Available online at www.sciencedirect.com
ScienceDirect
Remote Sensing of Environment 108 (2007) 74–81
Remote Sensing of Environment
www.elsevier.com/locate/rse

Mapping an invasive plant, *Phragmites australis*, in coastal wetlands using the EO-1 Hyperion hyperspectral sensor

Bruce W. Pengra [a,1], Carol A. Johnston [a,*], Thomas R. Loveland [b]

[a] *Box 2207B, South Dakota State University, Brookings, SD 57007, United States*
[b] *U.S. Geological Survey, National Center for EROS, Sioux Falls, SD 57198, United States*

Received 28 February 2006; received in revised form 29 October 2006; accepted 4 November 2006

Abstract

Mapping tools are needed to document the location and extent of *Phragmites australis*, a tall grass that invades coastal marshes throughout North America, displacing native plant species and degrading wetland habitat. Mapping *Phragmites* is particularly challenging in the freshwater Great Lakes coastal wetlands due to dynamic lake levels and vegetation diversity. We tested the applicability of Hyperion hyperspectral satellite imagery for mapping *Phragmites* in wetlands of the west coast of Green Bay in Wisconsin, U.S.A. A reference spectrum created using Hyperion data from several pure *Phragmites* stands within the image was used with a Spectral Correlation Mapper (SCM) algorithm to create a raster map with values ranging from 0 to 1, where 0 represented the greatest similarity between the reference spectrum and the image spectrum and 1 the least similarity. The final two-class thematic classification predicted monodominant *Phragmites* covering 3.4% of the study area. Most of this was concentrated in long linear features parallel to the Green Bay shoreline, particularly in areas that had been under water only six years earlier when lake levels were 66 cm higher. An error matrix using spring 2005 field validation points ($n=129$) showed good overall accuracy—81.4%. The small size and linear arrangement of *Phragmites* stands was less than optimal relative to the sensor resolution, and Hyperion's 30 m resolution captured few if any pure pixels. Contemporary *Phragmites* maps prepared with Hyperion imagery would provide wetland managers with a tool that they currently lack, which could aid attempts to stem the spread of this invasive species.
© 2006 Elsevier Inc. All rights reserved.

Keywords: Phragmites; Remote sensing; Hyperspectral; Hyperion; Invasive species; Wetlands; Great Lakes; Green Bay

1. Introduction

Phragmites australis is an invasive species of special concern in the wetlands of the North American Great Lakes (Marks et al., 1994; Galatowitsch et al., 1999). It is a cosmopolitan species of wetland grass native to every continent but Antarctica. The genotypes native to North American have been joined by an aggressive introduced genotype, apparently brought by European settlement (Lynch & Saltonstall, 2002; Saltonstall, 2002). The primary concern relative to *Phragmites* invasion is the loss of species richness and consequent potential for extinction and loss of biodiversity (Havens et al., 1997; Chambers et al., 1999).

To understand the magnitude and distribution of *Phragmites* invasions requires research in a spatial context at several scales. At the scale of individual stands, field survey is essential to understand the myriad factors which drive invasions and the environmental changes that follow. It is equally important to determine if these small scale patterns are consistent across a wider geographic scale and to tie these small scale patterns to the larger patterns of invasion. The appropriate scale, uniformity, and timeliness required of these data are all but impossible to acquire with only field assessment and monitoring (Heywood, 1995).

Airborne hyperspectral sensors such as AVIRIS, CASI, HyMap and PROBE-1 have been somewhat successful in mapping vegetation at the species level (Bachmann et al., 2002; Lopez et al., 2004; Parker Williams & Hunt, 2002; Schmidt & Skidmore, 2001). However, the high cost and limited coverage of airborne sensors is a barrier to their use for wider scale management and study. Space platform remote sensors such as

* Corresponding author.
E-mail addresses: Carol.Johnston@sdstate.edu (C.A. Johnston), bpengra@usgs.gov (B.W. Pengra).
[1] Now with Science Applications International Corporation (SAIC), U.S. Geological Survey (USGS), Center for Earth Resources Observation and Science (EROS), Sioux Falls, SD 57198, USA.

0034-4257/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.rse.2006.11.002

Landsat provide the wider view and lower cost needed for practical applications, but have shown less facility for distinguishing species than have hyperspectral sensors including Hyperion (Goodenough et al., 2002; Thenkabail et al., 2004). The Hyperion sensor aboard the EO-1 satellite is an experimental sensor that merges the spectral resolution of airborne hyperspectral instruments with the practicality of satellite remote sensing (USGS, 2006a). We utilized Hyperion imagery to determine if a space platform hyperspectral sensor could distinguish invasive wetland plant species, as has already been done using more expensive airborne hyperspectral sensors (Bachmann et al., 2002; Lopez et al., 2004).

In this study we tested the applicability of Hyperion for mapping monodominant stands of *Phragmites* in the coastal wetlands of the Great Lakes. We acquired a Hyperion image of the west coast of Green Bay, Wisconsin and predicted the occurrence of large stands of monodominant *Phragmites* by creating a two-class thematic map. We evaluated the *Phragmites* map with field data.

2. Methods

2.1. Study site

The Hyperion image used in this study was a 7.5 km swath within Wisconsin extending northeast from the City of Green Bay to the Michigan state border. The image intersected a 30 km section of the Green Bay shoreline centered on the Pensaukee and Oconto Rivers (Fig. 1). Hyperion imagery must be tasked in advance, and image acquisition was requested to occur between September 1 and October 15, 2004 to take advantage of more pronounced spectral differences between *Phragmites* and other wetland species in mid to late autumn (Bachmann et al., 2002; Bernthal & Willis, 2004). Actual image acquisition took place on September 4, 2004 at 11:23:53 a.m. CDT., during the first satellite overpass within this time window. Cloud cover was less than 25% of the image, and was mostly west of coastal areas.

The entire coastal study area has very level terrain, and average land elevation in the study area (178.3 m) rises only 2 m above Green Bay (USGS, 2006b). The study area contains extensive areas of wetland and hydric soils (National Cooperative Soil Survey, 1988), and coastal wetlands expand and contract as water levels in Green Bay fluctuate in response to climate change (Harris et al., 1977). Lake levels at the time of image acquisition were 176.3 m, close to their historic low extremes (U.S. Army Corps of Engineers, 2006). Upland land cover consists primarily of hardwood forests, agricultural crops, pasture, and hayland (Wolter et al., 2006).

2.2. Data preprocessing

Hyperion has a low signal to noise ratio in comparison to airborne hyperspectral sensors, the result of signal lost to atmospheric absorption and the reduced energy available from surface reflectance at orbital altitude. In addition, detector arrays used in the Hyperion sensor were “spares” originally designed

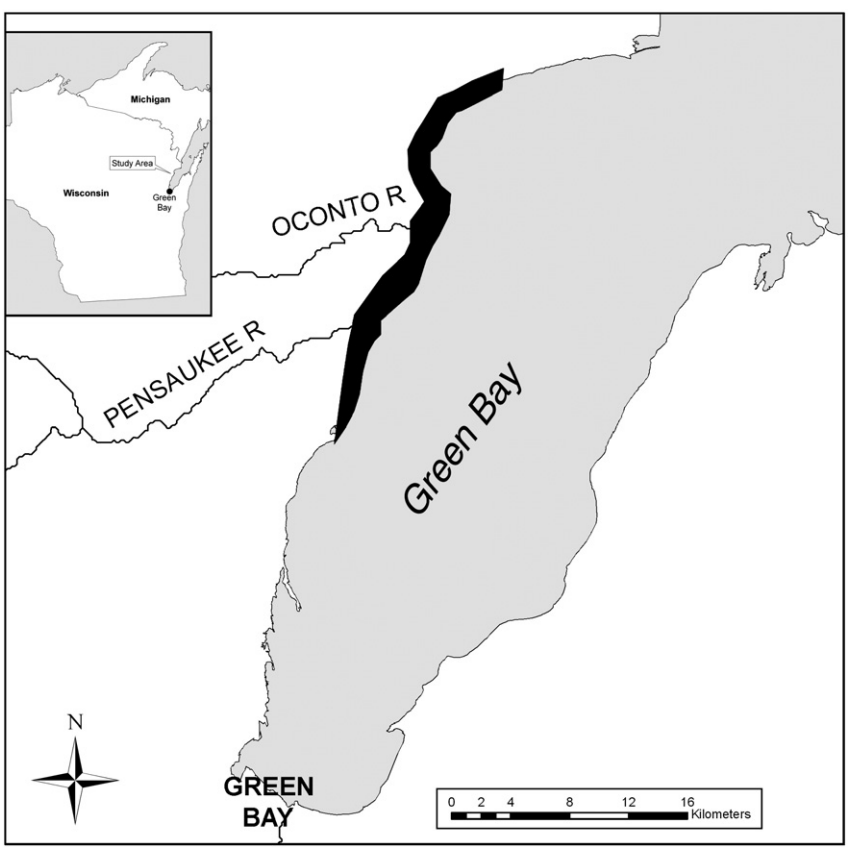

Fig. 1. Study site on the west coast of Green Bay.

Table 1
Bands removed prior to image analysis

	VNIR	SWIR
Bands removed USGS level 1 R data	1,2,3,4,5,6,7,58,59, 60,61,62,63,64,66, 67,68,69,70	71,72,73,74,75,76,226,227, 228, 229, 230,231,232, 233, 234,235,236, 237, 238,239, 240,241,242
Additional bands removed prior to image analysis due to noise		77,79,80,81,98,99,100,120, 121,122, 123,124,125,126, 127,128,129,130, 131,132, 133,134,165,166,167,168, 169,170,171,172,173,174, 175,176, 177,178,179,180, 181,182,184,185, 186,187, 190,219,220, 221,222,223, 224,225

for another purpose, which further decreases the signal to noise ratio (Jupp & Datt, 2004). As a result, Hyperion data require preprocessing to address problems of miscalibration and noise before use (Datt et al., 2003; Goodenough et al., 2002; Han et al., 2002). Full column atmospheric effects necessitate atmospheric correction for many applications (e.g. multi-temporal studies, multi-scene studies, and use of spectral libraries), but Datt et al. (2003) state that "it is not necessary to atmospherically correct image data for a single observation." Our choice of preprocessing techniques was based on validation and calibration study recommendations (Jupp & Datt, 2004) and widespread application in other studies using Hyperion (Datt et al., 2003; Kruse et al., 2003; Apan et al., 2004) and airborne hyperspectral data (Lopez et al., 2004; Underood et al., 2003).

Only 147 of the 242 total Hyperion bands were used in the analysis, because many of the bands exhibited low signal to noise ratio or other problems. Several of the Hyperion sensor's bands are unusable, particularly those at wavelengths most prone to atmospheric scattering, so that USGS delivers level 1R Hyperion data with 44 bands set to 0 values (Table 1). An additional 50 bands with excessive noise were identified by inspecting for speckling that lacked a systematic relationship to features visible in other bands (Datt et al., 2003) and removed prior to image analysis. One redundant band in the overlap between the Short Wavelength Infrared (SWIR) and Visible and Near Infrared (VNIR) sensors was also removed, leaving 147 bands to be stacked for image processing.

Hyperion is a push broom sensor, collecting simultaneously across 256 detectors along the track of the satellite's orbit (USGS, 2006a). Miscalibrated and dead detectors show up as vertical stripes in the image (Han et al., 2002; Kruse et al., 2003). In many of the striped columns, digital number (DN) values followed the general patterns of land cover reflectance but were consistently lower than those in the adjacent columns (Fig. 2). These columns were judged to contain meaningful information and were corrected by normalizing the striped columns to the mean of the two adjacent columns using a simple additive offset that was applied to each pixel in the striped column. This localized normalization was preferable to global normalization (Datt et al., 2003) because column means systematically varied across the image as increasing proportions of each column were measuring the water of Green Bay. A smaller number of the striped columns had DNs which were constant—clearly containing no meaningful information (Han et al., 2002). These columns were corrected by interpolating values from the corresponding pixels in the immediately adjacent columns (Datt et al., 2003; Pearlman, 2003).

The panchromatic band of a Landsat 7 Enhanced Thematic Mapper Plus (ETM+) orthorectified image was used as a base map for georectification. The 15 m cell size provided adequate resolution and the orthorectification insured that the geographic fidelity was consistent throughout the image. A 2nd order polynomial transformation was used, requiring a minimum of 6 control points (LEICA Geosystems, 2003a, pg. 339). Nine reliable control points were located with good distribution throughout the image. The resulting transformation had an estimated total root mean square error of .2962 pixels or about 8.9 m, providing adequate spatial accuracy for the small features being mapped. Nearest neighbor resampling was used to retain as much spectral fidelity as possible.

To minimize the influence of systematic sensor noise during image analysis, the minimum noise fraction (MNF) transformation was used (Green et al., 1988; Lee et al., 1990). The MNF transform identifies systematic noise (which is presumed to

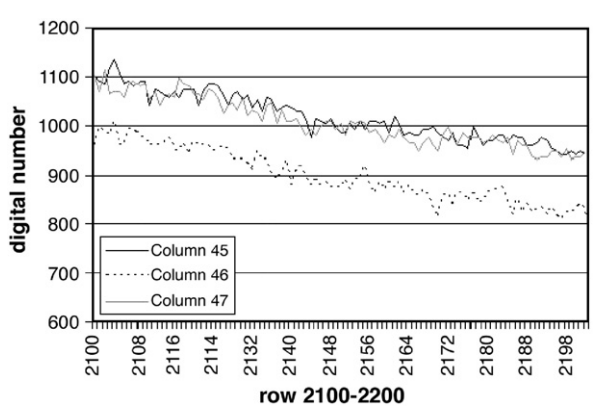

Fig. 2. Digital number data in image column 46 are consistently lower than data in adjacent columns due to a miscalibrated sensor.

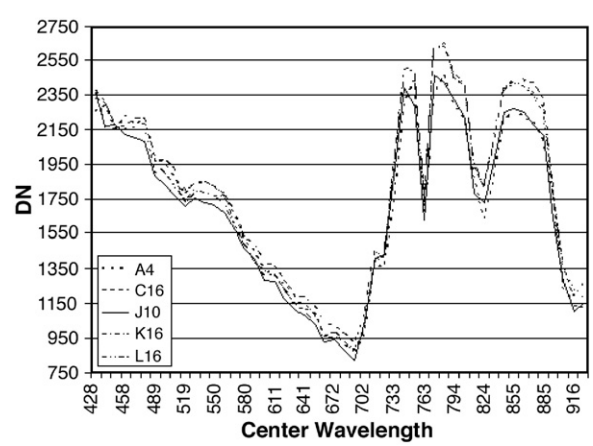

Fig. 3. Five training pixel candidates were chosen based on field data. Their spectra plotted through the bands of Hyperion's VNIR sensor showed them to be similar with J10 being the most anomalous.

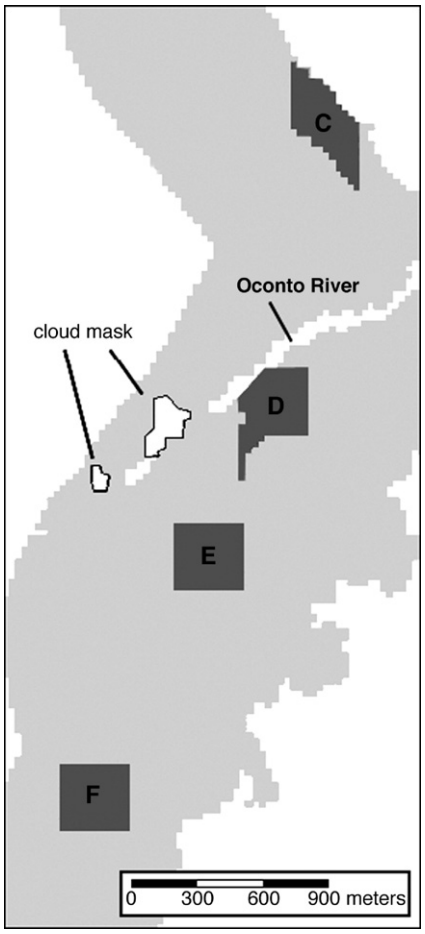

Fig. 4. Four of the twelve 121-pixel sampling windows (dark grey) used in the two-stage cluster sampling of the coastal study area (light grey). Windows E and F are 11 × 11 pixels; the shapes of C and D have been altered to accommodate lake and river shorelines that are outside of the study area.

arise from sensor and processing anomalies) and can then segregate it from the meaningful signal. Covariance for the MNF transformation was calculated from a subset of the image centered over a large, spectrally homogeneous area of open water to minimize the possible contribution of systematic variations in the valid signal from being confused with sensor noise (LEICA Geosystems, 2003b, pg. 163). The transform was applied separately to each trial of the target detection process.

Clouds were masked out of the image because they obscured ground conditions, and the waters of Green Bay too deep to support *Phragmites* growth were also masked out.

2.3. Image analysis

Image analysis was performed with the ERDAS IMAGINE™ 8.7 Spectral Analysis module (LEICA Geosystems, 2003b). We used a target detection mode of operation, which searched the input image for target material suspected to be present in very low concentration (in this case, *Phragmites*). A reference spectrum representing *Phragmites* was developed for use in target detection by extracting from the Hyperion image five spectral data sets centered on pure *Phragmites* stands identified during field reconnaissance (Fig. 3). The initial five spectra appeared to be similar, but two of the spectra (J10 and L16) were subsequently eliminated based on field data and visual comparison of each sample spectrum to the mean spectrum of all five *Phragmites* samples (Fig. 3). The three retained samples (A4, C16, K16) were distributed across the study site, with a minimum distance between points of 7.5 km. Spectra for these three samples were merged into a single reference spectrum. The target detection operation was then run with the Spectral Correlation Mapper (SCM: de Carvalho & Meneses, 2000) spectral analysis methodology, which created a continuous value raster with values ranging from 0 to 1, where 0 represented the greatest similarity between the reference spectrum and the image spectrum, and 1 the least similarity. The SCM is a refinement of the Spectral Angle Mapper (SAM) algorithm, which determines the spectral similarity between two spectra by calculating the angle between the two spectra, treating them as vectors in a space with dimensionality equal to the number of bands (Kruse et al., 1993). SAM (and by inference SCM) is designed for use primarily with hyperspectral data, and have broad application in hyperspectral remote sensing of vegetation (Artigas & Yang, 2004; de Lange et al., 2004; Eckert & Kneubühler, 2004; Lumme, 2004; Ustin et al., 2002). Although both spectral mapper algorithms are supported in ERDAS IMAGINE (LEICA Geosystems, 2003b), we chose the SCM algorithm because it standardizes vectors prior to calculating the spectral angles such that positive and negative correlations between samples can be distinguished, whereas

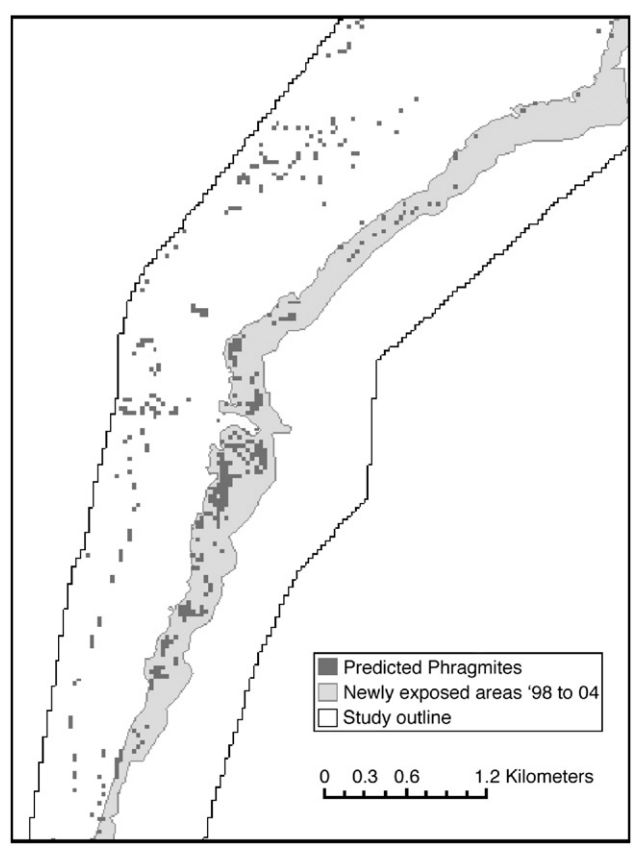

Fig. 5. Approximately 52% of the predicted *Phragmites* was on the newly exposed mudflat between the 1998 and 2004 shorelines.

Table 2
Error matrix using spring 2005 reference points

Hyperion class	Field survey data			User's accuracy
	Non-Phragmites	Phragmites	Total	
Non-*Phragmites*	**83**	10	93	89.2%
Phragmites	14	**22**	36	61.1%
Total	97	32	129	
				Overall accuracy
Producer's accuracy	85.6%	68.8%		81.4%

SAM does not distinguish between positive and negative correlation (Lumme, 2004).

The continuous value raster layer created with the merged spectra of these three sample points was used to create a final binary thresholded (yes/no) image of predicted monodominant *Phragmites*. This map was created using an iterative process whereby the threshold for the *Phragmites* class was adjusted and then tested against the field validation data until the maximum overall accuracy was achieved.

2.4. Field data collection

Field validation data were collected in Autumn 2004 and Spring 2005. Autumn field data were collected at 38 sample locations shortly after image acquisition, but before image processing. Additional data were acquired in Spring 2005 ($n=129$) to better randomize the validation data and to increase the total number of validation sites to 167. *Phragmites* culms from one season will generally remain standing through the next growing season (Haslam, 1969), which allowed reliable field data collection despite the temporal offset between image acquisition and Spring 2005 field work. Autumn field samples targeted prospective locations of large *Phragmites* stands, pre-selected based on aerial photography, the Wisconsin Wetlands Inventory, United States Geological Survey (USGS) quadrangle maps, and the Hyperion image itself. Spring field data were collected based on a two-stage cluster sampling scheme in which 121-pixel sampling windows (usually 11×11 pixel squares) were defined around twelve points randomly selected on the image (Fig. 4). Within those 12 sampling windows, a total of 129 points was selected by stratified sampling of the two strata, *Phragmites* and not-*Phragmites*. This two-stage clustering scheme thus provided randomization but reduced expensive field travel time by limiting sampling to the twelve sampling windows. These randomized spring data were used to assess the accuracy of the binary map of predicted monodominant *Phragmites*. The full data set of 167 sites was used to evaluate confusion between *Phragmites* and nine other land cover classes: cattail (*Typha*), mixed emergent vegetation (MixedEm), meadow, conifer forest (ForestC), deciduous forest (ForestD), scrub, impervious, shallow water, and "other." The purpose of utilizing these classes was not to map them, but to classify error, and the full data set of 167 sites was needed to increase representation of several of the land cover classes.

Methods for recording field observations were the same for both sets of samples. The sample area was defined as a 60 m × 60 m square, the equivalent site of four adjacent pixels, centered on the sample point. Data collected for each sample point included a description of the vegetation, a sketched map of vegetation distribution, digital photographs, GPS location, and estimated GPS accuracy. Relative vegetation cover values between 1 (sparse) and 6 (very dense), corresponding to Braun-Blanquet cover class ranges (ASTM, 1997) were recorded for major polygons of vegetation associations drawn on the map. Based on field data, a sample area was classified as *Phragmites* when *Phragmites* of density class 5 or 6 (70% cover) filled a majority of the sample area or where *Phragmites* of density class 4 (50 to 70% cover) filled the entire sample area. Other vegetation classes were similarly defined.

3. Results

The final classification predicted monodominant *Phragmites* covering 3.4% of the study area. Most of this was concentrated in long linear features parallel to the Green Bay shoreline. In many cases these were areas that had been under water only six years earlier when lake levels were 66 cm higher. On the strip of land between the 1998 water level and the 2004 water level, the final classification predicted 13.3% of the area to be covered by dominant *Phragmites* (Fig. 5). Just over half of the study area's predicted *Phragmites* was concentrated within this lakebed exposure zone, which constituted less than 15% of the study area.

An error matrix using the spring 2005 field validation points showed good overall accuracy—81.4% (Table 2). Of the 36 points predicted to be dominated by *Phragmites*, 22 had been accurately predicted for a user's accuracy of 61.1% for the *Phragmites*-dominant class. Of the 32 validation points where the land cover was actually dominated by *Phragmites*, 22 points had been correctly predicted for a producer's accuracy of 68.8%.

An error matrix between *Phragmites* and other land cover types identified specific sources of confusion (Table 3). The conifer (*ForestC*) class, which was predominantly tree plantations, was the most anomalous area of confusion. The largest contiguous area of erroneously predicted *Phragmites* was

Table 3
Errors distributed across land cover classes of all field validation points

Hyperion class	Field survey data											User's accuracy
	ForestC	ForestD	Imperv	Meadow	MixedEM	Other	*Phrag.*	Scrub	*Typha*	Water	Total	
Non-*Phragmites*	**3**	**11**	**1**	**25**	**18**	**2**	17	**18**	**14**	**14**	123	86.2%
Phragmites	4	0	1	0	6	0	**28**	0	4	1	44	63.6%
Total	7	11	2	25	24	2	45	18	18	15	167	
Producer's accuracy	42.9%	100.0%	50.0%	100.0%	75.0%	100.0%	62.2%	100.0%	77.8%	93.3%		Overall accuracy 80.24%

actually a conifer plantation located approximately 1 km from the shoreline. Other similar sized stands of conifers in the study area also contained scattered *Phragmites* false positives. In the error matrix which included the complete set of reference points (Table 3), 4 of the 7 *ForestC* reference points were erroneously predicted to be *Phragmites*. None of these 4 sample sites contained any *Phragmites*.

More predictably, there was also some confusion between the mixed emergent vegetation class (*MixedEm*) and the *Phragmites* class, with 6 of the 24 mixed *MixedEm* points being incorrectly predicted to be dominated by *Phragmites*. In this case 5 of the 6 points did contain *Phragmites,* but at densities that did not meet the class criteria. Four of the 18 *Typha* reference points (Table 3) were erroneously predicted to be monodominant *Phragmites*. Although many of the sample points for *Typha* and for *Phragmites* contained some of both genera, only one of these four incorrectly predicted points contained any *Phragmites*.

4. Discussion

The *Phragmites* maps prepared using the Hyperion imagery represent a substantial advance over existing wetland maps for the western coast of Green Bay. Wisconsin Wetlands Inventory maps for the region were prepared using aerial photos from 1978 and 1979 (Johnston, 1984), when water levels in Green Bay were nearly 70 cm higher than they were in 2004 (U.S. Army Corps of Engineers, 2006). Water level fluctuations on this scale have occurred in Lake Michigan regularly over time spans of 10 to 30 years since at least 1918. Because of the gradually sloping bathymetry along the west coast of Green Bay these fluctuations lead to dramatic changes in the extent and location of the coastal wetlands (Harris et al., 1977). It is not surprising that there was almost no match between areas mapped as *Phragmites* by this project and areas mapped as emergents by the Wisconsin Wetlands Inventory because of the large change in environmental conditions that occurred over the 25 years between the two inventories (Pengra, 2005).

The *Phragmites* maps prepared by this project could alert local resource managers to "hotspots" of *Phragmites* invasion that could be targeted for early control efforts. When asked to identify large *Phragmites* stands as potential field reconnaissance sites in Fall 2004, local employees of the Wisconsin Department of Natural Resources identified a stand that was close to a highway, but were apparently unaware of much larger *Phragmites* stands subsequently found by this project's field researchers that were less visible from major highways. Reasonably accurate, contemporary *Phragmites* maps prepared with Hyperion imagery would provide wetland managers with a tool that they currently lack, one that could aid attempts to stem the spread of this invasive species. The dynamic nature of these wetlands requires more frequent mapping than once every 25 years.

User's accuracy for predicted *Phragmites* (61.1%) compares favorably with previous attempts at remote sensing of *Phragmites* with airborne hyperspectral sensors, despite Hyperion's much coarser spatial resolution and reduced signal to noise ratio. Bachmann et al. (2002) distinguished *Phragmites* from other wetland species with 68% accuracy using the HyMap airborne sensor with a spatial resolution of 4.5 m. Lopez et al. (2004) were able to distinguish between monodominant stands of *Phragmites*, Typha and mixed wetland vegetation with 91% accuracy using the PROBE-1 airborne hyperspectral sensor with 5 m resolution. In addition to the difference in sensors and sensor platforms used, our study differed from theirs by covering a much longer swath of coastline (28 km vs. 6 km) containing multiple wetland complexes, presumably encompassing a wider range of environmental conditions.

Phragmites false positives occurring at MixedEm and *Typha* reference points were expected due to the intermixing of these vegetation classes and the similar structure and context of the areas where these points were located. The *Phragmites* false positives occurring in the ForestC areas were not expected. However, similar errors of commission were experienced by Ernst-Dottavio et al. (1981) albeit with Landsat multispectral data rather than hyperspectral data. They found, "The mixing of water and vegetation in deep marshes results in a spectral response pattern that is similar to conifers …this may result in errors of commission where conifers are classified as deep marshes." Comparison of the plotted mean spectra of *Phragmites* training samples and the mean spectra of the ForestC reference points shows them to be similar in hyperspectral data as well (Fig. 6). Through the visible portion of the spectrum, reflectance is consistently lower for the ForestC reference samples. However, from the red edge through most of the bands of the NIR and SWIR the mean spectra are very similar. In spite of removal of 51 bands from the SWIR range, the number of SWIR bands (wavelengths N1000 nm) used was roughly twice the number of VNIR bands (wavelengths b 1000 nm) used. The SCM algorithm gives equal weight to all of the bands included in the data; therefore the image analysis relied heavily on the SWIR data. Spectral differences between *Phragmites* and 26 other saltmarsh vegetation associations have been found to occur at wavelengths below 1100 nm (Schmidt & Skidmore, 2003). If this holds true for freshwater wetland vegetation

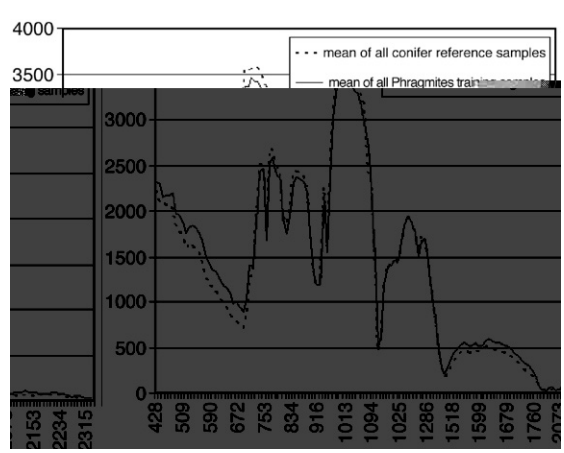

Fig. 6. Spectral plots of the mean of all conifer reference samples and the mean of the *Phragmites* training samples suggests that most of the similarity occurs in the SWIR wavelengths.