

A Robust Statistical-Based Estimator for Soil Moisture Retrieval from Radar Measurements

Michael S. Dawson, *Member, IEEE*, Adrian K. Fung, *Fellow, IEEE*, and Michael T. Manry, *Senior Member, IEEE*

Abstract—We examine the use of a robust statistical inversion approach to the estimation of soil moisture and roughness statistics from backscatter measurements. Two sets of basis functions are examined; the first is a set of basis functions from multinomial combinations of the inputs (termed the MBF) while the second is a set of basis functions generated by a multilayer perceptron referred to as MLPBF. We discuss potential sources of training patterns upon which to base these estimators, including empirical forward models and more rigorous theoretical scattering models such as the IEM. These estimators are applied to a set of measured POLARSCAT data from Oh *et al.* [1992]. Comparisons are made with other inversion methods including neural networks.

I. INTRODUCTION AND BACKGROUND

ALTHOUGH significant progress has been made in the ability to acquire remotely sensed data with the addition of new airborne and spaceborne sensors, extracting soil moisture from this data has been problematic for several reasons, including:

- 1) the soil moisture profile is usually not available from ground truth measurements and the moisture measurement acquired at a point may not be representative of the region observed by radar;
- 2) a good closed-form signal model $S(\Theta)$ relating the unknown surface parameters Θ such as the volumetric soil moisture m_v to the measured backscattering coefficient σ_{pp}° has not been available until recently;¹
- 3) the measured backscatter signal is dependent on several parameters including the surface roughness, the presence of vegetation, and the soil moisture content. We may not be able to isolate the contributions from each of these sources; and
- 4) the relationship between soil moisture and the received backscattering coefficient is nonlinear and the problem of retrieving parameters may be ill-posed [36], [34]. The detection and design of estimators for these possibly ill-posed problems are considered elsewhere [7], [8].

The work here presents the background and motivation of a multidimensional statistical estimation method for soil moisture based on a set of “training” patterns. This set of training

Manuscript received January 8, 1996; revised September 28, 1996. This work was supported by the NASA Global Change Fellowship NGT-30245.

The authors are with the Department of Electrical Engineering, University of Texas at Arlington, Arlington, Texas, USA.

Publisher Item Identifier S 0196-2892(97)00856-5.

¹Here, the closed-form signal model used to relate the backscattering coefficient to the surface characteristic is the IEM [Fung *et al.* 1992; Fung, 1994] where we have used the dielectric model of Hallikainen *et al.* [1985] to convert the physical characteristics of the soil into electrical parameters for input into the electromagnetic model (the IEM).

patterns may be from a large experiment in which ground truth is available, or synthetically generated using empirical or theoretical surface scattering models, or a combination of these three sources. Following a review of current empirical methods used to retrieve soil moisture from active measurements in Section I, we present a robust statistical inversion technique along with two different sets of basis functions in Section II. In Section III we consider the impact of different scattering models for the creation of training patterns for this estimator including the IEM model [12], [13] and we show comparisons between the results of this and other inversion techniques.

A. Empirical Models for Soil Moisture Estimation

The sensitivity of remotely sensed data to changes in soil moisture has been demonstrated in a number of previous studies [2], [30], [26], [23]. The measurement of soil moisture by microwave instruments is based on the large contrast in permittivities for wet and dry soil for both active and passive measurements. Dry soil has a real dielectric constant (ϵ_r') of about 2.5 and water has an ϵ_r' of about 80. When soil and water are mixed, the resulting effective permittivity can range from 2.5 for very dry soil to over 25 for very moist soil. This contrast is dependent on both the composition of the soil and the frequency.

The quantitative estimation of soil moisture from backscatter measurements is currently obtained by employing one of several empirical relationships that map the measured normalized backscatter coefficient $\sigma_{qq}^{\circ}(\text{freq}, \theta)$ into the volumetric soil moisture denoted by m_v . The notation $\sigma_{qq}^{\circ}(\text{freq}, \theta)$ indicates that the backscatter coefficient at polarization qq is a function of frequency and angle from nadir θ .

The development of empirical relations (models) has been carried out by many investigators. Shi *et al.* [31], [32] used a polarization signal ratio developed by observing theoretical signatures of a first order small perturbation model to predict soil moisture from L-band SAR measurements. Oh *et al.* [26] arrived at an empirical forward scattering model of the form

$$\sigma_{vv}^{\circ}(\theta, \epsilon_r', k\sigma) = \frac{g(\text{Cos } \theta)^3}{\sqrt{p}} [\Gamma_v(\theta) + \Gamma_h(\theta)] \quad (1a)$$

$$\sigma_{hh}^{\circ}(\theta, \epsilon_r', k\sigma) = g\sqrt{p}(\text{Cos } \theta)^3 [\Gamma_v(\theta) + \Gamma_h(\theta)] \quad (1b)$$

and

$$\sigma_{hv}^{\circ}(\theta, \epsilon_r', k\sigma) = q\sigma_{vv}^{\circ}(\theta, \epsilon_r', k\sigma) \quad (1c)$$

where $\Gamma_v(\theta)$ and $\Gamma_h(\theta)$ are power reflection coefficients for vertical and horizontal polarizations. The quantity, $q = \sigma_{hv}^{\circ}/\sigma_{vv}^{\circ}$ is the cross polarization ratio and $p = \sigma_{hh}^{\circ}/\sigma_{vv}^{\circ}$

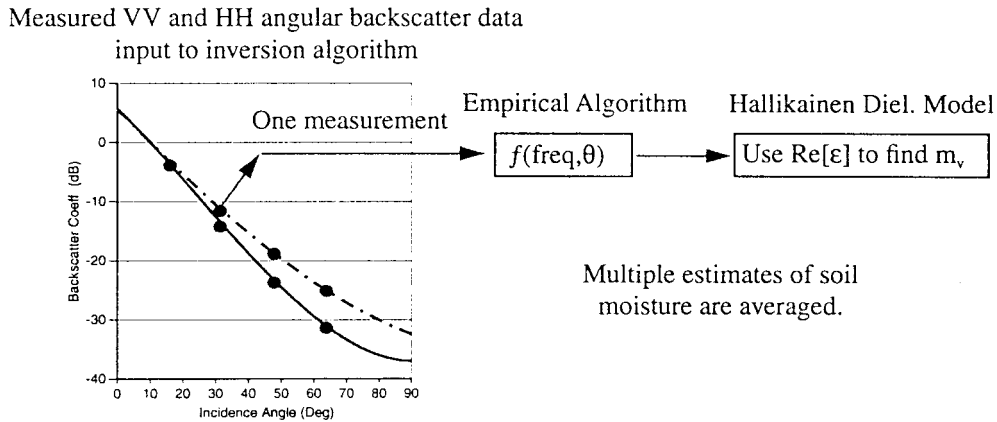


Fig. 1. Simplistic representation of parameter retrieval using some empirical models in which we get one estimate for each angular measurement per polarization and then average.

is the like polarization ratio. These quantities along with g are determined by regression on a set of previously measured data. Because the data set used by Oh *et al.* are from surfaces with approximately the same correlation length l , their model shows no dependence on this surface parameter and hence is restricted in its validity to surfaces of the type that generated such data.

In a later paper, Oh *et al.* [1994] developed yet another empirical forward model. In this newer set of relations, the VV-polarized backscattering coefficient was found to be

$$\sigma_{vv}^o = 13.5e^{-1.4(k\sigma)^{0.2}} \frac{1}{\sqrt{p}} \Gamma_h(k\sigma)^2 (\text{Cos } \theta)^{3.25-0.05kl} \times e^{-(2k\sigma \text{Cos } \theta)^{0.6}} W_k \quad (2)$$

where W_k is the roughness spectrum, $k\sigma$ and kl are normalized rms height and correlation length of the surface, and p is an empirically determined function from Oh *et al.* [26]. Although the explicit dependence on surface correlation is included in (2), its variation is incorrect in that a larger correlation length does not cause a faster angular dropoff of the backscattering coefficient at large angles of incidence.

The soil moisture inversion technique of Haddad and Dubois [19] used a catalogue of backscatter measurements and soil parameters in a Bayesian estimation scheme. However the effect of the parameters are site and time dependent. In general, a large amount of experimental measurements is needed in order to derive general statistical laws along with their associated uncertainties. In addition, their inversion scheme ultimately relied on the set of empirical relations from Oh *et al.* [26] in their *a priori* densities.

In another soil moisture inversion model, Dubois *et al.* [9], [10] created a semi-empirical model of the form

$$\sigma_{hh}^o(\theta, \lambda, \epsilon_r, \sigma) = 10^{-2.75} \frac{\text{Cos}^{1.5}\theta}{\text{Sin}^5\theta} 10^{0.028\epsilon_r \text{ Tan } \theta} (k\sigma \text{ Sin } \theta)^{1.4} \lambda^{0.7} \quad (3)$$

$$\sigma_{vv}^o(\theta, \lambda, \epsilon_r, \sigma) = 10^{-2.35} \frac{\text{Cos}^3\theta}{\text{Sin}^3\theta} 10^{0.046\epsilon_r \text{ Tan } \theta} (k\sigma \text{ Sin } \theta)^{1.1} \lambda^{0.7} \quad (4)$$

These two relations are valid for frequencies varying between 1.5 and 11 GHz for surfaces ranging from 0.3–3 cm RMS height and for incidence angles between 30–65°. Dubois *et al.* presented this inversion algorithm with a discussion of the effects of calibration errors and presented a simple yet effective method to limit the applicability of the algorithm to bare soil or lightly vegetated areas. Here again, no dependence on the surface correlation length appears.

B. A Summary Comment on Some Existing Empirical Forward Models

These empirical models between radar backscatter and the spatial and temporal variations of soil moisture have been developed and successfully applied to several specific data sets. They are, however, generally valid only for conditions under which those measured data were taken and some of the models ignore one of the surface parameters; namely, the surface correlation length which is a much more difficult parameter to estimate. The end effect is that these empirical models have limited range of applicability (as any model does).

Also note that the same retrieval algorithm applied to a given set of data may yield several estimates depending on which data points from that set are selected as inputs to the inversion scheme. For example, in the empirical-model retrieval approach by Oh *et al.* [26] and Dubois *et al.* [9], several single-point estimates for the soil moisture from the same data grouping are possible because only a *single* VV or HH measured value at a given angle is used as input to the retrieval (see Fig. 1). Thus, several estimates of soil moisture from the same data set are obtained (although the surface roughness and soil moisture are constant for the entire set). The final estimate for the soil moisture depends on how one combines these single-point estimates via some linear or weighted average.

In addition, the accuracy of inversion is heavily dependent on the initial estimate for the Fresnel reflectivity at nadir $\Gamma(\theta=0)$. It is important to point out that this approach ignores any contribution from the imaginary part of the dielectric

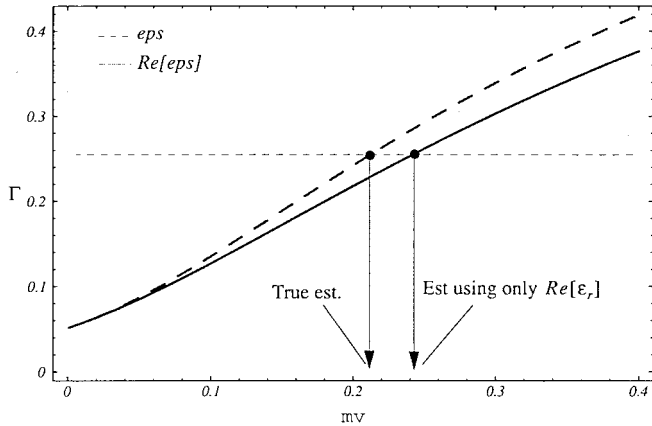


Fig. 2. A plot of the Fresnel reflectivity at nadir $\Gamma_{(\theta=0)}$ using the real and imaginary parts of ϵ_r (dotted line) and the real part only of $\text{Re}[\epsilon_r]$ (solidline) as a function of volumetric soil moisture m_v calculated at 18 GHz. The effect at lower frequencies is negligible.

$\text{Im}[\epsilon_r] = \epsilon_r''$ in the Fresnel reflectivity given by

$$\Gamma_{(\theta=0)} = \left| \frac{1 - \sqrt{\epsilon_r}}{1 + \sqrt{\epsilon_r}} \right|^2 = \left| \frac{1 - \sqrt{\epsilon_r' - j\epsilon_r''}}{1 + \sqrt{\epsilon_r' - j\epsilon_r''}} \right|^2. \quad (5)$$

For low frequencies the assumption that the omission of ϵ_r'' in (5) will have little effect is valid, however, at higher frequencies, the effect of ϵ_r'' is no longer negligible and can lead to over estimates of soil moisture by 5–15% as indicated by example in Fig. 2.

C. Soil Moisture Inversion Using Neural Networks

Traditionally, inversion methods such as those discussed in the previous section have been developed to extract a single parameter at a time. However, because of the complexity and nonlinearity of problems, it may not be possible to separate the contributions from different mechanisms thus making the retrieval of several parameters simultaneously necessary. Le Toan *et al.* [23] assessed the retrieval of soil moisture from ERS-1 SAR data in which it was noted that inversion of soil moisture is not possible without information on the surface roughness. Thus, it is necessary to perform the inversion of these parameters simultaneously.

A method suitable for multidimensional retrieval is the neural network [5], [6], [35]. In the neural network based inversion, a complete set of backscatter data is fed into a trained network (see Fig. 3). The training patterns used to train this network can be developed using a theoretical forward scattering model to estimate the mean while the data set taken from a known surface condition is used to estimate the variance of the training patterns. A main advantage in the use of such training patterns is that we can control the content and ranges of the surface parameters. Note that in practice radar measurements are available only for a very limited discrete set of surface parameters. Thus, it is nearly impossible to have enough training patterns based on measurement alone. Dawson *et al.* developed an inversion method combining the fast learning algorithm of Manry *et al.* [25] and the IEM scattering model [12] to estimate not only the layer permittivity, but

also the surface roughness parameters and volume scattering characteristics. Dawson *et al.* [6] successfully applied this inversion technique to experimental data.

The advantage in using the neural network is that all surface parameters are included and the trained neural network acts as an empirical mapping relation between the radar measurement and the surface parameters. The disadvantage of such an approach is that this empirical relation is so involved that it is not practical to write it down explicitly. As a result, there is no mathematical expression available for examination after a network is trained. Furthermore, there is a variety of neural networks available in the literature. A user not in the field will feel uncomfortable using such a tool because it appears to him or her as a ‘black box’ although the feed-forward neural networks used today are based more heavily in statistics than in biology from which they found their name. To overcome the above stated difficulties, we present in the next section a method where the entire retrieval technique is transparent. The statistical properties of the method shown are analogous to the deterministic equation used in the context of adaptive filtering.

Note that as in the purely-empirical based inversions the quality of the data used to ‘train’ the processor (neural network) will determine the overall quality of the inversion. Even the methods presented here are limited in validity to the data sets and/or model(s) upon which they are based either directly or indirectly. These models include both the backscattering signal model and the soil dielectric model which will be discussed in Section III.

II. A ROBUST MULTI-DIMENSIONAL REGRESSION TECHNIQUE

A. Statement of the Inversion Problem

To motivate the development of a multidimensional regression technique, we assume that we have a set of training patterns which characterize the forward mapping $F(\Theta) \rightarrow \mathbf{Z}$ from a set of surface parameters Θ and their associated noisy measured signal \mathbf{Z} . This set of N_v training patterns each consisting of M ground parameters and the corresponding N radar observations have the form $\{\mathbf{Z}; \Theta\}$ or more explicitly $\{z_1, z_2, \dots, z_N; \theta_1, \theta_2, \dots, \theta_M\}$. Statistical inversion schemes based on a set of training patterns $\{\Theta; \mathbf{Z}\}$ can be expected to work well when the measured signal denoted by \mathbf{Z} is expressible as the addition of a noise free signal model $S(\Theta)$ plus some additive noise n as

$$\mathbf{Z} = S(\Theta) + n. \quad (6)$$

If we have no *a priori* information about the distribution of the noise component n , we usually assume it to be Gaussian. The approximate variance of the noise component can be computed from measured data.

For the current problem, Θ represents the set of unknown surface parameters denoted by $\{\sigma, l, m_v\}$; the surface rms roughness, the correlation length, and the volumetric soil moisture, respectively. The set of noisy observations \mathbf{Z} contains the available sensor channels at the available polarizations, incidence angles and frequencies, for example $\mathbf{Z} =$

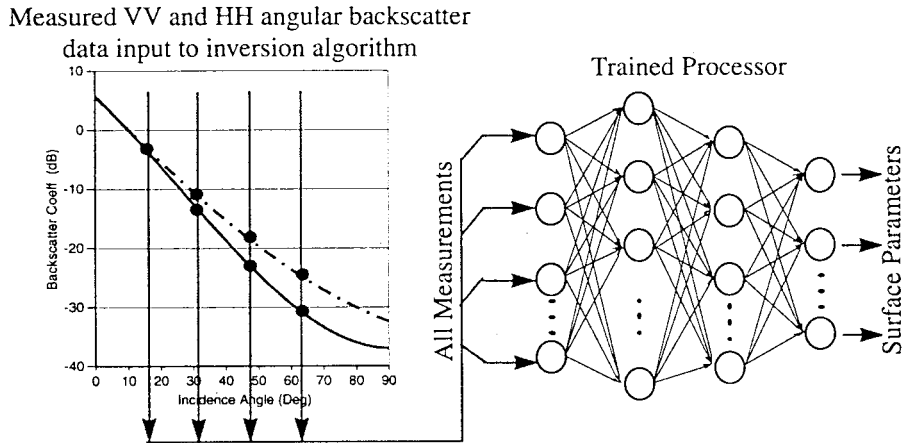


Fig. 3. The approach of Dawson [1992] uses a neural network based inversion in which all measured data (both VV and HH) is input at once with m_v as a direct output.

$\{\sigma_{vv}^o(10^\circ, 5 \text{ GHz}), \sigma_{hh}^o(50^\circ, 18 \text{ GHz}), \dots\}$. The objective of the problem is to find the inverse mapping $\Theta \leftarrow F^{-1}(\mathbf{Z})$.

B. Properties of the Training Patterns

We cannot use just any set of training patterns when designing an inversion. A desired set of training patterns for inversion should have the following properties:

1. It should span the entire observation and parameter space of interest or else the derived inversion algorithm will be restricted in its applicability to only those conditions represented in the training data;
2. The training patterns must be self-consistent and have the same trends in polarization, frequency, level and angles as field measurements;
3. Its noise should have the same variance as the measured data to be input to the estimator.

Indeed, there is a dearth of such a set of experimental data which contain not only a good set of calibrated measurements but also the appropriate ground-truth data. Generally, the ranges in the surface parameters from many experimental data sets are small and it is necessary to augment these measurements using an empirical or theoretical forward scattering model as performed here.

C. Development of the Regression Model

Once a set of training patterns are known, an inverse mapping relation from the measured data to the ground parameters can be developed with a regression technique. The structure of such a technique is shown in Fig. 4(a) in which the j th output $\tilde{\theta}(j)$ (an estimate for the j th unknown parameter $\theta(j)$) for the p th input vector \mathbf{Z}_p consisting of N elements is expressed as the weighted sum of a set of basis functions ϕ_p due to the p th input vector as

$$\tilde{\theta}(j, \mathbf{Z}_p) = \sum_{i=1}^L w(j, i) \cdot \phi_p(i) \equiv \phi_p^T \mathbf{w}(j). \quad (7)$$

In (7), the summation is for all L members of the basis function ϕ which is evaluated for the p th input vector \mathbf{Z}_p

and $w(j, i)$ is the weight connecting the i th basis function $\phi_p(i)$ to the j th output. Each of the desired outputs denoted by $\theta(j, \mathbf{Z}_p)$ represents a different approximating function; that is $\tilde{\theta}(j, \mathbf{Z}_p) \approx F_j^{-1}(\mathbf{Z}_p)$. This representation in which the output is a weighted sum of a set of linear or nonlinear basis functions is standard for most regression problems. The basis function $\phi = \{\phi_1, \phi_2, \dots, \phi_L\}$ may represent the original N system inputs $\mathbf{Z} = \{z_1, z_2, \dots, z_N\}$, in which case the length of the basis vector $L = N$, or a function of \mathbf{Z} or a combination of both. Two such basis functions are discussed in more detail in the next subsections. The form in (7) is also attributed to the fact that linear activations (rather than sigmoidal) are used in the last layer.

We now derive the expression for the vector of weighting coefficients $\mathbf{w}(j)$ for all the M outputs for a given ϕ . Assuming that we have a total of N_v training patterns, the total error for the j th output E_j is

$$E_j = \sum_{p=1}^{N_v} [\theta_p(j) - \tilde{\theta}_p(j)]^2 = \sum_{p=1}^{N_v} [\theta_p(j) - \phi_p^T \mathbf{w}(j)]^2. \quad (8)$$

The total mean square error (MSE) for the system is

$$\text{MSE} = \frac{1}{N_v} \sum_{j=1}^M E_j = \frac{1}{N_v} \sum_{j=1}^M \sum_{p=1}^{N_v} \left[\theta_p(j) - \sum_{i=1}^L w(j, i) \cdot \phi(i) \right]^2. \quad (9)$$

The next step is to determine the best set of coefficients $\mathbf{w}(j)$ for the j th output which minimize the mapping error in (8). This is done by taking the partial derivative of the expression for error with respect to each unknown coefficient and equating the result to zero. This operation [Manry *et al.*, 1994] gives M sets of L linear equations; one set of equations for each output containing L unknowns. Minimizing the error for the j th output E_j with respect to the n th weighting

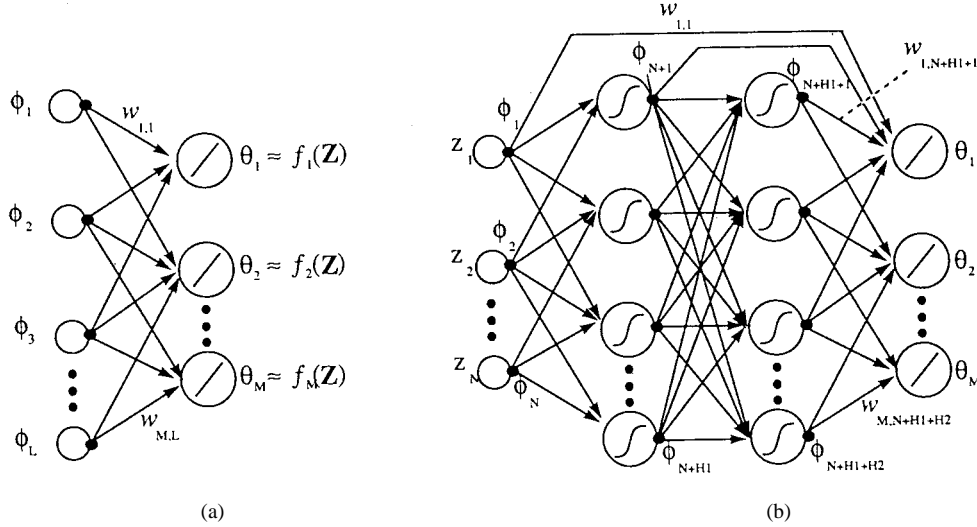


Fig. 4. Illustration of (a) general regression and (b) MLPBF structures. The MLPBF structure includes links to *all* preceding units and uses linear (rather than sigmoidal) activations in the output layer. Note that the MLPBF structure can be expressed in the form of the general regression structure with the basis functions being equal to the input and hidden unit activations as given in (14).

coefficient $w(j, n)$ yields

$$\frac{\partial E_j}{\partial w(j, n)} = -2 \sum_{p=1}^{N_v} [\theta_p(j) - \tilde{\theta}_p(j)] \left\{ \frac{\partial}{\partial w(j, n)} \tilde{\theta}_p(j) \right\} = 0 \quad (10)$$

where $n = 1$ through L . The partial derivative in (10) is equal to $\phi_p(n)$. After substituting $\tilde{\theta}_p(j)$ using (7) in (10) we get the following set of equations

$$\sum_{p=1}^{N_v} \theta_p(j) \phi_p(n) = \sum_{p=1}^{N_v} \phi_p(n) [\phi_p^T \mathbf{w}(j)]. \quad (11)$$

Equation (11) can be expressed more compactly using the definition of the cross correlation vector $\mathbf{R}_{\Theta\phi}$ and the autocorrelation matrix of the basis functions $\mathbf{R}_{\phi\phi}$ as

$$\mathbf{R}_{\Theta\phi} = \mathbf{R}_{\phi\phi} \mathbf{w}(j). \quad (12)$$

The linear equation in (12) yields $\mathbf{w}(j) = [\mathbf{R}_{\phi\phi}]^{-1} \mathbf{R}_{\Theta\phi}$. If the autocorrelation matrix $\mathbf{R}_{\phi\phi}$ is ill-conditioned, a constrained minimization technique for the linear system such as the biconjugate gradient technique [11], [17] can be applied to find $\mathbf{w}(j)$.

In what follows we shall discuss two possible sets of basis functions ϕ : the first set is called multinomial basis functions (MBF) and the second multilayer perceptron basis functions (MLPBF).

D. The MBF

The first basis function discussed is the MBF which represents a transformed version of the original set of N inputs $\mathbf{Z} = \{z_1, z_2, \dots, z_N\}$ into an expanded multinomial combination of \mathbf{Z} . This basis function follows the approach of Gaut *et al.*, [16]. The transformed vector $\mathbf{Y} = [y_{01}, y_{11}, \dots, y_{1N}, y_{22}, \dots, y_{31}, \dots]^T$ has elements y_{km} which denote the m th one-term polynomial of degree k in the input variables z_j for $j = 1 \dots N$. For example, for

($N = 3$) and a highest degree expansion of $D = 2$, we have $\mathbf{Z} = \{z_1, z_2, z_3\}$, $y_{2m} = \{z_1^2, z_1 z_2, z_1 z_3, z_2^2, z_2 z_3, z_3^2\}$, $y_{1m} = \{z_1, z_2, z_3\}$, and $y_{01} = 1$. The transformed input vector \mathbf{Y} contains the elements in the basis function as $\phi = \mathbf{Y} = \{1, z_1, z_2, z_3, z_1^2, z_1 z_2, z_1 z_3, \dots\}$. In the more general case, given the dimension N of the input vector \mathbf{Z} and the highest degree D , the number of degree- k terms is $(k + N - 1)! / (k!(N - 1)!)$ and the total number of terms in the multinomial basis ϕ is

$$L = \sum_{k=0}^D \frac{(k + N - 1)!}{k!(N - 1)!} = \frac{(D + N)!}{D!N!}. \quad (13)$$

E. The Multilayer Perceptron Basis Function (MLPBF)

One consideration in the design and use of the MBF technique described in the previous subsection is the selection of the highest degree D in (13). This selection is analogous to the selection of the basis functions used in the inverse function expansion. The length L of the basis vectors in the MBF grows almost exponentially with increased order D and number of system inputs N . The solution to the resulting $L \times L$ set of linear equations becomes more unstable with increased order D . An alternative is to use the multilayer perceptron (MLP) to generate a set of basis functions which is more compact in that the number of equations in the linear system will be smaller while providing the needed combinations required to build complex basis functions. This second basis function from the MLP (termed the MLPBF) is from Manry *et al.* [24], [25] which is an extension of the network representation as given by Sartori and Antsaklis [29] and Barton [1].

Several researchers [15], [21], [22] have shown that the MLP with a single hidden layer having nonlinear activation is capable of approximating any real-valued continuous function, provided a sufficient number of units within this hidden layer exists. In this sense, multilayer feedforward networks form a class of *universal approximators* and can be used to supply a

set of basis functions which are the nonlinear outputs to each node within the network [Fig. 4(b)]. While a general MLP can form arbitrary mappings, in practice, finite network size and limited training demand that only an approximation to the optimal solution can ever be achieved.

The improvement this technique provides is a more general set of basis functions which can be adjusted when we change the inner hidden weights. In the MBF, no further refinement of the basis functions are possible. The structure of the MLPBF is shown in Fig. 4(b). The network is fully connected in which the j th output is expressed as a linear combination of $L = 1 + N + N_h$ total terms which include the N inputs, all $N_h = N_{h_1} + N_{h_2}$ hidden units with nonlinear activations in each of the inner layers (for simplicity the figure does not show all connections), and a bias on the output. This fully connected network used in the MLPBF is an improved version of the simple feed-forward MLP network presented by Rumelhart and McClelland [28] as they have more free parameters (weights) and thus a higher pattern storage capacity [18]. Thus the basis functions are no longer multinomial combinations of the inputs as in the MBF but are the set given by

$$\phi(i) = \begin{cases} z_i & i = 1 \cdots N \\ g\left(\sum_{k=1}^{N_f} w_h(k) \cdot \phi_p(k)\right) & i = (N+1) \cdots N_h \\ 1 & \text{Last basis element} \end{cases} \quad (14)$$

In (14) $g(\bullet)$ represents the sigmoidal or other nonlinear activation function as commonly found in neural network MLP's trained using back propagation (BP) [Werbos, 1972]. In the same equation, w_h denotes the weights connecting a hidden unit to the preceding layers; this weight is different from the output weights in (7). The j th output $\tilde{\theta}_p(j)$ for the p th pattern is the inner product of the previous network activations vector ϕ_p and weight vector for the output $w_o(j)$ as

$$\tilde{\theta}_p(j) = \sum_{i=1}^L w_o(j, i) \cdot \phi_p(i), \quad (15)$$

We use the added \bullet_o subscript for the weighting coefficients to indicate that we are optimizing only the weights connection to the output in the MLPBF technique and not all weights within the network as done in the neural network MLP. This MLPBF is advantageous over those found in the MBF structure because the basis functions are actually optimally changed during training.

After randomly selecting random initial weights in the interval $(-1, 1)$, we can solve the linear set of equations $\mathbf{R}_{\Theta\phi} = \mathbf{R}_{\phi\phi}\mathbf{w}(j)$ for the output weights. The inner weights feeding the hidden units can be optionally updated via standard back-propagation techniques to further reduce the total RMS error, although Dawson and Fung [5], [6] have demonstrated that updating the output weights via (12) alone offers the largest advantage in terms of computational time. This optional step of updating the inner hidden weights (and thus the basis functions which they form) can be thought of as 'fine tuning' the basis functions. In simulations, we have found that the training time for the MBF, which requires only a single pass through the linear-equation algorithm of (12), is about the same

as several passes through the FMLP algorithm. Our approach differs from a neural network approach because we use the MLP structure to generate basis functions only. In conventional MLP's used in a neural network setting, *all* weight coefficients within the network are updated via a sequential training method which generally requires a very large number of iterations. The advantage of this network representation is in the understanding of the network representation and the training speed. Manry et al. [24], [25] have shown that networks of this form allow much faster training than MLP neural networks trained using back propagation.

III. SOIL MOISTURE INVERSION EXAMPLE

To illustrate the inversion techniques described in the Section II, we apply them to data given in [26] collected by the POLARSCAT L, C, and X band scatterometer system [33]. The data set consists of seven angular (10 – 70°) measurements at VV and HH polarizations for four bare soil sites with different degrees of surface roughness statistics under two wetness conditions. Thus, there is a total of eight sets of multifrequency and multiangle data. This data is particularly useful because of the ground truth information which includes surface roughness statistics and a soil moisture profile. Correlation length for the data lies in the range of 8–10 cm. This POLARSCAT data has been previously examined using the empirical models described in Section I [9], [26]. Therefore, for comparing different inversion methods this data set is the most convenient.

The first step in the creation of an inversion algorithm is the generation of a set of training patterns as mentioned in Section II. This training set is usually made up of measured data and "synthetic" data created using an appropriate signal model. The signal model to be used here is the IEM [12].

A. A Comparison Between an Empirical Signal Model and the IEM Model

In this subsection we want to compare the performance of an estimator based on training patterns created using the empirical signal model [10] (hereafter denoted by the Dubois Empirical or simply by DE) and the IEM scattering model. A total of $N_v = 15000$ training patterns were generated by employing each of the signal models $S(\Theta)$ and then adding white Gaussian noise as $\mathbf{Z} = S(\Theta) + n$. The noise level was selected as to be consistent with that found in the experimental POLARSCAT data. The parameters of interest Θ used to generate the training patterns were randomly selected from within a bounded set whose limits were expected to cover the full ranges of the surface parameters in the scene.

Fig. 5 shows a comparison of these two signal model predictions to measured POLARSCAT data. Note that the DE model curves upward as the incident angle increases.

The DE training set contained VV and HH polarizations at L, C, and X frequency bands at 30 , 40 , 50 , and 60° (24 inputs) along with the corresponding unknowns $\Theta = \{\sigma, m_v\}$ used in the computation of $S(\Theta)$. The inputs for the DE

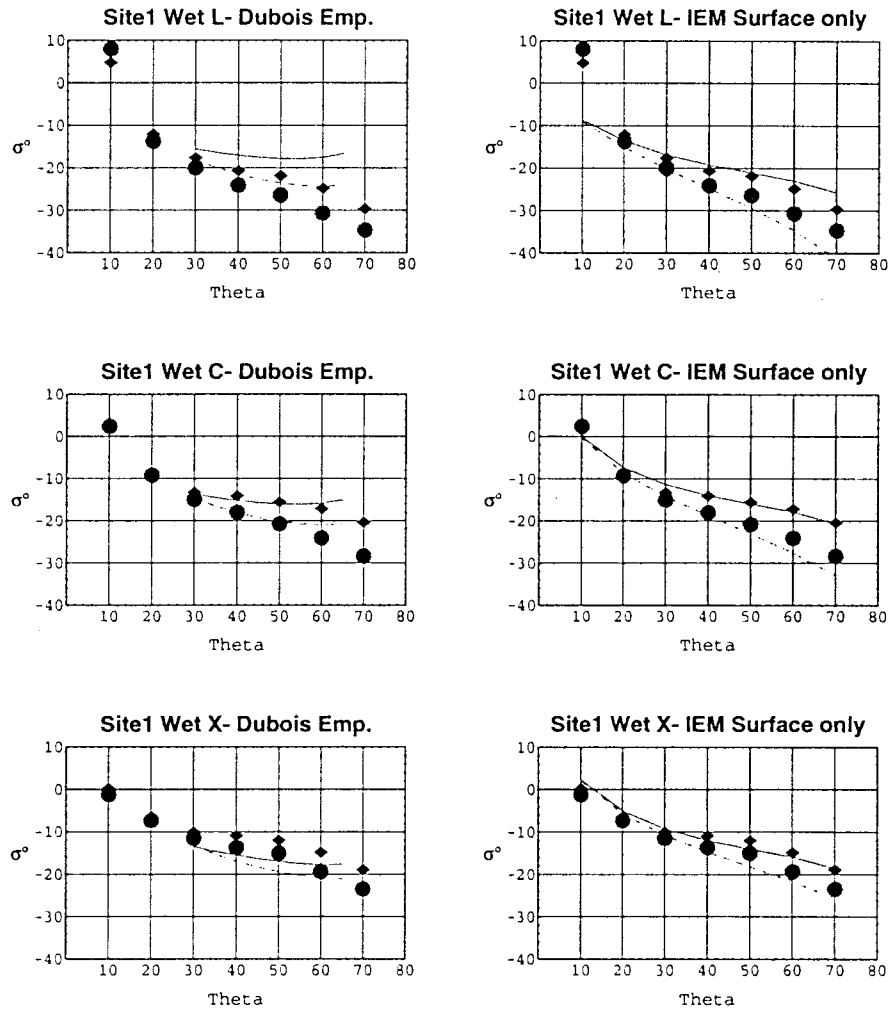


Fig. 5. Comparison of DE and IEM signal models with POLARSCAT measured data from Site 1 (wet). The diamond and circles represent measured data for VV and HH polarizations, respectively. The lines indicate Dubois empirical model fits (left column plots) and IEM surface data fits (right column) for the measured data points.

case were constrained by the validity of the DE model. To compare under identical conditions, the IEM training set used the same 24 inputs except for the correlation length which is an additional output parameter as $\Theta = \{\sigma, l, m_v\}$. Note that one more output parameter represents one more source of error and usually leads to a larger total RMS error. Careful attention was made during the generation of the parameters to keep the values for the surface correlation length and the surface rms height realizable in the natural scene by requiring the rms slope $\sqrt{2}(k\sigma)/(kl) < 0.4$ in using the IEM model. For both the DE and IEM, the permittivity model of Hallikainen *et al.* [20] was used to relate the physical soil moisture to the electrical parameter of the surface (ϵ_r). A single training set was used for all sites. The results for the inversion are given in Table I. As a sidenote, we have used a constant dielectric value for the dielectric layer although a more realistic method is to use a variable *dielectric transition layer* such as suggested by Fung *et al.* [14].

The two columns in Table I indicate the RMS errors for the inversion estimates using the POLARSCAT data. In general, the retrieval based on the IEM signal model performed very

TABLE I
COMPARISON OF TRAINING AND INVERSION BASED ON DE (DUBOIS EMPIRICAL) AND IEM SIGNAL MODELS. FOR THESE COMPARISONS, THE SAME 24 MODEL INPUTS WERE USED IN THE MBF TRAINING

Signal Model $S(\Theta)$	POLARSCAT data m_v inversion RMS Error	POLARSCAT data σ inversion RMS Error
DE	0.7838	0.6474
IEM	0.1931	0.1277

well relative to the DE. This is expected in view of the comparisons shown in Fig. 5.

B. An Analysis of MLPBF and MBF Performance

We now consider the performance of MLPBF and MBF in the regression algorithm. For MBF, we restricted the upper polynomial degree to $D = 2$ which has a total of 231 free parameters [using (13)] for each output. The training set contained VV and HH polarizations at L 30, 40°, C 10, 30, 40, 50, 60°, and X 30, 40, 50° along with the corresponding

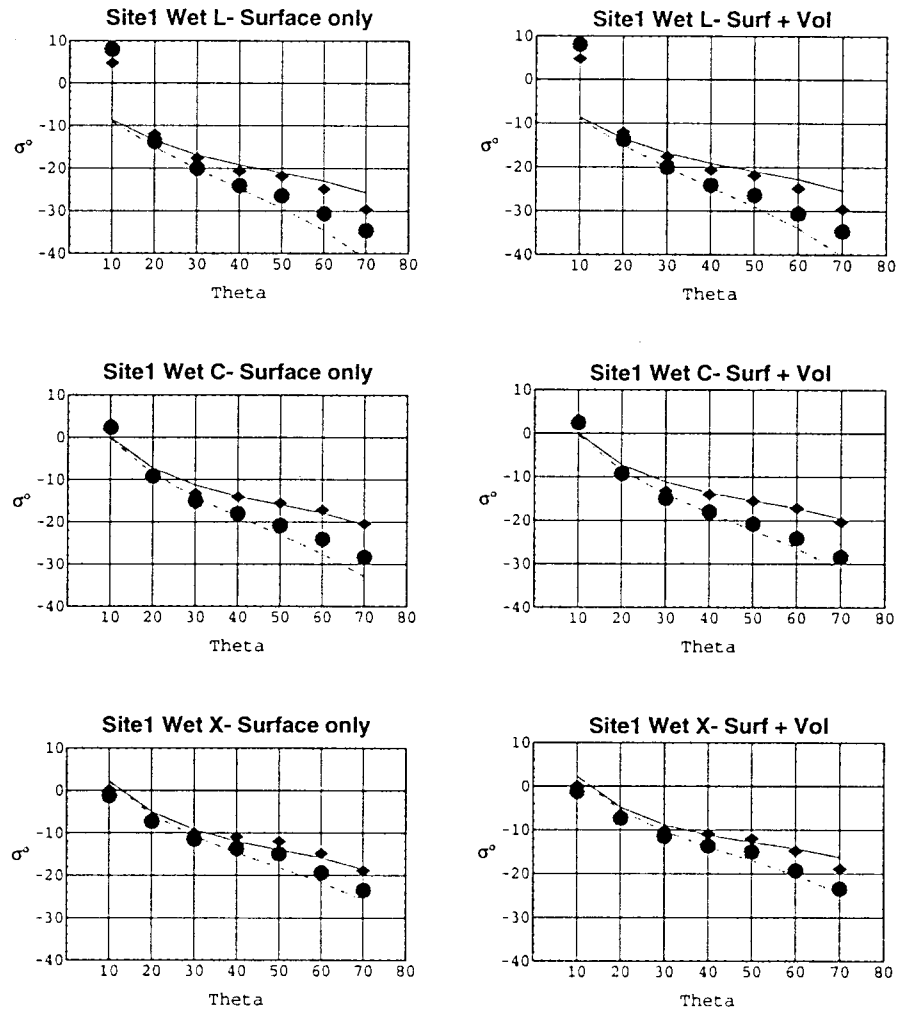


Fig. 6. Comparison of IEM signal model with and without volume scattering component with POLARSCAT measured data. The diamond and circles represent measured data for VV and HH polarizations, respectively. The lines indicate IEM surface-only fits (left column plots) and IEM surface + volume data fits (right column) for the measured data points.

unknowns $\Theta = \{\sigma, l, m_v\}$ computed using the IEM as the signal model. This set of inputs for inversion was selected during an initial comparison of $S(\Theta)$ and the measured data in which we verified the validity of the signal model. Thus, we have a total of 20 inputs to the IEM inversion scheme and three outputs. For consistency in the comparisons, we selected the number of hidden units in the one and two hidden-layer MLPBF structure such that the number of multiplies for the MLPBF and MBF are approximately equal. The MBF and MLPBF inversion techniques were trained using the patterns generated with the IEM model. For MLPBF, fifty choices of basis functions were made and the lowest RMS error is shown in Table II along with the total MSE from one pass of using MBF.

C. Comparison of the IEM With and Without Volume Scattering Model

Next, we consider the difference in the results when $S(\Theta)$ represents not only contribution from the surface scattering model (the IEM) but also the first-order radiative transfer vol-

TABLE II
TRAINING RESULTS FOR INVERSION OF SOIL MOISTURE
AND ROUGHNESS STATISTICS USING MBF AND MLPBF

Basis Function	Network Structure	Number of free parameters in network	Multiplies per pattern	Total Training MSE
MBF	D=2	Nout*L=693	903	1.582
MLPBF	20-34-3	879	910	1.558
MLPBF	20-15-12-3	891	915	1.579

ume scattering model [Fung, 1993]. Fig. 6 shows a comparison of the signal model with and without the effects of volume scattering. Note that when the effect of volume scattering is added to the IEM signal model, the HH backscattering at larger incidence angles rises. As the albedo increases, the separation between VV and HH decreases. The angles selected are those used in the comparison described in the previous subsection. The volume scattering component was selected to be consistent with the ground truth information. Inversion errors for the estimates based on MLPBF networks trained with 15 000 theoretical patterns and tested

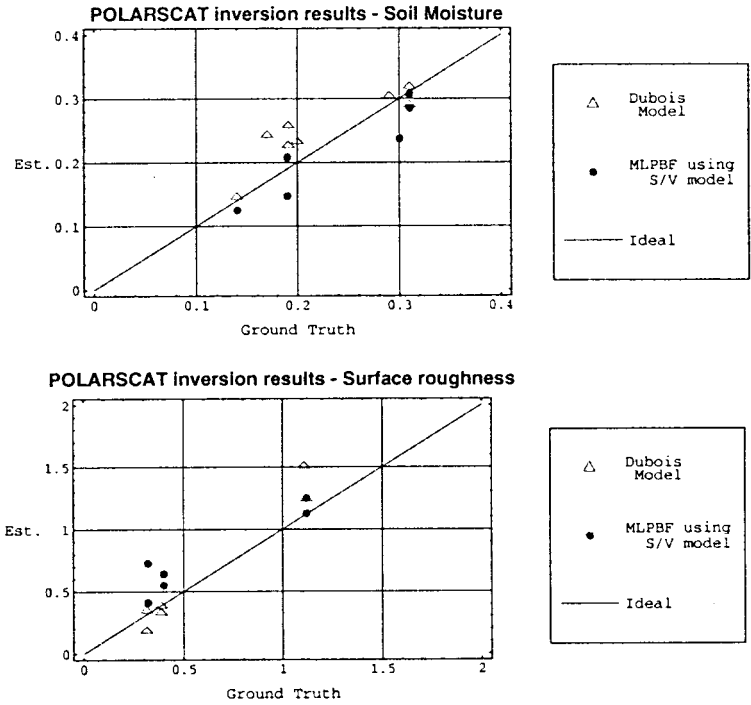


Fig. 7. Analysis of the POLARSCAT inversion using the empirical model of Dubois *et al.* [1995] and the MLPBF technique.

with POLARSCAT data are summarized in Table III. The signal model which includes the volume scattering contribution can be seen to slightly improve our soil moisture estimate.

D. Comparison of Parameter Inversion from MPLBF and Empirical Models of Section 1

The final comparison shown is for inversion based on MLPBF and the empirical models of Section I. The results reported for the inversion errors of the empirical models reported by Oh *et al.* [26] and Dubois *et al.* [10] are given for comparison. Our regression model employs a MLPBF and is trained with the IEM + Volume training set described in the previous section. The inversion results are tabulated in Table IV. Fig. 7 gives a comparison of the empirical inversion results of Dubois and our estimates. The straight diagonal line describes a perfect inversion algorithm where the estimated values match the values contained in ground truth data. Deviations from this line represent the cumulative error of the algorithm. Based on this information, we can say that the MLPBF trained with IEM + Volume compares well with that achieved using empirical models. This approach gives one more option which may be useful for the inversion of soil moisture and surface roughness statistics.

IV. SUMMARY AND DISCUSSION

There are several difficulties in the retrieval of soil moisture from active radar measurements including a dearth of backscatter measurements and the associated ground truth. In addition, partly because of the scarcity of appropriate data, current empirical forward scattering signal models are

TABLE III
COMPARISON FOR POLARSCAT INVERSION BASED ON THE IEM AND IEM + VOLUME FOR THE SIGNAL MODEL

	<i>POLARSCAT data m_v inversion RMS error</i>	<i>POLARSCAT data σ inversion RMS error</i>
MLPBF inversion based on IEM signal model	0.081	0.178
MLPBF inversion based on IEM + Volume signal model	0.034	0.213

TABLE IV
COMPARISON FOR DIFFERENT MODELING APPROACHES FOR SOIL MOISTURE INVERSION FROM POLARSCAT DATA (MLPBF RESULTS DIRECTLY FROM TABLE III)

	<i>POLARSCAT data m_v inversion RMS error</i>	<i>POLARSCAT data σ inversion RMS error</i>
MLPBF based on IEM + Volume signal model	0.034	0.213
Oh [’92] Empirical Results	0.040	0.410
Dubois [’95] Emp. Results	0.045	0.340

inadequate for other surface conditions which are not well represented in the training data. These weaknesses motivate our use of theoretically-based training data in the design of an estimator.

We have presented a multidimensional regression technique with two sets of basis functions; the MBF which uses multinomials as the basis functions for expansion and the MLPBF. Although the overall training time for the MLPBF is longer than the MBF, it offers the ability to create more general basis functions for remote sensing and other areas.

While the network structure to generate the MLPBF basis function is related to conventional neural networks, we did not use conventional neural network technique where one must sequentially update all weights within the network by using a gradient descent method repeatedly. The study indicates that the multidimensional regression method yields comparable results as the empirical regression models of Oh *et al.* [26] and Dubois *et al.* [10]. A major advantage over the empirical regression models is that the applicability of the method is not dependent on any specific set of field data.

The work presented here is in no way meant to be a final word on schemes for soil moisture retrieval. Clearly the limitations to this and other inversion techniques include the validity of the models upon which the training data is based (such as the IEM and dielectric models) and/or the consistency of the experimental data used to augment the training data set. Another problem, and one of more concern, is the potentially nonunique (ill-posed) nature of the soil moisture inversion problem under some conditions which is considered in other work [Dawson, 1996]. This ill-posedness of the problem is a limiting factor in the retrieval process particularly in the presence of vegetation where the radar return is not necessarily attributed to the effects from the soil.

When SAR data with adequate (and consistent) ground truth information are available, it would be desirable to test the many existing retrieval algorithms. Additional tests should concentrate on the retrieval of soil moisture which may include the effects of a profile as proposed by Fung *et al.* [1996].

ACKNOWLEDGMENT

The authors would like to thank Prof. Sarabandi at the University of Michigan for providing the POLARSCAT data and the reviewers for their constructive comments and taking the time to review this work.

REFERENCES

- [1] S. A. Barton, "A matrix method for optimizing a neural network," *Neural Computation*, vol. 3, no. 3, pp. 450–459, Fall 1991.
- [2] G. Bradley and F. T. Ulaby, "Aircraft radar response to soil moisture," *Remote Sensing Envi.*, vol. 11, pp. 419–438, 1981.
- [3] M. S. Dawson, "Remote sensing classification and inversion techniques using neural networks," masters thesis, The University of Texas at Arlington, 1992.
- [4] ———, "An improved estimation method with applications to soil moisture retrieval from remotely sensed data," Ph.D. dissertation, The University of Texas at Arlington, 1996.
- [5] M. S. Dawson and A. K. Fung, "Neural networks and their applications to parameter retrieval and classification," *IEEE Geosci. Remote Sensing Society Newsletter*, pp. 6–14, Sept. 1993.
- [6] M. S. Dawson, A. K. Fung, and M. T. Manry, "Surface parameter retrieval using fast learning neural networks," *Remote Sensing Reviews*, vol. 7, pp. 1–18, 1993.
- [7] M. S. Dawson, M. T. Manry, and A. K. Fung, "Information retrieval from remotely sensed data and a method to remove parameter estimator ambiguity," in *Proc. IGARSS'95*, Firenze, Italy, July 10–14, 1995, pp. 691–693.
- [8] ———, "Analysis of the sensitivity of active microwave measurements to soil moisture using the cramer-rao lower bound," in *Proc. PIERS 1995 (Progress in Electromagnetics Research Symp.)*, Seattle, WA, July 24–28, 1995.
- [9] P. C. Dubois, J. van Zyl, and T. Engman, "Measuring soil moisture with imaging radars," *IEEE Trans. Geosci. Remote Sensing*, vol. 33, July 1995.
- [10] ———, "Corrections to measuring soil moisture with imaging radars," *IEEE Trans. Geosci. Remote Sensing*, vol. 33, Nov. 1995.
- [11] R. Fletcher, "Conjugate Direction Methods," *Numerical Methods for Unconstrained Optimization*. W. Murray, Ed. New York: Academic, ch. 5, 1972.
- [12] A. K. Fung, Z. Li, and K. S. Chen, "Backscattering from a randomly rough dielectric surface," *IEEE Trans. Geosci. Remote Sensing*, vol. 30, Mar. 1992.
- [13] A. K. Fung, *Microwave Scattering and Emission Models and Their Applications*. Norwood, MA: Artech House, 1994.
- [14] A. K. Fung, M. S. Dawson, K. S. Chen, A. Y. Hsu, E. T. Engman, P. O. O'Neil, and J. Wang, "A modified IEM model for scattering from soil surfaces with application to soil moisture sensing," in *Proc. IGARSS'96*, Lincoln, NB, May 27–31, 1996.
- [15] A. R. Gallant and H. White, "There exists a neural network that does not make avoidable mistakes," in *Proc. IEEE Second Int. Conf. Neural Networks*, 1988, pp. 657–664.
- [16] N. E. Gaut, E. C. Reifstein, and D. T. Chang, "Microwave properties of the atmosphere, clouds and the oceans," *Environmental Research and Technology Final Report*, Project P-280, Prepared for NASA GSFC under Contract NAS 5-21624, Mar. 1992.
- [17] P. E. Gill, W. Murray, and M. H. Wright, *Practical Optimization*. New York: Academic, 1981.
- [18] A. Gopalakrishnan, X. Jiang, M. Chen, and M. T. Manry, "Constructive proof of efficient pattern storage in the multi-layer perceptron," in *Conf. Record 27th Annu. Asilomar Conf. Signals, Syst. Comput.*, Nov. 1993, vol. 1, pp. 386–390.
- [19] Z. S. Haddad and P. Dubois, "Bayesian estimation of soil parameters from remote sensing data," in *Proc. IGARSS'1994*, vol. III, pp. 1421–1423.
- [20] M. T. Hallikainen, F. T. Ulaby, M. C. Dobson, M. A. El-Rayes, and L. WU, "Microwave dielectric behavior of wet soil—Part 1: Empirical models and experimental observations," *IEEE Trans. Geosci. Remote Sensing*, vol. GE-23, pp. 25–34, Jan. 1985.
- [21] E. J. Hartman, J. D. Keeler, and J. M. Kowalski, "Layered neural networks with Gaussian Hidden Units as universal approximators," *Neural Computation*, vol. 2, pp. 210–215, 1990.
- [22] K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators," *Neural Networks*, vol. 2, pp. 359–366, 1989.
- [23] T. A. Le Toan, P. Smacchia, J. Souyris, A. Beaudoin, M. Merdas, M. Wooding, and J. Lichteneger, "On the retrieval of soil moisture from ERS-1 SAR data," in *Proc. Second ERS-1 Symp.*, Hamburg, Germany, Oct. 11–14, 1993 (ESA SP-361).
- [24] M. T. Manry, M. S. Dawson, S. J. Apollo, L. S. Allen, W. D. Lyle, W. Gong, and A. K. Fung, "Fast training of neural networks for remote sensing," *Remote Sensing Reviews*, vol. 9, pp. 77–96, 1994.
- [25] M. T. Manry, X. Guan, S. J. Apollo, L. S. Allen, W. D. Lyle, and W. Gong, "Output weight optimization for the multi-layer perceptron," *Conf. Record 26th Annu. Asilomar Conf. Signals, Syst. Comput.*, Oct. 1992, vol. 1, pp. 502–506.
- [26] 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 Sensing*, pp. 370–381, 1992.
- [27] ———, "An inversion algorithm for retrieving soil moisture and surface roughness from polarimetric radar observations," in *Proc. IGARSS'1994*, vol. III, pp. 1582–1585.
- [28] *Parallel Distributed Processing*, D. E. Rumelhart and J. L. McClelland, Eds. Cambridge, MA: MIT Press, vols. 1 and 2, 1986.
- [29] M. A. Sartori and P. J. Antsaklis, "A simple method to derive bounds on the size and to train multilayer neural networks," *IEEE Trans. Neural Networks*, vol. 2, pp. 467–471, July 1991.
- [30] C. Schmillius and R. Furrer, "Frequency dependence of radar backscattering under different moisture conditions," *Int. J. Remote Sensing*, vol. 13, pp. 2233–2245, 1992.
- [31] J. Shi, J. V. Soares, E. T. Engman, and J. J. van Zyl, "Soil moisture measurements from airborne SAR," in *Proc. Third Airborne AIRSAR Workshop 1991*, pp. 68–77.
- [32] J. Shi, J. J. van Zyl, J. V. Soares, and E. T. Engman, "Development of soil moisture retrieval algorithm for L-band SAR measurements," in *Proc. IGARSS'1992*, pp. 495–497.
- [33] M. A. Tassoudji, K. Sarabandi, and F. T. Ulaby, "Design consideration and implementation of the LCX polarimetric scatterometer (POLARSCAT)," *Report 022486-T-2*, Radiation Laboratory, Univ. of Michigan, Ann Arbor, June 1989.
- [34] A. N. Tikhonov and V. Y. Arsenin, *Solution of Ill-Posed Problems*. New York: Wiley, 1977.
- [35] L. Tsang, Z. Chen, R. Marks, and A. T. C. Chang, "Inversion of snow parameters from passive microwave remote sensing measurements by a neural network trained with a multiple scattering model," in *Proc. IGARSS'1991 Symp.*, pp. 1965–1976.
- [36] S. Twomey, *Introduction to the Mathematics of Inversion in Remote Sensing and Indirect Measurements*. Oxford, UK: Elsevier Scientific, 1977.



Michael S. Dawson (S'90) was born in Alameda, CA, in 1965. He received the B.S. (with high honors), M.S., and Ph.D. degrees in electrical engineering from the University of Texas at Arlington in 1990, 1992, and 1996, respectively.

From 1989 to 1996, he was a Research Associate at the Wave Scattering Research Center where he worked on theoretical modeling of emission from dense media and classification and inversion of remotely sensed microwave data. He has served as a reviewer for several technical journals in the areas of applied estimators and neural networks for classification and inversion. His current interests include inverse statistical estimation methods for ill-posed problems and inversion error bounds. He is currently working as a Senior Analyst at E-Systems in Garland, Texas.

Dr. Dawson is a NASA Global Change Fellow and a NASA Graduate Researcher Fellow and has received numerous academic awards. He is a member of Tau Beta Pi, Eta Kappa Nu, and Alpha Chi.



Adrian K. Fung (S'60–M'66–SM'70–F'85) received the B.S. degree from Cheng Kung University, Taiwan, R.O.C. in 1958, the M.S.E.E. degree from Brown University, Providence, R.I. in 1961, and the Ph.D. degree from the University of Kansas, Lawrence, in 1965.

From 1965 to 1984 he was on the faculty of the Department of Electrical Engineering, University of Kansas. In the fall of 1984 he joined the University of Texas at Arlington where he is now a Jenkins Garrett Professor with the Electrical Engineering Department and Director of the Wave Scattering Research Center. His research interests have been in the areas of radar wave scattering and emission from earth terrains and sea, and radar and ISAR image generation, simulation, and interpretation. He has contributed to many book chapters including *Manual of Remote Sensing*. He is the author of *Microwave scattering Models and Their Applications* (Norwood, MA: Artech House, 1994) and is a coauthor of the three-volume book *Microwave Remote Sensing* (Norwood, MA: Artech House, vol. 1, 1981; vol. 2, 1982; vol. 3., 1986.)

Dr. Fung has served as an Associate Editor of *Radio Science* and the *IEEE Journal of Oceanic Engineering* and is an editor of the *Journal of Electromagnetic Waves and Applications*. He was the recipient of the Halliburton Excellence in Research Award in 1986, the 1989 Distinguished Research Award of the University of Texas at Arlington, and the 1989 Distinguished Achievement Award from the IEEE Geoscience and Remote Sensing Society.



Michael T. Manry (S'73–M'75–SM'90) was born in Houston, Texas, in 1949. He received the B.S., M.S., and Ph.D. degrees in electrical engineering in 1971, 1973, and 1976, respectively, from The University of Texas at Austin.

After working at the university for two years as an Assistant Professor, he joined Schlumberger Well Services, Houston, where he developed signal processing algorithms for magnetic resonance well logging and sonic well logging. He joined the Department of Electrical Engineering at the University of Texas at Arlington in 1982, and has held the rank of Professor since 1993. He is currently the Director of the Image Processing and Neural Networks Laboratory in the Department of Electrical Engineering. In the summer of 1989, he developed neural networks for the Image Processing Laboratory of Texas Instruments in Dallas. His recent work, sponsored by the Advanced Technology Program of the state of Texas, E-Systems, Mobil Research, Teledyne, Bailey Network Management, the NSF, and NASA, has involved the development of techniques for the analysis and fast design of neural networks for image processing, parameter estimation, and pattern classification. He has served as a consultant for the Office of Missile Electronic Warfare at White Sands Missile Range, MICOM at Redstone Arsenal, Texas Instruments, Geophysics International, Halliburton Logging Services, Mobil Research, and Verity Instruments.