Coastal wetland vegetation classification with a Landsat Thematic Mapper image

YINLONG ZHANG, DENGSHENG LU, BO YANG, CHUNHUA SUN and MING SUN

In: International Journal of Remote Sensing
Coastal wetland vegetation classification with a Landsat Thematic Mapper image

YINLONG ZHANG†, DENGSHENG LU*‡, BO YANG†, CHUNHUA SUN† and MING SUN§
†College of Forest Resources and Environment, Nanjing Forestry University, Nanjing, 210037 China
‡Anthropological Center for Training and Research on Global Environmental Change (ACT), Indiana University, Bloomington, Indiana, 47405, USA
§Yancheng National Nature Reserve, Yancheng, Jiangsu Province, China

(Received 14 April 2008; in final form 6 August 2009)

Coastal wetland vegetation classification with remotely sensed data has attracted increased attention but remains a challenge. This paper explored a hybrid approach on a Landsat Thematic Mapper (TM) image for classifying coastal wetland vegetation classes. Linear spectral mixture analysis was used to unmix the TM image into four fraction images, which were used for classifying major land covers with a thresholding technique. The spectral signatures of each land cover were extracted separately and then classified into clusters with the unsupervised classification method. Expert rules were finally used to modify the classified image. This research indicates that the hybrid approach employing sub-pixel information, an analyst’s knowledge and characteristics of coastal wetland vegetation distribution shows promise in successfully distinguishing coastal vegetation classes, which are difficult to separate with a maximum likelihood classifier (MLC). The hybrid method provides significantly better classification results than MLC.

1. Introduction

Wetlands are important landscapes that can hold and slowly release flood water, recharge groundwater, improve water quality by acting as filters to cleanse water pollution, and provide habitats for fish and wildlife including many threatened and endangered species (Steven and Toner 2004, Lagos et al. 2008). The important ecological role of wetlands and their economic benefits have been increasingly recognized (Bridgham et al. 2006, Canepuccia et al. 2007). Accurately mapping wetland land cover types and monitoring their dynamic changes provide the scientific foundation for wetland protection and restoration, as well as biodiversity protection (McAllister et al. 2000, Ozesmi and Bauer 2002). As an important stopover for seasonal bird migration along the East Asian–Australasian Flyway, Yancheng National Nature Reserve (YNNR) in Jiangsu province, China provides critical habitat for different kinds of endangered species (Xu et al. 2005).

Every spring and autumn, about 3 000 000 seasonal birds stop by this wetland region for food, and about 200 000 water birds stay for the winter season (Xu et al. 2005). The muddy shoals provide rich foods for different kinds of birds during their

*Corresponding author. Email: dlu@indiana.edu
migration. For example, about half of the world’s population of red-crowned cranes (\textit{Crocus japonensis}) overwinter in YNNR. Habitat requirements for red-crowned cranes and other wildlife include food, water, and shelter. Water in the YNNR is not a constraining factor; vegetation is the key constraint affecting the red-crowned crane population and other wildlife species, as food and shelter are mainly dependent on vegetation in the coastal wetland. Different vegetation types in the coastal wetland provide different kinds of foods and quality of shelter for birds. Much previous research has been conducted in this region, but it mainly focuses on physiology, ecology, environmental carrying capability, and wildlife habitat based on field surveys in a limited area along the transition where human beings can access (Liu et al. 2003, Li et al. 2004b, 2005, 2006, Xu et al. 2005, Lü et al. 2006, Ou et al. 2006). One of the most important data types, coastal wetland vegetation distribution, is not available. Due to the difficulty in accessing coastal wetland areas, field surveys for mapping vegetation distribution in a large area are very difficult to undertake. Remote sensing techniques offer promising solutions to this problem.

Remote sensing techniques provide unique tools for mapping wetland land cover distribution in a large area. Rundquist et al. (2001) reviewed the remote sensing approaches for wetland classification and change detection, and Ozesmi and Bauer (2002) reviewed the satellite remote sensing of wetlands, including a summary of different satellite sensor data and techniques used for wetland identification and classification. However, in practice, wetland land cover classification, especially vegetation classification, is a challenge because of the complexity of wetland landscapes and the impacts of biophysical variables such as water levels, salt content, vegetation rigour and density. Therefore, much previous research used hyperspectral data to conduct wetland vegetation classifications due to its specific advantage in spectral resolution (Schmidt et al. 2004, Pengra et al. 2007). For example, Rosso et al. (2005) and Li et al. (2005a) used multiple endmember spectral mixture analysis to unmix hyperspectral AVIRIS (Airborne Visible/Infrared Imaging Spectrometer) data for mapping \textit{Spartina} in San Francisco Bay marshes and for mapping \textit{Spartina}, Salicornia, and Grindelia in China Camp, CA, USA, respectively. Vegetation species could be identified from hyperspectral image using spectral libraries that were measured in the field (Schmidt and Skidmore 2003). However, cloudy conditions and much moisture in the wetland landscape make it difficult to collect high quality hyperspectral images. Thus, radar data are another important data source for wetland classification (Augustijn and Warrender 1998, Baghdadi et al. 2001, Hess et al. 2003, Martinez and Toan 2007). The use of multi-temporal radar data (Hess et al. 2003, Martinez and Toan 2007) or a combination of radar and Landsat/Satellite Pour l’Observation de la Terre (SPOT) multispectral data improved wetland classification performance (Toyra and Pietroniro 2005).

For many study areas, the use of hyperspectral or radar data may be difficult due to data availability, cost constraints, and the requirement of advanced software and skills in processing hyperspectral and/or radar data. Therefore, optical-based multispectral data such as Landsat Thematic Mapper (TM) images are still the most common and important data source for wetland classification and monitoring (Pope et al. 1994, Sader et al. 1995, Ramsey and Laine 1997, Lunetta and Barlogh 1999, Harvey and Hill 2001, Townsend and Walsh 2001, Phillips et al. 2005, Baker et al. 2006, Wright and Gallant 2007). Literature reviews for mapping coastal wetland using optical sensor data such as SPOT, Landsat Multispectral Scanner (MSS) and TM images (Hardisky et al. 1986) and/or other optical and active sensor data (Ozesmi and Bauer 2002) were provided. Previous research has shown that it is difficult to produce
satisfactory wetland land cover classification with pure spectral signatures. Incorporation of ancillary data into spectral images and use of advanced classifiers can improve classification performance (Baker et al. 2006, Wright and Gallant 2007). For example, Wright and Gallant (2007) used a classification tree approach to combine TM imagery and ancillary environmental data to improve wetland classification in Yellowstone National Park. In wetland classifications, selection of a suitable image acquisition date is especially important because of the impacts of seasonal changes in water content and phenological variations in vegetation. The use of multitemporal optical sensor imagery has been proven useful to improve wetland vegetation classification (Ramsey and Laine 1997, Lunetta and Barlogh 1999, Townsend and Walsh 2001, Ghioca-Robrecht et al. 2008, Gilmore et al. 2008).

Although many classification approaches (Lu and Weng 2007), such as neural network (Augusteijn and Warrender 1998, Filippi and Jensen 2006), multiple end-member spectral mixture analysis (Li et al. 2005a, Rosso et al. 2005), and maximum likelihood (Laba et al. 2008) have been explored for wetland land cover classifications, coastal wetland vegetation classification with Landsat TM imagery remains a challenge, especially for regions such as YNNR where hyperspectral or radar data are not available. The availability of time series of Landsat data since the 1970s makes it the primary data source for land cover classification and change detection. It is an urgent task to develop a suitable procedure for wetland vegetation classification based on Landsat TM imagery. Hence, this paper aims to develop a suitable procedure to use Landsat TM imagery for improving coastal wetland vegetation classification by making full use of remote sensing techniques and expert knowledge.

2. Study area

Yancheng National Nature Reserve (YNNR) is located at the coastal region of Jiangsu Province, China, with longitude/latitude of 32° 20' ~34° 37' N and 119° 29' ~121° 16' E (see figure 1). The YNNR is the largest tidal flat wetland nature reserve in

Figure 1. The study area: Yancheng National Nature Reserve, Jiangsu province, China, with county boundaries and water bodies (grey).
the world, with a total area of approximately 4530 km\(^2\) and the coastal line of approximately 528 km. Established in 1983, YNNR was promoted as a national nature reserve by the Chinese government in 1992, and accepted as an international biosphere reserve by the United Nations Educational, Scientific and Cultural Organization (UNESCO) Biosphere Committee in 1993 (Ou et al. 2006). YNNR has become an important habitat for wildlife especially endangered birds (Zheng and Wang 1998). The weather in YNNR is a transition from warm temperate to subtropical, and has characteristics of a monsoon climate, i.e. warm and humid, rich sunshine, abundant precipitation and four distinct seasons. The annual average temperate is 14–15°C with an average temperature of 0–2.5°C in January and of 26.5–27.5°C in July. Annual precipitation ranges from 900 mm to 1050 mm (Lü et al. 2006).

In this research, the core region within YNNR was selected as the study area (see figure 1). Two rivers in north and south separate the core region from other areas. Most man-made wetlands (such as ponds for fishery) are located in north-west and west parts of the study area. The major vegetation types in this coastal wetland area are Phragmites australis (Phragmites hereafter), Suaeda Salsa (Suaeda hereafter), and Spartina alterniflora (Spartina hereafter). The soils here belong to salty soils, classified as tide salty according to the soil-formation process. Spartina, an invasive species, which was imported in the 1980s, is expanding rapidly and seriously affecting the expansion of Suaeda to the coastal direction.

3. Methods

3.1 Characteristics of wetland vegetation types

Sample plots with different vegetation types, such as Phragmites, Spartina and Suaeda, were identified during the fieldwork in June 2006. Because of the difficulty in accessing the coastal wetland area, typical sites that can be accessed were selected with special attention to the collection of different vegetation types. A comprehensive analysis of our field investigation and previous research work in this region (e.g. Li et al. 2005b, Wang and Chang 2005) show the typical land cover distribution and their spatial patterns. The distance from land to ocean reflects the change in tidal levels and represents the difference in environmental conditions which affect the vegetation distribution. A conceptual model illustrating the relationship between vegetation distribution and coastal wetland environment conditions is expressed in figure 2. Spartina grows well in salty and wet soil conditions with low tidal areas close to the shoal, whereas other vegetation types such as Phragmites cannot. As salt content is reduced in the areas away from the ocean, other vegetation types such as Phragmites and Suaeda occupy land with high and medium tidal levels. Phragmites has a relatively extensive adaptation to different environmental conditions and can grow in e.g. fresh water or water with low salinity levels, or areas with wet or dry soils. Based on the analysis of vegetation distribution, the objectives of this research, and the characteristics of the study area, a classification system consisting of seven classes was used in this research: forest, Phragmites, Suaeda, mixed Phragmites and Suaeda, Spartina, mixed Suaeda and Spartina, and others (e.g. shoals, water, road and bare land).

3.2 Image preprocessing

A Landsat 5 TM image (path/row: 119/37) acquired on 23 May 2004 was used for this study. The TM image was geometrically rectified into the Universal Transverse
Mercator coordinate system with 35 control points collected from 1:50,000 topographic maps with a root mean square error of less than 0.5 pixels. The nearest neighbour resampling technique was used to resample the TM image into a pixel size of 30 m by 30 m during image rectification. An apparent reflectance model was used to convert TM digital number to at-satellite reflectance (Markham and Barker 1986, Lu et al. 2002). This model corrects the effects caused by the solar radiance and sun zenith angle, but ignores the effects caused by atmospheric scattering and absorption.

3.3 Wetland vegetation classification with the maximum likelihood classifier

The maximum likelihood classifier (MLC) is a parametric classifier that assumes normal or near-normal spectral distribution for each feature of interest. An equal prior probability among the classes is also assumed. This classifier is based on the probability that a pixel belongs to a particular class. It takes the variability of classes into account by using the covariance matrix. MLC requires a sufficient number of training sample plots for each class with representative spectral signatures in order to accurately estimate the mean vector and covariance matrix needed by the classification algorithm. A detailed description of MLC can be found in many textbooks (Richards and Jia 1999, Lillesand and Kiefer 2000, Jensen 2004). MLC may be the most common classifier used in practice because of its sound theory and availability in any commercial image processing software, thus MLC is used in this research to examine the separability of vegetation classes based on per-pixel spectral signatures.

Selection of a sufficient number of representative training sample plots for each class is one of the critical steps during the land cover classification with MLC. In this study, a total of 135 training sample plots were collected, with approximately 10–20 plots (except for forest and the mixed Suaeda and Spartina class, because forest was
only distributed in a small area along the road in north part of the study area and the mixed *Suaeda* and *Spartina* class has limited and narrow areas within the wetland region where it is difficult for human beings to access) used for each class. A window size of $3 \times 3$ or $5 \times 5$ pixels or lines were collected for each plot, depending on the homogeneity of the land cover and the shape (such as roads), and used for training samples. Transformed divergence analysis was used to examine the separability of the training sample classes. The sample plots were further refined based on the separability analysis. Finally, the maximum likelihood classifier was used to classify the Landsat TM image into a thematic map.

### 3.4 Wetland vegetation classification with a hybrid approach

A hybrid approach based on linear spectral mixture analysis (LSMA), thresholding technique, unsupervised classification and post processing was developed, as illustrated in figure 3. The major steps included (1) unmixing the TM multispectral image
into four fractional images; (2) classifying the fraction images into major land-cover classes with a thresholding technique; (3) extracting the spectral signatures of each land cover class and conducting unsupervised classification, respectively; and (4) modifying the classified image with expert knowledge.

3.4.1 Development of fractional images from the Landsat TM image.  The LSMA is regarded as a physically based image processing tool and supports repeatable and accurate extraction of quantitative subpixel information (Adams et al. 1995, Mustard and Sunshine 1999). This approach assumes that the spectrum measured by a sensor is a linear combination of the spectra of all components (endmembers) within the pixel and the spectral proportions of the endmembers represent proportions of the area covered by distinct features on the ground (Mustard and Sunshine 1999). Much previous literature has detailed this approach (e.g. Adams et al. 1995, Mustard and Sunshine 1999, Lu et al. 2003). One critical step is to select suitable endmembers. In practice, image-based endmember selection approaches are often used because image endmembers can be easily obtained and they represent the spectra measured at the same scale as the image data. Image endmembers can be derived from the extremes of the image feature space, assuming that they represent the pure pixels in the images (Lu et al. 2003, Theseira et al. 2003). In order to effectively identify image endmembers and to achieve high-quality endmembers, principal component analysis (PCA) and minimum noise fraction (MNF) are often used to transform the multispectral images into a new dataset (Green et al. 1988, Boardman and Kruse 1994). Endmembers are then selected from the feature spaces of the transformed images (Garcia-Haro et al. 1996, Cochrane and Souza 1998, Van der Meer and De Jong 2000). In this study, MNF was used to convert the TM image into a new dataset, in which the first three components accounted for the majority of information (approximately 98%) from the Landsat TM image. Since fresh water and salt water are regarded as important factors affecting vegetation distribution in coastal wetland areas, four endmembers—vegetation, salt water, fresh water, and soil—were identified based on the scatterplots of MNF components and visual interpretation of Landsat TM colour composite. The four endmembers were also identified from the scatterplots of the first three MNF components. The spectral signatures of the four endmembers from the TM colour composite and from MNF components were compared and finally the endmembers representing the extreme spectral signatures were selected. An unconstrained least squares solution was then used to unmix the six TM reflective bands into four fraction images.

3.4.2 Land-cover classification with a thresholding technique.  After four fraction images were developed with the LSMA approach, a thresholding technique was used to classify the fraction images into five classes, as illustrated in figure 3. The thresholds were initially selected through the examination of sample plots with window size of 3 × 3 pixels on the fraction images, i.e. based on the statistical analysis of mean and 2.5 standard deviation by considering a minimum and maximum from the selected samples for each land cover class. The thresholds were then modified with the trial-and-error procedure. In this way, a threshold of 0.6 was used for both saltwater and freshwater fraction images to generate saltwater and freshwater dominated classes. A threshold of 0.5 was used for both vegetation and soil fraction images to produce vegetation and soil dominated classes respectively. Pixels not falling within the above thresholds were grouped into the ‘other’ class.

Each classified class was then combined with the TM spectral image for separate extraction of spectral signatures. For each extracted spectral image, an unsupervised
classification was used to classify the spectral signatures into 30 clusters. The analyst was responsible for merging the clusters into meaningful classes for each image according to the defined classification system. Finally, the five classified images were combined to generate a new image for a complete land-cover classification image for the coastal wetland study area.

3.4.3 Modification of classified image with expert knowledge. Some vegetation classes have similar spectral signatures because of the impacts of the complexity of biophysical environments. For example, Phragmites can grow in different environments, such as uplands, fresh water and salty water, resulting in significantly different spectral features in Landsat TM data for the same Phragmites class. The different vigour and density of Spartina and its mixture with residuals also generate different spectral signatures, resulting in confused spectral signatures among Phragmites, Spartina and other land covers. The classification approaches cannot automatically separate them based on their spectral signatures. However, Spartina is distributed in low tidal areas with high salt water content and Phragmites is distributed away from the shoal, as shown in figure 2. Based on this knowledge, the confusion between Phragmites and Spartina in the classified image can be corrected through the use of expert knowledge of vegetation distribution.

3.5 Accuracy assessment

Accuracy assessment of the classified image is an important part in the image classification procedure. The error matrix is the most frequently used method in accuracy assessment (Foody 2002) and previous literature has detailed this method (Congalton and Mead 1983, Hudson and Ramm 1987, Congalton 1991, Congalton and Green 1999, Congalton and Plourde 2002, Foody 2002). One critical step is to select a sufficient number of test sample plots for each class with a suitable sampling method. Random or stratified random sampling approaches are often used for a robust assessment of classification results. In this study, collection of test samples based on a random sampling technique is impossible because of the difficulty accessing the sites in the coastal wetland region and lack of high spatial resolution images such as aerial photographs to assist the identification of land covers. Hence, test sample plots for typical vegetation classes (e.g. forest, Phragmites, Suaeda, Spartina, mixed Phragmites and Suaeda, and mixed Suaeda and Spartina) were collected based on fieldwork and with the assistance of local experts. Other non-vegetation types (e.g. water, shoal, bare lands and roads) were directly selected from the TM image.

Although random sampling was not used during the collection of reference data, much attention was paid to avoiding spatial autocorrelation and bias. Based on our field survey and assistance from local experts, the sample size for each class roughly matched the area amount of the corresponding class and the sample plots were selected in different locations. Some land covers such as forest and the two mixed vegetation classes have limited areas in this study area, so a limited number of test sample plots were also collected in order to examine their classification accuracy. Therefore, the sample sizes for Phragmites, Suaeda, Spartina and the ‘other’ class are relatively large, and for forest and the two mixed classes are relatively small. Finally, a total of 151 test sample plots were collected and overlain on the classified images for examining the classification performance. The same test sample plots were used to evaluate both classified images from MLC and hybrid methods and the corresponding error matrix was developed independently. The producer’s accuracy, user’s accuracy and overall accuracy were calculated from the error matrix. Because the objective of this research is to identify a suitable classification method for wetland vegetation classification, it is
important to examine if the new method is significantly better than MLC. Therefore, kappa analysis, a powerful technique used for analysing a single error matrix and for comparing the differences between different error matrices (Congalton 1991, Smits et al. 1999, Foody 2009), was used in this paper. Kappa statistic, kappa variance and Z statistic were used to compare the performance between MLC and hybrid methods.

4. Results and discussion

4.1 Analysis of fractional images

Four fractional images were developed from the Landsat TM image with LSMA (see figure 4). Fresh water and salt water were selected as two endmembers because of

![Fractional images](image)

Figure 4. The four fraction images (a – fresh water, b – salty water, c – soil, and d – vegetation) derived from the spectral mixture analysis of the Landsat 5 TM image.
their specific roles in affecting vegetation type and distribution in the coastal wetland environment. The freshwater fraction image highlights the proportion of fresh water, which is mainly located in fishing ponds away from the ocean. The saltwater fraction mainly shows saltwater information in the ocean and rivers directly connecting to the ocean. The soil fraction image indicates the soil or vegetation residuals in the study area. In May, vegetation leaves had not fully developed to cover the ground, thus bare soils and vegetation residuals appear as high values in the soil fraction image because of the similar spectral features. Roads also appear as high values in the soil fraction image. The vegetation fraction image highlights the proportion of green vegetation information in a pixel. The *Phragmites* in upland and in freshwater areas and some *Spartina* areas appear as high values in the vegetation fraction image due to their growth vigour. Other vegetation types such as *Suaeda* and some *Spartina* areas have relatively low values in the vegetation fraction image due to their density and vegetation residuals. The different characteristics in the fraction images provide the basis for stratification of major land cover types with a thresholding technique on the fraction images.

### 4.2 Comparison of coastal wetland vegetation classification accuracies

The classification results indicate that the hybrid method improved overall classification accuracy by 11.26% and kappa coefficient by 0.15, compared with MLC (see table 1). The hybrid method is significantly better than ML at 99.5% confidential level. In the MLC-based classification result, the *Phragmites* class is confused with other land covers, especially with *Spartina*. The spectral signature of *Phragmites* is very complex because *Phragmites* can grow in different environments such as soils of varying moisture content and salinity levels. The vegetation density and water content are important factors affecting the spectral signatures. Therefore, *Phragmites* has a high chance of being confused with other land covers. Another major misclassification error is between the *Suaeda* and bare land due to the transition of sparse *Suaeda* vegetation and bare lands. Also the producer’s errors for the relatively small areas such as the mixed *Phragmites* and *Suaeda* class or the mixed *Suaeda* and *Spartina* class are relatively high because of their transition of different land covers. The MLC cannot effectively handle the spectral confusion among different land covers and transition of land covers, resulting in relatively poor performance. Also MLC relies highly on the quality of training samples. The vegetation samples for the same vegetation type could have significantly different spectral signatures due to the impacts of the complexity of biophysical conditions in the environment such as different moisture or salinity levels in this study area. Also the collection of a sufficient number of training samples for some vegetation classes, such as both mixed classes, is often difficult because access to the coastal wetlands is very limited.

MLC requires the data normal distribution and needs to extract mean and covariance parameters from the training samples, so it requires representative training samples for each land cover class. The hybrid method developed in this paper does not need the assumption of data normal distribution and does not extract the parameters from the training samples. This characteristic is especially important for wetland vegetation classification because of the difficulty in collecting training sample data. As table 1 shows, the result from the hybrid method provides much better classification accuracy than that from MLC. Because saltwater content is an important factor affecting vegetation type and distribution, spectral confusion between vegetation types often results in poor classification performance when traditional
Table 1. A comparison of accuracy assessment results of coastal wetland vegetation classification between maximum likelihood classifier and hybrid method.

<table>
<thead>
<tr>
<th>Method</th>
<th>Land cover type</th>
<th>Reference data for each class</th>
<th>Row total</th>
<th>Column total</th>
<th>User’s accuracy</th>
<th>Producer’s accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>(1) (2) (3) (4) (5) (6) (7)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maximum likelihood</td>
<td>Forest (1)</td>
<td>3 1</td>
<td>4</td>
<td>4</td>
<td>75.00</td>
<td>75.00</td>
</tr>
<tr>
<td>classifier</td>
<td><em>Phragmites</em> (2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mixed <em>Phragmites</em> and</td>
<td>2 1</td>
<td>2</td>
<td>5</td>
<td>95.65</td>
<td>80.00</td>
</tr>
<tr>
<td></td>
<td><em>Suaeda</em> (3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>Suaeda</em> (4)</td>
<td>1 9</td>
<td>2</td>
<td>12</td>
<td>75.00</td>
<td>50.00</td>
</tr>
<tr>
<td></td>
<td>Mixed <em>Suaeda</em> and</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>Spartina</em> (5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>Spartina</em> (6)</td>
<td>8 23</td>
<td>1</td>
<td>32</td>
<td>71.88</td>
<td>95.83</td>
</tr>
<tr>
<td></td>
<td>Other (7)</td>
<td>8 40</td>
<td>40</td>
<td>45</td>
<td>83.33</td>
<td>88.89</td>
</tr>
<tr>
<td></td>
<td>Total number of test samples: 151; overall accuracy: 82.12%; kappa: 0.76; kappa variance: 0.001635</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hybrid method*</td>
<td>Forest (1)</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>100.00</td>
<td>100.00</td>
</tr>
<tr>
<td></td>
<td><em>Phragmites</em> (2)</td>
<td>42 3 1</td>
<td>46</td>
<td>42</td>
<td>91.30</td>
<td>100.00</td>
</tr>
<tr>
<td></td>
<td>Mixed <em>Phragmites</em> and</td>
<td>3 2</td>
<td>5</td>
<td>6</td>
<td>60.00</td>
<td>50.00</td>
</tr>
<tr>
<td></td>
<td><em>Suaeda</em> (3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>Suaeda</em> (4)</td>
<td>11 1</td>
<td>12</td>
<td>15</td>
<td>91.67</td>
<td>73.33</td>
</tr>
<tr>
<td></td>
<td>Mixed <em>Suaeda</em> and</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>Spartina</em> (5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>Spartina</em> (6)</td>
<td>30 2</td>
<td>32</td>
<td>30</td>
<td>93.75</td>
<td>100.00</td>
</tr>
<tr>
<td></td>
<td>Other (7)</td>
<td>48 48</td>
<td>48</td>
<td>51</td>
<td>100.00</td>
<td>94.12</td>
</tr>
<tr>
<td></td>
<td>Total number of test samples: 151; overall accuracy: 93.38%; kappa: 0.91; kappa variance: 0.000696</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*The Z-statistic is 3.1259, significant at 99.5% confidential level.
supervised classification is used for wetland vegetation classification. The fraction images of salt water, fresh water, vegetation and soil are more useful in vegetation classification than spectral signatures. The stratification of major land cover types based on these fraction images is an effective approach in reducing the spectral confusion of vegetation classes. Moreover, the analyst can make full use of his/her knowledge about the characteristics of vegetation distribution in merging the clusters into a suitable vegetation class. Also, the expert knowledge can be used for further modifying the classified image. As figure 5 shows, forest areas were very limited and were mainly distributed along the roads and upland in the north side of this study area. *Spartina* was closely located along the shoals with conditions of high salinity.

Figure 5. Coastal wetland vegetation distribution developed with the hybrid approach from the Landsat TM image.
Phragmites occupied the area away from the coastal line, and Suaeda was distributed between Spartina and Phragmites. This research indicated that the hybrid approach combining the advantages of LSMA, thresholding techniques and expert knowledge can be successfully used for wetland vegetation classification.

During classification accuracy assessment, selection of a suitable sampling method for collection of reference data is important in reducing the bias in the assessment. However, in many situations, randomly selecting test sample plots is difficult due to being time-consuming and labour intensive, and also due to the difficulty in accessing sites such as this study area. The overall accuracies provided in table 1 may be overestimated because of the problem in collecting reference data. In this study, the training sample plots and test plots were selected from the typical locations where human beings can gain access. For sites away from roads, only limited training or test sample plots were collected, based on the vegetation distribution rules and the knowledge of local experts. Also, defining the transition classes such as mixed Phragmites and Suaeda or mixed Suaeda and Spartina was difficult, resulting in the problematic collection of test sample plots for these classes.

This study also indicated that there were still some misclassifications between vegetation classes and other non-vegetation classes because of the impacts of vegetation residuals and moisture. This research only used one date Landsat TM image for wetland vegetation classification. As previous research showed, the use of multi-temporal images (Townsend and Walsh 2001, Hess et al. 2003, Martinez and Toan 2007, Ghioca-Robrecht et al. 2008, Gilmore et al. 2008), the use of hyperspectral data (Schmidt and Skidmore 2003, Artigas and Yang 2006), the incorporation of ancillary data (Baker et al. 2006, Wright and Gallant 2007), as well as the use of advanced nonparametric classifiers (Baker et al. 2006, Wright and Gallant 2007) could further improve wetland vegetation classification performance.

5. Conclusions
Coastal wetland vegetation classification with Landsat TM images is a challenge because different water conditions in the wetland ecosystem generate complex spectral signatures of vegetation types. Traditional classification approaches cannot effectively deal with the spectral confusions of different vegetation types. A hybrid approach developed in this study, which combined LSMA, thresholding, unsupervised classification and expert knowledge, has proven effective in improving vegetation classification in a complex biophysical environment. The hybrid method improved wetland classification performance by 11.26% of overall classification accuracy and by 0.15 of the kappa coefficient and appeared significantly better performance than MLC. The hybrid approach can make full use of analysts’ experience in the study area and the specific features of vegetation distribution in coastal wetland ecosystems.

Acknowledgement
The authors would like to thank Shi-Chong, Hui Wang and Ke Dong for valuable assistance during the field work, and thank Scott Hetrick from Indiana University for correcting the English errors.

References


MARKHAM, B.L. and BARKER, J.L., 1986, Landsat MSS and TM post-calibration dynamic ranges, exoatmospheric reflectance and at-satellite temperatures. EOSAT Landsat Technical Notes, 1, pp. 3–8.


No. 10-03

No. 10-04

No. 10-05

No. 10-06

No. 10-09

No. 10-10

No. 10-11

No. 10-12

No. 10-13

No. 10-14

No. 10-15

No. 10-16