Automated Vegetation Mapping Using Digital Orthophotography

Roland J. Duhaime, Peter V. August, and William R. Wright

Abstract

We used near-infrared digital orthophotography and three collateral data sets to model ecological communities on Block Island, Rhode Island. Aerial photography of the island was taken on 19 May 1992 at a scale of 1:40,000. The photography was scanned and processed to remove distortions from terrain, aircraft tilt, and optical aberration. The resulting digital orthophotograph was comprised of three spectral bands representing the red, green, and blue colors of the scanned photography and had a pixel dimension of 1.27 m. Three textural variables were developed by calculating the standard deviation within a 10-m radius of every pixel for each of the three spectral bands in the image. The terrain model that was used to create the orthophoto was also used to derive slope and aspect for each pixel. Soil survey data were used to map the distribution of soil drainage classes to distinguish wetland from upland vegetation. We used linear discriminant analysis to develop a model to distinguish 11 vegetation and cover classes on the island. The full model consisted of nine independent variables derived from the orthophoto, the textural indices, terrain metrics, and soils. Classification accuracies ranged from 60 to 80 percent for an independent validation data set. The variable DRAINAGE CLASS dominated the model and explained the most variation in vegetation and cover class.

Introduction

Digital orthophotographs are photographic images from which distortion due to terrain, aircraft tilt, and optical aberration has been removed (Muehrcke and Muehrcke, 1992; Steiner, 1992). Typically, digital orthophotos are derived from aerial photographs that have been electronically scanned at a very high resolution, usually on the order of 25 micrometres (DeAngelis, 1993). Planimetric distortions are differentially rectified using a digital terrain model (DTM) and accurate ground control for the region. The resolution of an orthophotograph is determined by the scale of the photography and the scanning density. Large-scale orthophotos can easily have sub-metre pixel dimensions (Logan, 1993; Jadkowski *et al.*, 1994).

Using geographic information systems (GIS), digital orthophotography can be integrated with other cartographic data (Nellis *et al.*, 1990). Orthophotos can serve as an excellent reference backdrop when viewing other data, or they can provide new information that can be "heads-up" digitized into a GIS database (August, 1993). Because of their accuracy and overall utility in a GIS environment, digital orthophotos are emerging as a standard planimetric base for state or region-wide GIS systems (Parent, 1991; Jadkowski *et al.*, 1994; Nale, 1995). For example, many New England states (Vermont, Massachusetts, Connecticut) are developing large-scale, black-and-white digital orthophoto databases in conjunction with statewide GIS mapping efforts. The state of Maryland has developed statewide color infrared digital orthophotography to support wetland inventory and mapping (Logan, 1993). The U.S. Geological Survey Digital Orthophoto Quad (DOQ) project is an ambitious endeavor to produce 1:12,000-scale digital orthophotos for the entire nation (Ramey, 1992; DeAngelis, 1993; Light, 1993).

Digital orthophotography and satellite imagery are similar in that both are cell-based data with numeric values associated with each cell that represent the amount of light in a certain range of the electromagnetic spectrum that is reflected from that area on the ground (Wickland, 1991). Satellite data have been an excellent source of information for identifying land-cover and vegetation communities (Jensen, 1986; Roughgarden *et al.*, 1991; Lee and Marsh, 1995). A number of space-borne sensors provide data on the Earth's surface that have been used extensively in mapping applications. A thorough list of sensors, their spatial resolution, and spectral sensitivities are provided in Wickland (1991) and Jensen (1986). The NOAA AVHRR satellite (1-km pixel size), Landsat Thematic Mapper (30-m pixel size), and SPOT satellite (10-m pixel size) are frequently used for land-cover assessments (Wickland, 1991). Publicly available satellite data are well-suited for analysis of land cover and land-cover change over large areas (Zeff and Merry, 1993; Green *et al.*, 1994; Lee and Marsh, 1995). The spatial resolution of these data is, however, too coarse for most local or site-specific applications (Brannon *et al.*, 1994). In contrast, digital orthophotography can be produced at a very fine-grained resolution and should, in concept, be an excellent source data for large-scale land-cover information. The purposes of our research were, therefore, two-fold: (1) to evaluate the utility of using digital image processing on a color-infrared orthophoto for identification of ecological communities, and (2) to assess the importance of using collateral digital data (Hutchinson, 1982; Harris and Ventura, 1995) to enhance classification accuracy. We use fundamental methods developed by the remote sensing community (Jensen, 1986; Lillesand, 1994) to determine accuracy levels of a supervised classification of vegetation and cover types on Block Island. The focus of our study is to determine the utility of digital orthophotography to provide vegetation and cover information, especially in a GIS context.

Methods

Study Area

Block Island is on a terminal moraine 15.1 km south of mainland Rhode Island and 22.5 km northwest of Montauk

Department of Natural Resources Science, University of Rhode Island, Kingston, RI 02881.

Photogrammetric Engineering & Remote Sensing, Vol. 63, No. 11, November 1997, pp. 1295–1302.

0099-1112/97/6311-1295$3.00/0

Point, New York. The latitude and longitude coordinates of the center of the island (the State Airport) are 41° 10' 00" N and 71° 35' 00" W (Figure 1). The island is 25.7 km² and consists of gently rolling hills, dunes, and bluffs. The highest elevation is 62.8 m above sea level. The dominant land covers are shrublands of various types, fields and pasture land, freshwater and saltwater wetlands, residential development, and a small commercial district in the town of New Shoreham.

The parent geological material for Block Island is unconsolidated glacial outwash and till of Pleistocene and Cretaceous age (Hansen and Schiner, 1964). The dominant soil map units are Gloucester-Bridgehampton complex and Gloucester-Hinckley very stony sandy loam (Rector, 1981).

The vegetation of Block Island consists of dense shrublands of shadbush (*Amelanchier canadensis*), bayberry (*Myrica pennsylvanica*), and arrowwood (*Viburnum* spp.) (Dunwiddie, 1990). Japanese black pine (*Pinus thunbergii*) and red pine (*Pinus resinosa*) are the common tree species found on Block Island, but they are not native. Few forests have developed because the strong prevailing winds and salt spray severely hinder tree growth (Rector, 1981). Mowing or grazing maintains the island's pastures. Because of its unique landscape and biota, The Nature Conservancy has included Block Island in its world list of the "Last Great Places" (Ericson, 1992).

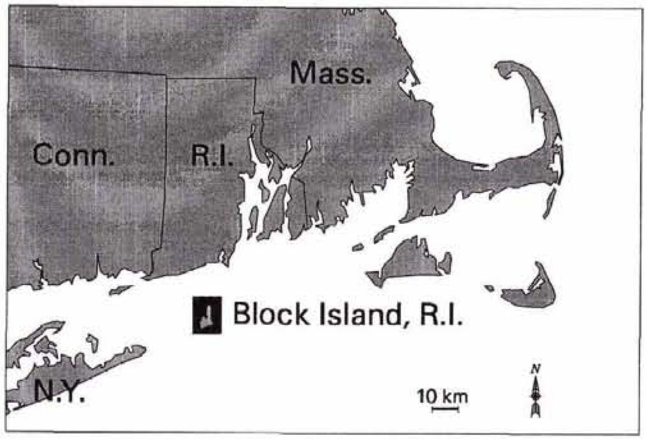

Figure 1. Location of Block Island off the southern coast of Rhode Island.

Digital Imagery

Color infrared aerial photography was taken of Block Island on 19 May 1992 at 11:35 AM. The photography was recorded on Kodak 2443 Aerochrome film at a scale of 1:40,000 from an altitude of approximately 6.1 km. A wratten filter was used on the camera to filter out blue light. Two overlapping stereo photos were scanned at a density of 315 lines/centimetre and the digital representation was processed to create a seamless image of the island (see cover). A digital terrain model (DTM) was developed from data obtained by a team of surveyors sent to the island to densify the ground control for the project. Twenty-five control points were used to rectify the orthophotography and the DTM (Novak, 1992). The DTM consists of elevation values spaced 90 m apart throughout the image and was produced to meet USGS Level 1 DEM standards (U.S. Geological Survey, 1987); the resulting RMSE was 2.42 m. The data were referenced to the vertical datum of NGVD 29 and a horizontal datum of NAD 1983. The orthophoto was created from the scanned imagery and DTM using a SUN Sparc Station 2, with algorithms developed by International Imaging Systems (Milpitas, California). Photography, scanning, and orthorectification were performed by Photo Science Inc., of Gaithersburg, Maryland. Surveying and ground control were established by Cherenzia & Associates Ltd. of Westerly, Rhode Island. The complete image is 4,591 by 7,754 pixels (106.8 Mb) in size where each pixel has a resolution of 1.27 m.

We field-checked the positional accuracy of four well-defined points on the imagery with a Trimble Pathfinder Basic GPS unit using averaging and differential correction (Welch *et al.*, 1992; August *et al.*, 1994). All of the points we tested fell within a single pixel of their true position.

The digital infrared image is stored with three bands of data. *BAND1* represents the infrared portion of the reflected light (spectral sensitivity peak at 740 nm). The visible red portion of the electromagnetic spectrum is represented in *BAND2* (spectral peak at 650 nm). *BAND3* represents the area of the electromagnetic spectrum that the eye perceives as green (spectral peak at 550 nm) (Holz, 1985). The use of the wratten filter limited the exposure of the film to the green, red, and infrared portion of the electromagnetic spectrum; any light with a wavelength less than 530 nm was filtered out and appeared black on the photograph.

Vegetation Communities

We identified 11 vegetation and cover types of six major classes on the island for use in the classification. The vegetation classes were chosen because they represent the dominant natural/semi-natural ecological communities on the island, or have been found to be significant habitat for rare species of plants and animals. Because the objective of this research was to assess the utility of color infrared orthophotography to discern specific vegetation types, rather than all possible cover types, we merged into a single OTHER class those unvegetated land covers (e.g., rocky shores) or those resulting from human disturbance (roads, buildings, pavement). The vegetation and cover types we identified in our analyses are as follows:

- SAND consists of gravel pits or beaches with little or no vegetation.
- DUNES are hills of wind-blown sand. This community is much like the sand category except that it is dominated by graminoids (*Ammophila* sp.).
- PASTURES and grass fields are farmed lands consisting of low herbaceous vegetation that is mowed or grazed.
- UPLAND PINE SHRUBLANDS occur as small patches of coniferous shrubs. Individual plants are less than 6 m tall. Japanese Black Pine (*Pinus thunbergii*) is the dominant species in this habitat.
- UPLAND DECIDUOUS SHRUBLANDS are the most common cover type on the island and consist of dense stands of shadbush (*Amelanchier* spp.), bayberry (*Myrica* sp.), and arrowwood (*Viburnum* spp.).
- FRESHWATER SHRUB WETLANDS are comprised of arrowwood (*Viburnum* spp.) and rosaceous shrubs in very poorly drained soils.
- FRESHWATER MARSHES are areas that are dominated by hydrophytic perennials, graminoids, and tussock sedges. Cattail (*Typha* spp.) is a dominant species in these communities. Water willow (*Decodon verticillata*) may also occupy a large percentage of these areas.
- SALT MARSH is characterized by perennial, salt tolerant graminoids, dominated by *Spartina alterniflora*, *Spartina patens*, and *Distichlis spicata*. These perennial species are present for most of the growing season.
- OCEAN covers a large area of the imagery and consists of standing water with no rooted vegetation.
- FRESHWATER includes palustrine and lacustrine open water.
- OTHER is a miscellaneous category comprised of roads, roofs of buildings, paved areas, or rocky shores.

Initially, we included a number of other vegetation classes of local ecological importance (e.g., morainal grassland) in the analyses (Duhaime, 1994). However, the areas encompassed by these additional vegetation classes were very small and it was difficult to obtain a complete sample of pixels from which we could model their distribution. Therefore, we limited the present analysis to the major vegetation classes listed above.

Classification Protocol

Because we had an *a priori* interest in specific vegetation classes, we used a supervised classification protocol for the study (Jensen, 1986; Lillesand and Kiefer, 1994). Training sites were discerned by making field visits to the island and locating representative areas of each of the vegetation types. The limits of each training area were entered into a vector GIS file using "heads-up" digitizing (i.e., by digitizing from the computer monitor) from the digital orthophotograph. The vector polygons were converted to raster format (1.27-m pixel size) and 1,000 pixels of each vegetation or cover type were randomly selected from the raster data set and used in developing the classification model. A second data file of "validation" points (Oreskes *et al.*, 1994) was created in the same manner as described above for the training data. Areas on the island with representative patches of the vegetation types were found. The boundaries for these sites were digitized into the GIS in vector format and converted into raster structure. A sample of 50 pixels of each vegetation or cover type was randomly selected for validation analysis. Habitat patches used in the validation analyses did not overlap with patches used to obtain classification data. All GIS data processing conducted in this study was done using vector and GRID modules of ARC/INFO software (Environmental Systems Research Institute, Redlands, California).

Discriminant analysis was used to develop a linear model that best distinguished each vegetation type (Jobson, 1992; Sokal and Rohlf, 1981; Williams, 1983). The assumptions of discriminant analysis are equality of dispersions (homoscedasticity), identifiability of prior probabilities, and precise and accurate estimation of means and dispersions (Williams, 1983). The distributions for each independant variable in the classification data set are not homoscedastic (Table 1), but meet the other assumptions of the method. Moreover, linear discriminant analysis is a robust statistic and performs reliably even when data do not meet the requisite assumptions (Lachenbruch, 1975).

The dependent variable for the model was vegetation/cover type and we used nine independent variables that were based upon ecological measurements or spectral properties of the image. The three image variables (*BAND1*, *BAND2*, *BAND3*) represent reflected light in the aforementioned wavelength classes. The amount of reflected light in each band is indicated in the data set by values ranging from 0 to 255. To estimate the microspatial variation in reflectance, we created a textural index (Briggs and Nellis, 1991; Musick and Grover, 1991; Benallegue *et al.*, 1995) for each of the three bands by calculating the standard deviation of reflectance in a 10-m (approximately 8 pixels radius) floating window around each pixel. Low textural values represent very little spatial variation in reflectance; for example, the center of a hay field would be quite homogeneous in reflectance of neighboring pixels. Large textural values are indicative of edge habitats or vegetation communities that exhibit considerable micro-variation in cover; for example, an upland shrubland is a mosaic of grass and shrubs. Such a cover type would appear mottled in the image and would yield high variation in reflectance. The textural variables for *BAND1*, *BAND2*, and *BAND3* are called *TEXT1*, *TEXT2*, and *TEXT3*, respectively.

The abiotic variables we used were slope, aspect, and depth to the seasonally high water table. Slope and aspect were derived directly from the DTM of the island. The digital elevation model was created by interpolating the 90-m DTM to a resolution of 1.27 m using an inverse distance weighted algorithm (ESRI, Redlands, California). Slope (degrees) was calculated for each pixel using the SLOPE utility of the GRID module of ARC/INFO. Vegetation communities tend to be more sensitive to north/south variation in aspect than east/west changes (Barbour *et al.*, 1987); therefore, we identified pixels that faced north or south. Flat areas were pooled with the north-facing (aspect $> 270°$ and $< 90°$) pixels and assigned a value of 1. South facing pixels ($\geq 90°$ and $\leq 270°$) were assigned a value of 0.

Depth to the seasonally high water table was extracted from the digital version of the Rhode Island Soil Survey (base scale 1:15,840; Rector, 1981). Because of the inherent error in defining "fuzzy" transitional features such as soil boundaries, only training pixels that fell further than 10 m from the edge of a soil polygon were used in the analysis. We categorized the soils data into five classes of soil moisture, from very dry (water greater than 91.4 cm from surface) to very wet. The wettest category was permanent standing water (Table 2). Soil moisture was included in the model to help discern upland from wetland habitats.

To develop the discriminant function that best distinguished vegetation types, we created an ASCII file that contained, for each pixel within all training sites, the vegetation type and the values for each of the nine independent variables. Each record in the file represented a single pixel in the

TABLE 1. DESCRIPTIVE STATISTICS (MEAN ± 1 SD) FOR CONTINUOUS VARIABLES AMONG VEGETATION AND COVER TYPES. ONE WAY ANALYSIS OF VARIANCE (ANOVA) WAS USED TO TEST THE NULL HYPOTHESIS THAT CLASSIFICATION VARIABLES WERE THE SAME AMONG VEGETATION TYPES. N = 1,000 FOR ALL VEGETATION CLASSES.

	Classification Variables						
Vegetation type	*BAND 1*	*BAND 2*	*BAND 3*	*SLOPE* (Degrees)	*TEXT1*	*TEXT2*	*TEXT3*
SAND	197.7 ± 19.7	184.8 ± 19.1	189.8 ± 16.3	2.4 ± 3.9	569.8 ± 799.6	419.3 ± 830.7	304.7 ± 635.8
DUNE	167.9 ± 23.3	148.7 ± 26.8	162.9 ± 24.4	1.3 ± 3.8	175.2 ± 132.8	306.5 ± 242.9	224.0 ± 202.6
PASTURE	171.1 ± 9.1	64.9 ± 20.1	98.1 ± 20.4	3.6 ± 4.4	128.1 ± 257.0	228.8 ± 269.8	208.1 ± 195.0
UPLAND PINE SHRUB	66.8 ± 30.2	18.7 ± 9.6	34.1 ± 10.2	5.5 ± 7.5	1075.9 ± 572.5	262.1 ± 514.7	250.1 ± 465.2
UPLAND DECIDUOUS SHRUB	110.9 ± 35.2	37.8 ± 19.2	56.6 ± 21.0	4.1 ± 8.0	555.9 ± 394.8	216.5 ± 264.9	246.8 ± 251.7
FRESHWATER SHRUB WET	40.7 ± 28.8	16.6 ± 10.1	36.4 ± 11.2	2.7 ± 3.8	591.2 ± 301.0	98.4 ± 84.5	123.7 ± 101.7
OCEAN	28.9 ± 43.1	37.0 ± 43.3	62.5 ± 41.8	0 ± 0	503.6 ± 867.0	398.6 ± 681.2	298.1 ± 499.2
FRESHWATER	169.3 ± 50.5	156.5 ± 45.8	164.0 ± 42.0	1.5 ± 3.6	86.0 ± 328.5	31.1 ± 55.1	31.5 ± 78.1
SALT MARSH	127.7 ± 29.1	103.3 ± 30.8	122.0 ± 27.1	.7 ± 1.7	832.3 ± 664.4	762.0 ± 497.4	588.5 ± 381.7
FRESHWATER MARSH	63.5 ± 35.6	48.2 ± 28.6	67.2 ± 26.2	2.2 ± 4.7	524.1 ± 328.2	300.9 ± 189.9	256.5 ± 169.9
OTHER	82.1 ± 58.9	81.7 ± 53.7	104.1 ± 46.9	.9 ± 3.5	1042.4 ± 1132.1	879.7 ± 889.5	653.4 ± 651.3
ANOVA F statistic	2,735	3,579	3,559	125	314	271	224
p	< 0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.0001

TABLE 2. DISTRIBUTION OF TRAINING CELLS AMONG VEGETATION AND COVER TYPES FOR *ASPECT* AND *DRAINAGE CLASS*. WE USED A LOG-LIKELIHOOD TEST (G TEST) TO EVALUATE THE NULL HYPOTHESIS THAT THE DISTRIBUTIONS AMONG VEGETATION AND COVER TYPES WERE THE SAME FOR EACH VARIABLE. VALUES PRESENTED ARE THE PERCENT OF THE TRAINING PIXELS (N = 1,000) USED FOR EACH VEGETATION TYPE.

		DRAINAGE CLASS					
Vegetation Type	% of South Facing Slope (*ASPECT*)	0 Variable	1 Well Drained to Excessively Well Drained	2 Moderately Well Drained	3 Poorly Drained	4 Very Poorly Drained	5 Open Water
SAND	40	94	6	0	0	0	0
DUNE	66	100	0	0	0	0	0
PASTURE	29	1	96	3	0	0	0
UPLAND PINE SHRUBLAND	48	0	100	0	0	0	0
UPLAND DECIDUOUS SHRUBLAND	45	1	99	0	0	0	0
FRESHWATER SHRUB WETLAND	0	0	6	0	29	65	0
OCEAN	9	0	0	0	0	0	100
FRESHWATER	53	0	3	0	0	0	97
SALT MARSH	90	13	2	0	0	85	0
FRESHWATER MARSH	23	0	4	5	11	80	0
OTHER	4	91	9	0	0	0	0
G	3,758	27,316					
p	< *0.0001*	< *0.0001*					

training data set. These data were analyzed with PROC STEPDISC, CANDISC, DISCRIM, and CANCORR routines in the software package PC-SAS (Cary, North Carolina) to calculate discriminant scores and classification matrices.

We used the "validation" data set to assess the quality of the model. The known vegetation type from the validation data set was compared to the predicted type derived from the discriminant model. The KHAT statistic (Congalton, 1991) was used to test the null hypothesis that the distribution of observations within a classification was random. KAPPA analyses were used to test the null hypothesis that any two classification matrices had equal distributions of observations (Congalton, 1991; Fitzgerald and Lees, 1994).

GIS computations were done using a Data General 5220 workstation. Statistical analyses were done using a 486-66 MHz microcomputer.

Results

There are significant differences (ANOVA, $p < 0.0001$) among the 11 land-cover types for all the classification variables (Table 1). Inspection of the SAND, DUNE, and FRESHWATER categories show high mean reflectance for each of the three spectral variables. The mean texture values for SALT MARSH were very high for all textural variables. The variable TEXT1 was very large for UPLAND PINE SHRUBLAND. Overall, OTHER had the largest mean textural values and was consistently the most variable of all the measures.

There were significant differences among cover classes for each of the categorical variables; however, some inconsistencies were found (Table 2). Areas presumed to be flat had relatively high frequencies of south facing aspect. For example, 9 percent of OCEAN category contained south facing pixels. However, the mean slope for OCEAN was zero (Table 1); thus, the oceanic pixels facing south are nearly flat or very shallow in slope. The course resolution of the original terrain model probably accounts for some of the non-zero slope pixels in the OCEAN category as well.

Greater than 90 percent of the training cells from upland vegetation were in the excessively drained to well drained categories. Most of the wetland classes were classified in the very poorly drained and poorly drained soils (Table 2). The three spectral bands were highly correlated among themselves but were, for the most part, uncorrelated with the other independent variables (Table 3). Likewise, the three textural variables were correlated among themselves, but much less so with the spectral, terrain, and soils data.

The Basic Model

The canonical structure of the full model shows the variables that have the strongest discriminatory power. The first ca-

TABLE 3. SPEARMAN RANK COEFFICIENTS OF CORRELATION AMONG INDEPENDENT VARIABLES. (N = 11,000) N.S. = P > 0.05, * = P ≤ 0.05, ** = P < 0.001, *** = P < 0.001

	BAND2	BAND3	TEXT1	TEXT2	TEXT3	SLOPE	ASPECT	DRAINAGE CLASS
BAND1	0.90 ***	0.88 ***	−0.24 ***	0.09 ***	0.02 n.s.	0.09 ***	−0.24 ***	−0.33 ***
BAND2		0.98 ***	−0.16 ***	0.23 ***	0.12 ***	−0.15 ***	−0.20 ***	−0.28 ***
BAND3			−0.22 ***	0.21 ***	0.10 ***	−0.18 ***	−0.19 ***	−0.27 ***
TEXT1				0.72 ***	0.73 ***	0.13 ***	−0.01 n.s.	−0.16 ***
TEXT2					0.97 ***	−0.10 ***	−0.10 ***	−0.29 ***
TEXT3						−0.06 ***	−0.09 ***	−0.25 ***
SLOPE							−0.01 ***	−0.17 ***
ASPECT								0.05 ***

Table 4. Standardized Canonical Coefficients of the First Four Canonical Variates (CV1, CV2, CV3, and CV4) Using all Training Data and all 11 Vegetation and Cover Types

	CV1	*CV2*	*CV3*	*CV4*
BAND1	−0.64	−3.36	1.90	−1.33
BAND2	0.55	5.11	−6.26	−8.59
BAND3	−0.43	−0.52	5.64	9.08
TEXT1	0.10	−0.01	−0.44	−0.80
TEXT2	−1.06	0.62	0.15	2.62
TEXT3	0.93	−0.67	0.00	−1.64
SLOPE	−0.02	−0.14	−0.05	−0.14
ASPECT	0.07	−0.17	−0.23	0.26
DRAINAGE CLASS	3.21	−0.22	0.82	−0.33
Eigenvalue	11.02	5.41	2.21	1.73
Proportion of Variance Explained	.53	.26	.11	.08

nonical variate (CV1) is dominated by the variable *DRAINAGE CLASS* (Table 4) and explains 53 percent of the total variance among vegetation types. CV2 is dominated by strong coefficients for *BAND1* and *BAND2*. The textural variables *TEXT2* and *TEXT3* also contributed to CV2. CV1 and CV2 together account for 79 percent of the variance among vegetation types. CV3 is dominated by *BAND1*, *BAND2*, and *BAND3*, and the variables *BAND2*, *BAND3*, *TEXT2*, and *TEXT3* contributed most to CV4. The first four canonical variates account for 98 percent of the total variation among vegetation classes. ASPECT and SLOPE contributed very little to the discriminatory power of the model.

Accuracy Assessment

We used a validation data set that was gathered independently of the training areas to confirm the accuracy of the classification model. The overall accuracy was 60 percent (Table 5). Most misclassifications occurred among spectrally similar cover classes or physiognomically similar vegetation types. For example, FRESHWATER was classified as SAND almost half of the time; both are highly reflective surfaces. FRESHWATER MARSH was frequently misclassified into SALT MARSH. Upland shrub habitats were regularly misclassified between UPLAND PINE SHRUBLAND and UPLAND DECIDUOUS SHRUBLAND.

Generalizing the Model

Misclassification of the validation data commonly occurred among similar habitat types. To see if we could improve the classification accuracy of the model, we merged similar vegetation and cover types into six major classes: SAND (DUNE, SAND); PASTURE; SHRUB (UPLAND PINE, UPLAND SHRUB, FRESHWATER SHRUB); WATER (OCEAN, FRESHWATER); MARSH (SALT MARSH, FRESHWATER MARSH); and OTHER. The discriminant model based on the six generalized vegetation classes yielded a total correct classification rate of 80 percent (Table 6). By merging classes, we markedly improved the classification accuracy of the model but lost thematic detail.

We performed a sensitivity analysis on the six generalized vegetation classes to determine the performance of the model when selected suites of variables were eliminated from the training data set (Table 6). When the *DRAINAGE CLASS* variable was dropped from the model, there was a 9 percent decrease in pixels being correctly classified in the validation data set. When just the three imagery variables (*BAND1*, *BAND2*, *BAND3*) were used to create the model, there was only a 61 percent correct classification of validation data. To determine what would happen if we had only one band of data (e.g., what would be available from black-and-white orthophotography) as opposed to the three bands available for color infrared imagery, pairs of bands were dropped from the model along with their respective textural variables. Regardless of the single band chosen, approximately 70 percent of the validation pixels were correctly classified as compared to 80 percent correct classification for the full model. The classification accuracy obtained using terrain and soils data without spectral information was comparable to the classification accuracy obtained using spectral data without soil or terrain information as collateral data (approximately 60 percent).

Discussion

The overall classification accuracy of the six vegetation/cover classes using validation data was 80 percent. Only 60 percent of the validation data were correctly classified when 11 classes of vegetation or land cover were considered. Our classification accuracies are, at best, of modest proportion. They are, however, within the accuracy range (50 to 80 percent overall correct classification) reported in other studies of cover classification using satellite or multispectral imagery (Todd *et al.*, 1980; Stenback and Congalton, 1990; Herr and Queen, 1993; Kremer and Running, 1993; Lauver and Whistler, 1993; Rutchey and Les Vilcheck, 1994).

The percent of south-facing pixels was higher than we expected in areas we knew to be predominantly flat; e.g., FRESHWATER (53 percent), FRESHWATER MARSH (23 percent), and SALT MARSH (14 percent). Many of the ponds and

Table 5. Classification Matrix of the Validation Dataset. Discriminant Functions Were Derived from the Training Dataset. Values Presented Are the Percent of the Validation Samples (N = 50) for each Vegetation Type. Mean Classification Accuracy: 60.2 Percent, Overall Classification Accuracy: 60.2 Percent, KHAT = 0.56 $p < 0.001$.

	Observed Vegetation Type										
Predicted Vegetation Type	DUNE	FRESH WATER	FRESH WATER MARSH	FRESH WATER SHRUB WETLAND	OCEAN	OTHER	PASTURE	SALT MARSH	SAND	UPLAND PINE SHRUB-LAND	UPLAND DECIDUOUS SHRUBLAND
DUNE	**72**	0	0	0	0	14	0	0	14	0	0
FRESHWATER	0	**50**	0	0	42	4	0	2	2	0	0
FRESHWATER MARSH	0	0	**40**	28	8	2	0	2	0	8	12
FRESHWATER SHRUB WETLAND	0	0	2	**32**	0	0	0	10	0	30	26
OCEAN	0	8	0	0	**92**	0	0	0	0	0	0
OTHER	4	2	2	4	0	**64**	0	8	12	4	0
PASTURE	0	0	0	0	0	0	**96**	0	0	0	4
SALT MARSH	8	12	22	2	0	4	0	**50**	0	2	0
SAND	2	42	0	0	0	0	0	0	**56**	0	0
UPLAND PINE SHRUBLAND	0	0	0	0	0	0	2	0	0	**54**	44
UPLAND DECIDUOUS SHRUB	2	0	0	2	0	10	8	0	0	22	**56**

TABLE 6. SENSITIVITY ANALYSIS OF THE CLASSIFICATION MODEL. WHEN SPECTRAL BANDS WERE DROPPED FROM THE FULL MODEL, THE CORRESPONDING TEXTURE VARIABLES WERE ALSO DROPPED. KAPPA ANALYSIS INDICATES THAT PAIRS OF MODELS THAT DIFFER MORE THAN 5 PERCENT IN TOTAL PERCENT CORRECT CLASSIFICATION ARE SIGNIFICANTLY DIFFERENT ($P < 0.05$).

Model	Number Of Independent Variables	Percent Correctly Classified PASTURE	SHRUB	MARSH	SAND	WATER	OTHER	Total % Correctly Classified
1) Full Model	9	94.0	74.0	70.0	74.0	100.0	68.0	80.0
2) No Drainage Variable	8	94.0	64.0	72.0	80.0	46.0	70.0	71.0
3) Only BAND 1, 2, 3	3	96.0	56.0	66.0	78.0	28.0	40.0	60.7
4) No BAND 2 and BAND 3	5	88.0	66.0	70.0	84.0	100.0	30.0	73.0
5) No BAND 1 and BAND 3	5	58.0	54.0	70.0	76.0	100.0	56.0	69.0
6) No BAND 1 and BAND 2	5	76.0	66.0	70.0	78.0	100.0	52.0	73.7
7) Slope, Aspect, Drainage	3	68.0	40.0	70.0	42.0	100.0	24.0	57.3

marshes on the island are kettle holes from the last glaciation (Wright and Sautter, 1988; Hansen and Shiner, 1964) and are bounded by gently sloping hills. The coarse resolution of the DTM produced slopes extending into the edges of ponds and marshes. This artificially inflated the slope and aspect values for marsh and pond vegetation types.

The classification matrices indicated that misclassified pixels were being assigned into logical but erroneous categories. The model did not always distinguish between subtly different types of vegetation or among cover classes that had similar spectral signatures. Pettinger (1982) found most omission and commission errors occurred in sagebrush/perennial-grass classes. We had similar results within the WETLAND and SHRUB classes.

DRAINAGE CLASS was an extremely important variable in our analyses and distinguished the water and wetland habitats from the upland habitats. Likewise, Levine *et al.* (1994) found that drainage class is the critical soil property controlling vegetation composition and productivity in northern deciduous forests. Other collateral data only slightly improved the model, as has been observed in other studies (Hutchinson, 1982; Janssen *et al.*, 1990). Slope and aspect explained only a small proportion of the variation among vegetation classes. Janssen *et al.* (1990) and Gemmell (1995) found that terrain data significantly improved land-cover identification; however, other studies have found that slope and aspect do not significantly improve overall classification accuracy (Lowell, 1990; Davis and Dozier, 1990). The terrain data used in this study had a source resolution of 90 m. To match the resolution of the other data sets, we interpolated the DTM to a raster grid with a resolution of 1.27 m. It is not clear why the terrain variables failed to contribute to the model. It is possible that the difference in resolution between the DTM and the imagery (90-m pixel versus 1.27-m pixel) rendered too coarse a representation of slope and aspect to provide significant discriminatory power. For example, boundaries between open water and the adjacent landscape were not used as breakline features when creating the DTM; therefore, cells near the edge of ponds and marshes could have been assigned slope and aspect values when they should have been flat. Furthermore, the Block Island landscape consists of gently rolling hills. There may be insufficient variation in slope and aspect to be of ecological importance to the plant communities of the island.

Eliminating variables from the model caused a decrease in correct classification of validation cells. By dropping two spectral bands from the model, we approximated the information content of a single band of 8-bit, black-and-white imagery as might be obtained from scanning black-and-white photos. There was approximately a 10 percent loss in classification accuracy in the single-band models. If this is acceptable, then less expensive black-and-white film could effectively be used for vegetation mapping instead of the more expensive color infrared imagery. Our single-band data sets do not, however, include reflected light from the blue portion of the spectrum which might contain ecologically important information.

Conclusions

Digital infrared orthophotography can be used as an alternative to satellite data for assessing vegetation and cover in a region. Orthophotography can have much finer spatial resolution as compared to satellite data, and (with appropriate collateral data) produces habitat classifications as accurate as those typically obtained with multispectral space-borne imagery. Our classification accuracies are, however, not exceptionally high. The vegetation and cover classes derived from a classified image would be a valuable guide for preparing detailed vegetation maps using "heads up" digitizing by persons familiar with the ecology of the study area (Roughgarden *et al.*, 1991). The classified images we derived are not, however, adequate to stand on their own for applications requiring high thematic accuracy.

Considerable research on the utility of digital orthophotography remains to be done. The positional accuracy of digital orthophotographs can be extremely high, and this makes it possible to use these data as a planimetric base for other data sets in a GIS system. The integration of multispectral satellite data with digital orthophotography promises to be an interesting experiment (Munechika *et al.*, 1993). We used basic linear models in our analyses. More elaborate classification protocols, such as neural network analysis (Anand *et al.*, 1993; Li *et al.*, 1994), should prove insightful. Integrating statistical methods to detect cover classes with detailed data on spectral reflectance patterns of vegetation and cover types (Goward *et al.*, 1994; Price, 1994) will be essential in fully assessing spectral and spatial elements of orthophoto imagery. We expect significant variation in what constitutes the most suitable suite of collateral data. For example, slope and aspect might be very important variables in montane regions where exposure significantly affects plant distributions. Furthermore, soil properties other than (or in addition to) drainage class (e.g., organic matter, pH, texture, nutrients) likely influence plant distributions in different ecosystems, and these relationships need to be investigated. The importance of different collateral data with respect to distinguishing cover classes can be hierarchical. Soil drainage properties, for example, are critically important in distinguishing between wetland and upland habitats (Golet *et al.*, 1993; National Technical Committee for Hydric Soils, 1991), but are sometimes less important in identifying different vegetation types within wet or dry regions.

Acknowledgments

This research was supported by grants from the Environmental Protection Agency, the Town of New Shoreham, The Nature Conservancy, and the University of Rhode Island. Ginger Carpenter, Dennis Wolkolf, David Holt, Don and Dorothy McClusky, Rick Enser, Chris Barnard, Mike MacDougall, Jim Fritz, Chris Diehl, Jim Kowalski, Matt Nicholson, Carol Baker, Chuck LaBash, and Dennis Myshrall provided essential technical or logistic support over the course of this project. Jim Yoder, Howard Ginsberg, and three anonymous reviewers provided excellent editorial advice on the manuscript. This is contribution number 3103 of the Rhode Island Agricultural Experiment Station.

References

Anand, R., K.G. Mehrotra, and C.K. Mohan, 1993. An improved algorithm for neural classification of imbalanced training sets, *IEEE Transaction Neural Networks*, 4(6):962–969.

August, P.V., 1993. GIS in Mammalogy: Building a Database, *GIS Application In Mammalogy* (S.B. McLaren and J.K. Braun, editors), Oklahoma Museum of Natural History, Norman, Oklahoma, pp. 11–26.

August, P.V., J. Michaud, C. LaBash, and C. Smith, 1994. GPS for environmental applications: accuracy and precision of locational data, *Photogrammetric Engineering & Remote Sensing*, 60(1):41–45.

Barbour, M.G., J.H. Burk, and W.D. Pitts, 1987. *Terrestrial Plant Ecology*, Benjamin/Cummings Publishing Company, Inc., Menlo Park, California, 634 p.

Benallegue, M., O. Taconet, D. Vidal-Madjar, and M. Normand, 1995. The use of radar backscattering signals for measuring soil moisture and surface roughness, *Remote Sensing of Environment*, 53: 61–68.

Brannon, D.P., C.L. Hill, B.A. Davis, and R.J. Birk, 1994. Commercial remote sensing program, *Photogrammetric Engineering & Remote Sensing*, 60(3):317–330.

Briggs, J.M., and M.D. Nellis, 1991. Seasonal variation of heterogeneity in the tallgrass prairie: A quantitative measure using remote sensing, *Photogrammetric Engineering & Remote Sensing*, 57(4): 407–411.

Congalton, R.G., 1991. A review of assessing the accuracy of classifications of remotely sensed data, *Remote Sensing of Environment*, 37:35–46.

Davis, F.W., and J. Dozier, 1990. Information analysis of a spatial database for ecological land classification, *Photogrammetric Engineering & Remote Sensing*, 56(5):605–613.

DeAngelis, R.H., 1993. Digital orthophotoquads for enhanced surface representation, *Proceedings of the Thirteenth Annual ESRI User Conference*, ESRI, Redlands, California, 1:527–536.

Duhaime, R.J., 1994. *The Use of Color Infrared Digital Orthophotography to Map Vegetation on Block Island, Rhode Island*, Masters Thesis, University of Rhode Island, Kingston.

Dunwiddie, P.W., 1990. Postglacial vegetation history of coastal islands in southeastern New England, *National Geographic Research*, 6(2):178–195.

Ericson, J., 1992. Island of hope, *Nature Conservancy*, January/February, pp. 14–21.

Fitzgerald, R.W., and B.G. Lees, 1994. Assessing the classification accuracy of multisource remote sensing data, *Remote Sensing of Environment*, 47:362–368.

Gemmell, F.M., 1995. Effects of forest cover, terrain, and scale on timber volume estimation with thematic mapper data in a Rocky Mountain site, *Remote Sensing of Environment*, 51:291–305.

Golet, F.C., A.J.K. Calhoun, W.R. DeRagon, D.J. Lowry, and A.J. Gold, 1993. *Ecology of Red Maple Swamps in the Glaciated Northeast: A Community Profile*, U. S. Fish and Wildlife Service Biological Report 12, Washington, D.C., 151 p.

Goward, S.N., K.F. Huemmrich, and R. Waring, 1994. Visible-near infrared spectral reflectance of landscape components in western Oregon, *Remote Sensing of Environment*, 47:190–203.

Green, K., D. Kempka, and L. Lackey, 1994. Using remote sensing to detect and monitor land-cover and land-use change, *Photogrammetric Engineering & Remote Sensing*, 60(3):331–337.

Hansen, A.J., and G.R. Schiner, 1964. *Ground-Water Resources of Block Island, Rhode Island*, Rhode Island Geological Bulletin No. 14, U.S. Geological Survey, Rhode Island Water Resources Coordinating Board, Providence, Rhode Island, 35 p.

Harris, P.M., and S.J. Ventura, 1995. The integration of geographic data with remotely sensed imagery to improve classification in an urban area, *Photogrammetric Engineering & Remote Sensing*, 561(8): 993–998.

Herr, A.M., and L.P. Queen, 1993. Crane habitat evaluation using GIS and remote sensing, *Photogrammetric Engineering & Remote Sensing*, 59(10):1531–1538.

Holz, R.J., 1985. *The Surveillant Science: Remote Sensing of the Environment*, John Wiley & Sons, New York, New York, 413 p.

Hutchinson, C.F., 1982. Techniques for combining Landsat and ancillary data for digital classification improvement, *Photogrammetric Engineering & Remote Sensing*, 48(1):123–130.

Jadkowski, M.A., P. Convery, R.J. Birk, and S. Kuo, 1994. Aerial image databases for pipeline rights-of-way management, *Photogrammetric Engineering & Remote Sensing*, 60(3):347–353.

Janssen, L.L.F., M.N. Jaarsma, and E.T.M. van der Linden, 1990. Integrating topographic data with remote sensing for land-cover classification, *Photogrammetric Engineering & Remote Sensing*, 56(11):1503–1506.

Jensen, J.R., 1986. *Introductory Digital Image Processing: A Remote Sensing Perspective*, Prentice Hall, Inc., Englewood Cliffs, New Jersey, 278 p.

Jobson, J.D., 1992. *Applied Multivariate Data Analysis Volume II: Categorical and Multivariate Methods*, Springer-Verlag, New York, New York, 732 p.

Kremer, R.G., and S.W. Running, 1993. Community type differentiation using NOAA/AVHRR data within a sagebrush-streppe ecosystem, *Remote Sensing of Environment*, 46:311–318.

Lachenbruch, P.A., 1975. *Discriminant Analysis*, Hafner, New York, New York, ix+128 p.

Lauver, C.L., and J.L. Whistler, 1993. A hierarchical classification of Landsat TM imagery to identify natural grassland areas and rare species habitat, *Photogrammetric Engineering & Remote Sensing*, 59(5):627–634.

Lee, C.T., and S.E. Marsh, 1995. The use of archival Landsat MSS and ancillary data in a GIS environment to map historical change in an urban riparian habitat, *Photogrammetric Engineering & Remote Sensing*, 61(8):999–1008.

Levine, E.R., R.G. Knox, and W.T. Lawrence, 1994. Relationship between soil properties and vegetation at the Northern Experimental Forest, Howland, Maine, *Remote Sensing of Environment*, 47: 231–241.

Li, Q., D.W. Tufts, R.J. Duhaime, and P.V. August, 1994. Fast training algorithms for large data sets with application to classification of multispectral images, *IEEE Asilomar Conference Proceeding*, Pacific Grove, California, 31 October - 2 November 1994, pp. 1–5.

Light, D.L., 1993. The national aerial photography program as a geographic information system resource, *Photogrammetric Engineering & Remote Sensing*, 59(1):61–65.

Lillesand, T.M., and R.W. Kiefer, 1994. *Remote Sensing and Image Interpretation*, 3rd edition, John Wiley & Sons, Inc., New York, 750 p.

Logan, B.J., 1993. Digital orthophotography bolsters GIS base for wetlands project, *GIS World*, 1:58–60.

Lowell, K.E., 1990. Differences between ecological land type maps produced using GIS or manual cartographic methods, *Photogrammetric Engineering & Remote Sensing*, 56(2):169–173.

Muehrcke, P.C., and J.O. Muehrcke, 1992. *Map Use: Reading, Analysis, and Interpretation*, 3rd Edition, JP Publications, Madison, Wisconsin, 631 p.

Munechika, C.K., J.S. Warnick, C. Salvaggio, and J.R. Schott, 1993.

Resolution enhancement of multispectral image data to improve classification accuracy, *Photogrammetric Engineering & Remote Sensing*, 59(1):67–72.

Musick, H.B., and H.D. Grover, 1991. Image textural measures as indices of landscape pattern, *Quantitative Methods in Landscape Ecology* (M.G. Turner and R.H. Gardner, editors), Springer-Verlag, New York, pp. 77–103.

Nale, D., 1995. Digital orthophotography: The foundation of GIS, *American City & County*, 110(8):26–38.

National Technical Committee for Hydric Soils, 1991. *Hydric Soils of the United States*, U.S.D.A. Soil Conservation Service, Miscellaneous Publication Number 1491, Washington, D.C.

Nellis, M.D., K. Lulla, and J. Jensen, 1990. Interfacing geographic information systems and remote sensing for rural land-use analysis, *Photogrammetric Engineering & Remote Sensing*, 56(3):329–331.

Novak, K., 1992. Rectification of digital imagery, *Photogrammetric Engineering & Remote Sensing*, 58(3):339–344.

Oreskes, N., K. Shrader-Frechette, and K. Belitz, 1994. Verification, validation, and confirmation of numerical models in the earth sciences, *Science*, 263(4):641–646.

Parent, P., 1991. Digital orthophotography provides a new tool for GIS database integration, *GIS World*, 4(11):48–49.

Pettinger, L.R., 1982. *Digital Classification of Landsat Data for Vegetation and Land-Cover Mapping in the Blackfoot River Watershed, Southeastern Idaho*, USGS Professional Paper 1219, pp. 1–33.

Price, J., 1994. How unique are spectral signatures, *Remote Sensing of Environment*, 49:181–186.

Ramey, B.S., 1992. U.S. Geological Survey national mapping program, digital mapmaking procedures for the 1990's, *Photogrammetric Engineering & Remote Sensing*, 58(8):1113–1116.

Rector, D.D., 1981. *Soil Survey of Rhode Island*, U.S.D.A. Soil Conservation Service, West Warwick, Rhode Island, 200 p. + maps.

Roughgarden, J., S.W. Running, and P.A. Matson, 1991. What does remote sensing do for ecology? *Ecology*, 72(6):1918–1922.

Rutchey, K., and Les Vilcheck, 1994. Development of an Everglades vegetation map using a Spot image and the global positioning system, *Photogrammetric Engineering & Remote Sensing*, 60(6): 767–775.

Sokal, R.R., and F.J. Rohlf, 1981. *Biometry*, W. H. Freeman and Co., San Francisco, California, xxi + 776 p.

Steiner, D.R., 1992. The integration of digital orthophotographs with GIS's in a microcomputer environment, *The ITC Journal*, 1:65–77.

Stenback, J.M., and R.G. Congalton, 1990. Using Thematic Mapper imagery to examine forest understory, *Photogrammetric Engineering & Remote Sensing*, 56(9):1285–1290.

Todd, W.J., D.G. Gehring, and J.F. Haman, 1980. Landsat wildland mapping accuracy, *Photogrammetric Engineering & Remote Sensing*, 46(4):509–520.

U.S. Geological Survey, 1987. *Digital Elevation Models: Data Users Guide 5*, U.S. Dept. of the Interior, U.S. Geological Survey, National Cartographic Information Center, Reston, Virginia, 38 p.

Welch, R., M. Remillard, and J. Alberts, 1992. Integration of GPS, remote sensing, and GIS technique for coastal resource management, *Photogrammetric Engineering & Remote Sensing*, 58(11): 1571–1578.

Wickland, D.E., 1991. Mission to planet earth: The ecological perspective, *Ecology*, 72(6):1923–1933.

Williams, B.K., 1983. Some observations on the use of discriminant analysis in ecology, *Ecology*, 64(5):1283–1291.

Wright, W.R., and E.H. Sautter, 1988. *Soils of Rhode Island Landscapes*, U.S. Soil Conservation Service, University of Rhode Island Agricultural Experiment Station Bulletin 429, Kingston, Rhode Island, 42 p.

Zeff, I.S., and C.J. Merry, 1993. Thematic mapper data for forest resource allocation, *Photogrammetric Engineering & Remote Sensing*, 59(1):93–99.

(Received 30 January 1995; revised and accepted 18 April 1996)