

# ART-MMAP: A Neural Network Approach to Subpixel Classification

Weiguo Liu, Karen C. Seto, Elaine Y. Wu, Sucharita Gopal, and Curtis E. Woodcock

**Abstract**—Global or continental-scale land cover mapping with remote sensing data is limited by the spatial characteristics of satellites. Subpixel-level mapping is essential for the successful description of many land cover patterns with spatial resolution of less than  $\sim 1$  km and also useful for finer resolution data. This paper presents a novel adaptive resonance theory MAP (ARTMAP) neural network-based mixture analysis model—ART mixture MAP (ART-MMAP). Compared to the ARTMAP model, ART-MMAP has an enhanced interpolation function that decreases the effect of category proliferation in ART<sub>a</sub> and overcomes the limitation of class category in ART<sub>b</sub>. Results from experiments demonstrate the superiority of ART-MMAP in terms of estimating the fraction of land cover within a single pixel.

**Index Terms**—Adaptive resonance theory (ART) mixture MAP (ART-MMAP), adaptive resonance theory MAP (ARTMAP), mixture analysis, neural network, subpixel classification.

## I. INTRODUCTION

ACCURATE mapping of land cover at continental or global scales is currently limited by the spatial and temporal characteristics of the available satellite data. Many landscape features occur at spatial scales much finer than the resolution of the primary satellites used for continental or global land cover mapping [e.g., the Moderate Resolution Imaging Spectrometer (MODIS) and the Advanced Very High Resolution Radiometer (AVHRR)]. Methods for mapping land cover at subpixel level are thus of great interest. Several approaches have been used to estimate subpixel fractional cover. Most of them use various models which are implemented during the image classification procedure, including linear spectral unmixing [1]–[3], neural networks [4]–[8], fuzzy classifier [9], regression and decision trees [10]–[12], Gaussian mixture discriminant analysis [13] and maximum-likelihood classifier [14], [15]. The most common approach is linear unmixing models, which prior research has shown to generate acceptable results.

However, as demonstrated by Borel and Gerstl [16], linear unmixing models may not be suitable in cases when multiple

scattering results in nonlinear mixture. In this context, a nonlinear model can produce better results. Carpenter *et al.* [5] present a nonlinear algorithm for mixture estimation based on an adaptive resonance theory MAP (ARTMAP) neural network for identifying life form components of the vegetation mixture. Landsat Thematic Mapper (TM) imagery is used to estimate the subpixel information for life-form components. The ARTMAP-based mixture model is able to capture nonlinear effects and thus perform better than the conventional linear unmixing models. Atkinson *et al.* applied a multilayer perceptron (MLP)-based mixture model to decompose AVHRR imagery [7]. The “unmixture” information from the model was superior to results generated through linear unmixing model and fuzzy c-means classifier. Foody used a simple regression- and contouring-based approach to produce a subpixel land cover map based on fuzzy classification [9].

The objective of the paper is to present a novel mixture analysis model, ART mixture MAP (ART-MMAP). ART neural network-based ART-MMAP mixture model helps to estimate the proportion of land cover types within a single pixel of the coarse resolution image. Such an approach will support a wide range of applications within the National Aeronautics and Space Administration’s Earth Science Enterprise including global climate modeling, estimation of photosynthesis, more accurate estimates of global urban extent, and biophysical parameter estimation [17]–[19]. To evaluate the performance, ART-MMAP is compared to ARTMAP and the regression tree (RT) model with two datasets.

## II. ARTMAP NEURAL NETWORK

The ART family of pattern recognition algorithms was developed by Carpenter and Grossberg [20], [21]. ART is a match-based learning system, the major feature of which is its ability to solve the “stability-plasticity dilemma” or “serial learning problem,” where successive training of a network interferes with previously acquired knowledge. That is, learning a new pattern usually involves replacing or modifying the existing information base. The modification of training data can be done with relative ease if the network can learn all existing patterns in the training data. However, the real-world environment likely is more complex and dynamic than the training data. Training data are supposed to represent the possible range of variability within and among land cover types, but rarely do since they usually only include “pure” archetypes. ART networks maintain the stability of keeping previously learned patterns, while simultaneously being flexible, or plastic, enough to master new patterns.

Among the ART family models, fuzzy ARTMAP is a supervised learning system that has been used widely in many fields.

Manuscript received December 11, 2003; revised April 16, 2004. This work was supported in part by the National Aeronautics and Space Administration under Grant NAG5-10534 and in part by the National Science Foundation under Grant SES-9973474.

W. Liu is with ACI Worldwide, Incorporated, Riverside, RI 02915 USA (e-mail: wgliu@crsa.bu.edu).

K. C. Seto is with the Department of Geological and Environmental Sciences, Stanford University, Stanford, CA 94305 USA.

E. Y. Wu is with the Department of Ophthalmology, Medical University of South Carolina, Charleston, SC 29425 USA.

S. Gopal and C. E. Woodcock are with the Department of Geography, Boston University, Boston, MA 02215 USA.

Digital Object Identifier 10.1109/TGRS.2004.831893

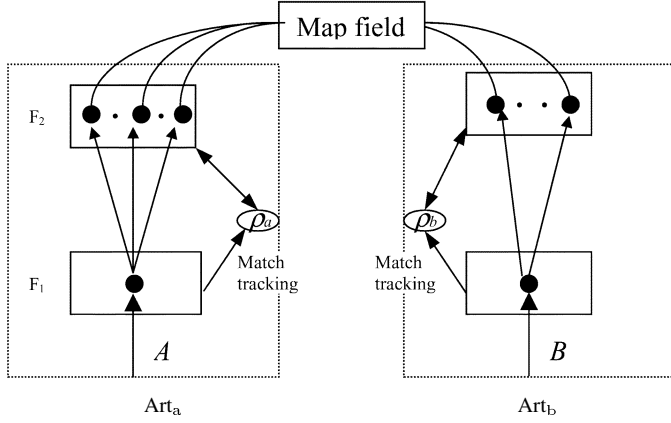


Fig. 1. Architecture of ARTMAP.

A comprehensive description of the model is detailed in Carpenter [22] from which the synopsis of the basic architecture of the fuzzy ARTMAP model is drawn (Fig. 1). It consists of a pair of fuzzy ART modules, ART<sub>a</sub> and ART<sub>b</sub>, connected by an associative learning network called map field. The architecture uses a learning rule that minimizes the training error (to reduce bias) while concurrently minimizing the testing error (to reduce the combination of bias and variance). The “hidden units” in ART<sub>a</sub> and ART<sub>b</sub> are called F<sub>2</sub> nodes, which represent learned recognition categories. Each category (cluster or F<sub>2</sub> node) extracts and generates common spectral properties from input training data.

During the training phase, the ART<sub>a</sub> and ART<sub>b</sub> modules are given input data A and desired output pairs B. As the two modules classify the A and B vectors into various map categories (clusters), the map field makes the association between ART<sub>a</sub> and ART<sub>b</sub> categories. If there is a discrepancy between the observed and predicted values of B, a memory search occurs in the ART<sub>a</sub> module. The match tracking component within the module increases the sensitivity of the ART<sub>a</sub> vigilance parameter,  $\rho_a$ , to activate a memory search. A new memory search increases the probability that an ART<sub>a</sub> category will generate a better predicted value of B. If none of the existing categories can minimize the predictive error or match the statistics of the input vector, a new category is generated. This allows intraclass variability to be captured through the creation of new classes. Alternatively, if there is a match, the spectral characteristics of the new input vector will be incorporated to redefine the attributes of the category. This weighting function in effect generalizes the category. Thus, an important attribute of fuzzy is that it can capture both intra and interclass variability, facilitating “many(cluster)-to-one (class)” mapping.

### III. ARTMAP ALGORITHM

The ARTMAP algorithm uses a choice parameter,  $\alpha > 0$ ; a learning parameter,  $\beta \in [0, 1]$ ; and a vigilance parameter,  $\rho \in [0, 1]$ . We summarize the ARTMAP algorithm by describing how procedures including category choice, reset and learning are executed. The choice function in ARTMAP, denoted by  $T_j$ , is defined for each input A and F<sub>2</sub> node j as follows:

$$T_j(A) = \frac{|A \wedge W_j|}{\alpha + |W_j|} \quad (1)$$

where  $W_j$  is the weight vector of the jth F<sub>2</sub> node, and the fuzzy AND operator “ $\wedge$ ” is defined by

$$(R \wedge Q)_i = \min(R_i, Q_i) \quad (2)$$

and the norm  $|\cdot|$  is defined by

$$|R| = \sum_{i=1}^L R_i \quad (3)$$

for any L-dimensional vectors R and Q.

At any given time, only one F<sub>2</sub> node can be active and the system is said to have made a category choice. Let the category choice be indexed by J, and let  $T_j(A)$  in the above equation be written as  $T_j$  for notational simplicity when the input A is fixed. Then

$$T_J = \max\{T_j, j = 1, \dots, M\} \quad (4)$$

where M is the number of category/F<sub>2</sub> node. Note that if more than one  $T_j$  is maximal, the category J with the smallest index is chosen.

**Resonance or Reset Operation:** Resonance occurs if the match function  $(|A \wedge W_j|)/(|A|)$  of the chosen jth category exceeds the vigilance criterion defined by  $\rho$ . A mismatch reset occurs when  $|A \wedge W_j|/|A|$  is less than  $\rho$ . As long as A remains constant, the same category j that was already selected as a category during search cannot then be selected.

**Learning:** Once the search terminates, the weight vector  $W_J$  is updated according to the following equation:

$$W_J^{\text{new}} = \beta \times (A \wedge W_J^{\text{old}}) + (1 - \beta) \times W_J^{\text{old}}. \quad (5)$$

In this equation, fast learning corresponds to setting  $\beta$  to 1. The mathematical and geometrical interpretation of fuzzy ART as well details of system dynamics and algorithm will not be discussed in this paper as they have been presented elsewhere in the literature [20]–[25].

**Prediction:** With (1)–(4) and a winner-take-all (WTA) rule, the F<sub>2</sub> node/category winner is assigned to the current input pattern. According to the association between ART<sub>a</sub> and ART<sub>b</sub> (one category in ART<sub>a</sub> is mapped to one category in ART<sub>b</sub>) through MapField, the prediction output of ART<sub>b</sub> is calculated based on the following equation:

$$O_i = \frac{\sum_{k=1}^{M_b} (W_{ki}^b \times M_{jk}^{ab})}{\sum_{m=1}^N \left( \sum_{k=1}^{M_b} W_{km}^b \times M_{jk}^{ab} \right)} \quad (6)$$

where N is the class number, i is the class index, j is the activated category in ART<sub>a</sub>,  $W_k^b$  is the weight vector of category  $F_2^{b_k}$ ,  $M_b$  is the number of categories in  $F_2^b$ , and

$$M_{jk}^{ab} = \begin{cases} 1, & \text{if } F_2^{a_j} \text{ maps to } F_2^{b_k} \\ 0, & \text{else} \end{cases} \quad (7)$$

For classification, the class label I is assigned to the input pattern with following equation:

$$O_I = \max\{O_i, i = 1, \dots, N\}. \quad (8)$$

For a mixture analysis, the mixture information is calculated as

$$P_i = O_i \quad (9)$$

where  $P_i$  is the probability/proportion of class i.

#### IV. ART-MMAP ALGORITHM

Current literature [26]–[29] shows the potential weakness and possible improvements of ARTMAP models. ARTMAP is sensitive to the noise and outliers, and this may cause over-fitting of the data and category proliferation. Several modified ARTMAP models (distributed ARTMAP (dARTMAP) [27], Gaussian ARTMAP [26], and  $\mu$ ARTMAP [28]) have been presented to partly resolve the category proliferation problem. The goal of the ART-MMAP method proposed in this paper is not to modify the learning process, but to decrease the effect of category proliferation in the testing processing for mixture analysis. The ART-MMAP model keeps the learning process of ARTMAP and changes the testing and recognition processes. During the testing process, ARTMAP selects the category to test samples using WTA rule. For the mixture analysis problem, because the categories in  $ART_a$  are similar, it is dangerous to take one winner over all others if the difference between the activation of the categories is indistinguishable. Instead of picking one winner, ART-MMAP selects the winners ( $ART_a$ ) based on one threshold parameter. If the activation value of the category is larger than the threshold value, the category is selected. If none of the categories are selected, the WTA rule will be activated. Below is the  $ART_a$  category selection rule for ART-MMAP

$$T_{j^a} = \begin{cases} \frac{|A \wedge W_{j^a}^a|}{\alpha + |W_{j^a}^a|}, & \text{if } \frac{|A \wedge W_{j^a}^a|}{\alpha + |W_{j^a}^a|} \geq \tau \\ 0, & \text{else} \end{cases} \quad (10)$$

where  $\tau \in [0, 1]$  is the threshold,  $T_{j^a}$  is the activation value of category  $j$  in  $ART_a$ , and  $W_{j^a}$  is the weight vector of the category in  $ART_a$ .

The interpretation of mixture information of ART-MMAP uses the weighted summation operator

$$P_i = \frac{\sum_{j=1}^{M_a} \sum_{k=1}^{M_b} (T_{j^a} \times M_{jk}^{ab} \times W_{ki}^b)}{\sum_{m=0}^N \left( \sum_{j=1}^{M_a} \sum_{k=1}^{M_b} (T_{j^a} \times M_{jk}^{ab} \times W_{km}^b) \right)} \quad (11)$$

where  $P_i$  is the proportion of class  $i$ , and  $j$  is the category index of  $ART_a$ ,  $k$  is the category index of  $ART_b$ ,  $W_k^b$  is the weight vector of  $F_2^{bk}$ , and  $M_a$  is the number of categories in  $ART_a$ .

With (10) and (11), ART-MMAP provides one enhanced function: interpolation. The learning of ARTMAP is data-driven, and the categories only store the information presented in the training data. Unlike the classification problem, the training data of mixture analysis problem cannot cover all combinations of class proportions. Thus, the learning process of ARTMAP will introduce the class category limitation in  $ART_b$ . For a two classes (0 and 1) mixture problem, if the training data only have pure class (either class 0 or class 1, no  $\langle 0.5, 0.5 \rangle$ -like combinations) samples, the prediction of the trained ARTMAP network will be either 0 or 1 regardless of the combination (i.e.,  $\langle 0.3, 0.7 \rangle$ ) of the testing sample. With (10), several feature categories ( $ART_a$ ) are selected if the input pattern is similar enough to the categories in the feature space. For example, three categories representing pure class 1 and 2 categories representing pure class 0 ( $ART_a$ ) are selected for a mixture pattern with (10). To simplify the computation, we

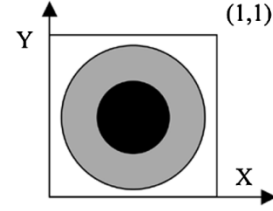


Fig. 2. Simulated dataset.

assume the activation value (1) of each selected winning category is equal. Following (11), the class proportion is calculated as 0.6 for class 0 and 0.4 for class 1. The mixture pattern is predicted as  $\langle 0.6, 0.4 \rangle$ , which does not exist in the training data. Thus, the interpolation function of ART-MMAP decreases the effect of class category limitation and smoothes the prediction error as well. Due to the learning rules, ARTMAP will generate categories representing the noise and outliers [26]. In the testing process, some input patterns will be classified into these noise categories with WTA selection rule and then may cause misclassification. With the interpolation function, the error caused by the noise category will be limited to a low level.

#### V. DATA

To evaluate the performance of ART-MMAP, two datasets are applied. One is a simulated dataset, and the other is derived from MODIS and Landsat images.

##### A. Simulated Dataset

The simulated two-dimensional (2-D) feature space has one square and two circles (Fig. 2). The points in the inner circle have class label 1, the points out of the outer circle are labeled as 0, and other points in the outer circle and out of the inner circle have mixture information. The class label of these mixed pixels in gray area (Fig. 2) ranges from 0 to 1 and is proportional to the reciprocal of the distance from the points to the center of the square. The dataset has randomly generated 900 training samples and 10 000 testing samples. Each sample has a 2-D feature vector  $\langle x, y \rangle$  (i.e.,  $\langle 0.1, 0.5 \rangle$ ) that is the input to  $ART_a$  and a class label (i.e.,  $\langle 0.3 \rangle$ ) that is the input to  $ART_b$  during the training process.

##### B. MODIS Dataset

The satellite data consists of a Landsat-7 Enhanced Thematic Mapper+ (ETM+) image (path 168 and row 032) acquired on August 19, 1999, and a MODIS image [20 horizontal and 04 vertical (20–4)] acquired on August 27, 2000, with a total area of  $139 \times 146$  (20 294) cells each having a resolution of  $1 \text{ km}^2$ . Landsat-7 ETM+ image has a resolution of  $30 \text{ m} \times 30 \text{ m}$ . The images cover north central Turkey. Three steps are applied to calculate the fractional land cover for each MODIS pixel at  $1\text{-km}^2$  resolution. First, the Landsat-7 ETM+ image is classified into four classes (forest, grass, water, and barren) with ART-VIP [25]. Second, the classified Landsat-7 ETM+ image is registered with MODIS images using ground control points (GCPs) and image-to-image registration with ERDAS Imagine software. Third, the classification map is associated with MODIS image to estimate the fraction of land cover

classes for each pixel at MODIS scale. For the purpose of training, a total of 1884 MODIS pixels were randomly selected from the image. The testing dataset includes all the pixels of the entire image. Each sample has a digital number value of seven spectral bands and the proportion of each land cover type. The spectral vector is the input to ART<sub>a</sub> and the class proportion vector (i.e.,  $\langle 0.3, 0, 0.7, 0 \rangle$ ) is the input to ART<sub>b</sub> for learning.

## VI. RESULTS AND DISCUSSION

Carpenter *et al.* [5] and Liu *et al.* [4] used the ARTMAP algorithm for mixture analysis, and the results are significant improvements to the linear unmixing model. In this paper, the performance of ART-MMAP is compared to that of the ARTMAP mixture algorithm described in Carpenter *et al.* [5] and Liu *et al.* [4] and the regression tree algorithm with two datasets. The regression tree can automatically generate a tree for class proportion estimation. Two measures of performance are used—root mean square (rms) errors and bivariate distribution functions (BDFs) between the real and estimated subpixel land cover proportions. The BDF is helpful to visualize the accuracy of prediction by mixture models. Finally, we calculate the Z-statistics using subpixel prediction error to test the statistical significance of differences in results between ART-MMAP and ARTMAP, ART-MMAP and the regression tree model [30]. The Z-statistics tests the null hypothesis that, for a particular class, the subpixel prediction error of the ART-MMAP model is equal to that using the ARTMAP/regression tree model. Values of Z that exceed the critical threshold indicate that the prediction errors are different. The less prediction error (more accurate prediction) is identified by the sign on Z. A negative sign indicates that the results from ART-MMAP model are more accurate (less error) than results from ARTMAP or the regression tree model.

### A. Simulated Dataset

With the simulated dataset, the parameters of ART-MMAP are set as  $\rho_a = 0.7$ ,  $\rho_b = 0.98$ , and  $\tau = 0.97$ . Testing with 10000 samples, the rms error is 0.06, 0.031, and 0.05 for ARTMAP, ART-MMAP, and the regression tree, respectively. The absolute value of the prediction error (difference between real and estimated class proportion) of each testing sample is shown in Fig. 3. The figure clearly shows the distribution of the prediction error that mainly exists in the buffer of two circles (gray area in Fig. 2). In terms of prediction error, ART-MMAP has smallest error range (0–0.17). The high error range of ARTMAP (0–0.65) is mainly caused by the two peak error areas. The error range of the regression tree is 0–0.23. The analysis results clearly illustrate that ART-MMAP is superior to ARTMAP and the regression tree with this simulated dataset.

Points along a 1 : 1 line on the BDF graph indicate a prediction that matches completely with the real proportion. The smaller the difference between the predicted and real proportion, the closer the point will lie to the diagonal 1 : 1 line. Fig. 4 illustrates that ART-MMAP predicts more points within the 10% error bound than ARTMAP and the regression tree. As discussed in Section IV, ARTMAP only stores the category information (both ARTs) of training data. Thus, the prediction of any sample will be one of the categories at ART<sub>b</sub> (class

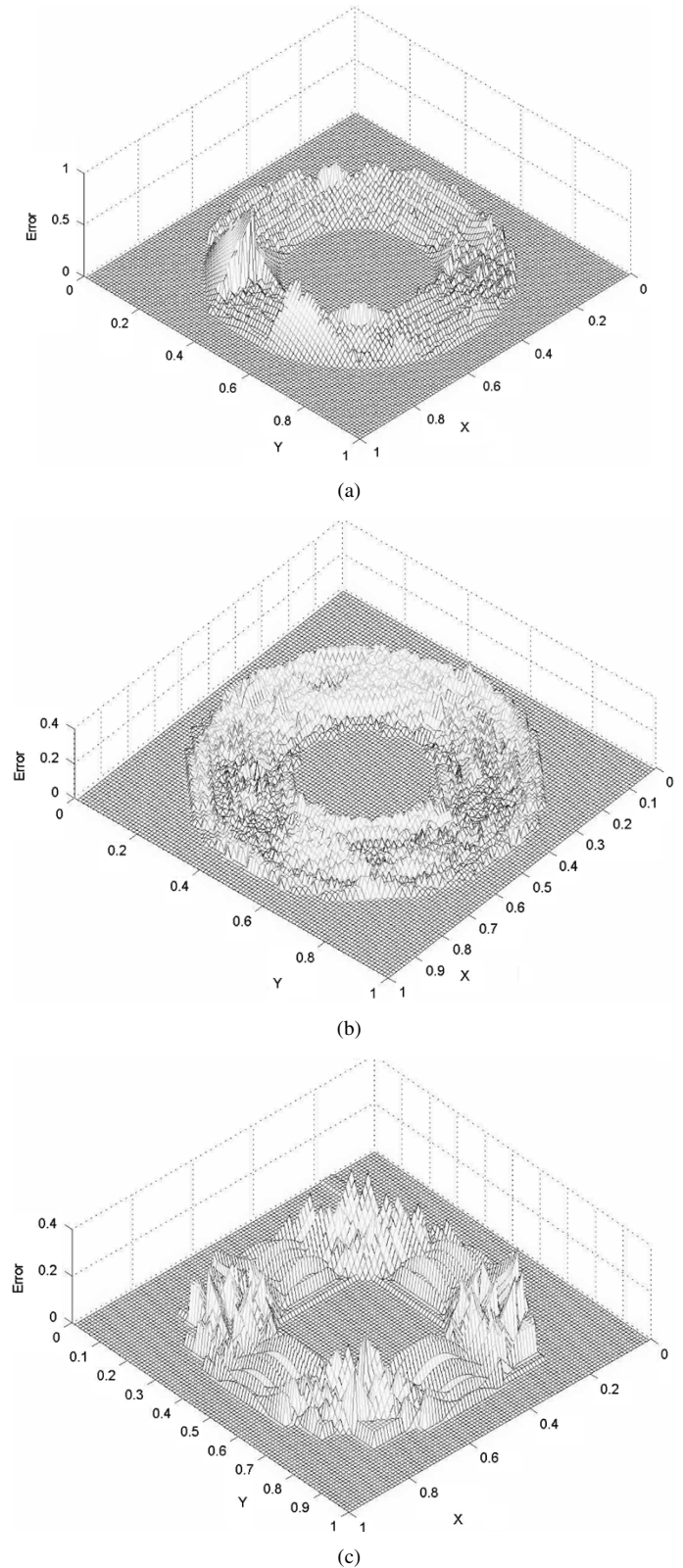


Fig. 3. Error distribution of test data (X and Y axis are the two dimensions in feature space, and Z axis is the absolute difference between real and estimated class proportion). (a) ARTMAP. (b) ART-MMAP. (c) Regression tree.

information). This is why we observe several horizontal lines in the BDF of ARTMAP. Each horizontal line represents all the samples predicted to one category of ART<sub>b</sub>. The regression tree BDF shows results similar to ARTMAP. With the enhanced

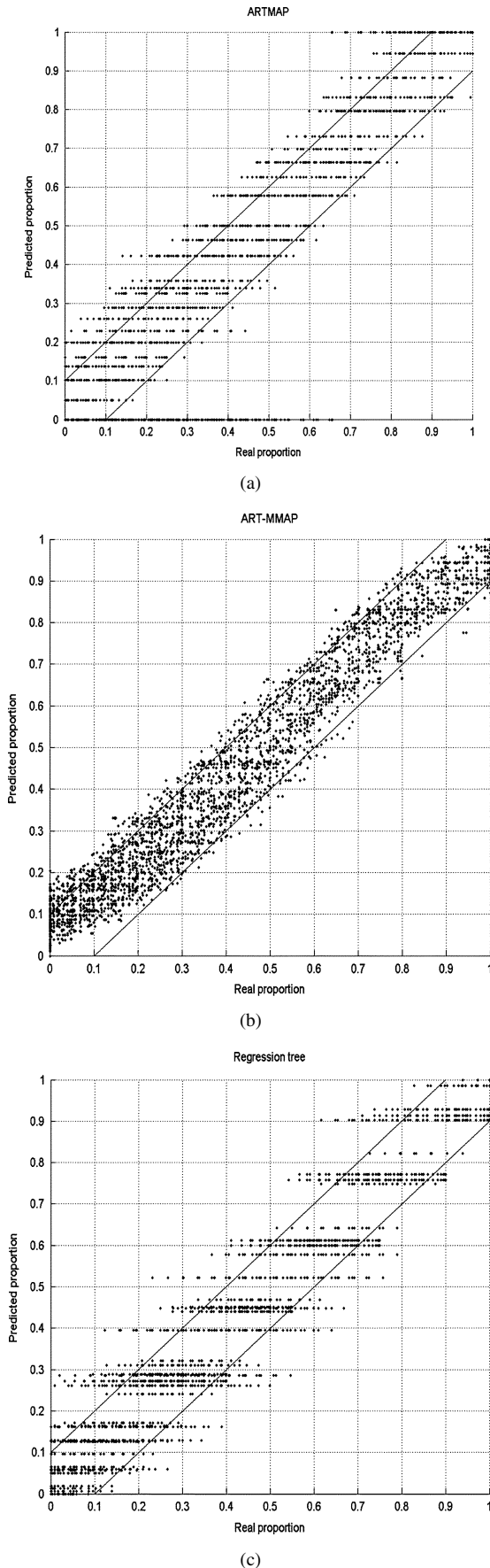


Fig. 4. BDFs of simulated test data. (a) ARTMAP. (b) ART-MMAP. (c) Regression tree.

interpolation function, the prediction of ART-MMAP overcomes the class category limitation of ARTMAP and is more accurate.

Table I shows the Z-statistics result of simulated dataset. The prediction accuracy of ART-MMAP model is significantly better than that for ARTMAP ( $Z = -2.75, p < 0.01$ ) and the regression tree ( $Z = -14.95, p < 0.001$ ).

### B. MODIS Dataset

With this dataset, the parameters of ART-MMAP are set as  $\rho_a = 0, \rho_b = 0.9$ , and  $\tau = 0.9$ . In Fig. 5, ART-MMAP has more points than ARTMAP and the regression tree drawn between the two 20% difference lines. Similar to Fig. 4(a), there are many horizontal lines in Fig. 5(a) presenting the class category limitation of ARTMAP. The prediction of ART-MMAP shown in Fig. 5(b) produces more different kinds of class combinations to match the test data. The comparison between the horizontal lines shown in ARTMAP and even distribution of all points in ART-MMAP provides another good example for ART-MMAP's interpolation function.

Table II shows the rms error of MODIS testing data using ARTMAP, ART-MMAP, and the regression tree models. There are fewer rms errors with ART-MMAP for three out of four classes (*barren*, *grass*, and *forest*) indicating its better performance than ARTMAP and the regression tree model. The performance of ARTMAP is similar to that of the regression tree for these three classes. Both ARTMAP and ART-MMAP are superior to the regression tree in terms of predicting *water* class proportion.

The Z-statistics result of prediction accuracy comparison between ART-MMAP and ARTMAP, ART-MMAP, and the regression tree is shown in Table III. In terms of individual class, ART-MMAP 1) has significant higher prediction accuracy than ARTMAP with class *grass* ( $Z = -6.597, p < 0.01$ ) and class *forest* ( $Z = -4.355, p < 0.01$ ), 2) has better accuracy with class *barren* ( $Z = -2.137, p < 0.05$ ), and 3) has similar accuracy to ARTMAP with class *water* ( $Z = 1.166, p < 0.5$ ). Compared to the regression tree model, the prediction accuracy of ART-MMAP is significantly higher with class *grass* ( $Z = -6.597, p < 0.01$ ) and *forest* ( $Z = -4.355, p > 0.001$ ). There is no significant difference between ART-MMAP and a regression tree for class *barren* and *water*.

## VII. IMPLEMENTATION

To implement ART-MMAP, several parameters need to be considered. For the vigilance parameter of both  $ART_a$  ( $\rho_a$ ) and  $ART_b$  ( $\rho_b$ ), the general rule is that the bigger the value, the more categories ARTMAP generates. For simulated dataset, we set  $\rho_a = 0.7, \rho_b = 0.98$  and  $\rho_a = 0, \rho_b = 0.9$  for the MODIS data and it produces good result. More details about setting parameters of ARTMAP model was discussed in the previous literature [5], [20]–[25].

The new parameter introduced in ART-MMAP is the threshold value ( $\tau \in [0, 1]$ ). Tables IV and V show the rms error with different threshold value for simulated and MODIS dataset respectively. With MODIS dataset, ART-MMAP produces the least rms error when threshold value is set as 0.90,

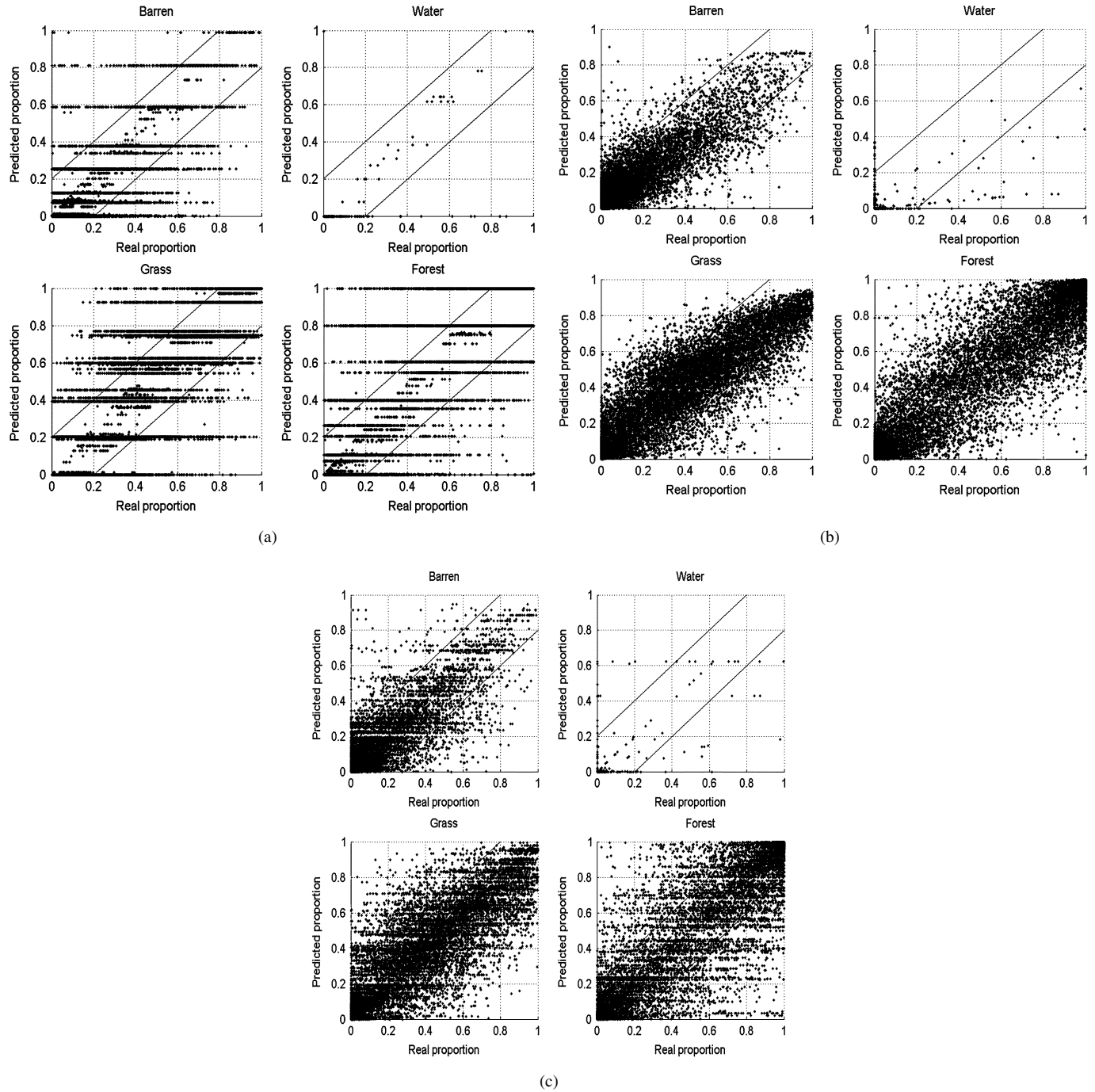


Fig. 5. BDFs between real and estimated subpixel land cover types of test data. (a) ARTMAP. (b) ART-MMAP. (c) Regression tree.

TABLE I  
COMPARISON BETWEEN ART-MMAP, ARTMAP, AND THE  
REGRESSION TREE WITH SIMULATED DATA

Model comparison	Z value	P value
ART-MMAP vs. ARTMAP	-2.75	0.0029
ART-MMAP vs. RT	-14.95	0.0001

TABLE II  
RMS ERROR OF TEST DATA

Mixture Models	<i>barren</i>	<i>water</i>	<i>grass</i>	<i>forest</i>
ART-MMAP model	<b>0.095</b>	<b>0.028</b>	<b>0.140</b>	<b>0.145</b>
ARTMAP model	0.111	0.028	0.163	0.169
Regression tree	0.106	0.040	0.161	0.174

but generates the best result with threshold value set as 0.97 for simulated data. With a higher threshold value, more samples will find fewer winning categories, and even no category will be selected, which will trigger WTA rule. Thus, interpolation function does not affect these samples and the prediction accuracy will be similar to that of ARTMAP. With a very low threshold value, the sample will select too many categories, which will also cause prediction error. Since MODIS dataset has more noise than simulated dataset, and thus the activation value is not very high, the threshold value is set lower for MODIS dataset. With real-world satellite images, we suggest that the threshold value be set up around [0.85, 0.95].

TABLE III  
COMPARISON BETWEEN ART-MMAP, ARTMAP, AND THE REGRESSION TREE WITH MODIS DATA

Model comparison	<i>barren</i>		<i>water</i>		<i>grass</i>		<i>forest</i>	
	Z value	P value	Z value	P value	Z value	P value	Z value	P value
ART-MMAP vs. ARTMAP	-2.137	0.016	1.166	0.122	-6.597	0.0001	-4.355	0.0001
ART-MMAP vs. RT	-0.686	0.246	-1.081	0.140	-2.401	0.008	-2.748	0.003

TABLE IV  
RMS ERROR WITH DIFFERENT THRESHOLD VALUES—SIMULATED DATASET

Threshold value	RMS error
0.90	0.176
0.93	0.103
0.95	0.060
<b>0.97</b>	<b>0.031</b>
0.98	0.032
ARTMAP	0.060
Regression tree	0.050

TABLE V  
RMS ERROR WITH DIFFERENT THRESHOLD VALUES—MODIS DATASET

Threshold value	RMS error			
	<i>barren</i>	<i>water</i>	<i>grass</i>	<i>forest</i>
0.70	0.134	0.031	0.234	0.241
0.80	0.124	0.031	0.195	0.187
0.85	0.119	0.030	0.178	0.169
<b>0.90</b>	<b>0.095</b>	<b>0.028</b>	<b>0.140</b>	<b>0.145</b>
0.95	0.097	0.030	0.141	0.147
0.98	0.110	0.029	0.160	0.166
ARTMAP	0.111	0.028	0.163	0.169
Regression tree	0.106	0.040	0.161	0.174

## VIII. CONCLUSION

Moderate- and coarse-resolution satellite images are widely being used for continental and global land cover mapping. Since the spatial resolution of these satellites is much coarser than the scale of ecological land cover features, many of the image pixels may be mixed. Mixture analysis with linear and nonlinear algorithms has been widely used to estimate land cover type proportions at subpixel levels.

Some nonlinear mixture analysis models, i.e., artificial neural networks, have shown improved performance compared to traditional linear models. The ART-MMAP mixture analysis model presented here is based on one of the most successful artificial neural networks, ARTMAP. Previous research has proved that the ARTMAP mixture model outperformed the linear mixture model in terms of subpixel-level interpretation [4], [5]. The goal of ART-MMAP is to decrease the effect of feature category proliferation in ART<sub>a</sub> and overcome the limitation of class category in ART<sub>b</sub> during the testing processing for mixture analysis. The ART-MMAP model keeps the learning process of ARTMAP and changes the testing and recognition processes. Instead of picking one winner, ART-MMAP selects the feature categories (ART<sub>a</sub>) based on the threshold parameter. With an enhanced interpolating prediction function that is based on a weighted summation operator, ART-MMAP overcomes the limitation of class category of ARTMAP and increases the prediction accuracy as well. Test results based on two datasets show the improvement of ART-MMAP over ARTMAP and the regression tree model.

## ACKNOWLEDGMENT

The authors would thank the five anonymous reviewers for the helpful comments.

## REFERENCES

- [1] J. Adams, D. Sabol, and V. Kapos, "Classification of multispectral images based on fraction endmembers, application to land cover change in the Brazilian Amazon," *Remote Sens. Environ.*, vol. 52, pp. 137–154, 1995.
- [2] D. Roberts, M. Gardner, R. Church, S. Ustin, G. Scheer, and R. Green, "Mapping chaparral in the Santa Monica mountains using multiple end-member spectral mixture models," *Remote Sens. Environ.*, vol. 65, pp. 267–279, 1998.
- [3] D. Roberts, M. Smith, and J. Adams, "Green vegetation, nonphotosynthetic vegetation, and soils in AVIRIS data," *Remote Sens. Environ.*, vol. 44, pp. 255–269, 1993.
- [4] W. Liu, S. Gopal, and C. Woodcock, "ARTMAP multisensor/resolution framework for land cover characterization," in *Proc. 4th Int. Conf. Information Fusion*, Montreal, QC, Canada, Aug. 7–10, 2001, pp. WeC2-11–WeC2-16.
- [5] G. Carpenter, S. Gopal, M. S. Macomber, and C. Woodcock, "A neural network method for mixture estimation for vegetation mapping," *Remote Sens. Environ.*, vol. 70, pp. 138–152, 1999.
- [6] G. Foody, "Sharpening fuzzy classification output to refine the representation of sub-pixel land cover distribution," *Int. J. Remote Sens.*, vol. 19, pp. 2593–2599, 1998.
- [7] P. Atkinson, M. Cutler, and H. Lewis, "Mapping sub-pixel proportional land cover with AVHRR imagery," *Int. J. Remote Sens.*, vol. 18, no. 4, pp. 917–935, 1997.
- [8] G. Foody, R. Lucas, P. Curran, and M. Honzak, "Non-linear mixture modeling without end-members using an artificial neural network," *Int. J. Remote Sens.*, vol. 18, no. 4, pp. 937–953, 1997.
- [9] G. Foody, "Approaches for the production and evaluation of fuzzy land cover classifications from remotely-sensed data," *Int. J. Remote Sens.*, vol. 17, no. 7, pp. 1317–1340, 1996.
- [10] R. Defries, J. Townshend, and M. Hansen, "Continuous fields of vegetation characteristics at the global scale at 1-km resolution," *J. Geophys. Res.*, vol. 104, pp. 16 911–16 923, 1999.
- [11] D. McIver, "Adapting machine learning approaches for coarser resolution land cover classification," Ph.D. dissertation, Boston Univ., Boston, MA, 2001.
- [12] D. McIver and M. Friedl, "Beyond accuracy: Advantages of boosting in the classification of satellite remote sensing data for earth science applications," in *Proc. Workshop Mining Scientific Datasets, 1st Society for Industrial and Applied Mathematics Conf. Data Mining*, Chicago, IL, pp. 37–44.
- [13] J. Ju, E. D. Kolaczyk, and S. Gopal, "Gaussian mixture discriminant analysis and sub-pixel land cover characterization in remote sensing," *Remote Sens. Environ.*, vol. 84, pp. 550–560, 2003.
- [14] G. Foody, N. Campbell, N. Trodd, and T. Wood, "Derivation and applications of probabilistic measures of class membership from maximum-likelihood classification," *Photogram. Eng. Remote Sens.*, vol. 58, no. 9, pp. 1335–1341, 1992.
- [15] R. Schowengerdt, "On the estimation of spatial-spectral mixing with classifier likelihood functions," *Pattern Recognit. Lett.*, vol. 17, no. 13, pp. 1379–1387, 1996.
- [16] C. Borel and S. Gerstl, "Nonlinear spectral mixing models for vegetative and soil surfaces," *Remote Sens. Environ.*, vol. 47, pp. 403–416, 1994.
- [17] P. Sellers, Y. Mintx, Y. Sud, and A. Dalcher, "A simple biosphere model (SiB) for use within general circulation models," *J. Atmos. Sci.*, vol. 43, no. 6, pp. 505–531, 1986.
- [18] Y. Tian, Y. Knyazikhin, R. Myneni, J. Glassy, G. Dedieu, and S. Running, "Prototyping of MODIS LAI and FPAR algorithms with LASUR and Landsat data," *IEEE Trans. Geosci. Remote Sensing*, vol. 38, pp. 2387–2401, Sept. 2000.

- [19] Y. Knyazikhin, J. Martonchik, D. Myneni, M. Verstraeter, B. Pinty, and N. Gobron, "Estimation of vegetation canopy leaf area index and fraction of absorbed photosynthetically active radiation from atmosphere-corrected MISR data," *J. Geophys. Res.*, vol. 103, pp. 32 239–32 256, 1998.
- [20] G. A. Carpenter, S. Grossberg, and J. H. Reynolds, "ARTMAP: Supervised real-time learning and classification of nonstationary data by a self-organizing neural network," *Neural Networks*, vol. 4, pp. 565–588, 1991.
- [21] G. A. Carpenter, S. Grossberg, and D. B. Rosen, "Fuzzy ART: Fast stable learning and categorization of analog patterns by an adaptive resonance system," *Neural Networks*, vol. 4, pp. 759–771, 1991.
- [22] G. A. Carpenter, S. Grossberg, S. Markuzon, N. Martens, J. Reynolds, and D. Rosen, "Fuzzy ARTMAP: A neural network architecture for incremental supervised learning of analog multidimensional maps," *IEEE Trans. Neural Networks*, vol. 3, pp. 698–713, Sept. 1992.
- [23] G. A. Carpenter and S. Grossberg, *Pattern Recognition by Self-Organizing Neural Networks*. Cambridge, MA: MIT Press, 1991.
- [24] S. Gopal, W. Liu, and C. Woodcock, "Visualization based on the fuzzy ARTMAP neural network for mining remotely sensed data," in *Geographic Data Mining and Knowledge Discovery*, H. J. Miller and J. Han, Eds. London, U.K.: Taylor & Francis, 2001, pp. 315–336.
- [25] W. Liu, S. Gopal, and C. Woodcock *et al.*, "Spatial data mining for classification, visualization and interpretation with ARTMAP neural network," in *Data Mining for Scientific and Engineering Applications*, R. Grossman *et al.*, Eds. Amsterdam, Netherlands: Kluwer, 2001, pp. 205–222.
- [26] J. R. Williamson, "Gaussian ARTMAP: A neural network for fast incremental learning of noisy multidimensional maps," *Neural Networks*, vol. 9, pp. 881–897, 1996.
- [27] G. A. Carpenter, B. L. Milenova, and B. W. Noeske, "Distributed ARTMAP: A neural network for fast distributed supervised learning," *Neural Networks*, vol. 11, pp. 793–813, 1998.
- [28] E. Gomez-Sanchez, Y. A. Dimitriadis, J. M. Cano-Izquierdo, and J. L. Lopez-Coronado, " $\mu$  ARTMAP: Use of mutual information for category reduction in fuzzy ARTMAP," *IEEE Trans. Neural Networks*, vol. 13, pp. 58–69, Jan. 2002.
- [29] A. Baraldi and E. Alpaydin, "Constructive feedforward ART clustering networks—Part I and part II," *IEEE Trans. Neural Networks*, vol. 13, pp. 645–677, May 2002.
- [30] K. C. Seto and W. Liu, "Comparing ARTMAP neural network with maximum-likelihood classifier for detecting urban change," *Photogramm. Eng. Remote Sens.*, vol. 69, no. 9, pp. 981–990, 2003.



**Karen C. Seto** received the B.A. degree from the University of California, Santa Barbara, and the M.A. and Ph.D. degrees from Boston University, Boston, MA.

Since 2000, she has taught at Stanford University, Stanford, CA, where she is currently an Assistant Professor in the Department of Geological and Environmental Sciences, School of Earth Sciences, and a Center Fellow in the Center for Environmental Science and Policy, Stanford Institute for International Studies. Her research interests include remote sensing of urban areas, modeling land use and land cover change, and the human dimensions of global environmental change.



**Elaine Y. Wu** received the M.S. degree in information systems from Northeastern University, Boston, MA, in 2001.

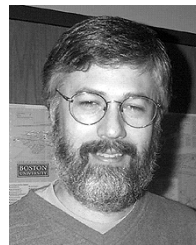
She is currently a Research Analyst with the Medical University of South Carolina, Charleston.

**Sucharita Gopal**, photograph and biography not available at the time of publication.



**Weiguo Liu** received the Ph.D. degree from Boston University, Boston, MA.

Since 2001, he has been a Senior Research Engineer with ACI Worldwide, Incorporated, Riverside, RI. His research interests include data mining, neural networks, financial fraud detection, mixture analysis, geocomputation, and remote sensing.



**Curtis E. Woodcock** received the B.A., M.A., and Ph.D. degrees from the University of California, Santa Barbara, all in geography.

He is currently a Professor of geography at Boston University, Boston, MA. His research interests are in the area of remote sensing of vegetation and monitoring of environmental change.