Classification of land-cover types in muddy tidal flat wetlands using remote sensing data

Cong Wang
Hong-Yu Liu
Ying Zhang
Yu-feng Li
Classification of land-cover types in muddy tidal flat wetlands using remote sensing data

Cong Wang, Hong-Yu Liu, Ying Zhang, and Yu-feng Li
Nanjing Normal University, Geographical Science Institute, 210023 China
liuhongyu@njnu.edu.cn

Abstract. Remote sensing classification of tidal flat wetlands is important for obtaining high-precision information on wetland features. In this study, Thematic Mapper (TM) images of the Yancheng National Reserves, Jiangsu Province, China, for the years of 1996, 2002, 2006, and 2010 were considered. First, the optimum combination of bands was chosen. Second, vegetation and nonvegetation regions of interest were established to investigate the spectral reflectance characteristics of the different ground objects. Then we used the knowledge-based decision tree method on different features, such as the normalized difference vegetation index and the spectral reflectance. In particular, the ancillary information is helpful to distinguish the vegetation classes. The results demonstrate that the classification system has advantages in identifying the types of vegetation in ecotones, and it is 4 percentage points higher than the maximum likelihood method in classification accuracy. This study is useful to discriminate vegetation, and it provides an important reference for the effective extraction of tidal flat land-cover information from TM images. © The Authors. Published by SPIE under a Creative Commons Attribution 3.0 Unported License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI. [DOI: 10.1117/1.JRS.7.073457]

Keywords: muddy tidal flat wetlands; spectral reflectance; normalized difference vegetation index; decision tree classification; ecotones.

Paper 13133 received May 29, 2013; revised manuscript received Oct. 21, 2013; accepted for publication Oct. 29, 2013; published online Jan. 2, 2014.

1 Introduction
Several researchers have mapped vegetation types using Landsat data, through methods such as supervised classification and the use of vegetation indices and decision trees.1–5 Pal and Mather have already described the effectiveness of decision tree methods for land-cover classification using Landsat data.6 There are also comparisons of Landsat Thematic Mapper (TM), multitemporal remote sensing image data sets, and high-resolution photography to identify changes in complex coastal wetlands.7–9 Coastal wetlands experience intense and sustained environmental pressures, which show multiscale dynamics and are likely to have considerable uncertainties in response to future environmental change.10–12 As an important part of the study of wetlands, methods have been applied to extract landscape information effectively from TM images. Research into methods to identify land-cover types has gained momentum but also has become increasingly difficult.13 TM images act as a source of data on wetlands, and such information can be extracted for a combination of multisubareas and multilayers on the basis of systemic analysis of spectrum characteristics of all types of wetlands.14 Decision tree methods for land-type classification are more accurate than traditional computer automatic classification methods.15–20 The SanJiang plain (China) freshwater wetland decision tree model was established by combining the band threshold method with visual interpretation in Chinese coastal zones.21 The aim of decision tree classification is to confirm a threshold value. Artificially seeking the threshold based on the distribution of the value in the feature space is convenient and efficient.22,23 Previously, several studies have employed Landsat TM images for mapping wetland vegetation types.24–26 They attempted to combine TM imagery and ancillary environmental data using classification trees.27,28 However, complete utilization of image and environmental characteristics to quickly and accurately extract tidal flat wetland vegetation information is still an urgent and important challenge. Our description is not only about the methods involved but also, more importantly,
about how we can map vegetation land cover using remote sensing approaches in coastal wetlands.

2 Study Area

The study area of Yancheng National Reserve Area is in Jiangsu Province, China. It is a typical muddy tidal flat coastal wetland due to the wasted Yellow River mouth and the Yangtze River, which outputs a lot of sediment in northern Jiangsu Beach and convergence deposition. Sediment particles are smaller because of the gentle tidal flow, the formation of the largest and the most original muddy tidal flat in China (Fig. 1). The study area is located in the latitude range of 33° 25' 0" to 33° 39' 4" N, and the longitude range of 120° 27' 40" to 120° 40' 40" E. Its total area is 191 ha, and it is located between Xinyang port and DouLong port, to the east of HuangHai, with the west bounded by a sea road. The geographical position of this wetland is very important. Rainfall is concentrated in the summer season, with less precipitation in winter, a submerged intertidal zone time interval of 7 to 12 h, and a high-water level of 1.27 to 3.60 m. The tidal flat wetland is considerably influenced by regular semidiurnal tides. The wetland is rich in vegetation, with a complete tidal flat vegetation succession; the vegetation types from land to sea can be mainly classified into the _Phragmites communis_ belt, the _Suaeda salsa_ belt, the _Spartina alterniflora Loisel_ belt, and the mudflats belt. Since the tidal flat wetland constantly silts up, the landscape pattern is continuously evolving.

3 Data Processing and Research Methods

3.1 Data Selection

The data source of the TM data is taken (days, months, years) as shown in Table 1. Images of the vegetation growing season and low water are selected since they are useful for improving the recognition accuracy.

![Fig. 1 The study area.](image-url)

<table>
<thead>
<tr>
<th>Type</th>
<th>Type</th>
<th>Type</th>
<th>Type</th>
<th>Type</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Path</td>
<td>119037</td>
<td>119037</td>
<td>119037</td>
<td>119037</td>
<td>119037</td>
</tr>
</tbody>
</table>

Table 1 Data source for the Thematic Mapper (TM).
3.2 Data Processing

Remote sensing image preprocessing includes atmospheric correction, geometric precision correction, and feature extraction. The FLAASH atmospheric correction module, which is based on MODTRAN4 and retrieves spectral surface reflectance from multispectral remote sensing images, is the preferred atmospheric correction model. This method is based on observations by Kaufman et al.\textsuperscript{32} of a nearly fixed ratio between the reflectance for pixels at 660 and 2100 nm.\textsuperscript{33–35} The FLAASH module in ENVI probably provides the most accurate means of compensating for atmospheric effects. For geometric correction, 12 to 20 ground control points are selected, usually including the locating object, which is explicitly identified.\textsuperscript{36} The quadratic polynomial resampling technology correction refers to the total error control in a given pixel. A training area and an appropriate band combination was selected by considering the band wavelength of TM imagery, the region of interest (ROI) for the spectral reflectance of different types, the normalized difference vegetation index (NDVI), and the ecological characteristics (ecological niche and the color of the vegetation leaves in September). We made an exploration of TM band reflectance characteristics of each band for each land-cover type through the ROI training area. The ROI includes 1316\textit{ Spartina alterniflora Loisel}, 633 mudflats, 1253\textit{ Suaeda salsa}, 1002\textit{ Phragmites cammunis}, 525 water bodies, and 4803 road samples (Fig. 2).

The NDVI and spectral reflectance values of different features were collected. The NDVI has been widely used for vegetation application, and the calculation of NDVI is as follows:

\[
\text{NDVI} = \frac{(\text{NIR} - \text{Red})}{(\text{NIR} + \text{Red})}. \text{(1)}
\]

Red and NIR stand for the spectral reflectance in the red and near-infrared bands, respectively. Normalized NDVI values lie between (−1, 1). Under normal circumstances, values for water are less than 0 and values for vegetation are larger during the growing season. A large amount of the biomass in this study has a large NDVI value.\textsuperscript{37,38}

3.3 Decision Tree Classification Method

Decision tree methods for land cover classification are more convenient and efficient to solve the basic classification problem. The decision tree is based on a multistage or hierarchical decision scheme or a tree-like structure. The tree is composed of a root node containing all data, a set of internal nodes (splits), and a set of terminal nodes (leaves). Each node of the decision tree’s structure makes a binary decision that separates either one class or some of the classes from the remaining classes. The processing is generally carried out by moving down the tree until the leaf node is reached. Thus, the basic concept of a decision tree is to split a complex decision into several simpler decisions, which may lead to a solution that is easier to interpret. In a decision tree approach, features of data (i.e., bands) are predictor variables, whereas the class to be mapped is referred to as the target variable.\textsuperscript{39} When the target variable is discrete (e.g., class attribute in a land cover classification), it is known as a decision tree classification. A decision tree classification system should confirm a threshold value based on the distribution of features in the feature space. The classification processing is conducted using ENVI4.7 software. The regional overlay is visualized using ArcGIS9.3 software.

3.4 Selection of the Remote Sensing Band Combination

The band selection is done through data exploration. It helps to display and analyze remote sensing data. Remote sensing data at different wavelengths can be used to distinguish different vegetation types.

The results of feature selection indicate the range (maximum minus minimum) of bands 1 and 2 are small; the spectral reflectance area is narrow, which is unfavorable for feature recognition. Moreover, the spectral reflectance of \textit{P. cammunis} and \textit{water bodies} have almost identical mean values and hence cannot be distinguished from one another. However, the standard deviation value of \textit{P. cammunis} is higher than that of \textit{water bodies}, revealing that feature separability is stronger for \textit{P. cammunis} than \textit{water bodies}. A comparison of the reflectance
deviation values of all seven bands shows that bands 1 and 2 have the lowest values, so these two bands should not be chosen. Band 3 (0.63 to 0.69 μm) reflects not only plant chlorophyll information but also beta-carotene and pigment information in autumn. Band 4 (0.76 to 0.90 μm) is in a high-reflection region for plants and is hence used for vegetation classification, biomass investigation, and determination of crop growth; accordingly, it is a general band in plant classification. Band 5 (1.55 to 1.75 μm) includes a large amount of information and thus has a high utilization rate. Bands 5 and 7, the spectral reflectance minima of water body, mudflat, and road, appear to be negative. These three bands (bands 5, 4, and 3) were used to separate vegetation from nonvegetation land-cover types.

3.5 Tidal Flat Wetland Hierarchical Classification

This research considers NDVI, the band reflectance characteristics, and ecological features (especially ecological niche and the color of the vegetation leaves in September) of six typical tidal flat items, namely, water body, roads, mudflat, S. alterniflora, S. salsa, and P. communis. The first three types are nonvegetation. The last three types belong to vegetation. In the tidal flat wetland in the study region, the environmental characteristics are that the plants have

Fig. 2 The statistics of Thematic Mapper (TM) band reflectance characteristics of each band for each type.
Table 2  Statistics of normalized difference vegetation index values of zonal vegetation distribution.

<table>
<thead>
<tr>
<th>Vegetation types</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>P. cammunis</td>
<td>0.172336</td>
<td>0.704160</td>
<td>0.509434</td>
</tr>
<tr>
<td>S. salsa</td>
<td>-0.180490</td>
<td>0.459669</td>
<td>0.084258</td>
</tr>
<tr>
<td>S. alterniflora</td>
<td>0.103365</td>
<td>0.753506</td>
<td>0.550454</td>
</tr>
</tbody>
</table>

predominant zonal distribution characteristics. Vegetation and nonvegetation land-cover information are precisely extracted by utilizing NDVI index segmentation. The process of selecting a value was repeated, and the value of NDVI was finally determined to be zero after comparing several results. This means that pixels with NDVI > 0 represent vegetation regions and those with NDVI < 0 represent nonvegetation regions.

3.5.1 Classification of wetland vegetation types

The tidal flat vegetation can be mainly divided into three types: S. alterniflora, S. salsa, and P. cammunis. Their NDVI values are shown in Table 2.

The NDVI mean value (0.084258) for S. salsa was significantly lower than the values (0.509434, 0.550454) for the other two types of vegetation in the TM images. This is mainly because S. salsa’s biomass was relatively small and its color was red, while the biomasses of S. alterniflora and P. cammunis were higher and were shown in green on the image of September 21, 2010a. S. salsa can therefore be extracted through its NDVI value.

Zonal distribution vegetation classification. The distributions of other types of vegetation in the area were not determined only on the basis of analysis of the NDVI values. Since the NDVI values for the other two plants are almost identical, it is difficult to distinguish between them only through the NDVI values; hence the plants must be identified by their different environmental characteristics and reflectance properties.

A spectral reflectance curve represents the spectral selective absorption of the incident light, the light scattering, and the comprehensive characteristics of specular reflection on the surface of the object. Figure 3 shows the spectral reflectance response values of different wave bands for all three plants, obtained from the reflectance image acquisition of pixels, comprising purely pixel of each plant.

The closer the spectral curves, the more difficult it is to distinguish the features. From the typical feature spectrum curve in Fig. 2, the three vegetation types are distinguishable. However, because of the phenomenon of “foreign body with same spectrum” and the ecotone that exists in the case of S. alterniflora, S. salsa, and P. cammunis, it is difficult to further distinguish between

Fig. 3 Spectral reflectance characteristics of the vegetation.
them. Meanwhile, it is important to find a way to more accurately interpret vegetation types in the ecotone.

**Ecotones vegetation classification.** The ecotones bandwidth for S. alterniflora and S. salsa is ∼2.503 km, while that for P. cammunis and S. salsa is ∼1.655 km. Pure pixels, i.e., those with only one vegetation type, were used to establish ROI. The spectral response curve was collected repeatedly from these pure pixels through the whole reflectivity spectrum to finally determine the value from the spectral response curve of vegetation (Figs. 4 and 5).

Since the separability of each plant’s spectral reflectance value in band 5 is strong, band 5 spectral reflectivity value was chosen to effectively separate S. alterniflora and P. cammunis. After repeated tests, it was found that the band 5 reflectance value should be 0.12 to effectively distinguish between these plant types (Fig. 6).

In the classification results, a small proportion of P. cammunis was classified as S. alterniflora. From the ecological feature point of view, this is located in the west side of S. salsa. The field test was used to verify that the S. alterniflora was P. cammunis in fact. Accordingly, the classification was changed to P. cammunis. Through the above methods, it is possible to identity the distribution of zonal vegetation types in ecotones and scattered patches of vegetation types.

### 3.5.2 Classification of nonvegetation land-cover types

The main nonvegetation land-cover types are water body, mudflat, and road. Mudflat is mainly found to the east of the area dominated by S. alterniflora. The waterbody and mudflat were
Wang et al.: Classification of land-cover types in muddy tidal flat wetlands…

![Fig. 6](image1.jpg)

**Fig. 6** The decision tree of vegetation classification.

separated using the same method as that used for vegetation division, namely, the pixel spectral acquisition method utilizing repeated NDVI and B3 band reflectance values (Fig. 7).

In the study area, identification of the road land cover is much more complex. One type is the sea road, which has trees on both sides and a ditch along the center, while other roads in aquaculture and waterfowl lake areas are covered by herbaceous vegetation. Furthermore, foreign bodies may be found across the middle of dirt roads. Through the integration of remote sensing image texture information with the vegetation classification information, road information can finally be completely extracted in the postprocessing stage.

4 Results

4.1 Classification Results

The other three phases of image classification in the study area are realized through hierarchical classification, with the results shown in Fig. 8.

Wetland land cover types are complex, especially for wetland vegetation. Because of “same object with different spectrum” and “different objects with same image” phenomenons, misclassification and leakage of the pixels occur more frequently. These factors lead to the low classification accuracy. There are different amounts of the pixels of vegetation types that have been misclassified only based on spectral characteristics. Therefore, ecological niche and vegetation growth characteristics in different belts as the ancillary information were used to analyze the image of TM. This can identify a scattered distribution of vegetation types and vegetation patches. It can distinguish between patches as small as 0.09 ha, of which there are 345, accounting for 27% of the total number of patches.

4.2 Classification Accuracy Test

A confusion matrix was established to analyze and evaluate the hierarchical classification accuracy. The ROI training method and 55 field data were used to perform a classification accuracy test. For this statistical classification and evaluation, there are six classes of land-cover type: water body, roads, S. alterniflora, S. salsa, P. cammunis, and mudflat. The sample space has a total of 3126 pixels in the 2010a image. After inspection, the overall accuracy was (2997/3126) 95.87%. The result from the same period as the image and the maximum likelihood method classification results (overall accuracy is 91.43%) are compared in Tables 3 and 4.

From a comparison of Tables 3 and 4, it can be seen that among the three types of vegetation, the user accuracy is highest for S. salsa, at 93.59%. The pixels correspond to the actual vegetation in Table 3 and to only 83.92% in Table 4. User precision indicates a high reliability of the classification diagram. Production shows that the actual accuracy for S. salsa is 94.62%, which is
properly assigned to this kind of vegetation in Table 3, while 4.59% were classified as *P. cammunis* and 1.81% as *S. alterniflora*. Meanwhile, 96.79% is properly assigned to *S. salsa* with 7.38% classified as *P. cammunis* in Table 4 and 6.41% as *S. alterniflora*. In the classification statistical classification accuracy, *mudflat* and *water body* have the highest classification accuracy. The number of false points for *road* classification was larger, mainly because dirt roads are covered with vegetation and their width is narrower than other land-cover types, and some of them have ditches.

![Classification map of landscape from TM images at different times.](https://www.spiedigitallibrary.org/journals/Journal-of-Applied-Remote-Sensing/7/73457-8)

**Table 3** Accuracy assessment of hierarchical classification of the 2010 TM image.

<table>
<thead>
<tr>
<th>Type</th>
<th><em>S. salsa</em></th>
<th><em>Mudflat</em></th>
<th><em>P. cammunis</em></th>
<th><em>Roads</em></th>
<th><em>S. alterniflora</em></th>
<th><em>Water body</em></th>
</tr>
</thead>
<tbody>
<tr>
<td><em>S. salsa</em></td>
<td>93.59</td>
<td>0</td>
<td>0</td>
<td>20</td>
<td>1.65</td>
<td>0</td>
</tr>
<tr>
<td><em>Mudflat</em></td>
<td>0</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><em>P. cammunis</em></td>
<td>4.59</td>
<td>0</td>
<td>99.3</td>
<td>5.93</td>
<td>0.78</td>
<td>0</td>
</tr>
<tr>
<td><em>Roads</em></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>65.19</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><em>S. alterniflora</em></td>
<td>1.81</td>
<td>0</td>
<td>0.62</td>
<td>0</td>
<td>97.57</td>
<td>0</td>
</tr>
<tr>
<td><em>Water body</em></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>7.41</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>Total</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>User p.</td>
<td>93.59</td>
<td>100</td>
<td>99.3</td>
<td>65.19</td>
<td>97.57</td>
<td>100</td>
</tr>
<tr>
<td>Production p.</td>
<td>94.62</td>
<td>100</td>
<td>91.26</td>
<td>65.19</td>
<td>97.95</td>
<td>97.34</td>
</tr>
</tbody>
</table>
5 Conclusions

1. The spectrum of the tidal flat wetland features in remote sensing images is very complex, especially for vegetation ecotones. Scattered distribution of vegetation patches makes it even more difficult to distinguish features. So the classification’s importance is in vegetation ecotones. It is the key to improve the accuracy.

2. The best band combination selection and feature extraction method selection is very important for accuracy. We find that the characteristics of TM bands 5, 4, and 3, NDVI, and the spectral reflectance of different bands further enable the construction of the feature space for classification.

3. The decision tree method divides the complicated problem into relatively simple questions. The advantage of this method is that it produces branch nodes that can use different attribute values in the process of establishing the decision tree. This can distinguish patches as small as 0.09 ha. There are 345 patches that are this small, accounting for 27% of the total number of patches.

4. Through the accuracy test, it was found that the user precision of all types of feature and the accuracy of the producer was generally >93%, and that the overall accuracy reached 95.87%. S. alterniflora, S. salsa, P. cammunis, mudflat, and water body classification accuracy is higher, with user precision and accuracy of producers always >95%. In comparison to the results of supervised classification, the overall accuracy was markedly improved.

Each classification method has the most comfortable application range and its own limitations. No one is the most common better method. It must be flexible in application. It is necessary to integrate application of many kinds of classification methods, and combine with other image processing technology to achieve maximum precision classification. In the future, we will focus on using more new theories and new technologies to improve the effect of the remote sensing image classification. We hope this article can provide certain reference in the tidal flat image classification.

Acknowledgments

This research was supported by the National Natural Science Foundation of China (41071119) Jiangsu Province University Natural Science(10kja170029) and Jiangsu Key Laboratory of Environmental Change & Ecological Construction.

References


**Wang Cong** is a doctor of science candidate in the Institute of Geographical Science at Nanjing Normal University of Jiangsu Province, China, at present. She received her master of science in physical geography from Hebei Normal University. Her current research interests include landscape ecology of wetlands, remote sensing classification and geographical information system applied study. She is a member of Jiangsu Key Laboratory of Environmental Change & Ecological Construction.

Biographies and photographs of the other authors are not available.