard," *IEEE Trans. Geosci. Remote Sens.*, vol. 28, pp. 477–491, 1990.
[26] K. C. McDonald and F. T. Ulaby, "MIMICS II: Radiative transfer modeling of discontinuous tree canopies at microwave frequencies," *IGARSS '90,* College Park, Maryland, 1990, pp. 20–24.
[27] M. A. Karam and A. K. Fung, "Scattering from randomly oriented circular discs with application to vegetation," *Radio Sci.*, vol. 18, pp. 557–565, 1983.
[28] T. B. A. Senior, K. Sarabandi, and F. T. Ulaby, "Measuring and modeling the backscattering cross-section of a leaf," *Radio Sci.*, vol. 22, pp. 1109–1116, 1987.
[29] K. Sarabandi, T. B. A. Senior, and F. T. Ulaby, "Effect of curvature on the backscattering from leaves," *J. Electromagnetic Waves Appl.*, vol. 2, pp. 653–670, 1988.
[30] F. T. Ulaby, R. K. Moore, and A. K. Fung, *Microwave Remote Sensing: Active and Passive, Vol. II—Radar Remote Sensing and Surface Scattering and Emission Theory.* Reading, MA: Addison-Wesley, 1982, p. 937.
[31] H. C. Van de Hulst, *Light Scattering by Small Particles.* New York: Wiley, 1957.
[32] G. T. Ruck, D. E. Barrick, W. D. Stuart, and C. K. Kirchbaum, *Radar Cross Section Handbook, Vol. 1.* New York: Plenum, 1970.
[33] M. A. Karam and A. K. Fung, "EM 'Scattering from a randomly oriented circular dielectric, finite-length cylinder," presented at Internat. Union Radio Sci. Commission F: Wave Propagation and Remote Sensing, University of New Hampshire, Durham, NH, pp. 4.1.1–4.1.3.
[34] F. T. Ulaby and M. A. El-Rayes, "Microwave dielectric spectrum of vegetation, Part II: Dual-dispersion model," *IEEE Trans. Geosci. Remote Sens.*, vol. 25, pp. 550–557, 1987.
[35] M. T. Hallikianen, F. T. Ulaby, M. C. Dobson, M. A. El-Rayes, and L. K. Wu, "Microwave dielectric behavior of wet soil—Part I: Empirical models and experimental observations," *IEEE Trans. Geosci. Remote Sens.*, vol. 23, pp. 25–34, 1985.
[36] P. P. Batlivala and F. T. Ulaby, "Feasibility of monitoring soil moisture using active microwave remote sensing," Univ. Kansas Remote Sensing Lab., Rep. 264-12, 1977.
[37] M. C. Dobson, K. McDonald, and F. T. Ulaby, "Modeling of forest canopies and anaylsis of polarimetric SAR data," Univ. Michigan Radiation Lab., Rep. 026143-1-F, 1988.
[38] E. Wilcox and M. C. Dobson, "Leaf area index data for the Michigan forest test sites," Univ. Michigan Radiation Lab., Rep. 025761-1-T, 1992.

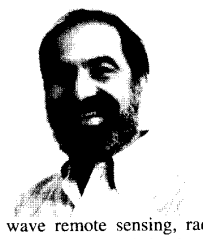
Paul F. Polatin was born in New York City, NY, on March 18, 1956. He received the B.A. and M.A. degrees in chemical physics from Columbia University, New York, NY in 1978 and 1979, respectively. From 1978 to 1983, he was involved in various research projects including nuclear quadrupole resonance spectroscopy and near-infrared reflectance spectroscopy. From 1984 to 1987 he was with the Upjohn Pharmaceutical Company as a research chemist and laboratory supervisor.

In 1987 he joined the Radiation Laboratory at the University of Michigan, Ann Arbor, MI. He has since received the M.S.E.E. degree and is currently working to complete the Ph.D. degree in electromagnetic engineering. At present he is engaged in studying the scattering of electromagnetic waves by random media and other problems related to remote sensing of the terrestrial environment.

Kamal Sarabandi (S'87–M'90–SM'93) received the B.S. degree in electrical engineering from Sharif University of Technology, Tehran, Iran, in 1980. From 1980 to 1984, he worked as a microwave engineer in the Telecommunication Research Center in Iran. He entered the graduate program at the University of Michigan, Ann Arbor, MI, in 1984 and received the M.S.E. degree in electrical engineering in 1986, the M.S. degree in mathematics, and the Ph.D. degree in electrical engineering in 1989.

He is presently an Assistant Professor in the Department of Electrical Engineering and Computer Science at the University of Michigan. His research interests include electromagnetic scattering, microwave and millimeter wave remote sensing, and calibration of polarimetric SAR systems.

Dr. Sarabandi is a Member of the Electromagnetics Academy and USNC/URSI Commission F.



Fawwaz T. Ulaby (M'68–SM'74–F'80) received the B.S. degree in physics from the American University of Beirut, Lebanon, in 1964 and the M.S.E.E. and Ph.D. degrees in electrical engineering from the Univerity of Texas, Austin, TX, in 1966 and 1968, respectively.

He is currently Professor of Electrical Engineering and Computer Science at The University of Michigan, Ann Arbor, MI, and Director of the NASA Center for Space Terahertz Technology. His current interests include microwave and millimeter wave remote sensing, radar systems, and radio wave propagation. He has authored several books and published over 400 papers and reports on these subjects. He is the recipient of numerous awards, including the IEEE Geoscience and Remote Sensing Distinguished Achievement award in 1983, the IEEE Centennial Medal in 1984, and the Kuwait Prize in applied science in 1986.