A Quantitative Comparison of Change-Detection Algorithms for Monitoring Eelgrass from Remotely Sensed Data

Robb D. Macleod and Russell G. Congalton

Abstract

*The eelgrass (*Zostera marina *L.) population in Great Bay, New Hampshire has recently undergone dramatic changes. A reoccurrence of the 1930s wasting disease and decreasing water quality due to pollution led to a reduction in the eelgrass population during the late 1980s. Currently, the eelgrass populations in Great Bay have experienced a remarkable recovery from the decline in the late 1980s. Eelgrass is important in our estuarine ecosystems because it is utilized as habitat by many commercial and non-commercial organisms and is a food source for waterfowl. In order to monitor the eelgrass populations in Great Bay, a change detection analysis was performed to determine the fluctuation in eelgrass meadows over time.*

Change detection is a technique used to determine the change between two or more time periods of a particular object of study. Change detection is an important process in monitoring and managing natural resources and urban development because it provides quantitative analysis of the spatial distribution in the population of interest. A large number of change-detection techniques have been developed, but little has been done to quantitatively assess the accuracies of these techniques.

In this study, post-classification, image differencing, and principal components change-detection techniques were used to determine the change in eelgrass meadows with Landsat Thematic Mapper (TM) data. Low altitude (1,000 m), oblique aerial photography combined with boat surveys were used as reference data. A proposed change-detection error matrix was used to quantitatively assess the accuracy of each change-detection technique. The three different techniques were then compared using standard accuracy assessment procedures. The image differencing change-detection technique performed significantly better than the post-classification and principal components analysis. The overall accuracy of the image differencing change detection was 66 percent with a Khat value of 0.43.

This study provided an application of Landsat Thematic Mapper to detect submerged aquatic vegetation and the methodology for comparing change detection techniques using a proposed change detection error matrix and standard accuracy assessment procedures. In addition, this study showed that image differencing was better than the post-classification or principal components techniques for detecting changes in submerged aquatic vegetation.

Introduction

Eelgrass (*Zostera marina* L.), a true flowering plant that completes its life cycle in shallow sea water, is a critical component of coastal and estuarine ecosystems (Milne and Milne, 1951; Ackleson and Klemas, 1987; Short, 1989; Ferguson *et al.*, 1993). It grows in bays, estuaries, and coastal oceans throughout the northern temperate regions of the world and can rival the productivity of agricultural crops (Thayer *et al.*, 1984). In addition, eelgrass meadows provide habitat for numerous organisms, including coastal fish, lobsters, crabs and scallops, and a food source for waterfowl. Eelgrass meadows also increase water quality by filtering sediments and nutrients within the water (Short, 1989). It is therefore important to maintain healthy populations of eelgrass in order to ensure the continuing prosperity of coastal and estuarine ecosystems.

Currently, the two major problems that are severely impacting eelgrass meadows throughout the world are the wasting disease (*Labyrinthula zosterae*) and pollution (Short *et al.*, 1991). As a result, the eelgrass population in Great Bay, New Hampshire has gone through dramatic changes in the last decade (Short *et al.*, 1993). Monitoring the spatial distribution in eelgrass habitat is an important part of understanding the changes in eelgrass meadows, which in turn will ensure the viability of coastal and estuarine ecosystems (Ferguson *et al.*, 1993). Historically, eelgrass and other submerged aquatic vegetation (SAV) have been monitored in the field with either permanent transects or stations (Ackleson and Klemas, 1987). However, the cost of field sampling has become expensive for large areas and is now used primarily to assess the accuracy of more efficient techniques such as aerial photography (Short *et al.*, 1986; Ackleson and Klemas, 1987; Ferguson *et al.*, 1993). More recently, satellite imagery has been used to detect eelgrass and other SAV (Ackleson and Klemas, 1987; Jensen *et al.*, 1993; Luczkovich *et al.*, 1993; Zainal *et al.*, 1993).

An increasingly popular application of remotely sensed data is for change detection. Change detection is the process of identifying differences in the state of an object or phenomenon by observing it at different times (Singh, 1989). Four aspects of change detection are important when monitoring natural resources: (1) detecting that changes have occurred, (2) identifying the nature of the change, (3) measuring the areal extent of the change, and (4) assessing the spatial pattern of the change (Brothers and Fish, 1978; Malila, 1980; Singh, 1986). Techniques to perform change detection with satellite imagery have become numerous because of increasing versatility in manipulating digital data and increasing

Department of Natural Resources, University of New Hampshire, Durham, NH 03824 (russ.congalton@unh.edu).

Photogrammetric Engineering & Remote Sensing, Vol. 64, No. 3, March 1998, pp. 207–216.

0099-1112/98/6403-207$3.00/0

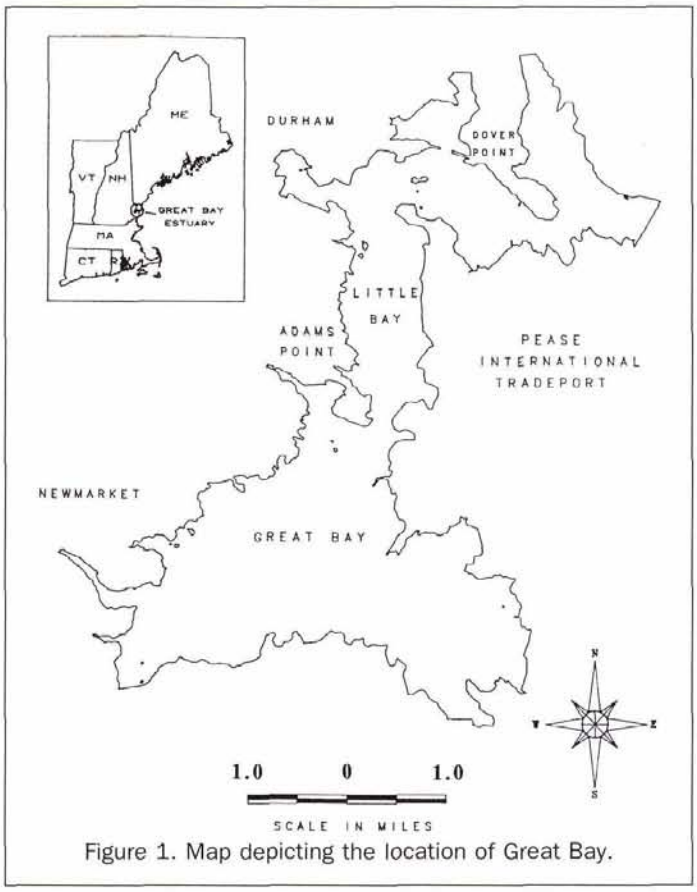

Figure 1. Map depicting the location of Great Bay.

computing power (Jensen, 1996). However, relatively little work has been conducted to determine the quantitative accuracy of the different change-detection techniques, and no standard techniques have been adopted. Singh (1989) provides an excellent review of all the digital change-detection techniques and presents a table modified from Nelson (1983) that lists the different techniques.

Standard accuracy assessment techniques have been developed for single-date remotely sensed data and are reviewed by Congalton (1991). However, the very nature of change detection makes quantitative analysis of the accuracy difficult. For example, how does one obtain reference data for images that were taken in the past? How does one sample enough areas that will change in the future to have a statistically valid assessment? Which change-detection technique will work best for a given change in the environment? Because most studies on change detection have not included quantitative accuracy assessments, it is difficult to determine which change-detection methods are most effective. To date, no standard accuracy assessment techniques or procedures for change detection have been developed. Studies to determine the optimal threshold value (Fung and LeDrew, 1988) and the accuracies between different change-detection techniques (Singh, 1986; Martin, 1989; Macleod, 1994) have made encouraging steps towards accomplishing standard accuracy assessment techniques for change detection. In addition, Congalton and Macleod (1994) have developed a modified form of the single-date error matrix to allow standard accuracy assessment techniques to be used with change-detection studies. Clearly, as change-detection studies become more popular, the urgency for developing procedures to determine the accuracy of the different techniques becomes increasingly important.

The primary objective of this research was to determine the change-detection technique that had the highest accuracy for identifying changes in eelgrass meadows. The specific objectives were to (1) determine the appropriate methods to detect eelgrass distributions with Landsat TM data, (2) perform three types of change-detection methods to map the changes in eelgrass distributions in Great Bay, New Hampshire using Landsat TM data, (3) perform an accuracy assessment on the change-detection methods using a newly devised change-detection error matrix, and (4) compare the change-detection methods using the change-detection error matrix and additional standard accuracy assessment techniques.

Methods

The methods section is divided into a description of the study area, satellite and reference data acquisition, image classification, change detection, and accuracy assessment.

Study Area

The study area used in this project was Great Bay, New Hampshire, which is part of a larger estuarine system called the Great Bay Estuary. The Great Bay Estuary is composed of the Piscataqua River, Little Bay, and Great Bay and has a total drainage area of 2,409 km^2. It is located on the New Hampshire-Maine border approximately 6.25 km inland from the Gulf of Maine (Figure 1). The Great Bay Estuary is a drowned river valley that was formed during the most recent deglaciation, approximately 14,500 years ago (Short, 1992). It supports 23 species of threatened or endangered plants and animals and supports a significant commercial fishery. On 3 October 1989, the Great Bay Estuary was officially designated a National Estuarine Research Reserve by the National Oceanic and Atmospheric Administration (NOAA).

Our study site is a subset of the Great Bay Estuary and is simply called the Great Bay. It begins at Adams Point where a 13.5-metre deep channel extends nearly shore to shore. The average depth of Great Bay is 3 metres, with the deepest part being 18 metres. At low tide about one half of Great Bay is exposed.

Satellite and Reference Data

In order to adequately detect eelgrass meadows and perform a change-detection analysis, two Landsat TM scenes acquired during low tide were needed. The tidal stage of the imagery was important to facilitate the classification of submerged vegetation. An image acquired at high tide may not be able to classify eelgrass habitat because of the spectral interference of water. At low tide the eelgrass plants are near the surface or floating on the surface, which helps separate the spectral signature of eelgrass from water.

The level of the low tide was also considered because of the fluctuations in low tide. A spring low tide was optimal because this is when the tide is at its lowest level. A spring low tide would allow for the maximum amount of eelgrass near the surface. The optimal time for the satellite data would be acquired just before low tide because of the increased turbulence that occurs after low tide. The turbulence would act like a dust storm, preventing an adequate reflectance from the eelgrass beds to be detected by the Landsat TM sensor. Table 1 shows the date and time of the Landsat

TABLE 1. DATE, TIME, AND TIDE LEVEL OF THE LANDSAT TM SCENES

Date	Satellite Overpass	Low Tide	Tide Level
08 Sep 1990	9:51 AM	9:57 AM	−0.4' MLW*
29 Sep 1992	9:51 AM	9:45 AM	−0.5' MLW

*Mean Low Water

TM scenes and the corresponding time of low tide and the height at low tide.

The reference data used to compare with the Landsat TM classifications were obtained from the University of New Hampshire's Jackson Estuarine Laboratory where aerial slides have been acquired yearly since 1986 to map and monitor the eelgrass distributions in Great Bay (Short, 1992; Short *et al.*, 1993). Low altitude (1,000 m), low oblique (near vertical) color slides taken at low tide using 35-mm Kodachrome 64 and 200 color slide film were used to map the eelgrass distribution. The slides were then projected and interpreted from various angles onto a scaled map containing the outline of Great Bay. The slide projector was tilted to compensate for the oblique photos and the photos were displayed on a wall containing a paper outline of the bay and the deep water channel. In this way, the slides were aligned exactly with the outline of the bay and channel and the eelgrass was mapped directly onto the paper outline. Although cumbersome, this technique has been used quite successfully to accurately map eelgrass distributions (Short and Burdick, 1996).

In addition, a field assessment was performed to verify the interpretation of the aerial slides. The field assessment was accomplished using boat surveys of the eelgrass beds. Locations were marked relative to abundant natural landmarks and using the base map of Great Bay with the outline of the bay, the deep water channel, and the eelgrass distributions all delineated on it. The reference data map was then entered into a geographic information system (GIS) by digitizing it using PC Arc/Info. The data were then ready for comparison with the results of the remotely sensed change detection analysis.

Since the beginning of mapping eelgrass from 35-mm slides in 1986, many technological advances have been made. With the advent of protocols for mapping SAV under the National Oceanic and Atmospheric Administration's (NOAA) Coastal Change Analysis Program C-CAP (Dobson *et al.*, 1992), the Jackson Estuarine Lab has modified its techniques. Large-scale, vertical aerial photos are used and Global Positioning System (GPS) units are employed to better locate the eelgrass from boat surveys. However, neither of these techniques were used in collecting the 1990 and 1992 reference data. Nonetheless, the techniques used were more than adequate to produce accurate maps delineating the simple classes required by this project of (1) dense eelgrass, (2) sparse eelgrass, and (3) no eelgrass (i.e., open water).

The 1990 Landsat TM imagery was acquired on 8 September and the aerial photography was acquired the exact same day. The boat survey to verify the photo interpretation was conducted later in September of the same year. The 1992 Landsat TM imagery was obtained on 29 September while the reference data were generated in August from aerial photos acquired on 1 August.

Classification

Prior to any change detection, it is imperative that the imagery be geometrically rectified so that the same pixel at one date overlaps the same pixel for the other date (Townshend *et al.*, 1992). In this study, the registration of each image was performed using the nearest-neighbor resampling algorithm. The 1990 image was registered to New Hampshire state plane coordinates with a root-mean-square (RMS) error of 0.49 pixels. The 1992 image was rectified to the 1990 image with an RMS error of 0.36 pixels. Both rectifications are within acceptable limits.

In addition to the registration, before the classification process was begun the land surface surrounding Great Bay was masked out of the image in order to reduce the computer processing time and enhance the classification process. The delineation between land and water surrounding Great Bay was determined using the digital hydrographic data from New Hampshire's Geographically Referenced Analysis and Information Transfer System (GRANIT). These data were selected to provide a consistent standard for all further GIS analysis. If the imagery itself had been used to delineate the water/land boundary, then the boundary could change for each new image depending on the water level in the bay. This is true even though both images were acquired at low tide. In addition to the land mask, Ackleson and Klemas (1987) and Zainal (1993) found that masking the deep water also improved the accuracy of the classification. In Great Bay, eelgrass generally will not grow in water deeper than about 2.5 metres below mean low water. Bathymetry data available from GRANIT were used to delineate and mask out the deep water (i.e., the channel).

Unsupervised and supervised classifications were performed separately on both the 1990 and 1992 Landsat TM images. The classifications were performed with the land surface removed to enhance the eelgrass classification process. In addition, the classification of each date was retried with the deep water removed from the image as well as the land surface. An initial unsupervised classification was completed using the ISODATA algorithm (Erdas, 1991) for 100 clusters. However, there was significant confusion and mixing of clusters between sparse eelgrass and water. Therefore, the unsupervised classification was rerun using 255 clusters, and a much better discrimination of the final three classes resulted. The supervised classification used a maximum-likelihood algorithm with a total of 18 training areas. Several diagnostics, including spectral pattern analysis and transformed divergence measures (Erdas, 1991), were used to determine the best combination of spectral bands to use in the classification. The final classifications were filtered with a 3 by 3 majority filter to eliminate some of the speckling that occurred during the classification (see Macleod (1994) for a more detailed description of the classification process).

Change Detection

Post Classification

In the post-classification change-detection technique, each image was classified separately using both supervised and unsupervised classification techniques and then compared to create a change image map (Figure 2). The 1990 image was classified first followed by the 1992 image according to the classification procedures described in the classification section. The classification (supervised or unsupervised) for each year with the highest overall accuracy was used in the change detection analysis. The results of the post-classification change-detection was a new classification with "from" and "to" identifiers. For example, if the best 1990 classification distinguished a pixel as eelgrass and the best 1992 classification distinguished the same pixel as water, then the result of the new classification would describe the pixel as "from eelgrass to water."

Image Differencing

Image differencing is performed by subtracting the DN (digital number) value of one date for a given band from the DN value of the same pixel for the same band of another date (Figure 3). Image differencing was performed on the first four raw bands (blue, green, red, and near infrared). For each band, the two images were subtracted from each other, resulting in a new image. A series of threshold values based on standard deviations from the mean were used on the new image to determine the changed from unchanged pixels. An accuracy assessment on the no-change/change pixels was performed to determine the threshold value with the highest accuracy. This process was repeated for each band. The band

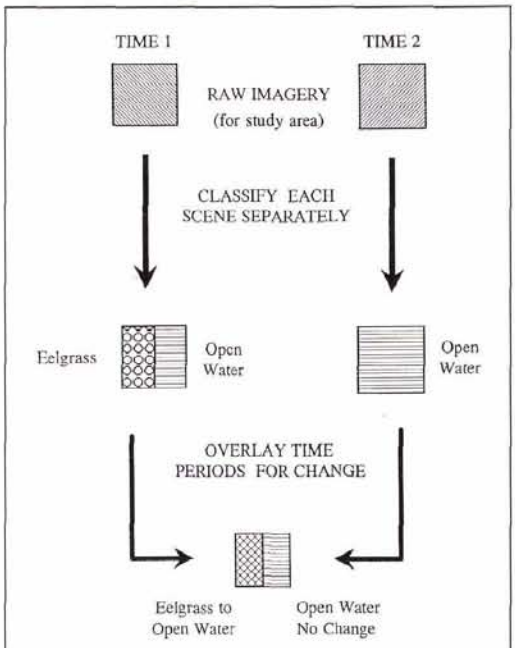

Figure 2. Post-classification change detection.

or combination of bands and threshold value that had the highest accuracy, computed by performing a quantitative accuracy assessment as described in this paper, were then used to produce the final change classification.

The results of the image differencing technique are simply the pixels that changed between the two time periods. A process of classification or labeling needs to occur to assign the appropriate "from" and "to" identifiers. The single-date

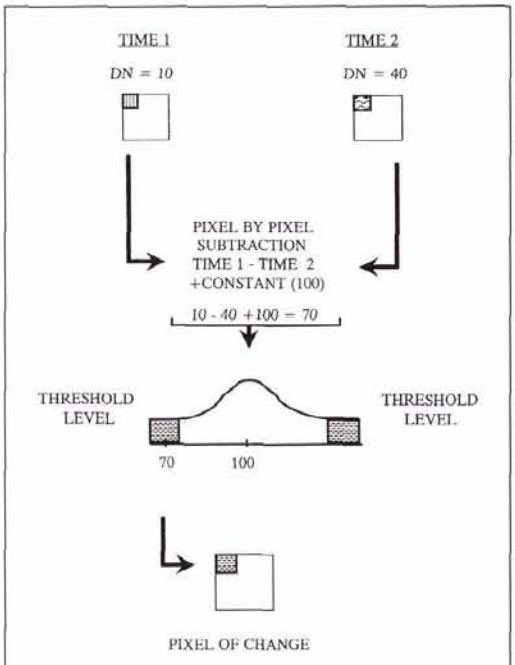

Figure 3. Image differencing change detection.

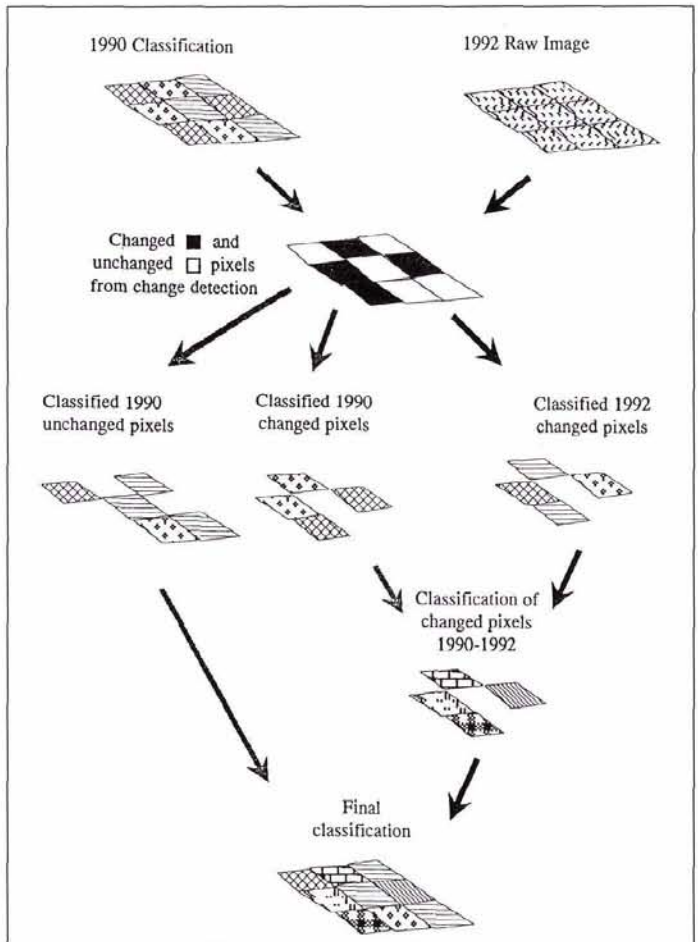

Figure 4. Classification process of the image differencing change-detection technique.

1990 classification previously performed was used to label the pixels that the image differencing technique determined had not changed between 1990 and 1992. Only the pixels that had changed were classified using the 1992 imagery. A combination of the 1990 classification and the resulting classification of the changed pixels from the 1992 image were used to label the change pixels in the final change classification (Figure 4). In other words, the final change classification for the image differencing technique consisted of the pixels labeled from the 1990 classification where the pixels were considered unchanged as determined by the image differencing technique. These unchanged pixels were then removed from the 1992 image and a classification was performed on just the changed pixels in the 1992 image. The classified change pixels in the 1992 image were combined with the same classified pixels of the 1990 classification to generate the "from" and "to" identifiers. The classified changed pixels were then combined with the unchanged pixels of the 1990 classification to create the final change classification. This procedure was created by Pilon *et al.* (1988) and adopted as the NOAA Coastal Change Analysis protocol by Dobson *et al.* (1992). Image differencing differs from the post-classification technique because it alleviates the need to fully classify both images and lowers the chance of misclassification that is inherent in the post-classification change-detection technique.

Principal Component Analysis

In principal component analysis (PCA), the first four bands of the 1990 image were combined with the first four bands of

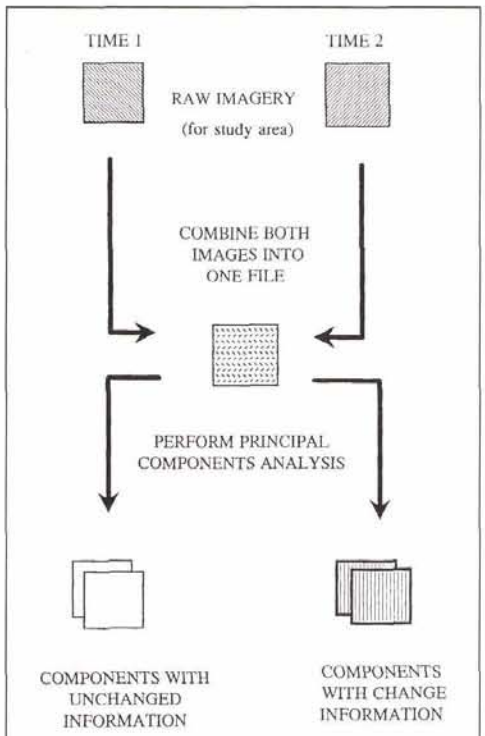

Figure 5. Principal components change-detection technique.

the 1992 image. The new multi-temporal image was used for both the standardized and unstandardized principal component analysis (Figure 5). The standardized principal components have been shown to provide better results then the unstandardized principal components (Fung and LeDrew, 1987; Eastman and Falk, 1993). In addition, just the visible bands (bands 1-3) from the 1990 image and 1992 image were combined to make a second multi-temporal image. This second image was created to determine if the infrared band enhanced or reduced the effectiveness of the principal components analysis. In theory, if a small percentage of the image has changed, the first few principal components will contain the unchanged information while the latter bands will contain the change information.

After the change bands in the principal components analysis have been selected, a threshold value must be used to separate the change pixels from the unchanged pixels. As in the image differencing threshold values, several different threshold values were used in the principal components analysis. The threshold value and analysis with the highest accuracy, computed by performing a quantitative accuracy assessment as described in this paper, were used in the classification process. The resulting change pixels were classified using the same procedure as described in the image differencing change detection section.

Accuracy Assessment

The accuracy of the supervised and unsupervised classifications for both the 1990 and 1992 Landsat TM images was obtained using standard, single-date, quantitative accuracy assessment procedures (i.e., an error matrix and KAPPA analysis) (Congalton, 1991). Because the reference data were in a GIS and because it was a complete coverage and not just the usual sample, the error matrix was automatically calculated by comparing the classification to the reference data pixel by pixel (i.e., a total enumeration). The selection of the most appropriate threshold levels for the image differencing and principal components analysis change-detection techniques were also assessed using these methods.

In order to perform an accuracy assessment on the change-detection techniques, the error matrix for the single-date classification had to be modified (Congalton and Macleod, 1994). The new matrix has the same characteristics as the single-date classification error matrix, except that it also assesses errors in changes between two time periods and not simply a single classification. Figure 6 shows a single classification error matrix for three vegetation/land-cover categories (A, B, and C) and a change-detection error matrix for the same three categories. The single classification matrix is of dimension 3 by 3, whereas the change detection error matrix is no longer of dimension 3 by 3 but, rather, 9 by 9. The dimensions have changed because we are no longer looking at a single classification but rather a change between two different classifications generated at different times. For both error matrices, one axis presents the three categories as derived from the remotely sensed classification and the other axis shows the three categories identified from the reference data. The major diagonal of the matrices indicates correct classification. Off-diagonal elements in the matrices indicate the different types of confusion (omission and commission error) that exist in the classification. This information is helpful in guiding the analyst to problems that exist in the classification.

In order to further analyze if the errors are due to the classification or the change detection technique, the change detection error matrix can be simplified or collapsed into a no-change/change error matrix. The no-change/change error matrix can be formulated by summing the cells in the four appropriate sections of the change-detection error matrix (Figure 6). For example, to get the number of areas that both the classification and reference data correctly determined that no change had occurred between two dates, one would simply add together all the areas in the upper left box (the areas that did not change in either the classification or reference data). The upper right box would indicate the areas that the classification detected no change and the reference data considered changed. From the change-detection error matrix and no-change/change error matrix, the analyst can determine if a low accuracy was due to a poor change-detection technique, misclassification, or both. If the accuracy in the no-change/change matrix increases significantly over the full

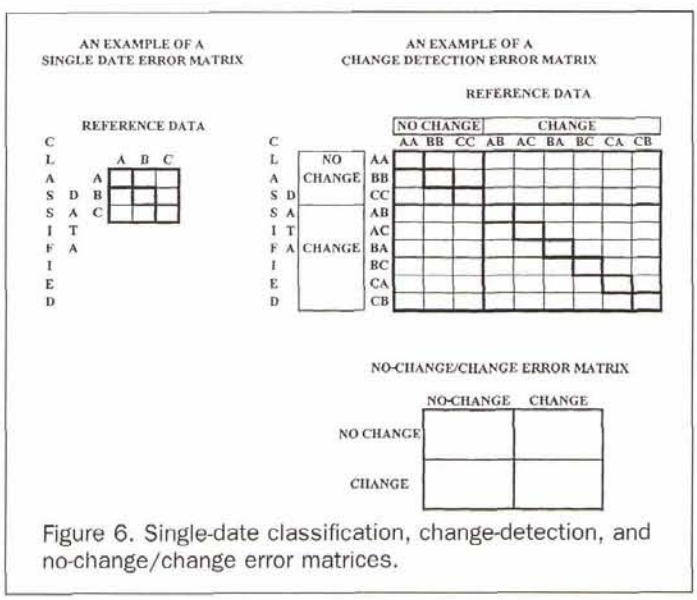

Figure 6. Single-date classification, change-detection, and no-change/change error matrices.

TABLE 2. NINE CHANGE-DETECTION CATEGORIES

		1990 to 1992	
[LL]	Low density eelgrass	to	Low density eelgrass
[HH]	High density eelgrass	to	High density eelgrass
[WW]	Water	to	Water
[LH]	Low density eelgrass	to	High density eelgrass
[LW]	Low density eelgrass	to	Water
[HL]	High density eelgrass	to	Low density eelgrass
[HW]	High density eelgrass	to	Water
[WL]	Water	to	Low density eelgrass
[WH]	Water	to	High density eelgrass

change-detection error matrix, then it is clear that general change is being detected but not the exact classes of change. If the two matrices produce similar accuracies, then not even the general changes are being detected.

In this study, the post-classification, image differencing, and principal components change-detection error matrices were compared using Kappa analysis to determine which method performed the best for detecting changes in eelgrass. The change-detection error matrix and no-change/change error matrix were compared for each method. In addition, a collapsed classification scheme consisting of just eelgrass and water was analyzed to determine if the change-detection methods had trouble detecting the densities of eelgrass.

Results

Single-Date Classification

The majority filtered, unsupervised classification had the highest accuracy for both the single-date 1990 and single-date 1992 images. The 1990 unsupervised classification had an overall accuracy of 77 percent and a Khat value of 0.43 while the supervised classification accuracy was 74 percent (Khat of 0.38). The 1992 unsupervised classification had an overall accuracy of 66 percent with a Khat of 0.39 and the supervised classification was 65 percent accurate (Khat of 0.36). The accuracy of the classification was on the low end of average for satellite classifications. The low accuracy could be partly attributed to some registration and interpretation problems due to the limitations of the way in which the reference data were collected. However, more error should be attributed to the fact that this accuracy assessment was a total enumeration and not just a sample of the map. When sampling for accuracy assessment, it is common to pick homogeneous sites and locate your sample in the middle of the site. Using this approach tends to inflate the accuracy by avoiding edges (i.e., boundaries) that may be off by a pixel or two because of a combination of factors including misregistration, variability in interpretation, and image filtering. It should also be remembered that fully detecting the eelgrass requires some penetration through the water, and turbidity and other factors can obscure this discrimination.

Change Detection

Post Classification

Post-classification change-detection was accomplished using the best classifications (i.e., the filtered unsupervised) from the 1990 and 1992 dates. These two classifications were combined together, resulting in nine change detection categories (see Table 2). The overall accuracy for the post-classification change-detection technique was 51 percent with a Khat value of 0.27. The change-detection error matrix and the collapsed no-change/change error matrix are presented in Figure 7.

Image Differencing

The first four Landsat TM bands were used in the image differencing process because the shorter wavelength light has better water penetration. For each band, every pixel from the 1990 image was subtracted from the same pixel from the 1992 image. A constant of 100 for the visible bands and 150 for the infrared band was added to the subtraction process in order to compensate for the negative numbers. The result was the creation of four new change bands consisting of the subtraction of bands 1, 2, 3, and 4 from the 1990 and 1992 images. Thresholds were chosen using 0.5, 1.0, 1.5, 2.0, 2.5, and 3.0 standard deviations from the mean (Table 3). To quantitatively assess which threshold level was the most accurate, both the overall accuracy and Khat measures were used to compare the different levels.

The standard deviation threshold did not perform well because most of the changes that occurred between 1990 and 1992 were an increase in eelgrass, either a gain in density or a change from water to eelgrass. Because the spectral response (DN) of eelgrass becomes lower as the density increases, most of the changes that occurred resulted in a decrease in digital value. Therefore, most of the change that occurred resulted in a higher digital number (DN) than the mean after the subtraction process. The threshold level was selectively modified by considering fewer of the pixels changed below the mean and more changed above the mean (i.e., using an asymmetrical approach). The accuracy of the new selective threshold value was significantly better than the standard deviation thresholds.

Considering that each band may determine different types of changes, the best (most accurate) thresholds for each image differencing band were combined and then compared to the reference data. As a result of this accuracy comparison, the band combinations that were chosen for the final analysis were all four bands; bands 1, 3, and 4; and bands 1

REFERENCE DATA

CLASSIFIED DATA		NO CHANGE			CHANGE						TOTAL
		LL	HH	WW	LH	LW	HL	HW	WL	WH	
NO CHANGE	LL	201	55	106	49	59	1	7	128	11	617
	HH	113	844	168	127	5	15	61	17	167	1517
	WW	106	923	8383	547	366	86	234	237	717	11599
CHANGE	LH	212	707	208	215	30	51	68	85	127	1703
	LW	28	147	108	31	4	14	5	21	37	395
	HL	21	4	1	60	0	1	0	1	10	98
	HW	12	139	27	14	0	11	1	2	32	238
	WL	82	98	398	71	55	11	37	330	127	1209
	WH	215	873	940	211	45	93	84	284	459	3204
TOTAL		990	3790	10339	1325	564	283	497	1105	1687	20580

OVERALL ACCURACY = 10438 / 20580 = 50.72 %

NO-CHANGE / CHANGE ERROR MATRIX

	NO CHANGE	CHANGE	TOTAL
NO CHANGE	10899	2834	13733
CHANGE	4220	2627	6847
TOTAL	15119	5461	20580

OVERALL ACCURACY = 13526 / 20580 = 65.7 %

Figure 7. Change-detection and no-change/change error matrices for the post-classification change-detection technique.

TABLE 3. ONE STANDARD DEVIATION THRESHOLD LEVEL FOR BAND 1.

	DN	Frequency
	91	1
Changed pixels	92	9
	93	43
	94	125
Gain in DN (90–92)	95	250
	96	450
	97	724
	98	924
1.0 standard deviation	99	1204
	100	1460
Unchanged pixels	101	1403
	102	1268
	103	1095
Changed pixels	104	823
	105	570
	106	373
Loss in DN (90–92)	107	227
	108	124
	109	34
	110	7
	110	4

and 4. Bands 1 and 4 and bands 1, 3, and 4 were significantly better than band 4 alone or bands 1, 2, 3, and 4. Because bands 1 and 4 had a slightly better overall accuracy, this combination was chosen for the classification process. Figure 8 shows the change-detection error matrix for the image differencing change-detection technique and the no-change/change error matrix. The overall accuracy for the classification was 55 percent with a Khat of 0.33. The overall accuracy for the no-change/change error matrix was 78 percent with a Khat of 0.41.

REFERENCE DATA

CLASSIFIED		NO CHANGE			CHANGE						
		LL	HH	WW	LH	LW	HL	HW	WL	WH	TOTAL
NO CHANGE	LL	366	592	375	165	70	36	52	142	112	1910
	HH	45	841	195	117	4	24	40	16	154	1436
	WW	308	1121	9284	393	136	60	82	334	630	12348
CHANGE	LH	139	654	242	319	208	69	215	192	308	2346
	LW	8	5	3	1	0	2	2	6	4	31
	HL	40	36	0	49	0	0	1	3	21	150
	HW	4	2	1	3	0	0	0	0	3	13
	WL	38	151	127	116	83	33	69	295	197	1109
	WH	42	388	112	162	63	59	36	117	258	1237
	TOTAL	990	3790	10339	1325	564	283	497	1105	1687	20580

OVERALL ACCURACY = 10438 / 20580 = 55.21 %

NO-CHANGE / CHANGE ERROR MATRIX

	NO CHANGE	CHANGE	TOTAL
NO CHANGE	13127	2567	15694
CHANGE	1992	2894	4886
TOTAL	15119	5461	20580

OVERALL ACCURACY = 13526 / 20580 = 77.8 %

Figure 8. Change-detection and no-change/change error matrices for the image differencing change-detection technique.

REFERENCE DATA

CLASSIFIED		NO CHANGE			CHANGE						
		LL	HH	WW	LH	LW	HL	HW	WL	WH	TOTAL
NO CHANGE	LL	412	792	387	248	89	61	71	219	158	2437
	HH	101	948	179	110	5	26	62	16	174	1621
	WW	298	1705	9230	732	459	151	349	702	981	14607
CHANGE	LH	28	114	25	46	4	5	7	11	17	257
	LW	1	3	10	1	0	0	2	4	0	21
	HL	45	37	16	91	0	1	0	4	33	227
	HW	0	2	1	0	0	0	0	0	2	5
	WL	73	47	246	42	5	30	2	118	195	758
	WH	32	142	245	55	2	9	4	31	127	647
	TOTAL	990	3790	10339	1325	564	283	497	1105	1687	20580

OVERALL ACCURACY = 10438 / 20580 = 52.9 %

NO-CHANGE / CHANGE ERROR MATRIX

	NO CHANGE	CHANGE	TOTAL
NO CHANGE	14052	4613	18665
CHANGE	1067	848	1915
TOTAL	15119	5461	20580

OVERALL ACCURACY = 13526 / 20580 = 72 %

Figure 9. Change-detection and no-change/change error matrices for the principal components analysis change-detection technique.

Principal Component Analysis

Four different principal component analyses were performed to determine the best method for detecting change in eelgrass. An unstandardized (covariance matrix) principal components analysis (PCA) was performed on two multi-temporal datasets (bands 1-4 for 1990 and 1992 and bands 1-3 for 1990 and 1992). In addition, a standardized (correlation matrix) principal components analysis was also performed on the same multi-temporal data sets.

The standardized PCA was significantly better than the unstandardized when using bands 1-4. There was no significant difference between the standardized and unstandardized technique when using just the visible bands. The standardized PCA using bands 1-4 was significantly better (statistically, significantly higher Khat statistic) than the other three methods and was used in the classification process. Figure 9 shows the change-detection error matrix for the PCA change-detection technique and the no-change/change error matrix. The overall accuracy for the PCA classification was 53 percent with a Khat of 0.24. The no-change/change accuracy was 72 percent with a Khat of 0.17.

Comparison of the Change-Detection Techniques

Table 4 compares the classification accuracy of the three change-detection methods. All three techniques were statistically, significantly different (Table 5) with the image differencing technique having the highest Khat value. Table 6 compares the no-change/change accuracy of the three change-detection methods. The image differencing technique also had a significantly higher accuracy, using the Khat statistic, than the other two techniques (Table 7).

TABLE 4. COMPARISON OF THE CLASSIFICATION ACCURACIES FOR THE THREE CHANGE-DETECTION METHODS

Method	Overall Accuracy	Khat
Post classification	51%	0.27
Image differencing	55%	0.33
Principal components	53%	0.24

TABLE 5. SIGNIFICANCE TEST FOR THE CLASSIFICATION OF THE CHANGE-DETECTION TECHNIQUES

Comparison	Z statistic
Post classification vs. Image differencing	−9.651
Post classification vs. Principal components	6.825
Imaging differencing vs. Principal components	16.465

TABLE 6. COMPARISON OF THE NO-CHANGE/CHANGE ACCURACIES FOR THE THREE CHANGE-DETECTION METHODS

Method	Overall Accuracy	Khat
Post classification	66%	0.30
Image differencing	78%	0.41
Principal components	72%	0.11

TABLE 7. SIGNIFICANCE TEST FOR THE NO-CHANGE/CHANGE CATEGORIES OF THE CHANGE-DETECTION TECHNIQUES

Comparison	Z statistic
Post classification vs. Image differencing	−11.003
Post classification vs. Principal components	18.909
Image differencing vs. Principal components	31.097

Discussion

Change Detection

There is no simple way to evaluate the errors in the classification associated with change detection. In addition to the errors generated in the single-date classification process, the analyst must contend with the propagation of errors in the second-date classification, the change-detection algorithm, and registration and radiometric differences between the two dates. Congalton (1994) details the possible propagation of errors in both single-date classifications and change-detection studies. As in complex GIS analysis with multiple layers, change-detection studies are difficult to assess the accuracy because of the large number of variables associated with the process.

Post-classification change detection is probably the most straightforward change-detection process, but it also has the most potential for being the least accurate. The accuracy of the post-classification change-detection technique may be poor because of combining the errors from both of the classifications. Other change-detection techniques like image differencing and principal components, can minimize some of the classification errors by only classifying the areas that have changed for the second date. In addition, for the post-classification technique there must be enough ancillary data to classify both dates. When using the image differencing or principal components change detection, the amount of ancillary data for the second date may be minimized greatly due to classifying only a portion of the scene.

Image differencing is also a very simple method to understand; however, the classification process creates some difficulty. Except for post classification and multi-date classification, change-detection techniques only determine if the area has changed or if it has not changed. In most instances, more information than just if the area has changed is required. Therefore, there needs to be some sort of classification process of the changed pixels. This classification process has the potential to eliminate some of the errors in the change-detection process due to outside factors, like changes in categories that are not of interest. For example, in this study the mud flats changed in digital number over time but remained mud flats. The difference in mud flats was not a change that was of interest for this study and was not considered changed in the final classification. However, labeling these changed pixels is a difficult task and can create additional errors.

The most critical part of the image differencing change-detection method is the placement of the threshold level. It should be noted that, when placing the threshold level, the type of change occurring should be determined in order to adequately place the threshold without a lot of trial and error. For studies that have equal amounts of changes occurring in both the gain and loss of digital value, a standard deviation threshold value may perform the best because it uses equal threshold values on each side of the mean. However, if the changes are weighted towards either a gain or loss in digital value, the threshold should be selectively modified to adjust for this unbalanced gain or loss. The Kappa analysis, as recommended by Fung and LeDrew (1988), should be the measure of accuracy for the threshold values.

The principal components analysis confirmed the results of Fung and LeDrew (1987) that the standardized principal components were significantly more accurate than the unstandardized principal components for change detection studies. In order to use PCA, an understanding of the spectral characteristics of the classes of interest and corresponding principal components must be understood. The bands to include in the PCA are also an important consideration. As in image differencing, the threshold level is very important and may be the determining factor as to the accuracy of the final change classification.

The question of whether different change-detection techniques actually detect different types of change may be of interest. The basic principle of all change-detection techniques is that the digital number of one date is different from the digital number of another date. Each change-detection technique has a different way of extracting this change, but they all try to detect the same change. This theory leads to the question of whether the different techniques actually detect different types of changes, or if the differences between the techniques are more a factor of the threshold placement and classification.

Change-Detection Error Matrix

The change-detection error matrix provided the means to determine the accuracy of the change-detection techniques but, more important, it allowed the utilization of standard accuracy assessment techniques already available for single-date accuracy assessments. Therefore, the change-detection error matrix allowed the change-detection techniques used in this study to be quantitatively assessed and compared.

The change-detection error matrix is a very powerful tool for change-detection studies. The classification in this study consisted of three categories, which resulted in a 9 by 9 change-detection error matrix. For classification schemes any larger than three classes, the change-detection error matrix becomes very awkward and difficult to obtain a suitable sample. For larger classification schemes, the change categories in the change-detection error matrix that are not present could be left out to simplify the matrix (Congalton *et al.*, 1993).

Conclusions

This study compared three techniques for detecting and mapping eelgrass with Landsat TM data and analyzed three different change-detection techniques using a change-detection error matrix. In order to detect eelgrass with Landsat TM data, the land surface and the deep water were removed from the image to determine if removing the land and deep water would enhance the separability of the spectral classes. A post-classification change detection was than performed using the best classifications from the 1990 and 1992 images. Image differencing and principal components change-detection techniques were also performed using the raw imagery. A change-detection error matrix was used in order to utilize standard accuracy assessment techniques to compare the change-detection techniques.

Image differencing produced the highest accuracy of the three change-detection techniques when compared to the reference data (55 percent). When determining the threshold values for the image differencing change-detection technique, Kappa analysis should be used instead of the overall accuracy. In certain cases, the overall accuracy may not be a reliable measure of the threshold level because of overestimating the accuracy. The standard deviation threshold level will work well if the change occurs equally between gain and loss. For example, if there is an equal amount of vegetation loss and vegetation gain throughout the image creating an equal amount of change on both sides of the mean. A user-defined threshold will work better if there is an uneven change, either more gain or loss in digital value.

The principal components analysis change-detection method had the lowest Khat accuracy of the three techniques. The overall accuracy for the principal components analysis was higher than the post classification. However, an examination of the error matrix and Khat accuracy of the principal components analysis shows the inability of the principal component method to detect a sufficient amount of change in the eelgrass meadows. The principal components analysis confirmed the results of Fung and LeDrew (1987) that the standardized principal components were significantly more accurate than the unstandardized principal components.

This study provides an application of Landsat Thematic Mapper to detect submerged aquatic vegetation and the methodology for comparing change-detection techniques using standard accuracy assessment procedures. In addition, this study showed that image differencing was better than the post-classification or principal components techniques for detecting changes in submerged aquatic vegetation.

Acknowledgments

Funding for this project was provided by NOAA's Coastal Change Analysis Program (C-CAP).

References

Ackleson, S.G., and V. Klemas, 1987. Remote sensing of submerged aquatic vegetation in lower Chesapeake Bay: A comparison of Landsat MSS to TM Imagery, *Remote Sensing of Environment*, 22:235–248.

Brothers, G.L., and E.B. Fish, 1978. Image enhancement for vegetation pattern change analysis, *Photogrammetric Engineering & Remote Sensing*, 44(5):607–616.

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

———, 1994. Accuracy assessment of remotely sensed data: Future needs and directions, *Proceedings of the Pecora 12 Symposium: Land Information from Space-Based Systems*, Sioux Falls, South Dakota, Am. Soc. Photo. and Remote Sensing, pp. 385–388.

Congalton, R.G., R.D. Macleod, and F.T. Short, 1993. *Developing Accuracy Assessment Procedures for Change Detection Analysis*, Final Report, submitted to NOAA's CoastWatch Change Analysis Project, 57 p.

Congalton, R.G., and R.D. Macleod, 1994. Change detection accuracy assessment on the NOAA Chesapeake Bay Pilot Study, *Proceedings of the International Symposium on the Spatial Accuracy of Natural Resources Data Bases*, ASPRS, Williamsburg, Virginia, 16–20 May, pp. 78–87.

Dobson, J.E., R.L. Ferguson, D.W. Field, L.L. Wood, K.D. Haddad, H. Iredale III, V.V. Klemas, and R.J. Thomas, 1992. *NOAA Coast-Watch Change Analysis Project: Guidance for Regional Implementation, Version 1.0*, NOAA Coastal Ocean Program, Washington, D.C., 128 p.

Eastman, J., and M. Fulk, 1993. Long sequence time series evaluation using standardized principal components, *Photogrammetric Engineering & Remote Sensing*, 59(8):1307–1312.

Erdas, 1991. *Erdas v7.5 Field Guide*, Erdas, Inc., Atlanta, Georgia, 394 p.

Ferguson, R.L., L.L. Wood, and D.B. Graham, 1993. Monitoring spatial change in seagrass habitat with aerial photography, *Photogrammetric Engineering & Remote Sensing*, 59(6):1033–1038.

Fung, T., and E. LeDrew, 1987. Application of principal components analysis to change detection, *Photogrammetric Engineering & Remote Sensing*, 53(12):1649–1658.

———, 1988. The determination of optimal threshold levels for change detection using various accuracy indices, *Photogrammetric Engineering & Remote Sensing*, 54(10):1449–1454.

Jensen, J.R., D.J. Cowen, J.D. Althausen, S. Narumalani, and O. Weatherbee, 1993. An evaluation of the CoastWatch Change Detection Protocol in South Carolina, *Photogrammetric Engineering & Remote Sensing*, 59:1039–1046.

Jensen, John, 1996. *Introductory Digital Image Processing: A Remote Sensing Perspective, Second Edition*, Prentice Hall, 316 p.

Luczkovich, J.J., T.W. Wagner, J.L. Michalek, and R.W. Stoffle, 1993. Discrimination of coral reefs, seagrass meadows, and sand bottom types from space: A Dominican Republic case study, *Photogrammetric Engineering & Remote Sensing*, 59:385–389.

Macleod, R.D., 1994. *Using a Quantitative Accuracy Assessment to Compare Various Change Detection Techniques for Eelgrass Distributions in Great Bay, NH with Landsat Thematic Mapper Data*, M.S. Thesis, University of New Hampshire, Durham, New Hampshire, 149 p.

Malila, W.A., 1980. Change vector analysis: an approach for detecting forest changes with Landsat, *Proceedings of the 6th Annual Symposium on Machine Processing of Remotely Sensed Data*, Purdue University, pp. 326–335.

Martin, L.R.G., 1989. Accuracy assessment of Landsat-based visual change detection methods applied to the rural-urban fringe, *Photogrammetric Engineering & Remote Sensing*, 55:209–215.

Milne, L.J., and M.J. Milne, 1951. The eelgrass catastrophe, *Scientific American*, 184:52–55.

Nelson, R.F., 1983. Detecting forest canopy change due to insect activity using Landsat MSS, *Photogrammetric Engineering & Remote Sensing*, 49:1303–1314.

Pilon, P.G., P.J. Howarth, R.A. Bullock, and P.O. Adeniyi, 1988. An enhanced classification approach to change detection in semi-arid environments, *Photogrammetric Engineering & Remote Sensing*, 54(12):1709–1716.

Short, F.T., 1989. *Eelgrass and the Wasting Disease*, UNH Cooperative Extension, Taylor Hall, University of New Hampshire, Durham, New Hampshire.

———, (editor), 1992. *The Ecology of the Great Bay Estuary, New Hampshire and Maine: An Estuarine Profile and Bibliography*, NOAA - Coastal Ocean Program Publication, 222 p.

Short, F.T., A.C. Mathieson, and J.I. Nelson, 1986. Recurrence of the eelgrass wasting disease at the border of New Hampshire and Maine, USA, *Marine Ecology Progress Series*, 29:89–92.

Short, F.T., G.E. Jones, and D.M. Burdick, 1991. Seagrass decline: Problems and solutions, reprint from Coastal Wetlands Coastal Zone '91 Conference-ASCE, Long Beach, California, July 1991, pp. 439–453.

Short, F.T., D.M. Burdick, J. Wolf, and G.E. Jones, 1993. *Eelgrass in Estuarine Research Reserves along the East Coast, U.S.A., Part I: Declines From Pollution and Disease and Part II: Management of Eelgrass Meadows*, NOAA - Coastal Ocean Program Publication, 117 p.

Short, F.T., and D.M. Burdick, 1996. Quantifying eelgrass habitat loss in relation to housing development and nitrogen loading in Waquoit Bay, Massachusetts, *Estuaries*, 19(3):730–739.

Singh, A., 1986. Change detection in the tropical rain forest environment of Northeastern India using Landsat, *Remote Sensing and Tropical Land Management* (M.J. Eden and J.T. Parry, editors), John Wiley & Sons, London, pp. 237–254.

———, 1989. Digital change detection techniques using remotely-sensed data, *International Journal of Remote Sensing*, 10(6):989–1003.

Thayer, G.W., W.J. Kenworthy, and M.S. Fonseca, 1984. *The Ecology of Eelgrass Meadows of the Atlantic Coast: A Community Profile*, U.S. Fish Wildl. Serv., FWS/OBS-84/24, 85 p.

Townshend, J.R.G., C.O. Justice, C. Gurney, and J McManus, 1992. The impact of misregistration on change detection, *IEEE Transactions on Geoscience and Remote Sensing*, 30(5):1054–1060.

Zainal, A.J.M., D.H. Dalby, and I.S. Robinson, 1993. Monitoring marine ecological changes on the East Coast of Bahrain with Landsat TM, *Photogrammetric Engineering & Remote Sensing*, 59: 415–421.

(Received 29 July 1996; revised and accepted 5 November 1997)

➤ ➤ FORTHCOMING ARTICLES ➤ ➤ ➤ ➤ ➤

Michel Arnaud and *Albert Flori*, Bias and Precision of Different Sampling Methods for GPS Positions.

Edward A. Ashton and *Alan Schaum*, Algorithms for the Detection of Sub-Pixel Targets in Multispectral Imagery.

Stéphane Chalifoux, François Cavayas, and *James T. Gray*, Map-Guided Approach for the Automatic Detection on Landsat TM Images of Forest Stands Damaged by the Spruce Budworm.

Warren B. Cohen, Maria Fiorella, John Gray, and *Karen Anderson*, An Efficient and Accurate Method for Mapping Forest Clearcuts in the Pacific Northwest Using Landsat Imagery.

F.M. Danson, Teaching the Physical Principles of Vegetation Canopy Reflectance Using the SAIL Model.

F. Deppe, Forest Area Estimation Using Sample Surveys and Landsat MSS and TM Data.

Sheldon D. Drobot and *David G. Barber*, Towards Development of a Snow Water Equivalence (SWE) Algorithm Using Microwave Radiometry over Snow Covered First-Year Sea Ice.

Hamid Ebadi and *Michael A. Chapman*, GPS Controlled Strip Triangulation Using Geometric Constraints of Man-Made Structures.

Jay Gao and *Matthew B. Lythe*, Effectiveness of the MCC Method in Detecting Oceanic Circulation Patterns at a Local Scale from Sequential AVHRR Images.

J.R. Harris, A.N. Rencz, B. Ballantyne, and *C. Sheridon*, Mapping Altered Rocks Using Landsat TM and Lithogeochemical Data: Sulphurets-Brucejack Lake District, British Columbia, Canada.

Stanley R. Herwitz, Robert E. Slye, and *Stephen M. Turton*, Co-Registered Aerial Stereopairs from Low-Flying Aircraft for the Analysis of Long-Term Tropical Rainforest Canopy Dynamics.

Ross S. Lunetta, John G. Lyon, Bert Guindon, and *Christopher D. Elvidge*, North American Landscape Characterization Dataset Development and Data Fusion Issues.

Kenneth C. McGwire, Mosaicking Airborne Scanner Data with the Multiquadric Rectification Technique.

Kenneth C. McGwire, Improving Landsat Scene Selection Systems.

Victor Mesev, The Use of Census Data in Urban Image Classification.

Jeffrey T. Morisette and *Siamak Khorram*, Exact Binomial Confidence Interval for Proportions.

S.V. Muller, S.A. Walker, F.E. Nelson, N.A. Auerbach, J.G. Bockheim, S. Guyer, and *D. Sherba*, Accuracy Assessment of a Land-Cover Map of the Kuparuk River Basin Alaska: Considerations for Remote Sensing.

Ram M. Narayanan and *Brian D. Guenther*, Effects of Emergent Grass on Mid-Infrared Laser Reflectance of Soil.

Elijah W. Ramsey III, Dal K. Chappell, Dennis Jacobs, Sijan K. Sapkota, and *Dan G. Baldwin*, Resource Management of Forested Wetlands: Hurricane Impact and Recovery Mapped by Combining Landsat TM and NOAA AVHRR Data.

Juliang Shao and *Clive S. Fraser*, Scale-Space Methods for Image Feature Modeling in Vision Metrology.

K.M.S. Sharma and *A. Sarkar*, A Modified Contextual Classification Technique for Remote Sensing Data.

E. Terrence Slonecker, Denice M. Shaw, and *Thomas M. Lillesand*, Emerging Legal and Ethical Issues in Advanced Remote Sensing Technology.

M. Stojic, J. Chandler, P. Ashmore, and *J. Luce*, The Assessment of Sediment Transport Rates by Automated Digital Photogrammetry.

Chuang Tao, Rongxing Li, and *Michael A. Chapman*, Automatic Reconstruction of Road Centerlines from Mobile Mapping Image Sequences.

Randolph H. Wynne, Thomas M. Lillesand, Murray K. Clayton, and *John J. Magnuson*, Satellite Monitoring of Lake Ice Breakup on the Laurentian Shield (1980-1994).

David A. Yocky and *Benjamin F. Johnson*, Repeat-Pass Dual-Antenna Synthetic Aperture Radar Interferometric Change Detection Post-Processing.