



U.S. Geological Survey Open-File Report OF 03-272

## Using the Landsat 7 Enhanced Thematic Mapper Tasseled Cap Transformation to Extract Shoreline *(March 2003)*

By Joseph W. Scott<sup>1</sup>, Laurence R. Moore<sup>1</sup>, William M. Harris<sup>1</sup>, and Matthew D. Reed<sup>2</sup>

### ABSTRACT

A semiautomated method for objectively interpreting and extracting the land-water interface has been devised and used successfully to generate multiple shoreline data for the test States of Louisiana and Delaware. The method is based on the application of tasseled cap transformation coefficients derived by the EROS Data Center for Landsat 7 Enhanced Thematic Mapper Data, and is used in conjunction with ERDAS Imagine software.

Shoreline data obtained using this method are cost effective compared with conventional mapping methods for State, regional, and national coastline applications. Attempts to attribute vector shoreline data with orthometric elevation values derived from tide observation stations, however, proved unsuccessful.

### INTRODUCTION

The 95,000 miles of U.S. shoreline data, because of its location and water-level measurement, form an important time-critical data layer that is widely used by coastal managers, mariners, the mapping community, and the general public for a variety of legal, resource, commercial, and planning applications.

A problem exists, however, in the mapping community because different government agencies have compiled and published shoreline delineations that are based on different shoreline definitions. This has created confusion and uncertainty for those who use shoreline information daily for decisionmaking, resource planning, and emergency preparedness. For example, the National Oceanic and Atmospheric Administration (NOAA) maps the national legal shoreline on its nautical charts and surveys, as required by the agency's mandate (Parker, 2001). The national legal shoreline is defined by the portrayal of the mean high water (MHW) level, which is based on a compilation of the mean lower low water (MLLW) level using tide-coordinated aerial photographs. In contrast, the U.S. Geological Survey (USGS) compiles shoreline data for the 1:24,000-scale topographic base map series from digital orthophoto quadrangles created from photographs that are not tide coordinated, thereby making the shoreline a snapshot in time. In yet another example, the National Imagery and Mapping Agency is examining a method to revise its World Vector Shoreline product from Landsat imagery. The proposed revision method defines the shoreline as the line separating "always dry" land from land that is inundated daily by tide (Brand and Hammack, 2002).

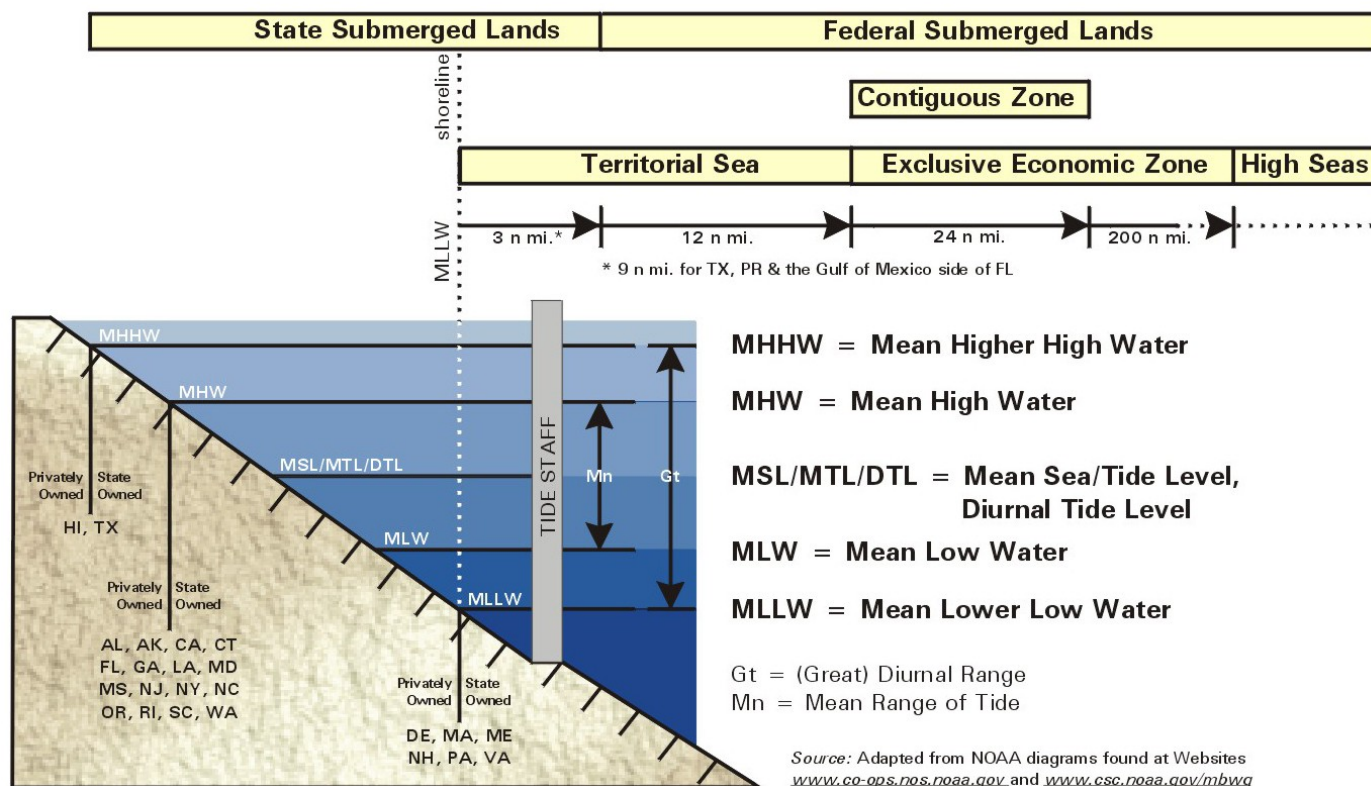
---

Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Although NOAA maps the national legal shoreline, not all States have adopted this MHW shoreline as their legal shoreline to separate privately owned land from State-owned land, as shown in figure 1. For example, Hawaii and Texas use the mean higher high water (MHHW) level to separate privately owned land from State-owned land, whereas Delaware, Maine, Massachusetts, New Hampshire, Pennsylvania, and Virginia use the MLLW level (Parker, 2001).



## TIDAL DATUMS AND THE IMPORTANCE ASSOCIATED WITH SHORELINE LOCATION



**Figure 1.** Diagram describing tidal datums and the relation of shoreline location with legal property ownership.

Regardless of how the shoreline data are defined and represented, a better approach is needed to link shoreline delineation derived from the most recent satellite imagery data with available tidal datum for graphic revision. Such an approach could provide shoreline data consistent with most coastal applications and meet many of the goals associated with *The National Map*, such as providing accurate, consistent, and continuously maintained geographic information. Such an approach also would provide current shoreline data in a more dynamic manner, perhaps several times per year.

This investigation researched the feasibility and cost-effectiveness of using Landsat 7 ETM scenes acquired on different dates to extract shoreline data with sufficient detail to contribute to a medium-scale graphic revision. From this overall purpose, the following four hypotheses were identified and tested:

- Shoreline data representing the land-water interface can be automatically interpreted from Landsat 7 ETM imagery;
- Shoreline data can be defined by increasing the temporal resolution of image data sources, even if

the spatial resolution is decreased;

- It is possible to accurately relate this interpreted shoreline data to elevation values obtained from coastal tide observation stations; and
- The interpretation can be done quickly and cost effectively, so that large lengths of shoreline data can be extracted at least several times per year.

## DATA AND METHODOLOGY

### Data sources and selection criteria

This investigation used only Landsat 7 ETM scenes as the source and Louisiana and Delaware as study areas.

#### Landsat 7 ETM Imagery

Landsat ETM imagery was selected because the 30- by 30-meter pixel size appeared to satisfy medium-scale graphic mapping requirements, the revisit cycle of 16 days yields 22 potentially useful scenes per year for a given area, and the acquisition cost is low compared with other platforms given the same coverage area. At the beginning of this investigation, we decided to use Landsat 7 ETM scenes rather than Landsat 5 TM because of the eventual demise of the aging Landsat 5 satellite.

To create a set of statewide coastline coverage of Louisiana, we needed seven Landsat 7 ETM scenes located along four paths and two rows; statewide coastline coverage of Delaware required only two scenes along one path. A review of the available scenes since the launch of Landsat 7 ETM in 1999 revealed sets of scenes from which shorelines could be extracted for Louisiana for 1999, 2000, and 2001, and for Delaware for 1999, 2001, and winter and summer, 2002. The scenes selected for the Louisiana study area are listed in table 1.

**Table 1. List of Landsat 7 ETM scenes used for Louisiana shorelines**

Shoreline	Path/Row	Image time
1999	P24, R39	10:37 AM CST November 16, 1999
	P23, R39 & R40	10:31 AM CST October 24, 1999
	P22, R39 & R40	10:25 AM CST November 18, 1999
	P32, R39 & R40	10:18 AM CST November 27, 1999
2000	P24, R39	10:35 AM CST December 20, 2000
	P23, R39 & R40	10:28 AM CST December 29, 2000
	P22, R39 & R40	10:22 AM CST December 22, 2000
	P21, R39 & R40	10:16 AM CST December 31, 2000
2001	P24, R39	10:32 AM CST November 5, 2001
	P23, R39 & R40	10:27 AM CST November 30, 2001
	P22, R39 & R40	10:20 AM CST November 7, 2001
	P21, R39 & R40	10:14 AM CST November 16, 2001

#### Study Sites

Louisiana was chosen as the primary study site because of its extensive and complex shoreline composed of numerous river deltas, large expanses of coastal marshes and submerged wetlands, a retreating shoreline because of land subsidence, the impact of the shoreline on the State's economy, and a comparatively small tide range present at this location within the Gulf of Mexico. Delaware was chosen as the second study site because of its relatively stable coastal configuration, numerous manmade shoreline features, and a larger tide range than in Louisiana.

### Tide Information

Tide information for both study States was obtained from NOAA's Center for Operational Oceanographic Products and Services (CO-OPS) Web site and is based on operating tide observation stations located along each State's coastline. For Louisiana, where only one tide station (Grand Isle) was operating, tide information had to be supplemented from two other tide stations located in neighboring States, one to the east (Waveland, Miss.), and one to the west (Sabine Pass, Texas). Tide information associated with those observation stations, as well as tide stage readings recorded at the time of Landsat 7 ETM imaging, is shown in tables 2 and 3.

**Table 2. Tide information of tide observation stations used for Louisiana shorelines**

Datum	Sabine Pass, TX	Grand Isle, LA	Waveland, MS
Mean Higher High Water (MHHW)	1.62 ft	1.11 ft	1.66 ft
Mean High Water (MHW)	1.50 ft	1.08 ft	1.59 ft
Mean Sea Level (MSL)	0.98 ft	0.56 ft	0.84 ft
Mean Lower Water (MLW)	0.47 ft	0.02 ft	0.12 ft
Mean Lower Low Water (MLLW)	0.00 ft	0.00 ft	0.00 ft
Diurnal Tide Range	1.62 ft	1.11 ft	1.66 ft

**Table 3. Tide stage readings associated with Louisiana Landsat 7 ETM scenes\***

Path	1999 shoreline	2000 shoreline	2001 shoreline
P24	0.90 ft @ Sabine Pass	0.09 ft @ Sabine Pass	0.21 ft @ Sabine Pass
P23	0.26 ft @ Sabine Pass	-1.40 ft @ Sabine Pass	0.15 ft @ Sabine Pass
	0.40 ft @ Grand Isle	-0.66 ft @ Grand Isle	0.44 ft @ Grand Isle
P22	0.62 ft @ Grand Isle	-0.58 ft @ Grand Isle	0.30 ft @ Grand Isle
P21	0.25 ft @ Waveland	-0.45 ft @ Waveland	0.07 ft @ Waveland
* values shown calculated with respect to MLLW			

### Methodology

The methodology used in this investigation consisted of the following processing steps:

- (1) Select Landsat 7 ETM scene(s). For this investigation, scene selection was based on the following criteria: multiple sets of nearly cloud-free statewide coverage were required, with each set temporally separated by at least 6 months; all scenes within each set must have been acquired within the narrowest possible temporal range of each other; and all necessary scenes located successively along the same Landsat 7 travel path must have been acquired on the same pass of Landsat 7.
- (2) Using ERDAS Imagine, register and mosaic scenes (if necessary). All necessary scenes located successively along the same Landsat 7 path were mosaicked together to form a continuous scene.
- (3a) Using ERDAS Imagine, transform scenes using the Landsat 7 ETM tasseled cap transformation to yield a three-layer, 24-bit tasseled cap raster file. This method for extracting a shoreline is based on the application of tasseled cap transformation coefficients derived for Landsat 7 ETM data by the EROS Data Center (Huang and others, 2002); these coefficients are used in conjunction with various ERDAS Imagine image manipulation modules and an ARCINFO vectorization AML. The tasseled cap transformation was chosen over other methods primarily because of the objective and consistent manner in which it classifies pixels and because its use allowed the creation of other useful raster byproduct files.

In operation, the tasseled cap transformation recombines spectral information of the six ETM bands into three principal view components through the use of coefficients derived by sampling known land cover spectral characteristics. Of the three principal view components created, i.e., brightness, greenness, and wetness, the wetness component is exploited to differentiate land from water. The objective and repeatable manner in which the transformation segregates (classifies) Landsat image pixels into the three components effectively eliminates tedious manual interpretation of the land-water interface. A full explanation of the basic tasseled cap concept, its development, characteristics, and applications are provided in Kauth and Thomas (1976), Crist and Cicone (1984), Crist (1985), and Crist and Kauth (1986). The tasseled cap coefficients used in this investigation are given in table 4.

**Table 4. Tasseled cap coefficients for Landsat 7 ETM+ at-satellite reflectance**

Index	Band 1	Band 2	Band 3	Band 4	Band 5	Band 7
Brightness	0.3561	0.3972	0.3904	0.6966	0.2286	0.1596
Greenness	-0.3344	-0.3544	-0.4556	0.6966	-0.0242	-0.2630
Wetness	0.2626	0.2141	0.0926	0.0656	-0.7629	-0.5388
Fourth	0.0805	-0.0498	0.1950	-0.1327	0.5752	-0.7775
Fifth	-0.7252	-0.0202	0.6683	0.0631	-0.1494	-0.0274
Sixth	0.4000	-0.8172	0.3832	0.0602	-0.1095	0.0985

(3b) Using ERDAS Imagine, reduce the three-layer, 24-bit tasseled cap file into a single layer, 8-bit pseudocolor image. This process is performed because the output 24-bit three-layer tasseled cap file is not easily recordable (classifiable) and is a very large file. Reducing this file to a single layer, 8-bit pseudocolor image file facilitates easy displaying, file sharing, and recoding of data into desired classes of information.

(3c) Using ERDAS Imagine, manipulate the single layer, 8-bit file to create a 2-bit raster file whose pixels are recoded into the categories land, water, and background. The tasseled cap file creates three principal view components: brightness, greenness, and wetness. The land-water interface is generated from a 2-bit raster file that results from recoding the 256 classes contained in the single-layer, 8-bit file down to three classes consisting of simply land, water, and background.

(4) Vectorize the 2-bit land-water raster file. An ARCINFO AML was created to generate a shoreline that used the largest scene edge-body of water and all water contiguous to it to create a vector shoreline from the 2-bit land-water raster file.

(5) Perform postvectorization editing to remove or correct artifacts. In some cases, minor interactive editing of the 2-bit raster file was done to overcome certain postvectorization artifacts that were observed after the vector shoreline and the original scene(s) were compared, such as truncation of the vector shoreline by the presence of wide bridges, large objects on the water (ships, ferries), and cloud “islands.”

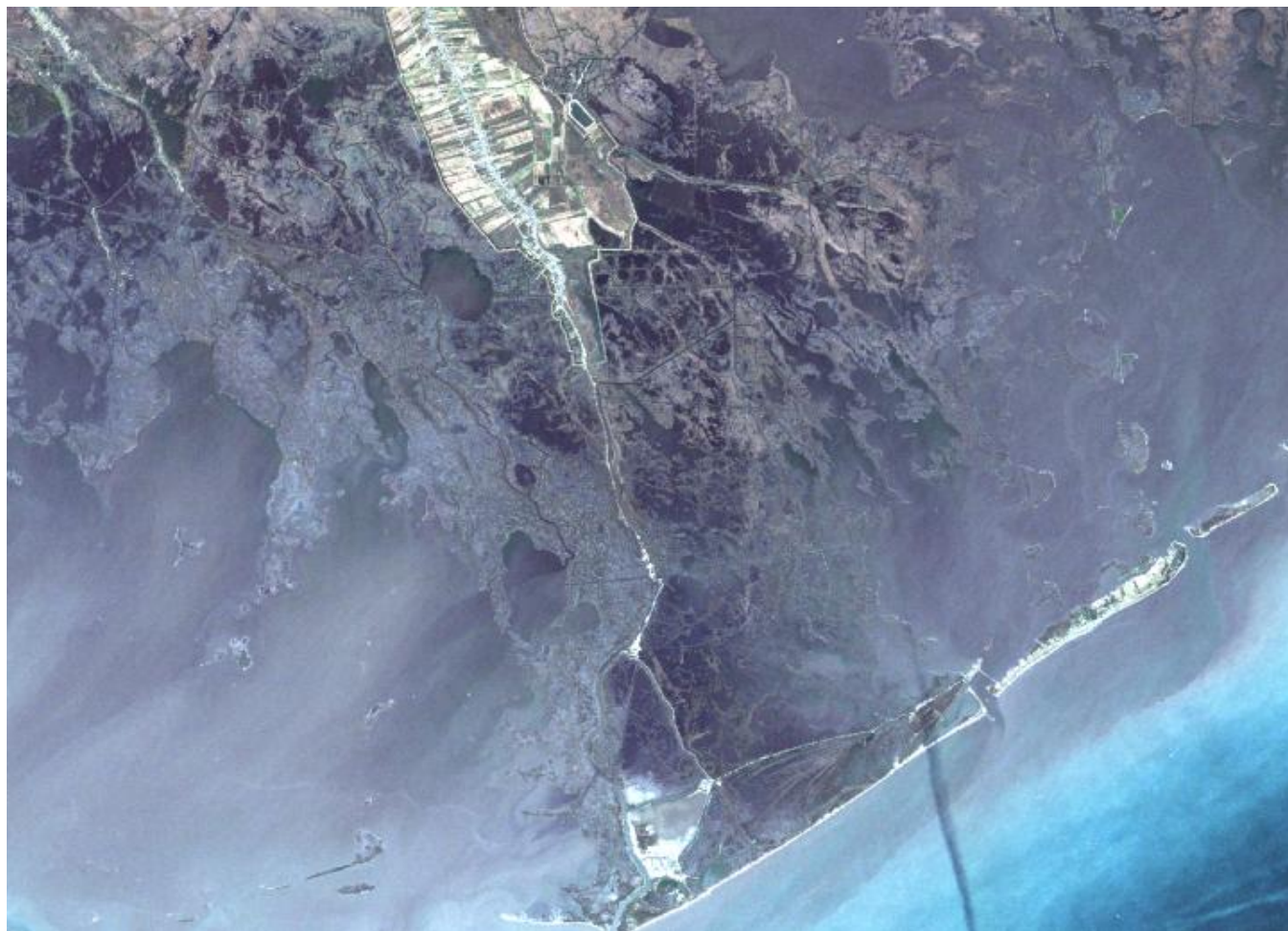
## RESULTS AND CONCLUSIONS

Hypothesis 1: Shoreline data representing the land-water interface can be automatically interpreted from Landsat 7 ETM imagery.

Mostly confirmed. A semiautomated method for objectively interpreting and extracting the land-water interface has been identified (as described above) and used successfully to generate three shorelines for the State of Louisiana and four shorelines for the State of Delaware; only the Louisiana results are presented in this report.

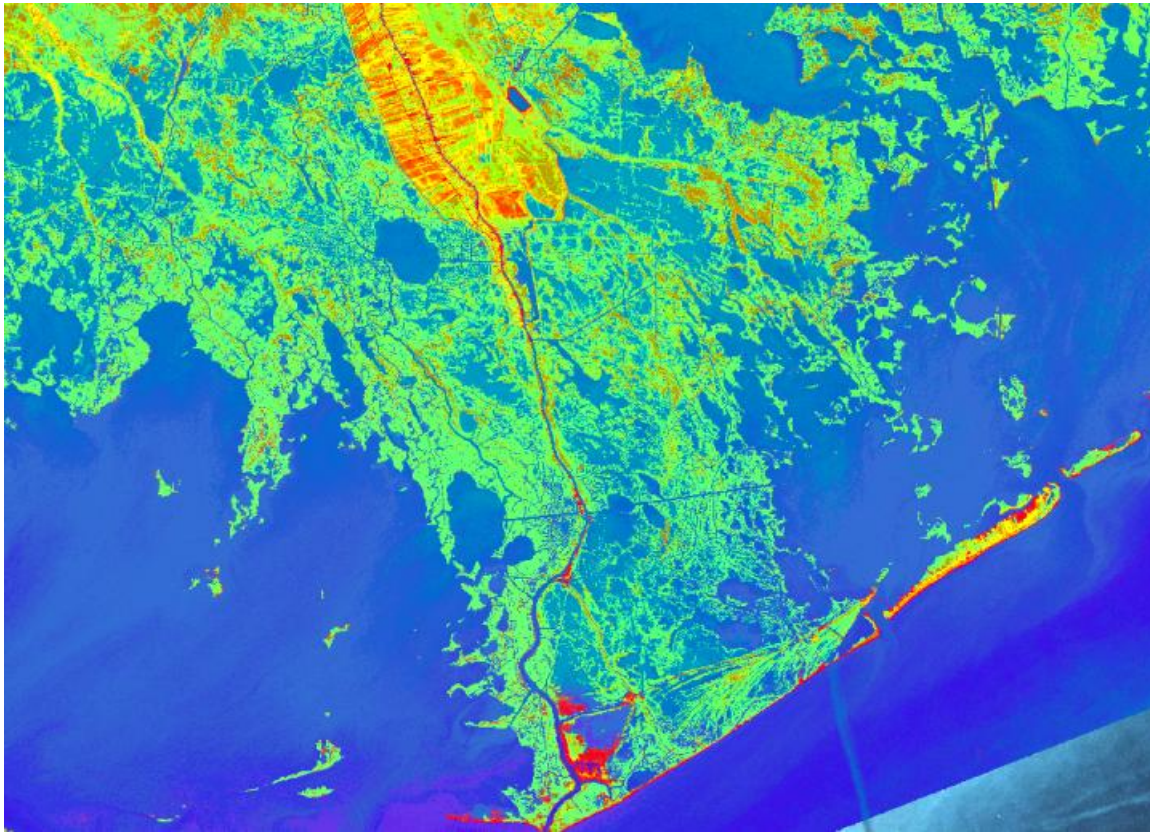


A part of a Landsat 7 ETM scene of south-central Louisiana in original true color (bands 3, 2, 1) is shown in figure 2. The tasseled cap transformation image of the same area is shown in figure 3. Pixels have been classified and rendered in 8-bit color into one of the three principal view components. Shown in figure 4 is the 2-bit, land-water classified raster file resulting from what had been a three-band, 24-bit tasseled cap transformation image file. The vector shoreline draped over the source 2-bit, land-water classified raster file and original Landsat 7 ETM image is shown in figures 5 and 6, respectively. Note the places where a vector shoreline was not generated. The ARC vectorization used in this investigation attempts to place a vector from the perspective of the largest aggregation of pixels classified as “water.” Thus, inland lakes and other bodies of water not directly linked to the ocean by water pixels will not show a vector delineation.

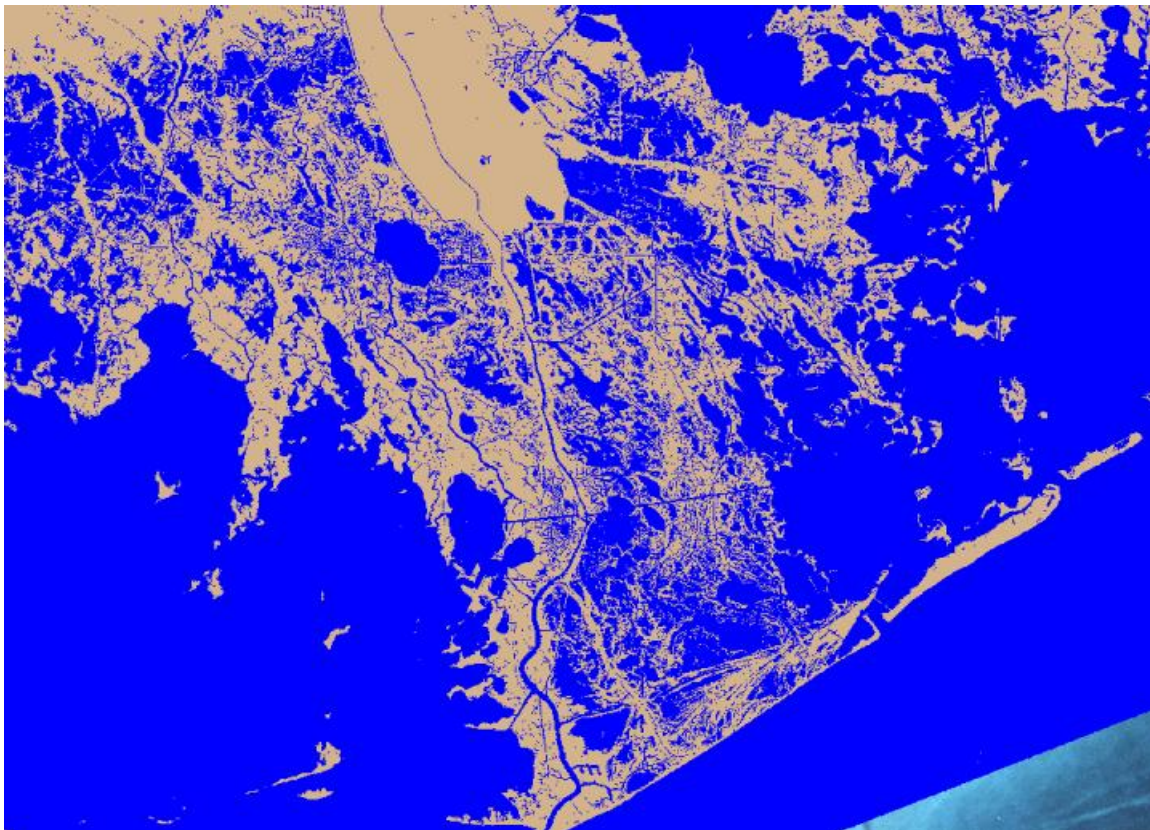


**Figure 2.** The north-central part of the original true-color Landsat 7 ETM scene P22 R40.



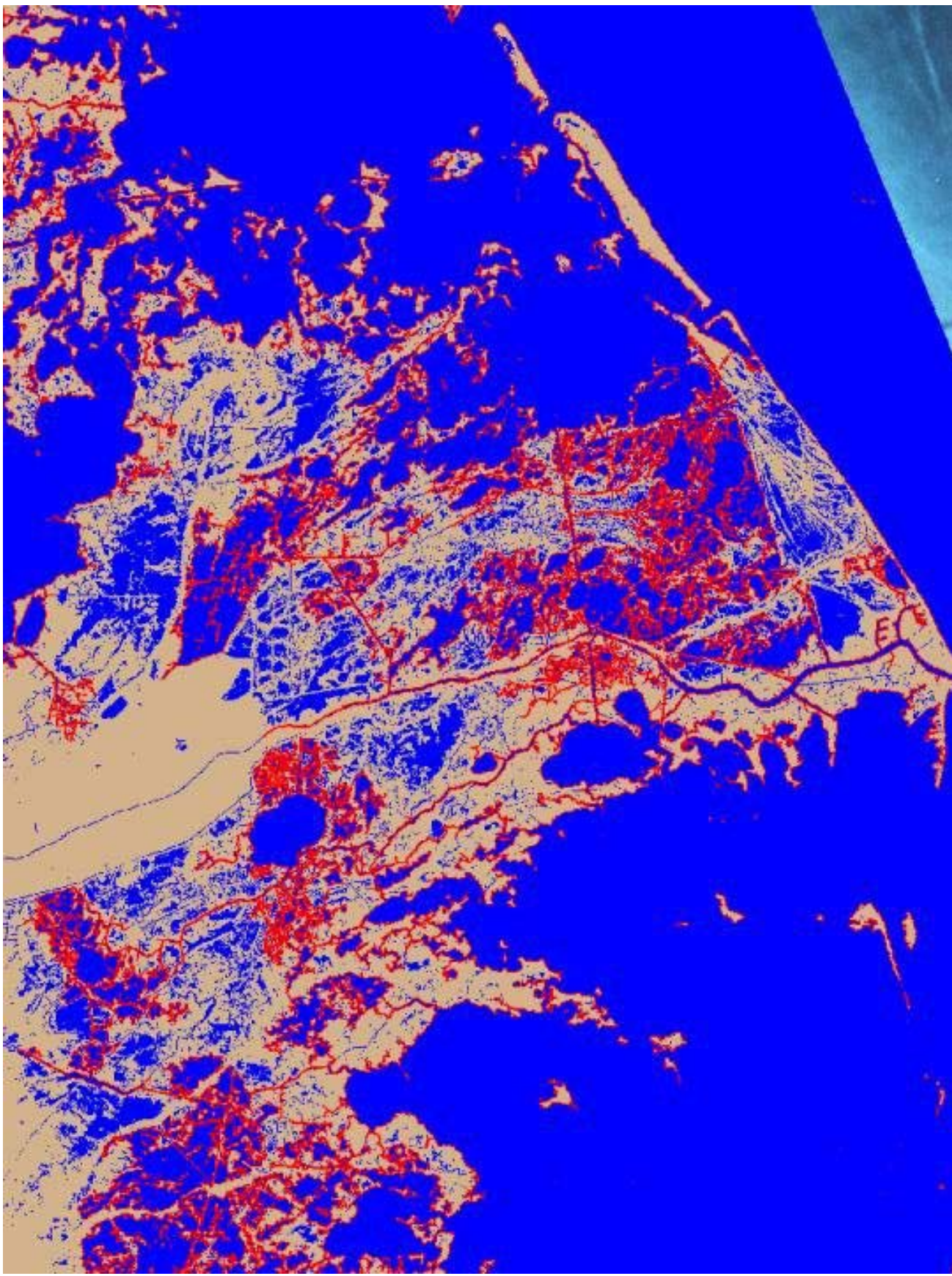


**Figure 3.** Three-band, 8-bit tasseled cap transformation image of the same Landsat 7 ETM scene from figure 2.



**Figure 4.** Two-bit raster file of the same Landsat 7 ETM scene showing pixel classification into land (brown) and water (blue).





**Figure 5.** Resulting vector shoreline (red) shown over 2-bit classified land-water raster file after vectorization





**Figure 6.** Vector shoreline (red) shown over original true-color Landsat 7 ETM scene.

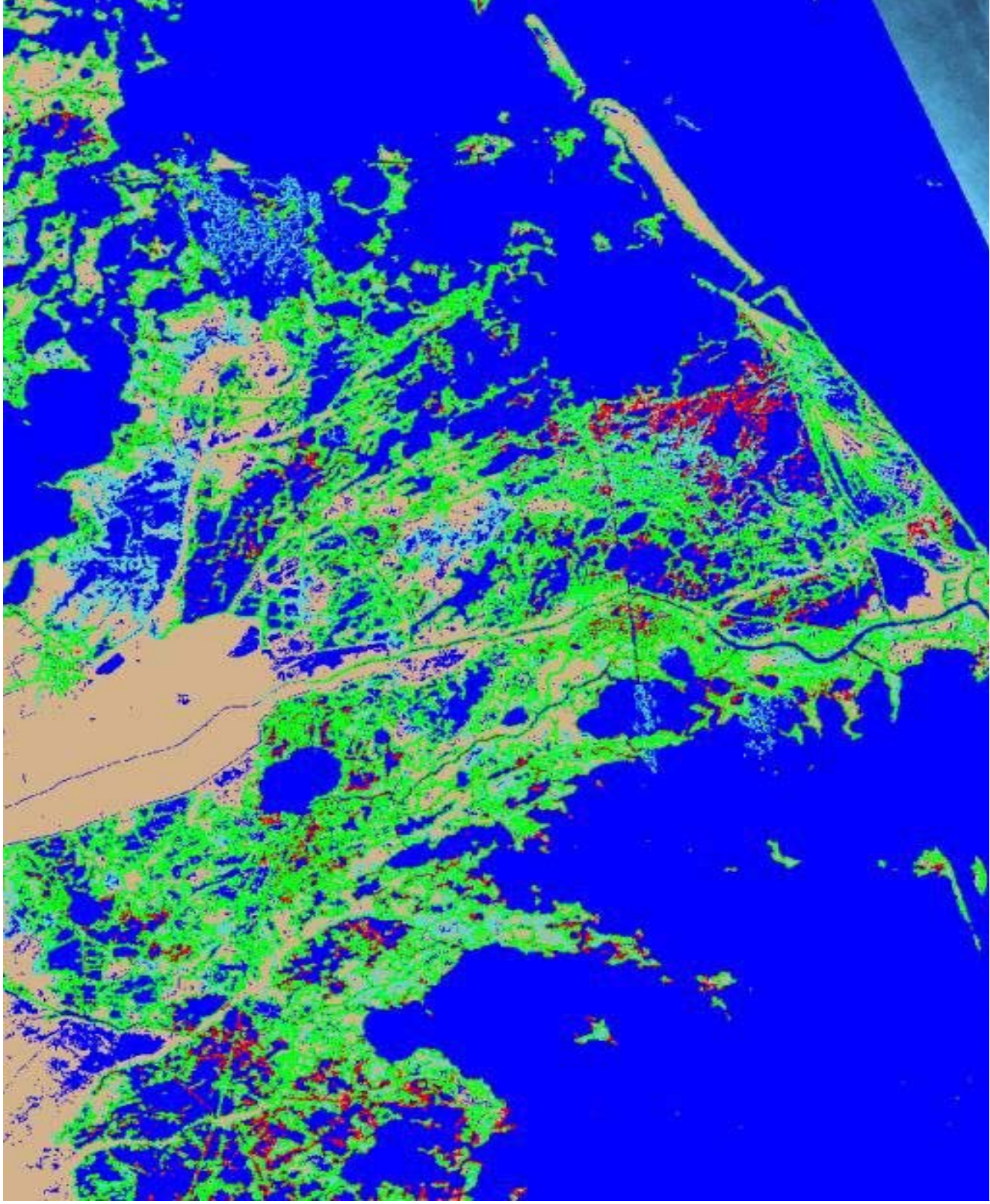
Research indicated that the following items should be noted to facilitate good shoreline extraction:

- Obtain scenes with a clear image of the shoreline. This method relies on a clear, unobstructed view of the shoreline as seen by the unaided eye and does not work if the shoreline is masked by clouds or is “extended” by ice or snow.
- Complexity in generating shorelines greatly increases with the number of scenes it takes to complete statewide/regional coverage. The State of Louisiana requires seven Landsat 7 ETM scenes to cover its shoreline with the Gulf of Mexico, which presents a challenge to the researcher when attempting to generate a shoreline that is a composite of a temporally distinct set of images. When attempting to generate multiple shorelines from multiple scenes of the same area, you must base the registration of the source images on the location of static, well-defined features. Scenes that lie successively along the same Landsat 7 path should be mosaicked.
- Acknowledge that all Landsat satellites (and most other currently operating satellites) do not acquire images in accordance with tide stages, and that whatever higher resolution ortho source is used as reference data are rarely in temporal sync with Landsat 7 ETM imaging.
- Accept the definition of shoreline as used in this investigation, which is simply the land-water interface as seen (classified) by the tasseled cap transformation. It is NOT a shoreline that represents NOAA’s “national legal” shoreline, nor does it follow other commonly used shoreline definitions. This research did not determine how the tasseled cap transformation handled turbid water, shallow water, and vegetation growing out of water, because reference data were unavailable. Therefore, it may not be suitable for detailed wetland studies that require highly accurate shoreline data.
- Based on a pixel size of 30 by 30 meters, the inherent positional accuracy of well-defined features extracted from Landsat 7 ETM scenes should meet the requirements for a 1:100,000-scale graphic map. Shoreline location, which is not considered a well-defined feature by Geography Discipline standards, would be valid for graphic products up to a graphic scale of 1:50,000. The accuracy of using the tasseled cap transformation to correctly classify Landsat 7 ETM pixels into the brightness, greenness, and wetness principal components is stated as being 97-percent correct for any one scene (Huang and others, 2002).

Hypothesis 2: A shoreline can be defined by increasing the temporal resolution of image data sources, even if the spatial resolution is decreased.

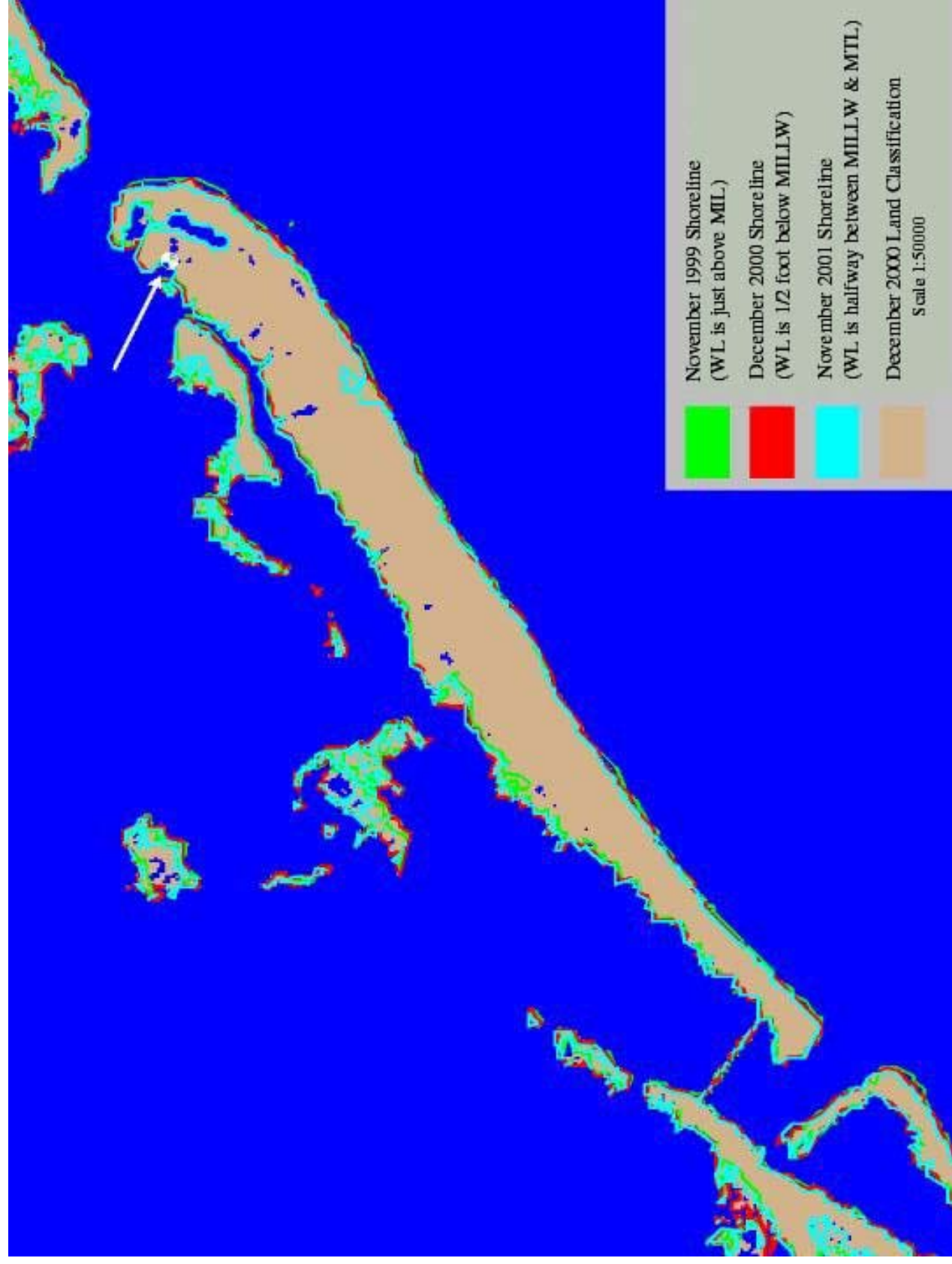
Achieved. A comparison of each of the Louisiana classified 2-bit raster files (and vector shoreline) with its source Landsat 7 ETM images for all 3 years displayed slight differences in position that corresponded well to water levels at the time of imaging (figures 7 and 8). This suggests that if more shorelines were to be generated from additional Landsat 7 ETM scenes, eventually the cumulative effect would be a display of the approximate range of shoreline position that could be occupied from low tide to high tide, even though Landsat 7 ETM scenes are imaged independently of tide stages.





**Figure 7.** Extracted 1999 (green), 2000 (red), and 2001 (lt. blue) shorelines shown over 2000 2-bit classified land-water raster file.





**Figure 8.** Enlargement of Grand Isle area of image showing more detail of all three shorelines' fit over 2-bit land-water classified raster file. White arrow and gauge symbol indicate location of NOAA's Grand Isle tide observation station.

Hypothesis 3: It is possible to accurately relate this interpreted shoreline to elevation values obtained from coastal tide stations.

A relation between tidal datums and the shoreline could not be established for the test areas. An investigation of the best available source for attributing the shoreline with an orthometric datum elevation value (National Vertical Datum of 1988) indicated that the “verified W1” water level observation holdings of NOAA’s (CO-OPS) National Water Level Observation Network System were the most viable. This was based on the fact that water-level observations were made every 6 minutes, which allowed good correlation to Landsat 7 ETM imaging time, and that the observations were made accessible on a Web site. However, several negative aspects were encountered in using the data, such as the lack of tide observation stations in remote, undeveloped stretches of the coast, questionable or obsolete tidal datum information in areas of known coastal subsidence and upheaval, and many reporting stations lacking the documentation to link their local tidal datum to an orthometric datum.

In addition, feedback from NOAA personnel regarding the proposed method of attributing shorelines with elevation values was that the hypothesis was unattainable at this time. The reasons included the need to incorporate tide modeling software between coastal tide observation stations (many of which either have not been developed or are obsolete) and the somewhat limited applicability of local tidal datum to surrounding areas. A potential future source for shoreline elevation values may be “VDatum,” a software tool currently being developed by NOAA. Although VDatum can convert tidal datum information to orthometric data for a specified location, it is currently limited to only a few coastal locations.

Hypothesis 4: The interpretation can be done quickly and cheaply enough that large lengths of shoreline can be extracted at least several times per year.

The investigation demonstrates that this goal is reasonable. In terms of labor hours, each of the Louisiana shorelines was created within 20 hours; two-thirds of this was machine processing time. This compares very favorably with labor hours of traditional manual compilation methods by an estimated factor of at least 4:1. In terms of source acquisition costs, each of the three Louisiana shorelines required seven Landsat 7 ETM scenes that were acquired at a total cost of \$4,200 (\$600/scene). When compared with the cost of acquiring aerial photographs to do similar shoreline mapping, the use of Landsat 7 ETM was again favored by a cost ratio of 8:1. The use of Landsat 7 ETM imagery as the data source also is very economical in comparison with other satellite acquired imagery, as demonstrated in table 5. Last, Landsat 7’s 16-day revisit cycle has the potential of yielding up to 22 usable scenes per year, providing a basis for rapid updating.

**Table 5. Cost of imaging comparison between various civil/commercial satellites**

	GSD	Scene size	Area	Scene price	Cost
Sensor	(m)	(km)	(km <sup>2</sup> )	(\$)	(\$ / km <sup>2</sup> )
AVHRR	1,000	2048 x 2048	4,194,304	0 – 50	0 - 0.000012
Landsat 7	15,30,60	170 x 185	31,450	480 – 600	0.02
Landsat 5	30,120	170 x 185	31,450	425 - 4400	0.02 - 0.14
IRS (LISS)	20	140 x 140	19,600	2500 - 3300	0.13 - 0.17
IRS (PAN)	5	70 x 70	4,900	2500 - 3300	0.51 - 0.67
SPOT 4	10,20	60 x 60	3,600	750 - 3400	0.21 - 0.04
SPOT 5	2.5,5,10	60 x 60	3,600	2450 - 5600	0.68 - 1.56
EROS1A	1.8(1)	12 x 12	10 - 144	250 - 1140	10.00 - 25.00
IKONOS	1,4	11 x 11	49 - 121	1210 - 14036	12.00 - 116.00
QuickBird	.61, 2.44	16.5 x 16.5	272	6120 - 25432	22.50 - 93.50

It should be noted that Louisiana is one of only six States that require seven or more scenes to achieve state-wide shoreline coverage, whereas most States require an average of two to three scenes. Many scenes contain parts of multiple States, thereby further reducing source acquisition costs when obtaining shoreline coverage for a region. For example, shoreline coverage for the Great Lakes involves twenty-one scenes, whereas the Gulf Coast requires twenty-two scenes. Creating complete shoreline coverage for the East Coast and West Coast requires twenty-four scenes and seventeen scenes, respectively.

In summary, a method of extracting shorelines that is based on exploiting the use of the tasseled cap transformation for Landsat 7 ETM imagery can be effectively and economically developed to provide State, regional, and national shoreline data. Such a method can provide a series of useful and practical vector and raster datasets containing 97-percent content accuracy, meet well-defined feature accuracy specifications for 1:100,000-scale graphic revision, and comply with many of *The National Map* requirements for consistency, content, and currentness. The user of this method, or of any of the resulting products of this method, must realize that the shoreline data do not follow any of the traditional shoreline definitions and have limitations when applied to coastal studies requiring a high degree of accuracy.

Recommendations for follow-on research include devising an attributing scene that would satisfy customer needs associated with a medium-scale shoreline map, researching the problem areas of shallow water and turbid water to determine how the tasseled cap transformation is classifying them, and determining if intermediate datasets used in creating land use/land cover products, such as the USGS National Land Cover Characterization 2001 project, could be substituted for Landsat 7 ETM images as source input to further reduce acquisition costs.

## REFERENCES

- Brand, R. and Hammack, J., 2002, A satellite image-based methodology for global shoreline mapping (*DRAFT*): National Imagery and Mapping Agency (NIMA).
- Crist, E.P., 1985, A TM tasseled cap equivalent transformation for reflectance factor data: *Remote Sensing of Environment*, 17: 301-306.
- Crist, E.P. and Cicone, R.C., 1984, A physically-based transformation of thematic mapper data -- the TM tasseled cap: *IEEE Trans. on Geosciences and Remote Sensing*, GE-22: 256-263.
- Crist, E.P. and Kauth, R.J., 1986, The tasseled cap demystified: *Photogrammetric Engineering & Remote Sensing*, 52: 81-86.
- Huang, C., Wylie, B., Homer, C., Yang, L., and Zylstra, G., 2002, Derivation of a tasseled cap transformation based on Landsat 7 ETM at-satellite reflectance: *International Journal of Remote Sensing*, v. 23, no. 8, p. 1741-1748.
- Kauth, R.J. and Thomas, G.S., 1976, The tasseled cap -- a graphic description of the spectral-temporal development of agricultural crops as seen in Landsat, in *Proceedings on the Symposium on Machine Processing of Remotely Sensed Data, West Lafayette, Indiana, June 29 -- July 1, 1976*, (West Lafayette, Indiana: LARS, Purdue University), 41-51.
- Parker, B., 2001, Where is the Shoreline?: *Hydro International*, July/August 2001, p. 6-9.

---

<sup>1</sup> U.S. Geological Survey, Mid-Continent Mapping Center, Rolla, Missouri, 573-308-3800

<sup>2</sup> Science Applications International Corporation (SAIC)