# Automatic Classification of Land Cover on Smith Island, VA, Using HyMAP Imagery

Charles M. Bachmann, *Member, IEEE*, Timothy F. Donato, *Member, IEEE*, Gia M. Lamela, W. Joseph Rhea, Michael H. Bettenhausen, *Member, IEEE*, Robert A. Fusina, *Member, IEEE*, Kevin R. Du Bois, John H. Porter, and Barry R. Truitt

***Abstract*—Automatic land cover classification maps were developed from Airborne Hyperspectral Scanner (HyMAP) imagery acquired May 8, 2000 over Smith Island, VA, a barrier island in the Virginia Coast Reserve. Both unsupervised and supervised classification approaches were used to create these products to evaluate relative merits and to develop models that would be useful to natural resource managers at higher spatial resolution than has been available previously. Ground surveys made by us in late October and early December 2000 and again in May, August, and October 2001 and May 2002 provided ground truth data for 20 land cover types. Locations of pure land cover types recorded with global positioning system (GPS) data from these surveys were used to extract spectral end-members for training and testing supervised land cover classification models. Unsupervised exploratory models were also developed using spatial–spectral windows and projection pursuit (PP), a class of algorithms suitable for extracting multimodal views of the data. PP projections were clustered by ISODATA to produce an unsupervised classification. Supervised models, which relied on the GPS data, used only spectral inputs because for some categories in particular areas, labeled data consisted of isolated single-pixel waypoints. Both approaches to the classification problem produced consistent results for some categories such as *Spartina alterniflora*, although there were differences for other categories. Initial models for supervised classification based on 112 HyMAP spectra, labeled in ground surveys, obtained reasonably consistent results for many of the dominant categories, with a few exceptions. For an invasive plant species, *Phragmites australis*, a particular concern of natural resource managers, this approach initially had an excessively high false-alarm rate. Increasing the number of spectral training samples by an order of magnitude and making concomitant improvements to the georectification led to dramatic improvements in this and other categories. The unsupervised spatial–spectral approach also found a cluster closely associated with *Phragmites* patches near the thicket boundary, but this approach did not identify the exposed *Phragmites*. Examples of *in situ* reflectance measurements obtained with an Analytical Spectral Devices FR spectrometer in early May 2001 are compared against HyMAP image spectra at model-predicted pixels and at validated GPS waypoints.**

***Index Terms*—Barrier islands, hyperspectral, in situ spectrometry, invasive plant species, land cover classification, neural networks, principle component analysis, projection pursuit, supervised classification, unsupervised classification.**

## I. Introduction: The Virginia Coast Reserve

A HYMAP [1], [31] scene of Smith Island, VA, acquired on May 8, 2000, served as the basis of the present study (Fig. 1). Smith Island is one of a series of barrier islands in the Virginia Coast Reserve (VCR) and the site of the University of Virginia's ongoing Long Term Ecological Research (LTER) program [32], [38]. The most extensive survey of the island dates from 1974 [32], [35] and was based on ground observations and interpretation of false-color infrared imagery for a set of 16 barrier islands that encompass the VCR. This historical reference data consisted of 26 land cover types. To develop our automatic land cover classification models, we chose a somewhat different approach, attempting to achieve species-level classification in many instances, while considering in some cases plant communities that were similar to those described in [35]. Our land cover classification models consisted of 16 to 19 categories. However, for purposes of this introduction, we have grouped the land cover into five or six principal categories, some of which equate to those described in [35], while others are aggregates of several of these categories. New definitions for coastal vegetation are presently under development by the state of Virginia [18].

Particularly in wetlands research and coastal applications, past emphasis has been on either 1) broad-band sensors such as Landsat TM [16], [28] or 2) hyperspectral sensors at lower spatial resolution [22]. For modeling regional scales with the former, the National Oceanic and Atmospheric Administration's C-CAP protocol has been widely used. A cornerstone of this has been the cluster-busting algorithm developed by Jensen [27], which is a labor-intensive, though highly accurate, approach. Likewise, in coastal applications, the dominant approach for hyperspectral modeling has been spectral linear mixing models (e.g., see [22]) applied to AVIRIS imagery at resolutions $\approx$20 m pixels. Other analyses have considered the use of vegetation indices for extracting biophysical parameters, comparing both Landsat TM and AVIRIS with *in situ* spectrometry measurements [48]. One of our goals in using a hyperspectral sensor with a higher spatial resolution of 4.5 m

Manuscript received October 25, 2001; revised August 11, 2002. This was work was supported in part by the Office of Naval Research under Contract N0001400WX40016, N0001401WX40009, N0001402WX30017, N0001401WX40009, and the National Science Foundation under Grant DEB-0080381.

C. M. Bachmann, T. F. Donato, G. M. Lamela, W. J. Rhea, and R. A. Fusina are with the Remote Sensing Division, Naval Research Laboratory, Washington, DC 20375 USA (e-mail: bachmann@nrl.navy.mil).

K. R. Du Bois is with the Bureau of Environmental Services, City of Norfolk, Norfolk, VA 23510 USA (e-mail: kdubois@city.norfolk.va.us).

M. H. Bettenhausen is with the Integrated Management Systems, Inc., Arlington, VA 22201 USA.

J. H. Porter is with the Department of Environmental Sciences, University of Virginia, Charlottesville, VA 22904-41231 USA (e-mail: jhp7e@virginia.edu; www.vcrlter.virginia.edu).

B. R. Truitt is with The Nature Conservancy, Virginia Coast Reserve, Nassawadox, VA 23413 USA (e-mail: btruitt@tnc.org).

Digital Object Identifier 10.1109/TGRS.2002.804834

# Report Documentation Page

*Form Approved*
*OMB No. 0704-0188*



| 1. REPORT DATE | 2. REPORT TYPE | 3. DATES COVERED |
|---|---|---|
| **2002** | | **00-00-2002 to 00-00-2002** |

| 4. TITLE AND SUBTITLE | 5a. CONTRACT NUMBER |
|---|---|
| **Automatic Classification of Land Cover on Smith Island, VA, Using HyMAP Imagery** | 5b. GRANT NUMBER |
| | 5c. PROGRAM ELEMENT NUMBER |
| 6. AUTHOR(S) | 5d. PROJECT NUMBER |
| | 5e. TASK NUMBER |
| | 5f. WORK UNIT NUMBER |
| 7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Naval Research Laboratory,Code 7213,4555 Overlook Avenue, SW,Washington,DC,20375** | 8. PERFORMING ORGANIZATION REPORT NUMBER |
| 9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) | 10. SPONSOR/MONITOR'S ACRONYM(S) |
| | 11. SPONSOR/MONITOR'S REPORT NUMBER(S) |

12. DISTRIBUTION/AVAILABILITY STATEMENT
**Approved for public release; distribution unlimited**

13. SUPPLEMENTARY NOTES
**The original document contains color images.**

14. ABSTRACT

15. SUBJECT TERMS

| 16. SECURITY CLASSIFICATION OF: | | | 17. LIMITATION OF ABSTRACT | 18. NUMBER OF PAGES | 19a. NAME OF RESPONSIBLE PERSON |
|---|---|---|---|---|---|
| a. REPORT | b. ABSTRACT | c. THIS PAGE | | | |
| **unclassified** | **unclassified** | **unclassified** | | **18** | |

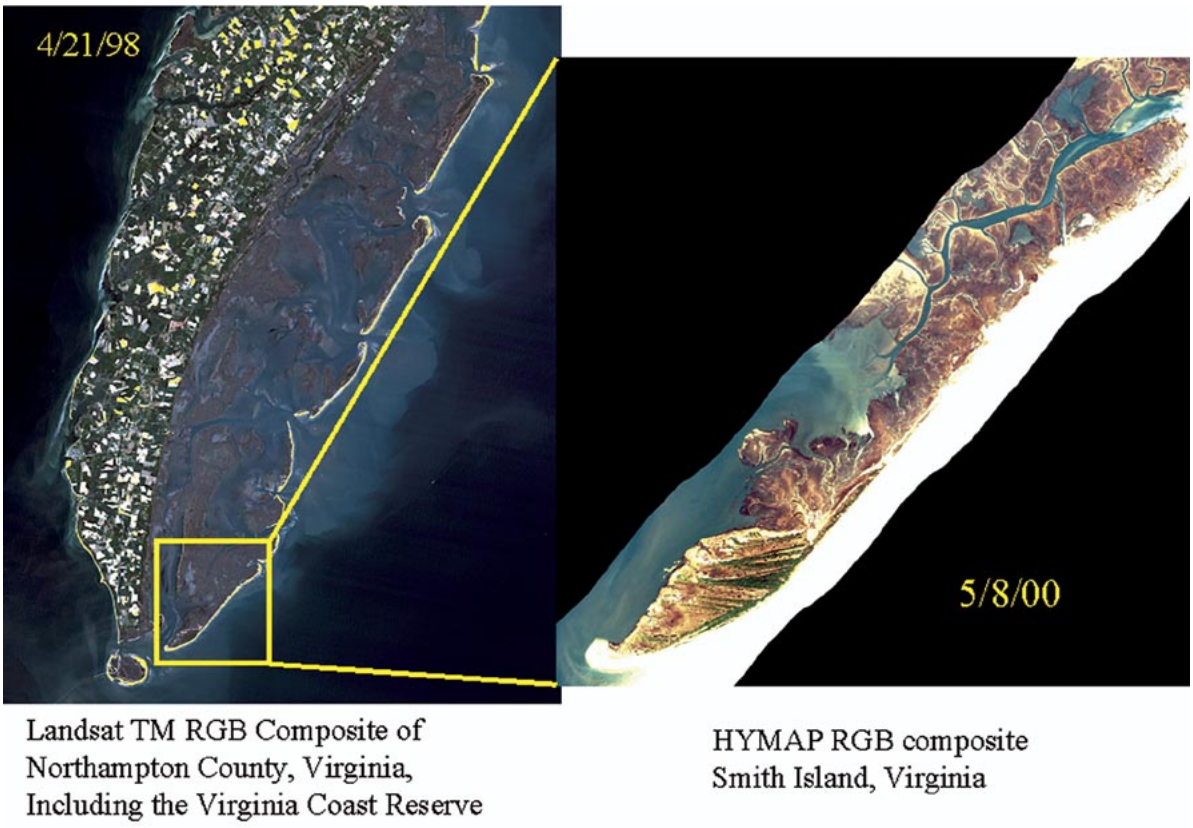



Fig. 1. (Left) RGB composite of the red, green, and blue channels from a Landsat Thematic Mapper (TM) image taken in April, 1998 of Northampton County, VA, showing a subset of the islands known as the Virginia Coast Reserve. Smith Island is highlighted in the box. (Right) RGB composite from 126-channel HyMAP imagery of Smith Island, VA and the southern portion of Myrtle Island, acquired May 8, 2000. A portion of Myrtle Island has been omitted.

was to be able to discriminate rapidly varying land cover types seen, for example, in the transition zone from the lagoonal shore to the upland. On Smith Island, six to seven distinct vegetation zones may occur in a distance as short as 50–75 m. Although we do not explore mixture models in the present study, they will be compared with the methods presented here in a future publication.

The spatial distribution of land cover types included in our models varied considerably. Categories such as *Myrica cerifera* (bayberry) thicket occur only in the southern end of Smith Island, striating the island in dense bands of vegetation. These thickets are typically tens of meters in width and can extend in some instances nearly the width of the island (about 2 km). Categories such as these, therefore, whose spatial extent is frequently greater than the resolution cell of the sensor (the H-resolution case described in [42]) are amenable to modeling that uses supervised classification, at least in the final stages of processing. In contrast, in other areas some vegetation categories have a spatial extent that is of the order of a pixel or less (the L-resolution case [42]). In some cases, the spatial extent in one dimension may be of the order of a pixel or less in one dimension, while having a length of several pixels or more in the other dimension. The latter occurs in some instances for the invasive plant species *Phragmites australis* in the southern end of Smith Island (not all stands are so narrow; the width varies considerably). In this part of the island, *Phragmites australis*, where it occurs, typically forms a narrow band of vegetation in the ecotone between the upland thicket and brackish and fresh water marshes in the swale immediately adjacent. In most cases for our validation testing, the width of the stands that we considered was at least a couple of pixels, although it is not uncommon to find stands whose width (extent perpendicular to the thicket line) is on the order of a pixel. In a sense, this is the ideal candidate for L-resolution methods, which assume that in at least one dimension the category extent may be a pixel or less, but the spectral pattern associated with the category is repeated in some regular spatial distribution that can be detected. This is our motivation for also considering spatial–spectral models that use unsupervised feature extraction and classification based on projection pursuit and pincipal component analysis.

As just mentioned, some vegetation communities have spatial extents that may be only a few pixels at the HyMAP spatial resolution of 4.5 m, so this resolution forms an upper bound on the ideal spatial resolution. The utility of land cover classification models is, of course, determined by the end-user [36], [37]. Beyond the narrow goal of achieving ecological modeling at high resolution, there are practical reasons for why using this kind of data will benefit natural resource managers. For example, as just described, the invasive plant species *Phragmites australis* may exist in patches whose spatial extent may be on the order of the pixel size of HyMAP in at least one dimension. Likewise, because it has spectral characteristics that are similar to other wetland plants, it is unlikely that systems with a few broad spectral channels would be able to discriminate it, especially when it occurs in close proximity, as it often does on Smith Island, to other vegetation types such as the *Myrica cerifera* thicket. Al-

though there is some debate as to how problematic *Phragmites* is [39], many natural resource managers agree that it supplants other wetland types, disrupting ecosystem balance, and *Phragmites* control and eradication programs are not uncommon.

Even within a single category, variations in spatial extent occur. For instance, one of the primary constituents of the low marsh vegetation (Fig. 2), *Spartina alterniflora* (Smooth Cordgrass), occurs in large monotypic stands in the northern end of Smith Island, while at the southern end of the island, it occurs in one or two narrow bands of vegetation at the water's edge on the lagoonal (western) shore, and in small zones in the brackish swales that cross the island.

High marsh species (Fig. 2) include *Salicornia virginica* (Perrenial Glasswort), *Limonium carolinianum* (sea lavender), *Borrichia frutescens* (Sea Ox-eye), *Iva frutescens* (Marsh-elder), *Sueada linearis and Sueada maritima* (Sea-blite), and *Spartina patens* (Salt-Hay or Saltmeadow Cordgrass).[1] The upper end of the high marsh frequently has a zone of "wrack," the dead, matted detritus of the previous year's growth, which typically marks the mean high-water line associated with tidal influences. The swales (Fig. 2) that cross the southern end of Smith Island contain brackish and fresh-water marshes. Swale vegetation includes *Distichlis spicata* (Saltgrass), *Spartina patens*, *Juncus roemerianus* (Needle Rush), *Scirpus robustus* (Saltmarsh Bulrush), and *Iva frutescens*.

Narrow upland zones (Fig. 2) alternate with swales across the southern end of the island. Here the typical vegetation consists of shrubs such as *Myrica cerifera* (Bayberry), the dominant vegetation, and *Baccharis halimifolia* (Groundsel-tree), with attendant vegetation such as *Smilax spp.* (Greenbriar). Stands of hardwoods and Pine, such as *Pinus taeada* (Loblolly pine), also occur in some of the upland zones. In these areas, it is common to find shrubs such as *Myrica cerifera* in the understory.

Flats (Fig. 2) appear throughout the island. These consist of mudflats, wash flats, and salt flats or salt pannes. Wash flats result, for example, from sudden storm surge events in which the dune line is breached. Salt pannes occur in places where water floods an area and evaporates, leaving behind a significant amount of salt. The high salinity tends to kill off most vegetation, and typically only the most salt-tolerant plants such as *Salicornia virginica* will survive in small clumps; wash flats are often predecessors of salt pannes [35].

The beach zone (Fig. 2) is highly variable. In the northern end of Smith Island, exposed peat outcrops are present in the surf zone. These are the decomposed residue of what was once salt marsh, and they serve as a reminder that the island is undergoing constant change. In the foredune zone (also Fig. 2), "wrack" is frequently found, and in summer, a low band of herbaceous vegetation, comprised principally of plants such as *Cakile edentula* (Sea Rocket) and *Salsola kali* (Russian Thistle). The dune line (Fig. 2) typically is comprised of plant species such as *Ammophila breviligulata* (American beachgrass), *Uniola paniculata* (Sea oats), *Salidago sempervirens* (Salt Marsh Goldenrod), and in some cases *Panicum amarum* (Seaside Panicum). The back dune is dominated by vegetation such as *Spartina patens*, *Ammophila breviligulata*, and *Andropogon spp.* (Broomsedge family).

[1]Common names of coastal vegetation may vary somewhat from author to author as do definitions of species names listed in italics; in this paper, we have used [15] and [43].

## II. HyMAP Data for Smith Island, VA

The HyMAP imagery was atmospherically corrected using ATREM/EFFORT [13] by Analytical Imaging and Geophysics LLC (AIG) prior to delivery. The Smith Island scene was acquired at 4.5-m resolution with 128 spectral channels; the final EFFORT product was the surface reflectance contained in 126 spectral channels ranging from 445–2486 nm. The image was acquired near high tide, so there is a significant degree of inundation in the wetlands, especially in the salt marsh. For the purposes of automated model development, we preprocessed the data on a per-sample basis in a number of different ways (Fig. 3). Since the data points labeled in the global positioning system (GPS) surveys consisted of a mix of both isolated points and areas, the supervised automatic classification models used only the single pixel spectrum as input, while the unsupervised models did not need to satisfy this constraint and, therefore, could ingest both single-pixel spectra and spatial–spectral windows.

## III. Field Observations: Geolocated Spectra and *In Situ* Spectrometry

We compared the unsupervised and supervised automatic classification category maps against *in situ* observations made during two days of field observations and GPS surveys conducted in October and December 2000, a week of surveys carried out with GPS between May 7–11, 2001, two weeks of differential GPS (DGPS) surveys conducted during August 20–23 and October 8–12 2001, and again between May 3–5 and May 13–15, 2002. During these trips, typical vegetation categories were identified, and positions were recorded using a GPS or DGPS. These same waypoints were also used to generate supervised classification maps. During the May 7–11, 2001 and May 13–15, 2002 field trips, we also measured *in situ* reflectance with an Analytical Spectral Devices (ASD) FR spectrometer, which covers a spectral range similar to that of HyMAP. Our DGPS survey equipment consisted of a Trimble Geoexplorer 3 and Beacon-on-a-Belt. During these weeks, we also surveyed four other islands to the extent that time permitted, taking data on Hog, Cobb, Wreck, and Myrtle, in addition to Smith. Equipment problems prevented additional spectral measurements; however, two weeks prior to the August 2001 survey at the VCR, we were able to acquire ASD measurements at another site in southern New Jersey. During the August field trip and one week after the October field trip, airborne hyperspectral data were acquired by PROBE2 [17] for all six of the VCR islands in our study area for comparison against our May 2000 HyMAP data. These PROBE2 data will be the subject of future papers. This is motivated by our desire to understand the effect that seasonal changes in the land cover have on spectral characteristics. As described below, models that were produced for the spring HyMAP data may not necessarily apply to data taken in the summer or fall. Likewise, tidal influences can have a significant impact on marsh vegetation and their associated spectra because of the degree of inundation present.


Fig. 2. Typical Smith Island land cover. (First row, left) Upland zone: *Myrica cerifera* thickets and some stands of hardwood and *Pinus taeda* (loblolly pine). (First row, middle) Typical mudflat near salt marsh edge. (First row, right) Peat outcrop in surf zone. (Second row, left) Foredune vegetation: primarily *Cakile edentula* (Sea Rocket) (inset) and *Salsola kali* (Russian Thistle). (Second row, middle) Dune vegetation and nearby backdune: primarily *Ammophila breviligulata* (American Beachgrass) (foreground), and ocassionally *Uniola paniculata* (Sea oats) (background). (Second row, right) Inland portion of backdune: predominantly *Andropogon spp.* (Broomsedge family). (Third row, left) *Spartina alterniflora* (Smooth Cordgrass), dominant in low marsh; (third row, right) *Borrichia frutescens*, typical high marsh plant; (third row, middle) "wrack." Brackish marsh dwellers: (fourth row, left) *Juncus roemerianus* (Needle Rush), (fourth row, middle and inset) *Scirpus robustus* (Saltmarsh Bulrush), and (fourth row, third column) *Phragmites australis*, an invasive plant species; (fourth row, fourth column) *Distichlis spicata* a dominant swale grass.

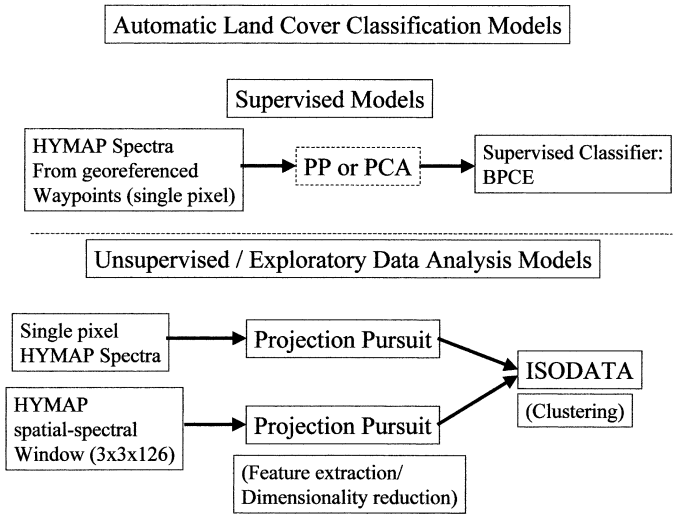



Fig. 3. Processing configurations for automated land cover classification models. (Top) Supervised models used georeferenced HyMAP spectra labeled during GPS and DGPS ground surveys. The supervised classifier was BPCE. Some models used PCA or PP for feature extraction/dimensionality reduction as a precursor to BPCE. (Bottom) Single-pixel and spatial–spectral windows were derived from a subset of the HyMAP data for the southern end of Smith Island. PP-filtered data were passed to ISODATA. For both unsupervised and supervised approaches, models were produced for the entire island.

TABLE I

| | |
|---|---|
| (1) *Phragmites australis* | (2) *Spartina alterniflora* |
| (3) *Spartina patens* | (4) *Salicornia virginica* |
| (5) *Borrichia frutescens* | (6) *Juncus roemerianus* |
| (7) Water | (8) *Distichlis spicata* |
| (9) *Scirpus spp.* | (10) "Wrack" |
| (11) Mudflat/saltflat | (12) *Ammophila breviligulata* |
| (13) Beach/sand | (14) *Uniola paniculata* |
| (15) Andropogon spp. | (16) *Myrica cerifera*-dominated Thicket |
| (17) Pine/hardwood complex | (18) Peat Outcrop |
| (19) *Iva frutescens* | (20) Foredune Vegetation |

We recorded the environment at many of the waypoints using digital still photographs and video. Based on these field observations, we initially defined a set of 16 categories, some of which appear or were aggregates of categories in Table I. We ended up using all but the foredune category in the final set as the basis of our supervised classification models (primarily because the foredune vegetation in early May will typically be nascent and sparse or completely absent). After the May and August 2001 surveys, two additional categories were added (Table I): Peat outcrop and *Scirpus robustus*, and we split two aggregate categories into their primary constituent plant species: backdune became *Andropogon spp.* and *Ammophila breviligulata*, and the thicket vegetation was separated into Pine/Hardwood complex and *Myrica cerifera*-dominated thicket. We created spectral libraries from individual HyMAP spectra extracted at the associated waypoint, or where appropriate, small regions of interest (ROIs) bounded by GPS waypoints. After the DGPS data were collected, points, lines, and areas were available with an accuracy estimated to be <1–5 m, similar to the spatial resolution of the May HyMAP data. These were used to train and test supervised automatic classification models more rigorously as described below. The DGPS ground data also were used to improve georectification of the imagery.

## IV. Methods

Both supervised and unsupervised classification models of the land cover were produced. In this section, we outline how the models were produced.

### A. Unsupervised Feature Extraction and Classification

Unsupervised feature extraction algorithms were used for two purposes in this study. In both cases, these fulfilled the role of dimensionality reduction as a precursor stage prior to the final classification algorithm, either unsupervised or supervised (Fig. 3). The two unsupervised feature extraction algorithms that we used were the projection pursuit (PP) algorithm described in [8] and the well-known principal component analysis (PCA) algorithm [47] that is popular in the remote sensing literature (e.g., see [23] and [44], and many others).

The underlying philosophies of PP and PCA are quite different. PCA uses the directions of maximal variance and derives an orthonormal set of basis vectors to identify significant structure in the data; these views of the data are not always easily interpreted with respect to specific underlying categories because of the orthogonality requirement [7], [8], [14]. Because PCA looks for directions of maximal variation in the data, it is incapable of detecting multimodal and other non-Gaussian departures that do not happen to be parallel, or nearly parallel, to the principal axes of the projected data distribution. In contrast, PP [7]–[10], [12], [20], [21], [46] uses higher order statistical information to overcome this difficulty and identify directions in which the projected data distribution (view) is non-Gaussian or multimodal.

Only within the last ten years has PP been applied in the field of remote sensing (see [2]–[5], [7], [8], [25], [29], and [30]) and in other disciplines (see [19], [26], [33], and [34]). The PP algorithm described in [8] (the PP algorithm used in this paper) is based on an algorithm originally proposed in [20]. However, in [8], projections are optimized simultaneously rather than in residual subspaces, as is sometimes the case in PP algorithms [21], [24], and projections are nonlinear, in order to remove sensitivity to outliers, rather than the linear form found in [20]. Although further details are provided in [8], the basic idea is that a cost function, emphasizing both intracluster spread and compactness within each cluster, is to be optimized. This function of the projected data distribution is the product of two functions, one measuring compactness of the data projection within a particular search scale and another measuring the spread of the data in that projection. The user defines a range of search scales, $\alpha_k$, that correspond to fractions of the standard deviation of projected data distributions onto initially selected random directions (the projection vectors) in pattern input space; $\alpha_k$ is

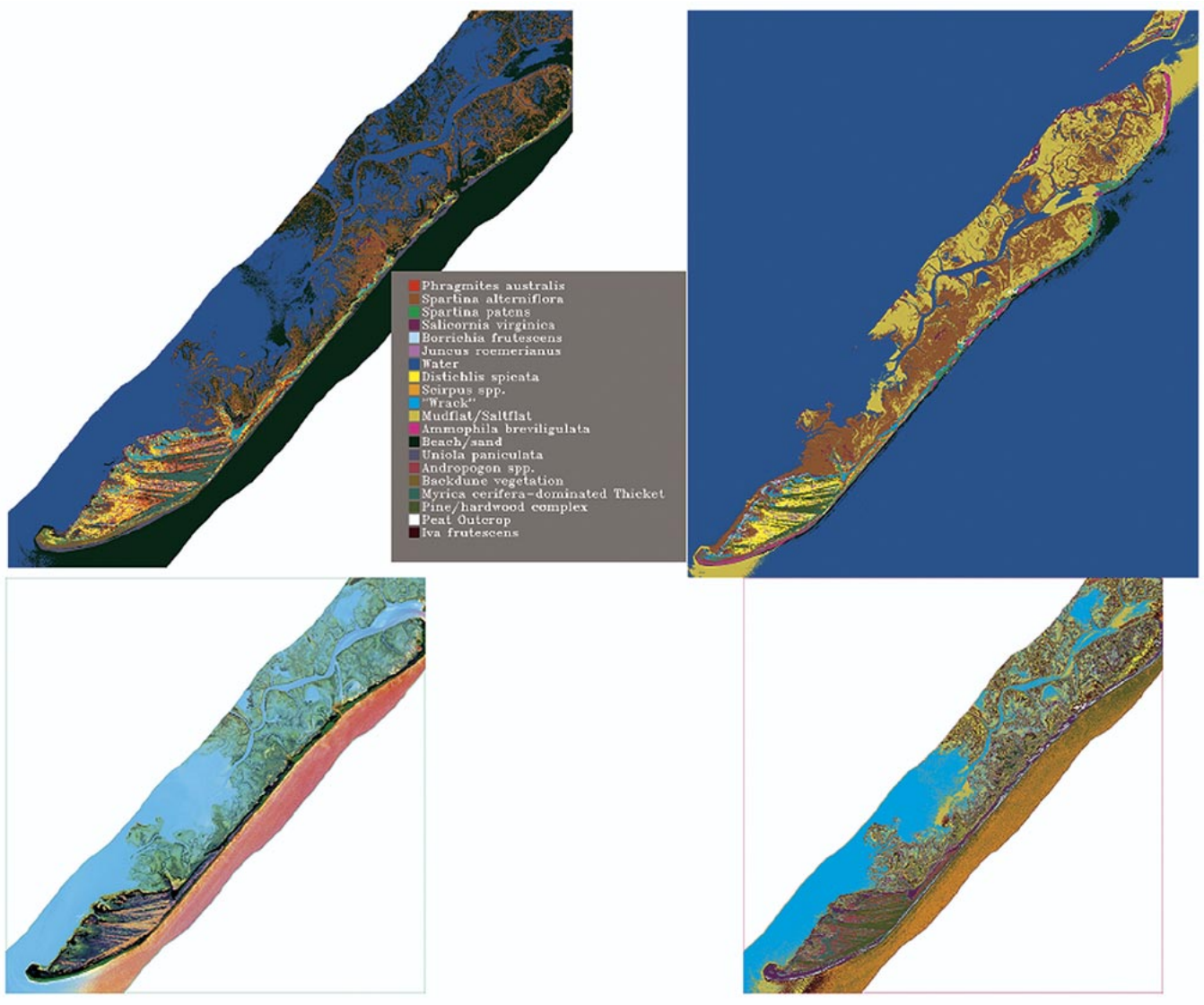



Fig. 4. (Top, left) Sixteen-category land cover supervised BPCE classification based on 112 georeferenced HyMAP spectral end-members labeled from GPS field surveys in October and December, 2000. Prediction of salt marsh vegetation in the north end, and many of the marsh and swale categories to the south appear consistent. Gaps in the center of salt marsh zones to the north are areas of heavy inundation which were declared as water by the model. Biggest errors occurred for *Phragmites australis*, *Juncus roemerianus*, and *Uniola paniculata*, all of which had high false-alarm rates. (Top, right) Nineteen-category supervised land cover classification based on 3656 HyMAP spectral end-members, labeled from DGPS and GPS surveys for training the model, showing dramatic reduction in false-alarm rate for these categories. (Bottom, left) RGB composite of three PP projections of HyMAP spatial–spectral windows; (bottom, right) 34-category land cover map produced by ISODATA clustering in a five-dimensional PP projection space, including the three PP projections shown in the RGB composite.

chosen at random within the user-specified range, and one $\alpha_k$ is associated with each data projection.

The Friedman–Tukey Projection Index [20], $I$, on which our projection index is based, was the product of a trimmed variance $S$ and a compactness function $N$

$$\text{Maximize: } I(c_k) = S(c_k)N(c_k) \tag{1}$$

$$S(c_k) = \sqrt{\frac{\sum_{\mu=1}^{M_p-m} (c_k(\mu) - E(c_k))^2}{M_p - m}} \tag{2}$$

$$N(c_k) = \sum_{\mu=1}^{M_p}\sum_{\nu=1}^{M_p} g(r_k(\mu,\nu)) \times \theta(R - r_k(\mu,\nu)) \tag{3}$$

$$\text{with } r_k(\mu,\nu) = \mid c_k(\mu) - c_k(\nu) \mid \tag{4}$$

$$c_k(\mu) = \hat{w}_k \cdot \vec{f}(\mu) \tag{5}$$

where $c_k(\mu)$ is the $k$th data projection of the $\mu$th sample vector, denoted $\vec{f}(\mu)$, and unit projection vector $\hat{w}_k$; $\theta$ is a step function; $R$ is a scalar compactness or cluster scale; $g(r_k(\mu,\nu))$ is a monotonically decreasing function of the distance between projected sample pairs $r_k(\mu,\nu)$; $M_p$ is the number of samples; and $m$ the number of outliers removed in the trimmed variance. We replaced their projection index with $I^*$

$$\text{Maximize: } I^*(\tilde{r}_k(\mu,\nu)) = \eta(\tilde{r}_k(\mu,\nu))D(\tilde{r}_k(\mu,\nu)) \tag{6}$$

$$\eta(\tilde{r}_k(\mu,\nu)) = E_{\text{pairs},(\mu,\nu)} g(\tilde{r}_k(\mu,\nu)) \tag{7}$$

$$D(\tilde{r}_k(\mu,\nu)) = E_{\text{pairs},(\mu,\nu)} \cdot [(1 - g(\tilde{r}_k(\mu,\nu)))] \tag{8}$$

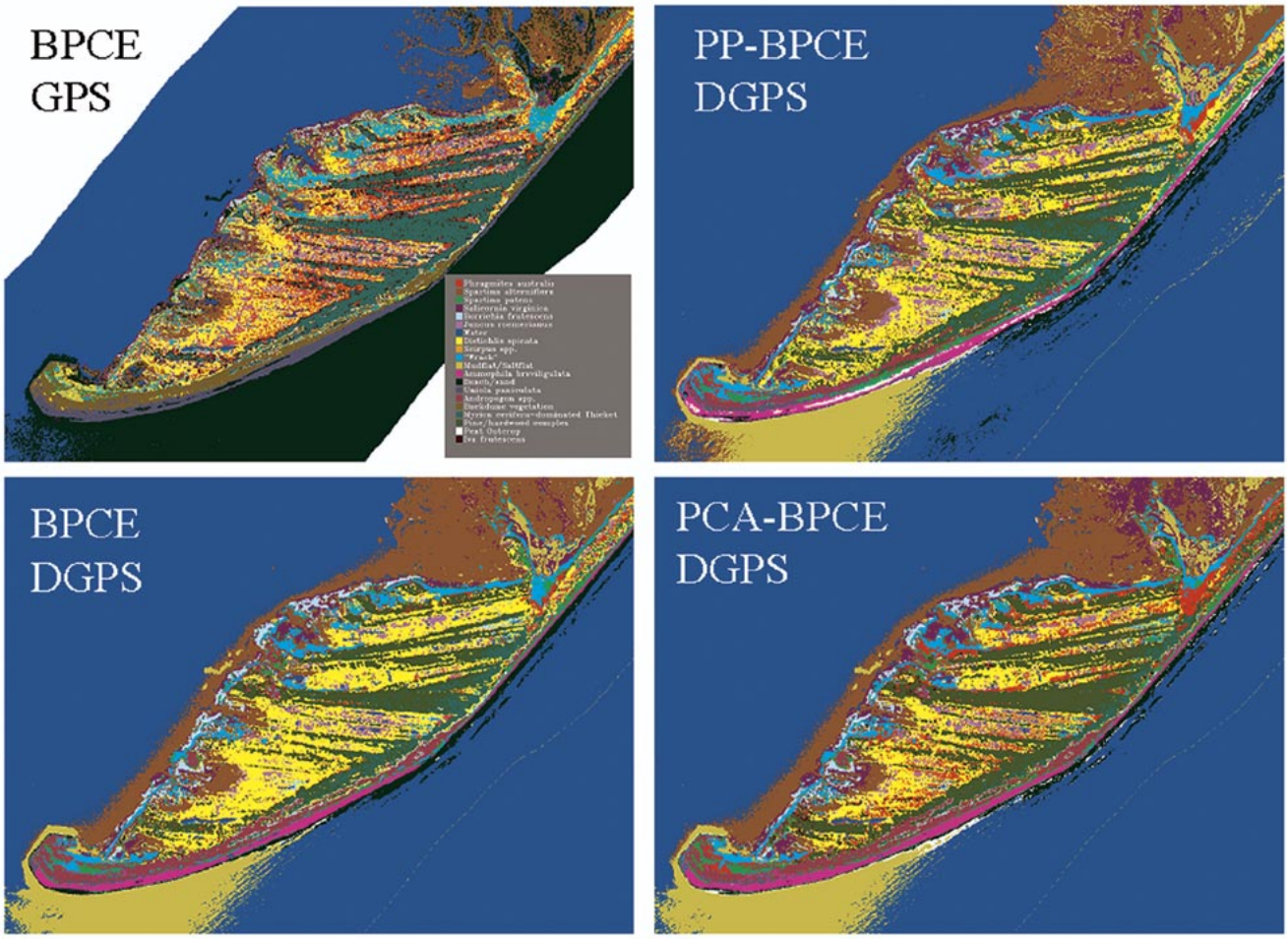



Fig. 5. Extract from four land cover classifications, showing results for the southern end of Smith Island. (Top, left) BPCE model based on original 112 HyMAP spectral end-members, shows high false-alarm rate for *Phragmites australis*, *Juncus roemerianus*, and *Uniola paniculata*. (Bottom, left) BPCE classification using 3656 spectral end-members with improved georeferencing, showing dramatic reduction in false alarms; model also discriminates some dune vegetation types, and separates *Myrica cerifera* from Pine/Hardwood complex. (Top, right) PP-BPCE composite, using expanded spectral set, and (bottom, right) PCA-BPCE composite trained on same. PCA-BPCE shows higher rate of false alarms for *Phragmites australis* than PP-BPCE.

$$\text{with} \quad \tilde{r}_k^2(\mu,\nu) = (\tilde{c}_k(\mu) - \tilde{c}_k(\nu))^2$$

$$g(\tilde{r}_k(\mu,\nu)) = e^{-(\tilde{r}_k^2(\mu,\nu)/\alpha_k)^2} \tag{9}$$

$$\tilde{c}_i(\mu) = \sigma\left(\sum_j L_{ij} c_j(\mu)\right) \tag{10}$$

$$\text{where} \quad \sigma(x) = a\tanh(a\lambda x)$$

$$(a, \lambda, \text{constants}) \tag{11}$$

$$c_j(\mu) = \vec{w}_j \cdot \vec{f}(\mu) + b_j \tag{12}$$

where $\eta$ is a continuous compactness function, of a nonlinear projection, $\tilde{c}_i(\mu)$, $D$ measures spread by sampling pairs of projections and approaches asymptotically a constant weight outside scale $\alpha_k$; and $E_{\text{pairs},(\mu,\nu)}$ signifies expected value over projected sample pairs. Other differences included 1) optimization of multiple projections at the same time, rather than serially, and the use of a coupling matrix $L_{ij}$ that is adjusted via gradient ascent to maximize the relative entropy of the data projections, and 2) our use of a saturating nonlinearity $\sigma$ to remove sensitivity to outliers, meaning that all data points can be included. Each compactness function $g(\tilde{r}_k(\mu,\nu))$ has a clustering search scale $\alpha_k$ associated with it. Each $\alpha_k$ is obtained by multiplying an estimate of the initial standard deviation of the projected data, with a random fraction drawn from a user-determined search range. We optimized $\vec{w}_j$ by stochastic gradient ascent in $I^*$.

While the approach that we defined in [8] does not specifically aim to derive an orthonormal PP filter set, it did incorporate a mechanism for decorrelating projections in the stochastic optimization process. Essentially, a coupling matrix, labeled $L_{ij}$ above, is defined between the projections, and this matrix is simultaneously optimized along with the projections in such a way that the relative entropy between the projections is maximized (decorrelation). The degree of decorrelation can be controlled by altering the size of the initial coupling and the relative rates of optimization of the relative entropy cost function used for the coupling and the PP cost function used for the projections. Additional implementation details can be found in [6] and [8].

To optimize the unsupervised PP and PCA filters, we used either the end-members associated with our GPS and DGPS ground data surveys or, in some instances, larger spectral sub-

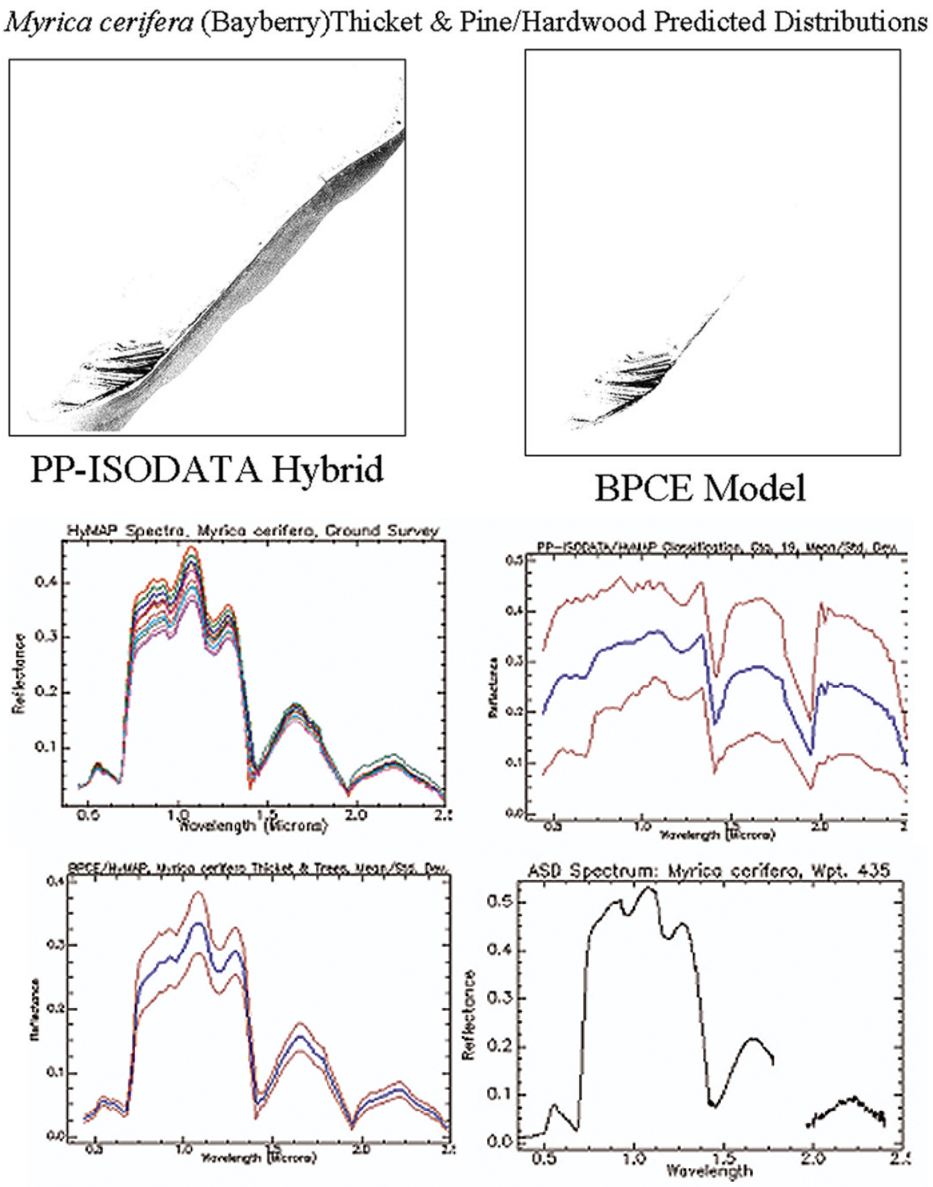



Fig. 6. Models associated with upland thickets and tree stands for (top row, left) PP-ISODATA and (top row, right) the first BPCE model based on 112 spectral samples. Distributions for the two are largely consistent, with the exception that the unsupervised approach has included an area of glint in the surf zone of the eastern shore. (Middle row, left) Spectral reflectance plots for HyMAP data at GPS waypoints associated with *Myrica cerifera*; (middle row, right) mean and standard deviation of PP-ISODATA category (includes glint zone); (bottom row, left) mean and standard deviation of upland thicket and tree stands distribution predicted by the BPCE model; (bottom row, right) ASD reflectance measurement of *Myrica cerifera* leaves taken on May 11, 2001. Note that the relative height of the first peak in the NIR is somewhat higher in the ASD measurment (gaps are removed atmospheric absorption windows, where spectrometer counts created numerical instabilities in the reflectance calculation), and overall reflectance is slightly higher in the ASD measurement.

sets derived from the southern end of the island that were representative of the typical spectral variation seen in the data. PP and PCA filters were derived from either $1 \times 1 \times 126$ or $3 \times 3 \times 126$ spatial–spectral windows. For the supervised classification models, described in Section IV-B, we always used the spectral end-members associated with our GPS and DGPS surveys. In the latter case, because the size and shape of these ROIs were quite variable, we restricted ourselves to inputs that were $1 \times 1 \times 126$ (single-pixel spectra).

The feature extraction stage of the unsupervised classification models considered in our experiments used either 1) projection pursuit or 2) pincipal component analysis. The final stage of the process was the ISODATA [45] algorithm.

### B. Supervised Classification Models

In all supervised clasification models considered in this paper, the final stage of classification was a variant of the backward propagation of error model [41] with a cross-entropy cost function (BPCE) [40]. The BPCE cost function is

$$\Xi(\vec{x}) = -\sum_i ((1 - d_i(\vec{x})) \ln(1 - \tilde{c}_i(\vec{x})) + d_i(\vec{x}) \ln(\tilde{c}_i(\vec{x}))) \tag{13}$$

where $d_i$ is the desired output, either 0 or 1, for one of the category nodes at the output of the model, and $\tilde{c}_i$ is the actual response of the output node to a particular input pattern propagated forward through the model. We use the cross-entropy

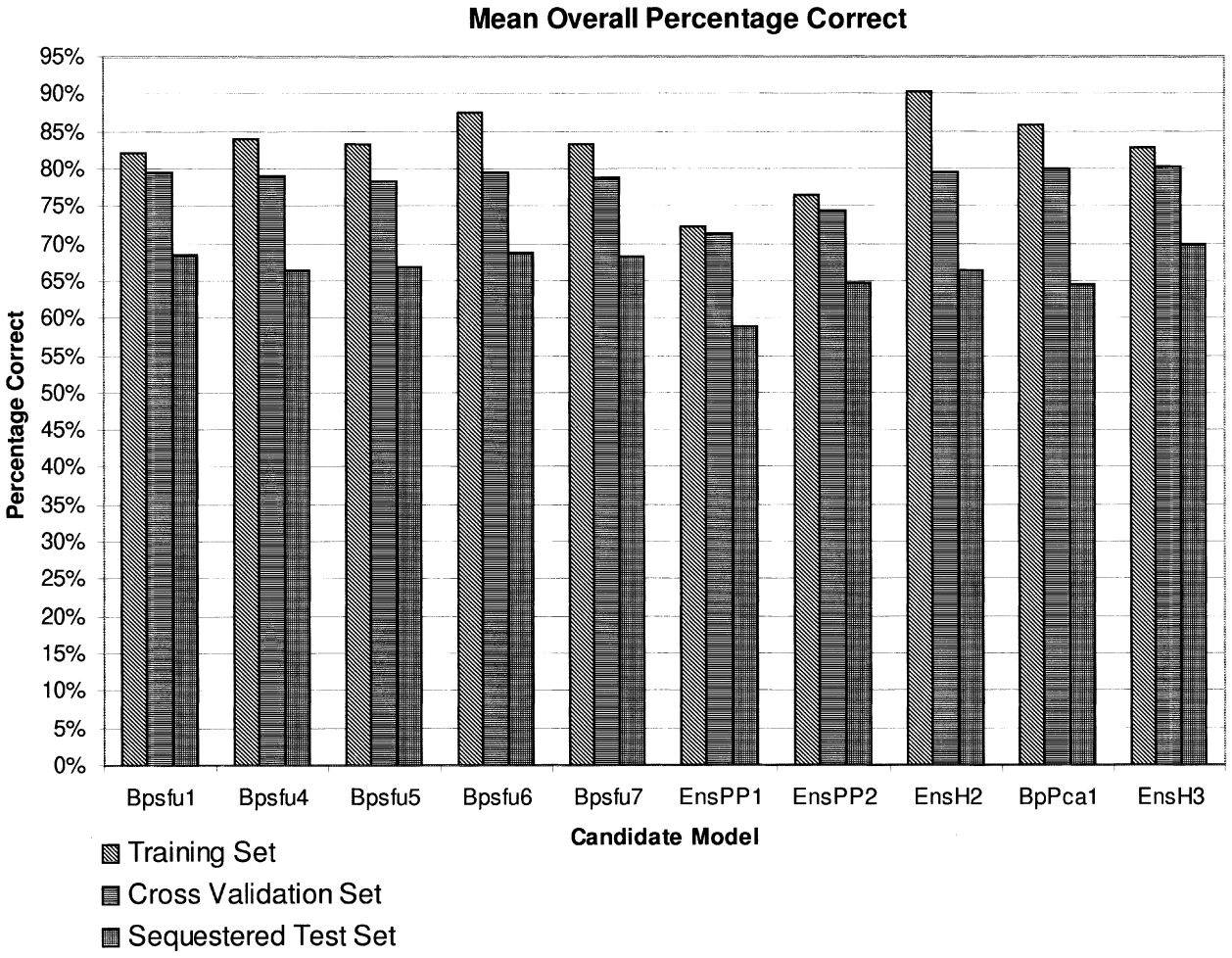


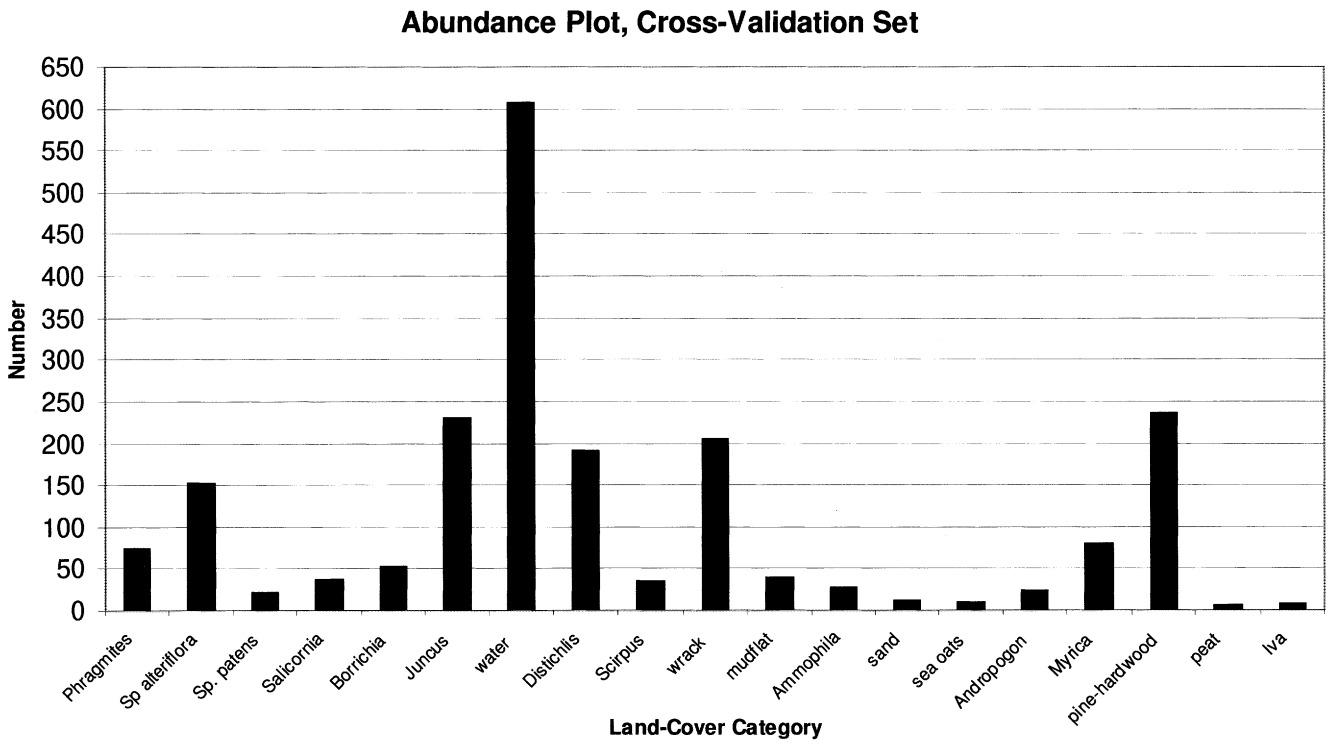

Fig. 7. (a) Overall classification accuracies for ten candidate classification models, showing relatively similar average performance across algorithms and model architectures.

cost function because it is less prone to local minima than the originally proposed least mean-square (LMS) error [40], owing to the form of the gradient used in the stochastic gradient descent. Comparing this with the more commonly used LMS error, $\Xi(\vec{x}) = \sum_i (\tilde{c}_i(\vec{x}) - d_i(\vec{x}))^2$, defined in [41], it can be seen that the cross-entropy cost function eliminates a factor in the gradient descent rule $\Delta(w_{ij}) = -\eta(t)\partial\Xi/\partial w_{ij}$ for the weight vectors $\vec{w}_i$. Specifically, for LMS, the derivative of the transfer function $\tilde{c}'(\vec{x}) = \tilde{c}_i(\vec{x})(1-\tilde{c}_i(\vec{x}))$ that appears in the gradient in the last layer weights, and in earlier layers through the backpropagation of error, can cause the updates to become "frozen" near zero when $\tilde{c}_i(\vec{x})$ is antipodal to the desired response. The latter occurs because the derivative of the transfer function has two zero crossings. The expression is also zero when the response is near the desired response, but it is the antipodal response that causes the undesirable behavior. The form of (13) eliminates the second zero crossing that causes this behavior because an extra factor appears in the gradient due to the presence of the logarithms.

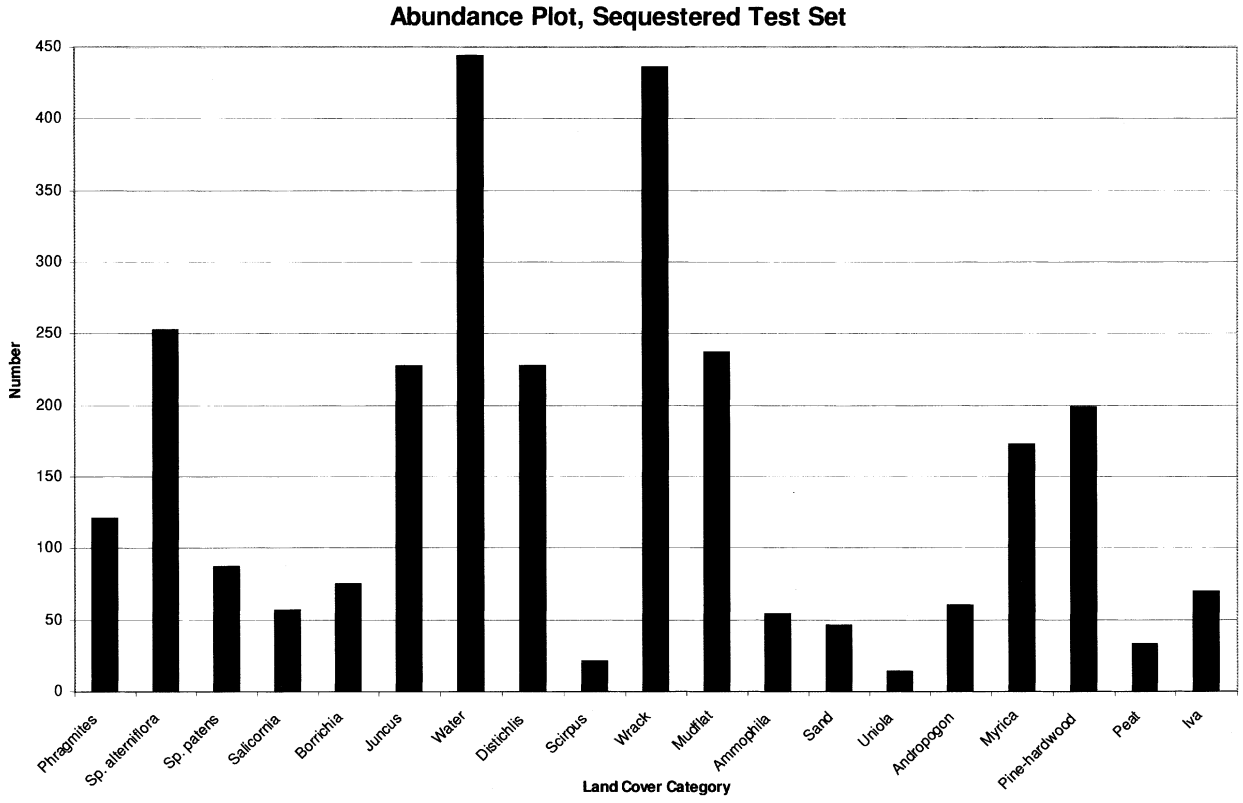
One additional feature of our supervised classification models was the use of an error-resampling buffer, which increased the frequency with which spectra-causing misclassifications were presented to the model. This forces filter adjustments to improve the model on boundaries between land cover categories where errors are more likely. This is particularly useful when some categories are sparsely represented, as is the case in this application. Details of this error-resampling buffer are beyond the scope of this paper, but this approach tends to accelerate model convergence and can lead to higher asymptotic classification rates [7].

HyMAP reflectance data corresponding to the spectral end-member sets delimited by the GPS and DGPS ground measurements for each category were the input to the model. These data were divided into three groups, one for training and two for testing generalization, as described in greater detail in Section V. A few unreliable bands were eliminated in the vicinity of the two major atmospheric absorption windows. In some of the models described in Section V, the data were first projected into a lower dimensional set of features using either preoptimized filters derived with the PP algorithm described in Section IV-A or PCA. In these models, the input to the BPCE model consisted of the lower dimensional feature vector (see Fig. 3); in other models, the spectral end-members were input directly to the BPCE model.

When PCA was the precursor stage of processing, we retained the first 42 eigenvectors. This number of features may have been excessive from the standpoint of noise reduction in most categories, since all but $1.7 \times 10^{-3}\%$ of the variance is explained




(b)




(c)

Fig. 7. *(Continued).* Relative category abundances for (b) cross-validation test set and (c) sequestered test set.

with this many components; however, it ensured that we would not be discarding small-scale spectral features that might permit discrimination of highly similar but distinct land cover types. For the results described in this study, when PP was the precursor stage, we used 32 PP projection vectors that were first optimized before insertion in the end-to-end classification model.

A variety of model architectures and complexities were explored using this framework, and the performance of a set of

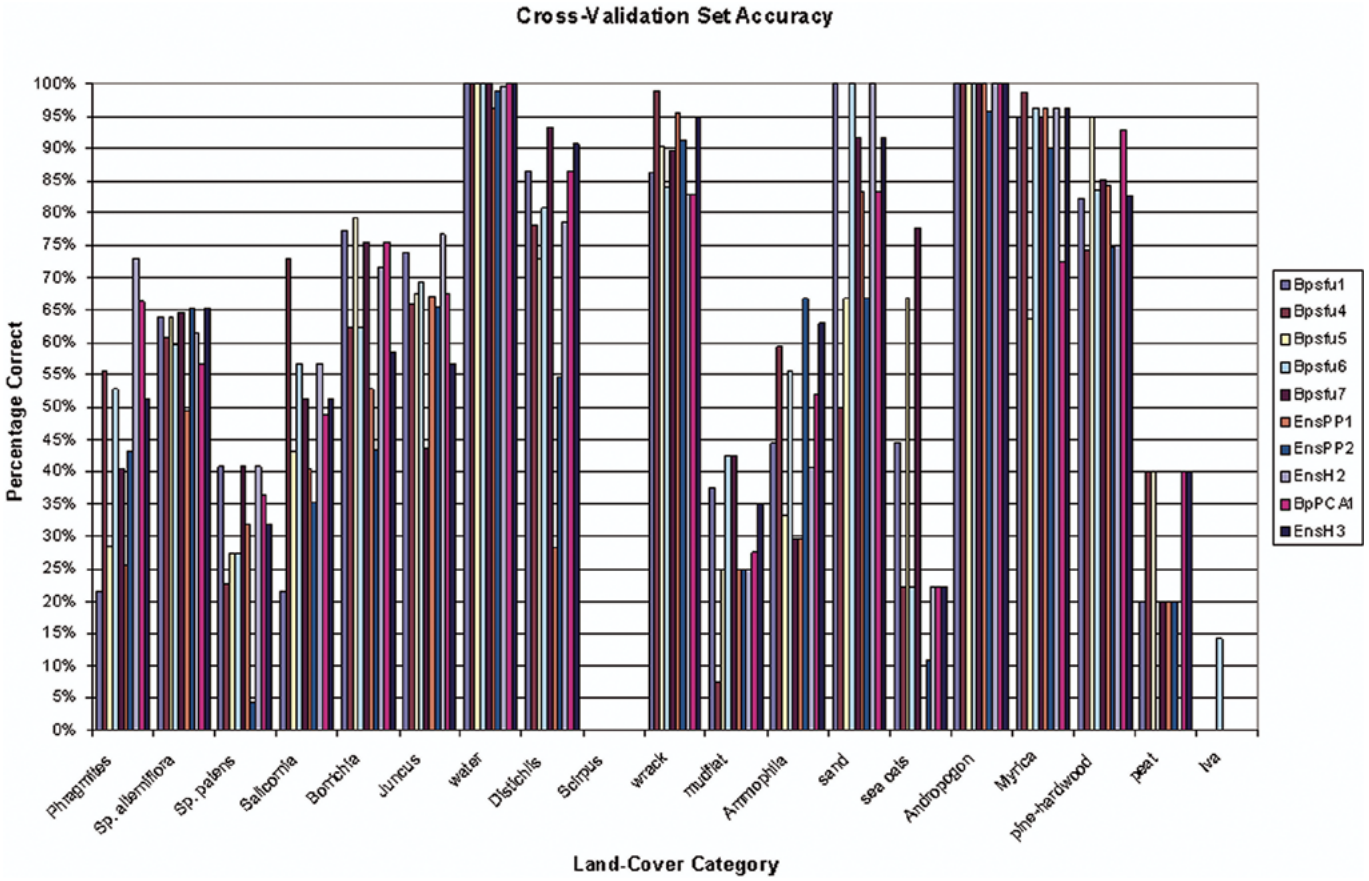



Fig. 7. *(Continued).* Performance versus category for the ten models: (d) cross-validation test set.

ten examplar models is shown in Section V. An analysis of the variability in the models suggested that smoothing in functional (classifier) space might achieve more reliable results.

## V. Results and Discussion

Both supervised and unsupervised classification models of the land cover were produced. Of these, the first supervised classification maps consisted of 16 of the 20 land cover categories in Table I, with one aggregate category that combined *Andropogon spp.* and *Ammophila breviligulata* into a "backdune" category. These first models used 112 ground-referenced spectral end-members (Fig. 4). Many of the categories, such as *Myrica cerifera* Thicket, *Distichlis spicata*, *Spartina alterniflora*, Backdune vegetation, "Wrack" (Fig. 2) appear to have produced consistent results based on our field observations and historical data [32], [35]. A few categories were problematic, however, and these included *Phragmites australis*, *Juncus roemerianus*, and *Uniola paniculata*, all of which had high false-alarm rates. Our first models did not attempt to distinguish the *Myrica cerifera* thicket from the pine-hardwood complex, particularly since *Myrica cerifera* typically appears in the understory of these tree stands (Fig. 2), and our field surveys had not at that point sufficiently documented the location of representative pine and hardwood stands. Subsequent models described below did include a distinction between these two land cover types after additional ground data had been acquired.

After the development of the first automatic land cover models, we visited the island to obtain additional survey data for validating the results. During the visit between May 7–11, 2001, a year from the time of the initial HyMAP data acquisition, we collected ASD FR *in situ* spectra and a large number of additional survey points (examples of ASD spectra appear in Figs. 6 and 8). Followup visits in August and October 2001 and May 2002 (while this paper was being revised) established more accurate ground data using DGPS as described above. Although the temporal gap between airborne and ground data acquisitions is not ideal, the interval is short enough for many of the categories that survey data would still be reliable. Exceptions to this are mudflats/salt pannes and wrack, although the dominant distribution of wrack at the mean high tide level is relatively stable.

The second set of supervised classification models that we produced was comprised of 19 of the 20 categories (Figs. 4 and 5) listed in Table I, omitting the foredune vegetation, which is often sparse or nascent in the early part of May in our study area. (Models described in future papers using PROBE2 data acquired in the summer, when this vegetation is fully present, will include this category.) In these experiments, we took advantage of the additional spectral data labeled during the DGPS surveys. The new experiments with the expanded set of labeled spectra included 3656 training samples spread across the 19 categories previously described. Additionally, two test sets were set aside, one for cross validation, which was used to determine a stopping point for optimization with the training set, and

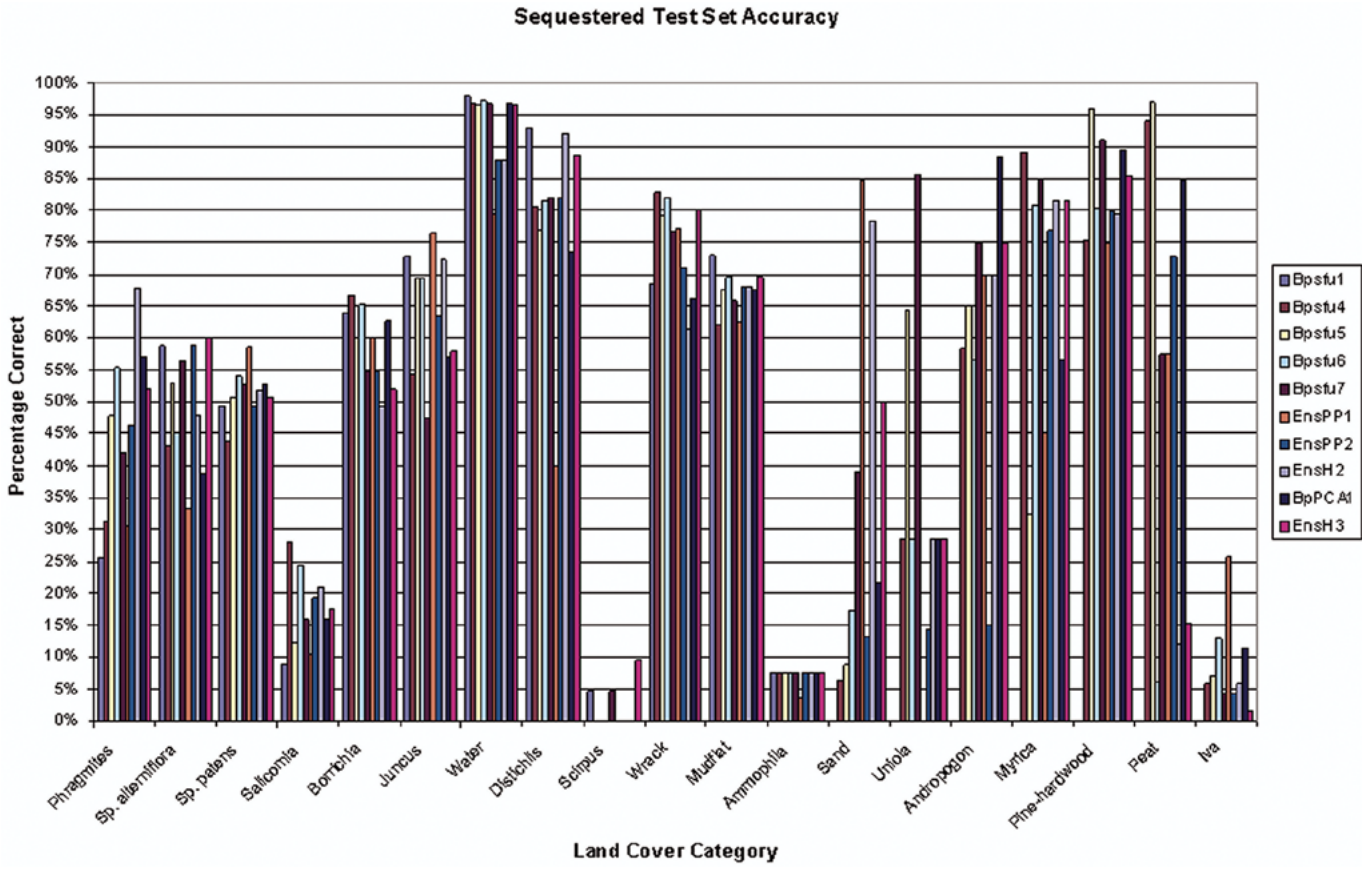



Fig. 7. *(Continued).* Performance versus category for the ten models: (e) sequestered test set.

one sequestered test set, used as a second estimate of expected generalization capability. The cross-validation test set contained 2049 spectral samples, while the sequestered test set consisted of 2836 spectral samples. These models showed immediate improvement because of the improved georectification and larger set of georeferenced spectral end-members that were used to train these models. Particularly noticeable was the large reduction in false-alarm rates for *Phragmites australis*, *Juncus roemerianus*, and *Uniola paniculata*. Heavily inundated portions of the northern salt marsh that were declared as water in the earlier models are now either declared as *Spartina alterniflora* or Mudflat, and the surf zone, where glint was present is now correctly labeled as Water, rather than Beach/sand (this is a high–tide result, so much of the beach is under water). The amount of Mudflat declared is probably too large and is the result of a number of factors, including the early stage of growth, the sparseness of the *Spartina alterniflora* in areas of heavy inundation, the fact that spectra used to model the Mudflat category may also have been partially inundated, and the occurence in mudflats of sparse vegetation or small deposits of wrack. Nevertheless, the overall categorization of *Spartina alterniflora* in the northern end of the island is significantly improved.

Fig. 4 illustrates the unsupervised approach, depicting an RGB composite of three PP projections and a 34-category PP-ISODATA category map derived from this and two other PP projections. In contrasting this with the supervised classifications shown in Fig. 4, it can be seen that a number of the categories in the two approaches are correlated, although in some cases the PP-ISODATA map has grouped two or more categories that are distinct BPCE categories, e.g., backdune vegetation and "wrack" are grouped in the PP-ISODATA map. In some instances, the opposite is true. For example, the *Spartina alterniflora* category was divided into two groups in the PP-ISODATA model, while it is, of course, a single distribution in the supervised BPCE classifications. The distinction made by the PP-ISODATA may be related to differences associated with short versus tall forms of the *Spartina alterniflora* (Fig. 2). Taller forms tend to be located near the berm edge of creeks and channels in the salt marsh at the northern end of Smith Island, while shorter forms are found in the interior of the salt marsh where elevation is lower and tidal inundation greater. Fig. 6 compares the distribution of *Myrica cerifera* and tree stands predicted by the original BPCE model using 112 spectral training samples and the PP-ISODATA category associated with these vegetation types. While the distributions are similar, one principal difference is that the PP-ISODATA has aggregated an area strongly affected by glint, and this leads to distortions when compared with the BPCE-predicted model and ASD FR *in situ* measurements. Several PP-ISODATA categories are coherent in structure but unlabeled at this point, pending further survey efforts.

A more exacting test of accuracy was made possible by the DGPS data that we collected on Smith Island in follow-on surveys. Although we have taken ground data on five of the Virginia Coast Reserve barrier islands during these subsequent visits, we have spent roughly half of that time on Smith Island. While

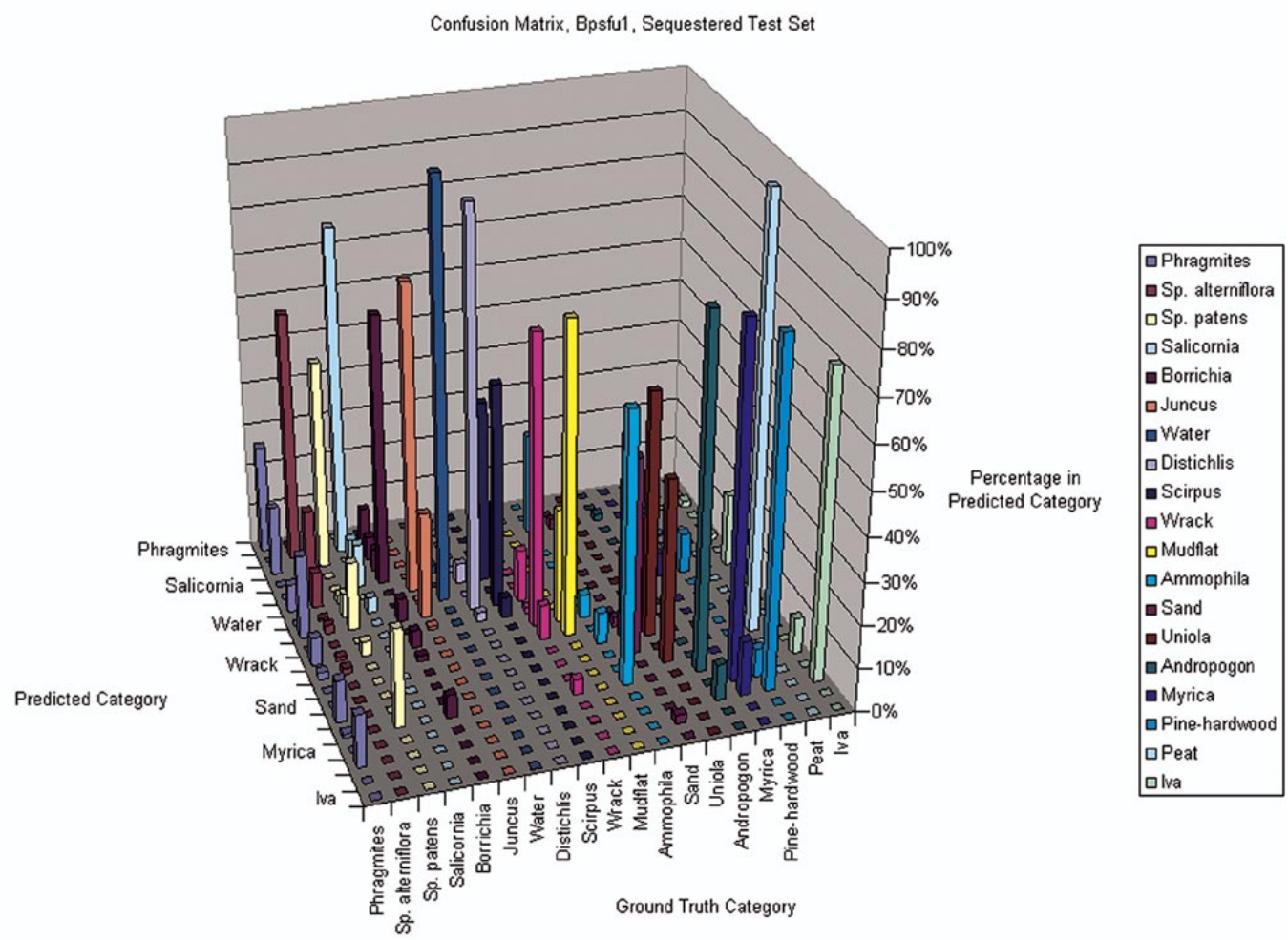



Fig. 7. *(Continued).* Performance versus category for the ten models: (f) typical confusion matrix from the first candidate model.

additional data will be needed to validate PROBE2 data acquired in other seasons, these DGPS data, nevertheless, provided us a much higher degree of precision in determining relative classification accuracies on Smith Island. As described above, data labeled during the GPS and DGPS surveys were divided into training data, cross-validation data (used to stop training of the supervised models), and sequestered test data, and 19 of the categories listed in Table I were used in the models. Ten candidate models were developed using the algorithms described in Section IV. Of the ten portrayed in Fig. 7, the first five were BPCE models of varying complexity; models six and seven were composite PP–BPCE models, models eight and ten were BPCE–BPCE composite models, pooling the results of several BPCE and BPCE composite models, and model nine was a PCA–BPCE composite. While overall accuracy for the training data reached as high as 90% for the training data in some of these classifications, a more important measure is the extent to which these models generalize to sequestered test data when challenged. Overall accuracy ranged between 72% and 90% for the training set, between 71% and 80% for the Cross-Validation Set, and between 58% and 69% for the sequestered test set. Fig. 7 compares the performance of a set of candidate models that were produced for Smith Island using the expanded spectral end-member sets derived from the DGPS and GPS surveys.

Relative abundance of categories in the training and test sets is also reported in Fig. 7. It shows that for the cross-validation set, in 13 of the 19 categories, one or more models lie within the range between 65% and 95% accuracy, while 14 fall within the range between 65% and 98% in the sequestered test set. Not surprisingly, some of the dominant categories such as *Distichlis spicata*, *Myrica cerifera* Thicket, Water, Pine/Hardwood Complex, and Wrack are at the top end of this range. At the same time, there is a high degree of variability in the models. While part of this is due to differences in algorithms, a significant contribution is due to the high degree of spectral overlap in many of the categories present in the early part of the growing season. One surprising result is the performance for the invasive plant *Phragmites australis*. In both test sets, for the category *Phragmites australis*, at least one model exceeds the 65% threshold, obtaining 73% and 68% accuracy respectively on the cross-validation and sequestered test sets. In the southern end of Smith Island, this invasive species typically grows in the ecotone between thicket and the marsh vegetation (Fig. 8) and is, therefore, difficult to detect due to mixing with other categories, such as *Myrica cerifera* Thicket. The left-hand column of spectral plots in Fig. 8 portrays this mixing, comparing the spectral response of HyMAP at areas known from our ground survey to consist of exposed *Phragmites*, *Myrica cerifera*, and *Phragmites aus-*

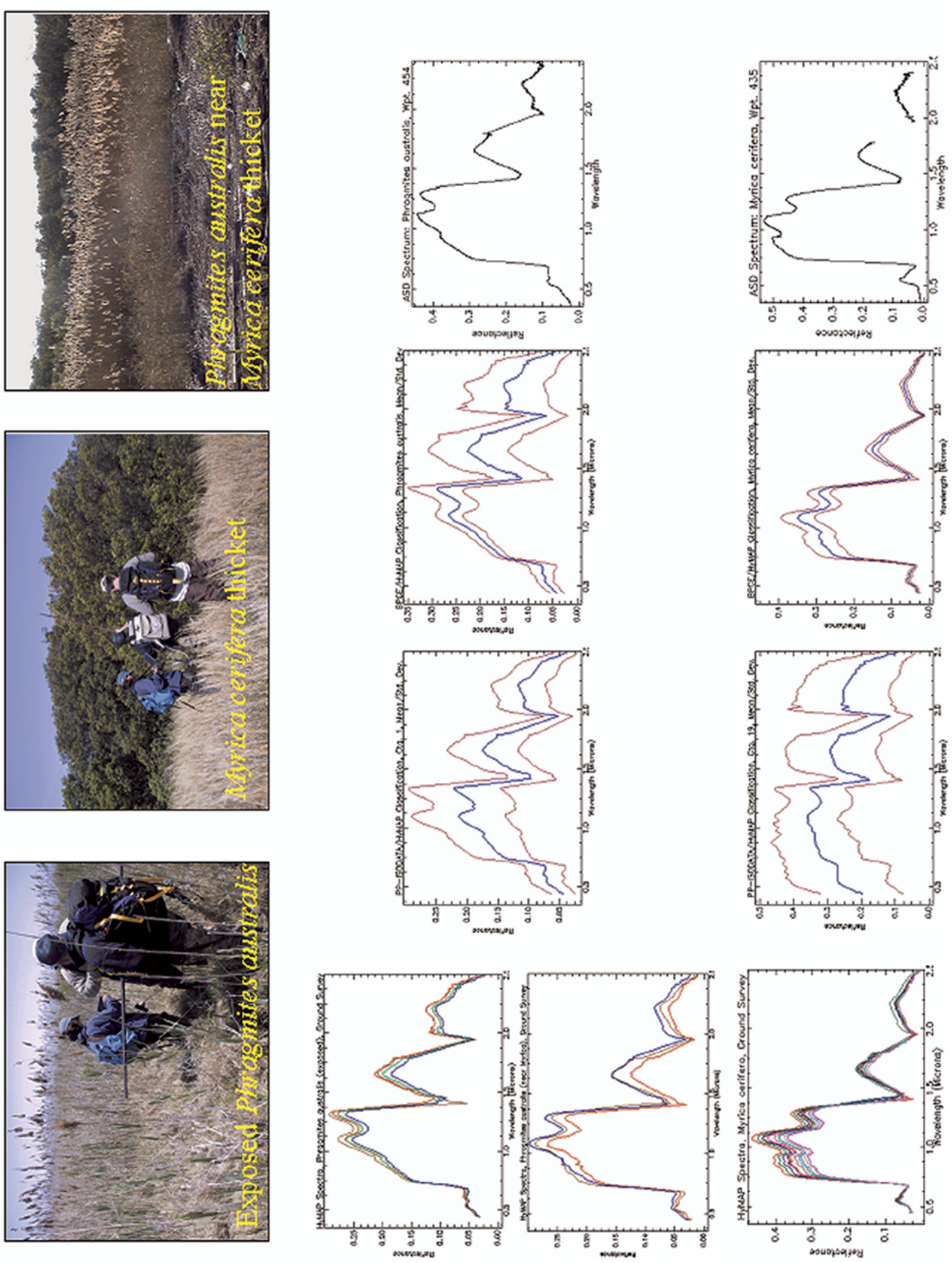



Fig. 8. (Top row photographs) Exposed *Phragmites australis*, *Myrica cerifera* thicket, and *Phragmites* near the thicket. Spectral plots, left column: (top) exposed *Phragmites australis*, (middle) *Phragmites australis* near *Myrica cerifera*, (bottom) *Myrica cerifera*. Spectral plots, top row: (left) Mean and standard deviation of PP-ISODATA category associated with *Phragmites australis* near thicket, and (middle) BPCE classification for all *Phragmites*, exposed and near thicket; (right) ASD FR spectrum of *Phragmites australis*. Spectral plots, bottom row: (left) PP-ISODATA category associated with *Myrica cerifera* and tree stands, distorted by glint grouped with the category, and (middle) BPCE clasification of *Myrica cerifera*; (right) ASD FR spectrum of lower canopy *Myrica cerifera* leaves.

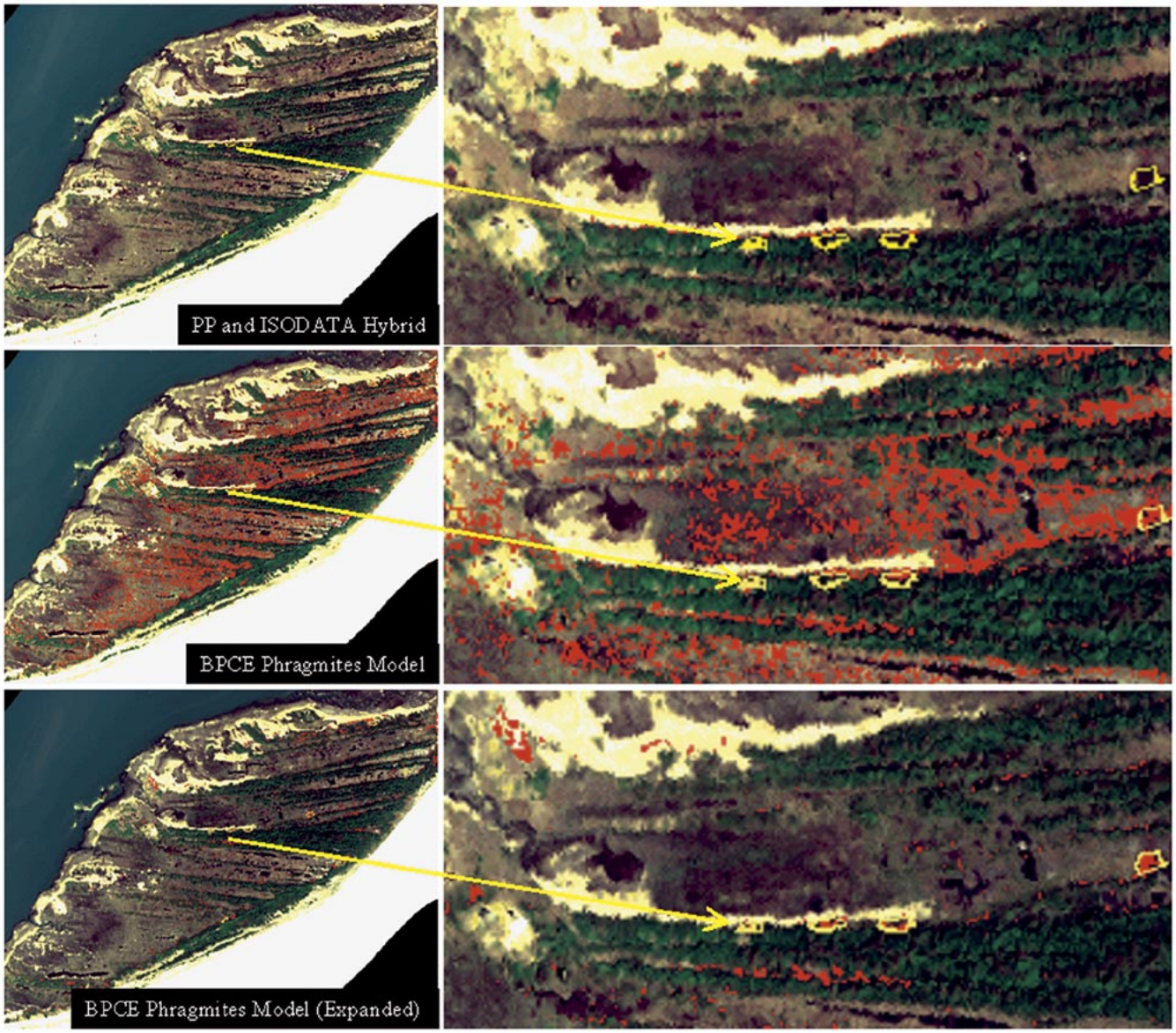



Fig. 9. Comparison of model performance for *Phragmites australis*: unsupervised versus supervised models (in red). (Top row) PP-ISODATA; (middle row) BPCE model based on original 112 spectral samples; (bottom row) BPCE model based on expanded spectral set of 3656 training samples and improved georectification. Also shown: areas identified as *Phragmites australis* during DGPS surveys. Zoomed areas show predictions in vicinity of *Phragmites* near thicket. PP-ISODATA and first BPCE model based on 112 spectra were prior to improved georectification, so the four *Phragmites* patches shown (areas in yellow from DGPS) appear shifted toward the bottom of the figure relative to the predicted distributions in the top and middle rows.

*tralis* adjacent to the *Myrica cerifera* thicket. *Phragmites australis* is not one of the dominant vegetation types on this island, which also makes it a challenging category to model. Inspection of the predicted distributions of *Phragmites australis* in Figs. 5 and 9 shows that the models based on the expanded set of spectral end-members has achieved a substantial reduction in false-alarm rate, when compared with the first set of models that used 112 spectral end-members. *Phragmites* is detected by this model both along the thicket edge and in areas where it is more exposed. The PP-ISODATA category most closely associated with *Phragmites australis* only detected *Phragmites* near the thicket; however, looking at Fig. 9, it can be seen that it lacked the desired specificity, tending to group other shrubs on the edge with the *Phragmites*. The best result was the BPCE model using the expanded set of spectral inputs, also shown in Fig. 9. While the false-alarm rate could still be further improved, it shows the most consistent prediction of *Phragmites* both in the open and along the thicket boundary. We conjecture that data acquired on dates later in the growth cycle may eventually allow us to reduce the false-alarm rate further. One additional observation concerning the supervised classification is that the *Phragmites* results for the PP-BPCE model and PCA-BPCE model had higher false-alarm rates than the BPCE model in isolation, but the PP-BPCE model did have a markedly better false-alarm rate than PCA-BPCE (Fig. 5), which is not surprising given earlier arguments developed in Section IV.

Looking at the results in Fig. 5, we note that similar problems that had existed for the category *Juncus roemerianus* in terms

of false-alarm rate in the original BPCE classification are now largely corrected in the models based on the expanded spectral set with improved georeferencing.

Categories such as Beach/sand, that would ordinarily not be so difficult to identify, were more problematic because the HyMAP scene had a significant area of glint in the beach and surf zone on the eastern shore of Smith Island, due to the high sun angle at the time of the data collection, and this contributed to the high degree of variability in performance across the ten candidate models. The presence of the glint also effected performance for categories such as Peat Outcrop, although this category can be challenging in and of itself because of its presence in the surf zone, depending on whether data are acquired near high or low tide.

Although in many areas the dune vegetation shows the proper delineation of *Andropogon spp.* toward the upland and *Ammophila brev.* toward the beach, the BPCE models based on the expanded spectral sets all tended to confuse *Ammophila brev.* with the *Andropogon spp.* in many of the specific ROIs used to evaluate accuracy. At this time of year, grasses and sedges such as these and *Spartina patens*, also found in the dune environment, are all tonally similar. For example, Fig. 7(e) shows that *Spartina patens* is most often confused with either *Distichlis spicata*, the dominant swale grass, or *Andropogon spp.* Data acquired in the early fall, when Andropogon is quite distinct visually from the other two, is likely to improve results. Thus, we expect that the October PROBE2 data acquisition will achieve higher accuracy when classification models are developed.

Other sources of difficulty for these models stem from the time of the year that the HyMAP data were acquired. At the beginning of May, it is early in the growing season in the VCR, so many vegetation communities contain a mixture of new growth and senescent or dead vegetation from the previous growth cycle. Distinguishing vegetation types such as, for example *Distichlis spicata* from *Scirpus spp.*, may be very difficult to achieve spectrally at this time of year, and this probably accounts for the fact that the majority of errors for the category *Scirpus spp.* are the result of confusion with *Distichlis spicata*. As we have noted earlier, tidal influences provide additional sources of spectral variability for many of the marsh vegetation communities because of variations in degree of inundation, and it obviously effects the beach zone, depending on the degree of inundation or wetting present. Many points that we acquired in the beach zone in the first surveys were effectively under water due to tidal stage or in an area of strong glint due to the time of data acquisition and, therefore, could not be used in the analysis. Although more than one model obtains a respectable score for the category Beach/sand, the poor performance for this category in the other models is almost certainly due to the presence of glint.

## VI. Conclusion

Our goal was to develop land cover maps that would be useful to natural resource managers at higher spatial resolution than has been available previously. Both unsupervised and supervised classification approaches were used to create these products and to evaluate their relative merits. We have seen that automatic land cover classification models can be developed successfully from HyMAP imagery, even in the early part of the growing season when spectral differences in vegetation may not be as pronounced. The expectation is that a more ideal data acquisition date in late summer or early fall would improve results further. PROBE2 imagery acquired during those intervals will be used to evaluate this conjecture. The use of a hyperspectral sensor with spatial resolution of 4.5 m was deemed necessary in order to be able to discriminate rapidly varying land cover types seen, e.g., in the transition zone from the lagoonal shore to the upland. On Smith Island, six to seven distinct vegetation zones may occur in a distance as short as 50–75 m.

Some technical difficulties such as extensive glint present in the HyMAP data in the beach and surf zone limited what could be achieved given a more ideal time of day for data collection. Other challenges stemmed from the fact that the data was acquired near high tide. Despite these difficulties and the fact that the early part of the growing season may not be the ideal time for distinguishing many types of vegetation, we have demonstrated success in identifying both plant communities and, in some instances, individual plant species from HyMAP through our field validation efforts with GPS, DGPS, and *in situ* reflectance measurements. Supervised classification models based on spectra labeled during GPS and DGPS surveys were used to demonstrate that models could discriminate 19 land cover types. Some of these categories were defined at the plant community level, with others being specific plant species.

Although there were differences between *in situ* measurements and the airborne hyperspectral data, there were strong correlations between spectral shape. Unsupervised models based on a PP-ISODATA hybrid were found to agree with the supervised models for a number of categories. In some cases, the exploratory PP-ISODATA approach may have identified subgroups within a major category such as *Spartina alterniflora*, for which it was observed that the unsupervised approach may be dividing the data into low and high vigor forms of the same species. Without *a priori* knowledge of pixel labels, the PP-ISODATA approach was found to be correlated with *Phragmites australis* that grows in the margin between marsh and upland; however, this approach did not identify exposed *Phragmites*. The partial success of this exploratory approach also benefited from the ability to input both spectral and spatial–spectral windows. Accuracy and specificity of supervised models based on BPCE and composite models, especially for *Phragmites australis*, were found to be highly dependent on the size of the labeled spectral training samples and on the accuracy of the georeferencing. Increasing this accuracy and expanding the number of spectral samples used in training provided a significant reduction in false-alarm rate for multiple categories, including *Phragmites*. The best model overall for *Phragmites* used BPCE and the expanded set of spectral training samples.

## Acknowledgment

The authors wish to acknowledge computing resources provided by the DOD High Performance Computing Modernization Program, including the Army Research Laboratory's Major Shared Resource Center, SMDC, and the NRL Origin.

## REFERENCES


[1] Analytical Imaging and Geophysics LLC (AIG) [Online]. Available: http://www.aigllc.com

[2] C. M. Bachmann, S. A. Musman, D. Luong, and A. Schultz, "Unsupervised BCM projection pursuit algorithms for classification of simulated radar presentations," *Neural Networks*, vol. 7, no. 4, pp. 709–728, 1994.

[3] C. M. Bachmann, E. E. Clothiaux, J. W. Moore, and D. Q. Luong, "Dynamically reconfigurable projection pursuit ensembles for cloud detection in AVIRIS imagery," in *Proc. IGARSS*, vol. I, Firenze, Italy, July 10–14, 1995, pp. 256–259.

[4] C. M. Bachmann, E. E. Clothiaux, J. W. Moore, and K. J. Andreano, "An ensemble approach to automatic cloud identification in AVIRIS imagery using BCM projection pursuit," in *Proc. Neural Networks for Signal Processing IV Workshop*, Ermioni, Greece, Sept. 6–8, 1994, pp. 394–4.

[5] C. M. Bachmann, E. E. Clothiaux, and D. Q. Luong, "Wavelet projection pursuit for feature extraction and cloud detection in AVISRIS and AVHRR imagery," in *Proc. IGARSS*, vol. I, Lincoln, NE, 1996, pp. 356–359.

[6] C. M. Bachmann, "Novel projection pursuit indices for feature extraction and classification: An intercomparison in a remote sensing application," in *Proc. Neural Networks for Signal Processing VII Workshop*, Amelia Island, FL, Sept. 24–26, 1997, pp. 54–63.

[7] C. M. Bachmann and T. F. Donato, "Mixtures of projection pursuit models: An automated approach to land-cover classification in landsat Thematic Mapper imagery," in *Proc. IGARSS*, Hamburg, Germany, 1999, pp. 339–341.

[8] ——, "An information theoretic comparison of projection pursuit and principal component features for classification of landsat imagery of central Colorado," *Int. J. Remote Sens.*, vol. 21, pp. 2927–2935, 2000.

[9] C. M. Bachmann, T. F. Donato, K. Du Bois, R. A. Fusina, M. Bettenhausen, J. H. Porter, and B. R. Truitt, "Automatic land-cover classification of a barrier island in the Virginia coast reserve using HyMAP imagery: An intercomparison of methods," in *Proc. IGARSS*, Sydney, Australia, 2001.

[10] ——, "Automatic detection of an invasive plant species on a barrier island in the Virginia coast reserve using HYMAP and IKONOS imagery," in *Proc. IGARSS*, Sydney, Australia, 2001.

[11] C. M. Bachmann, T. F. Donato, R. A. Fusina, and O. Weatherbee, "A methodology for automated land-cover classification in multi-sensor imagery of coastal environments," *IEEE Trans. Geosci. Remote Sensing*, submitted for publication.

[12] C. M. Bachmann, R. A. Fusina, and T. F. Donato, "Effects of time series imagery on automated classification of wetland environments using projection pursuit methods," in *Proc. IGARSS*, Sydney, Australia, 2001.

[13] J. Boardman, "Post-ATREM polishing of AVIRIS apparent reflectance data using EFFORT: A lesson in accuracy versus precision," in *Summaries of the 7th Annu. JPL Airborne Geoscience Workshop*, Pasadena, CA, 1998.

[14] L. Bruzzone and S. B. Serpico, "A technique for feature selection in multiclass problems," *Int. J. Remote Sens.*, vol. 21, pp. 549–563, 2000.

[15] W. H. Duncan and M. B. Duncan, *The Smithsonian Guide to Seaside Plants of the Gulf and Atlantic Coasts from Louisianna to Massachusetts, Exclusive of Lower Peninsular Florida*. Washington, DC: Smithsonian Institution Press, 1987.

[16] J. E. Dobson, E. A. Bright, R. L. Ferguson, D. W. Field, L. L. Wood, K. D. Haddad, H. Iredale III, J. R. Jensen, V. V. Klemas, R. J. Orth, and J. P. Thomas, "NOAA Coastal Change Analysis Program (C-CAP): Guidance for Regional Implementation," U.S. Department of Commerce, Seattle, WA, NOAA Tech. Rep. NMFS 123, 1995.

[17] ESSI [Online]. Available: http://www.earthsearch.com

[18] G. P. Fleming, P. P. Coulling, D. P. Walton, K. M. McCoy, and M. R. Parrish, "The natural communities of Virginia: Classification of ecological community groups. First approximation," Div. Natural Heritage, Dept. of Conservation and Recreation of the State of Virginia, Tech. Rep. 01-1, 2001.

[19] T. E. Flick, L. K. Jones, R. G. Priest, and C. Herman, "Pattern classification using projection pursuit," *Pattern Recognit.*, vol. 23, no. 12, pp. 1367–1376, 1990.

[20] J. H. Friedman and J. W. Tukey, "A projection pursuit algorithm for exploratory data analysis," *IEEE Trans. Comput.*, vol. C-23, pp. 881–890, 1974.

[21] J. H. Friedman, "Exploratory projection pursuit," *J. Amer. Stat. Assoc., Theory and Methods*, vol. 82, pp. 249–266, 1987.

[22] M. Garcia and S. L. Ustin, "Detection of interannual vegetation responses to climactic variability using AVIRIS data in a coastal savanna in California," *IEEE Trans. Geosci. Remote Sensing*, vol. 39, July 2001.

[23] D. N. H. Horler and F. J. Ahern, "Forestry information content of Thematic Mapper data," *Int. J. Remote Sens.*, vol. 7, pp. 405–428, 1986.

[24] P. J. Huber, "Projection pursuit," *Ann. Stat.*, vol. 13, pp. 435–475.

[25] A. Ifarraguerri and C. I. Chang, "Multispectral and hyperspectral image analysis with projection pursuit," *IEEE Trans. Geosci. Remote Sensing*, vol. 38, pp. 2529–2538, Nov. 2000.

[26] N. Intrator and L. N. Cooper, "Objective function formulation of the BCM theory of visual cortical plasticity: Statistical connections, stability conditions," *Neural Networks*, vol. 5, pp. 3–17, 1992.

[27] J. R. Jensen, *Introductory Digital Image Processing: A Remote Sensing Perspective*. Englewood Cliffs, NJ: Prentice-Hall, 1986.

[28] J. R. Jensen, D. J. Cowen, J. D. Althausen, S. Narumalani, and O. Weatherbee, "An evaluation of the coastwatch change detection protocol in South Carolina," *Photogramm. Eng. Remote Sens.*, vol. 59, no. 6, pp. 1039–1046, 1993.

[29] L. O. Jimenez and D. Landgrebe, "High dimensional feature reduction via projection pursuit," School of Elect. Comput. Eng., Purdue Univ., West Lafayette, IN, Tech. Rep. TR-ECE 96-5, 1996.

[30] L. O. Jimenez and D. A. Landgrebe, "Hyperspectral data analysis and supervised feature reduction via projection pursuit," *IEEE Trans. Geosci. Remote Sensing*, vol. 37, pp. 2653–2667, Nov. 1999.

[31] F. A. Kruse, J. W. Boardman, A. B. Lefkoff, J. M. Young, and K. S. Kierein-Young, "The 1999 AIG/HyVista HyMap group shoot: Commercial hyperspectral sensing is here," in *Proc. SPIE Int. Symp. on AeroSense*, Orlando, FL, 2000.

[32] Web Site for the University of Virginia's Long Term Ecological Research Program [Online]. Available: http://www.vcrlter.virginia.edu

[33] M. Maechler, D. Martin, J. Schimert, M. Csoppenszky, and J. N. Hwang, "Projection pursuit learning networks for regression," in *Proc. 2nd Int. IEEE Conf. Tools for Artificial Intelligence*, Herndon, VA, 1990, pp. 350–355.

[34] S. G. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Trans. Signal Processing*, vol. 41, pp. 3397–3415, Dec. 1993.

[35] C. A. McCaffrey and R. D. Dueser, "Plant associations on the Virginia barrier islands," *Virginia J. Sci.*, vol. 41, pp. 282–299, 1990.

[36] S. R. Phinn, D. A. Stow, and D. Van Mouwerik, "Remotely sensed estimates of vegetation structural characteristics in restored wetlands," *Photogramm. Eng. Remote Sens.*, vol. 65, pp. 485–493, 1999.

[37] S. R. Phinn, C. Menges, G. J. E. Hill, and M. Stanford, "Optimising remotely sensed solutions for monitoring, modeling and managing coastal environments," *Remote Sens. Environ.*, vol. 73, pp. 117–132, 2000.

[38] J. H. Porter, B. P. Hayden, and D. L. Richardson, "Information management at the Virginia coast reserve," in *Proc. Eco-Informa'96*, 1996.

[39] New Jersey Marine Sciences Consortium *et al.*, "Published abstracts," in *Phragmites Australis: A Sheep Wolf's Clothing: Technical Forum and Workshop*, Vineland, N.J, Jan. 6–9, 2002.

[40] M. D. Richard and R. P. Lippman, "Neural network classifiers estimate Bayesian a posteriori probabilities," *Neural Comput.*, vol. 3, pp. 461–483, 1991.

[41] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning internal representations by error propagation," in *Parallel Distributed Processing, Explorations in the Microstructure of Cognition*, D. E. Rumelhart and J. L. McClell, Eds. Cambridge, MA: MIT Press, 1986, vol. 1, Foundations, pp. 318–362.

[42] A. H. Strahler, C. E. Woodcock, and J. A. Smith, "On the nature of models in remote sensing," *Remote Sens. Environ.*, vol. 20, pp. 121–139, 1986.

[43] I. H. Stuckey and L. L. Gould, *Coastal Plants from Cape Cod to Cape Canaveral*. Chapel Hill, NC: Univ. North Carolina Press, 2000.

[44] P. A. Siljestrom, A. Moreno, K. Vikgren, and L. M. Caceres, "The application of selective principal components analysis (SPCA) to a Thematic Mapper (TM) image for the recognition of geomorphologic features configuration," *Int. J. Remote Sens.*, vol. 18, pp. 3843–3852.

[45] J. T. Tou and R. C. Gonzalez, *Pattern Recognition Principles*. Reading, MA: Addison-Wesley, 1974.

[46] D. Trizna *et al.*, "Projection pursuit classification of multi-band polarimetric SAR land image scenes," *IEEE Trans. Geosci. Remote Sensing*, vol. 39, no. 11, pp. 2380–2386.

[47] J. H. Tou and C. Reinsch, *Handbook for Automatic Computation*, 1971, vol. II, Linear Algebra.

[48] M. Zhang, S. L. Ustin, E. Rejmankova, and E. W. Sanderson, "Monitoring pacific coast saltmarshes using remote sensing," *Ecol. Appl.*, vol. 7, pp. 1039–1053, 1997.

**Charles M. Bachmann** (M'92) received the Ph.D. and Sc.M. degrees in physics from Brown University, Providence, RI, in 1990 and 1986, respectively, and the A.B degree in physics from Princeton University, Princeton, NJ, in 1984. While at Brown University, he participated in interdisciplinary research in the Center for Neural Science, investigating adaptive models related to neurobiology and to statistical pattern recognition systems for applications such as speech recognition.

In 1990, he joined the Naval Research Laboratory (NRL), Washington, DC as a Research Physicist in the Radar Division, serving as a Section Head in the Airborne Radar Branch between 1994–1996. In 1997, he moved to the Remote Sensing Division, where he worked in the Remote Sensing Hydrodynamics Branch. Since 2001, he has been with the Optical Sensing Section of the Radio, IR, and Optical Sensors Branch. He has been a Principal Investigator for projects funded by the Office of Naval Research, and more recently for an internal NRL project focused on coastal land cover from multisensor imagery. His research interests include basic and applied research in image and signal processing techniques and adaptive statistical pattern recognition methods and the instantiation of these methods in software. Applications have included multispectral and hyperspectral imagery, SAR, and multisensor data, as these apply to environmental remote sensing, especially coastal wetlands. Related interests include field spectrometry and the estimation of biophysical parameters.

He is a member of the American Geophysical Union, the Society of Wetland Scientists, and the Sigma Xi Scientific Research Society. He is the recipient of two NRL Alan Berman Publication Awards (1994 and 1996), and an Interactive Session Paper Prize at IGARSS '96.

**Timothy F. Donato** (M'99) was born in Washington, DC on August 8, 1961. He received the B.S. degree in biology with emphasis in ecology/environmental science and minors in mathematics and chemistry from Christopher Newport College, Newport News, VA, in 1986. Subsequently, he conducted postgraduate work in resource ecology while pursuing his Master's of Environmental Management Degree at Duke University, Durham, NC. He continued his graduate training at North Carolina State University (NCSU), Raleigh, NC, where he received his M.S. degree in physical oceanography, in 1994. He conducted research on the active microwave observations of the Gulf Stream frontal region at low grazing angles in support of a Naval Research Laboratory (NRL), Washington, DC, Advanced Research Initiative. He is also currently pursuing his Ph.D. degree in physical ocean sciences and engineering in the College of Marine Studies, University of Delaware, Newark, where he is also an active member of the Center for Remote Sensing.

He is currently a Geophysicist with the NRL for the Remote Sensing Division. He has been with NRL since 1995, and his current research involves quantitative interpretation and analysis of moderate- to high-resolution (spatial and spectral) satellite imagery (hyperspectral, multispectral and synthetic aperture radar imagery) in coastal environments, continental shelf plankton dynamics, hydrodynamic modeling of the coastal ocean, remote sensing data fusion, and the integration of hydrodynamic models and the landscape/ecosystems analysis of coastal wetlands. Concurrently, with his M.S. work at NCSU, he worked for Science Applications International Corporation (SAIC), Raleigh, NC as their satellite oceanographer. While at SAIC, he conducted work on a variety of coastal and open ocean environmental related projects for Mobile Oil, the Minerals Management Service, and the Environmental Protection Agency. In 1993, he joined Allied Signal Technical Services (now Honeywell) as a Research Scientist, performing work for the Remote Sensing Division at the NRL on the analysis of active microwave back scatter from open ocean environments.

**Gia M. Lamela** received the B.S. degree (with honors) in biological sciences from the University of Maryland, Baltimore County in 2000. She has been with the Naval Research Laboratory, Washington, DC since 1989 and joined the Optical Sensing Section in 1996.

**W. Joseph Rhea** received the B.S. degree in oceanography from the University of Washington, Seattle, WA, in 1986. From 1985 to 1988, he was an Assistant Scientist for the Oceanographic and Meteorological Science Group of Envirosphere Company, Bellevue, WA. From 1988 to 1994, he worked for the Biological and Polar Oceanography Group, Jet Propulsion Laboratory, Pasadena, CA. Since 1994, he has worked for the Optical Sensing Section of the Naval Research Laboratory, Washington, DC.

**Michael H. Bettenhausen** (S'93–M'95) received the B.S., M.S., and Ph.D. degrees in electrical engineering from the University of Wisconsin, Madison, in 1983, 1990, and 1995, respectively. His graduate research focused on theoretical and computational studies of radio frequency heating in plasmas.

He is currently with the Remote Sensing Division, Naval Research Laboratory, Washington, DC. He did software development and algorithm research for particle simulation while employed by Mission Research Corporation, Santa Barbara, CA, from 1997 to 2000. In 2000, he joined Integrated Management Services, Inc., Hershey, PA, where he worked on projects for analysis and processing of hyperspectral remote sensing data and inverse synthetic aperture radar data. His research interests include analysis of hyperspectral remote sensing data and high-performance computing.

**Robert A. Fusina** (M'01) received the B.S. degree from Manhattan College, NY, and the M.S. and Ph.D. degrees from the State University of New York, Albany, all in physics.

He has worked in the Remote Sensing Division, Naval Research Laboratory, Washington, DC, since 1993. His current research involves land cover classification, hyperspectral remote sensing, and data fusion. His previous work included calculation of radar scattering from ocean waves.

**Kevin R. Du Bois** received the M.S. degree from the Marine Science Research Center, State University of New York, Stony Brook, and is certified as a Professional Wetland Scientist.

He is currently employed as an Environmental Engineer with the City of Norfolk, Virginia, Bureau of Environmental Services, Norfolk, VA.

**John H. Porter** received the A.A. degree from Montgomery College, in 1974, the B.S. degree from Dickinson College, Carlisle, PA, in 1976, and the M.S. and Ph.D. degrees from the University of Virginia, Charlottesville, in 1980 and 1988, respectively.

He is currently a Research Assistant Professor in the Department of Environmental Sciences, University of Virginia. He is the Information Manager and one of the three lead Principal Investigators of the Virginia Coast Reserve Long-Term Ecological Research Project.

**Barry R. Truitt** was born October 8, 1948 in Norfolk, VA. He received the B.S. degree in biology from Old Dominion University, Norfolk, VA, in 1971.

He is currently Chief Conservation Scientist, responsible for the design and implementation of site conservation plans, research, and biological monitoring. He has been employed by The Nature Conservancy since 1976 at the Virginia Coast Reserve. His main professional interests include island biogeography, landscape ecology, conservation science, and marine and migratory bird conservation. He conducts and coordinates with other partners a 28-year-long colonial waterbird and shorebird monitoring program on the seaside. He is also involved in efforts to restore eelgrass and oyster reefs in the coastal bays. His interest in landscape ecology and barrier island history led to the publication, with Miles Barnes, of a book entitled "*Seashore Chronicles, Three Centuries of the Virginia Barrier Islands*"in 1998.