Houston-Galveston Area Council
2002 Land Cover Image Processing and Accuracy Assessment Protocol

Prepared by:

Roger Meyer
Houston-Galveston Area Council
Clean Rivers Program
3555 Timmons Lane, Suite120
Houston, Texas 77227

Prepared for:

Texas Commission on Environmental Quality
Texas Clean Rivers Program
12100 Park 35 Circle
Austin, Texas 78753

January 2003

Table of Contents

Figures

Tables

Houston-Galveston Area Council
2002 Land Cover Image Processing and Accuracy Assessment Protocol

Abstract

The image processing and accuracy assessment protocol used to produce the 2002 land cover data set for the Houston-Galveston Area Council (H-GAC) planning region are described in this document. The primary objective for the project was to develop a low cost classification and accuracy assessment methodology, which could produce a land cover product with an acceptable level of accuracy. An acceptable level of accuracy was defined as overall accuracy and overall Kappa values greater than 70%. The data set is to serve primarily as a component in watershed and water quality analysis within H-GAC's Clean Rivers Program assessment basins.

The classification methodology utilizes image stratification to reduce spectral mixing of the 9 target land cover classes over the entire region, and an iterative cluster busting process of multi-temporal ETM and TM satellite imagery to assign land cover labels to clusters produced by the unsupervised clustering algorithm. The accuracy assessment of the final land cover products consisted of a stratified random sampling scheme with a target of 75-100 sample points per land cover class. A total of 961 sample points have been accumulated thus far with a sample size greater than 75 for 8 of the 9 land cover classes. Current and historical aerial photography, and a limited number of field verification visits were used for developing the reference data. The accuracy assessments were done using different levels of post-processing, sampling restrictions, and class aggregation to better enable users to determine the efficacy of the data set.

Results of the conservative accuracy assessment produced a moderate overall Kappa of 71% and an overall accuracy of 75% for the raw pixel classification. The optimistic accuracy assessment, which limited reference sampling to 3x3 homogenous areas, produced an overall accuracy of 94% and a strong overall Kappa of 92%. Comparison of the conservative and optimistic accuracy assessments for the raw classification indicates that the classified map has strong agreement with the reference data (i.e., Kappa > .8 for all categories) in homogenous areas, but is less accurate in land cover transition zones and in other areas that exhibit heterogeneity of land cover types. Examination of commission and omission errors indicated that the primary source of classification errors were related to interpreting and assigning clusters to low or high intensity developed, spectral mixing in grassland and woodland transitional zones, spectral mixing of grassland and agriculture, and the omission of open woodland or scrub shrub transitional zones. Mode filtering of the classification appears to have positive effects on classification accuracy, slightly increasing the overall Kappa to 73% and overall accuracy to 77%. Merging of confused classes and reducing the total number of land cover classes to 6, increased overall accuracy to 82% and increased overall Kappa to 77%.

A comprehensive Kappa analysis indicated that the raw classification was 75% better than random in specifying location of the land cover classes (Kappa of Location = .75) given the fixed quantity of the accuracy assessment sample, and 76% (Kappa of Quantity = .76) better than random in specifying quantity of land cover quantities given the specified location of the sample data. Examination of the value of perfect information of location indicates that the overall classification accuracy could be increased by as much as 19% by improving the classification's ability to specify location, particularly for the grassland, agriculture, woodland and developed classes. Classification accuracy for developed areas may be improved by more diligent interpretation of clusters resulting from the iterative cluster busting process. Developing training signatures from additional temporal or hyperspectral satellite imagery, and inputting those signatures into a supervised classifier may improve classification accuracy of cultivated and transitional woodland areas. The 2002 aerial photography can also be used to identify and recode any major classification errors. Completion of the remaining accuracy assessment sample points will allow more rigorous statistical inferences to be made regarding overall and per category accuracies for the classified map.

1. Introduction

The image processing and accuracy assessment protocols used to produce the 2002 land cover data set for the Houston-Galveston Area Council (H-GAC) planning region are described in this document. The study area described herein includes the 15 counties encompassing the four assessment basins for the H-GAC Clean Rivers Program (CRP), approximately 17,814 square miles (Figure 1). The primary goal of the project was to develop a land cover data set that could be utilized as a component in watershed and water quality analysis within H-GAC's CRP assessment basins. Requirements for the project were a low cost classification and accuracy assessment methodology, which could be repeated every 3 to 5 years and produce a land cover product with an acceptable level of accuracy. An acceptable level of accuracy was defined as an overall accuracy and overall Kappa greater than 70%.

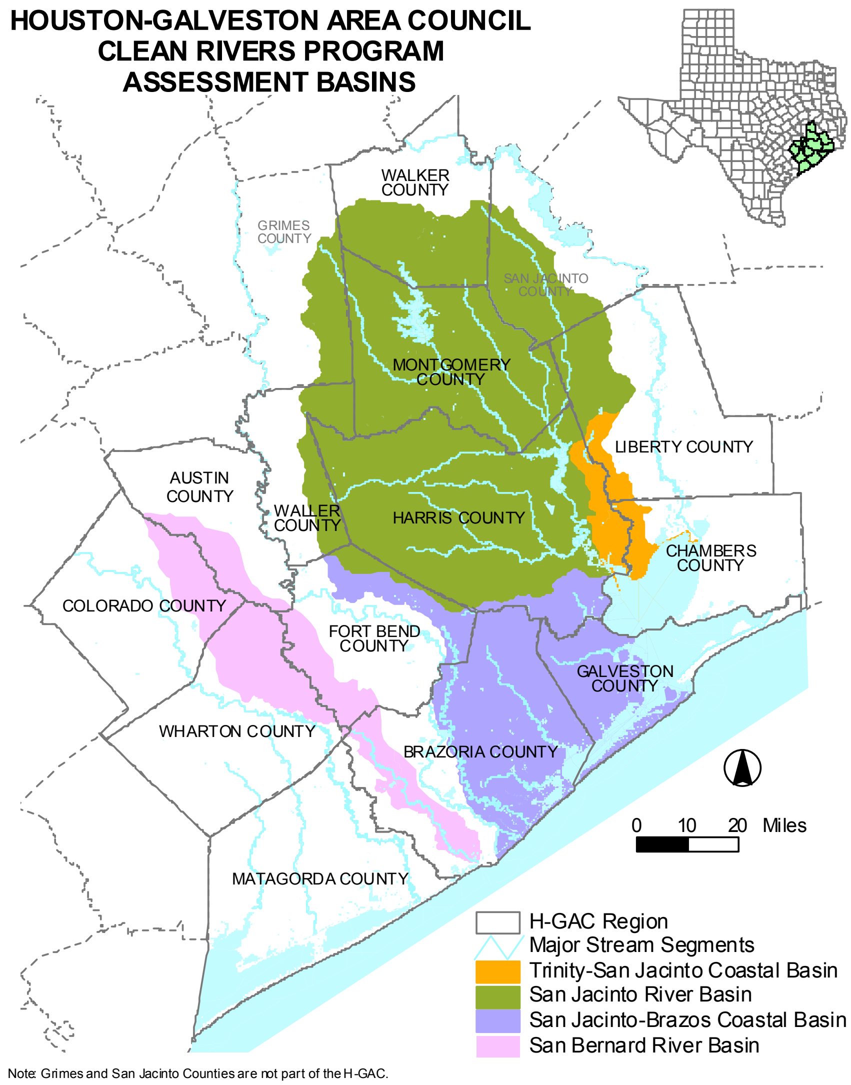

Figure 1: The 15 counties encompassing the study area and location of H-GAC's assessment basins

2. Image Selection and Acquisition

Portions of four Landsat scenes were required to provide full coverage of the study area (i.e., path 25-26 and rows 39-40). Multi-temporal imagery was purchased to facilitate the capture of land cover signatures that result from seasonal changes in vegetation phenologies. As seen in Table 1, a total of 8 scenes were purchased to capture summer "leaf on" and winter "leaf off" spectral signatures of our target land cover classes. All 8 scenes were purchased from the USGS EROS Data Center in NLAPS and FAST-L7A

Table 1: Satellite imagery purchased from the USGS EROS Data Center

Path - Row	Satellite	Acquisition Date
26 - 39	Landsat 5	22 July 2001
26 - 40	Landsat 5	22 July 2001
25 - 39	Landsat 7	25 September 2001
25 - 40	Landsat 7	25 September 2001
26 - 39	Landsat 7	23 February 2002
26 - 40	Landsat 7	23 February 2002
25 - 39	Landsat 7	15 January 2002
25 - 40	Landsat 7	15 January 2002

format and imported into IDRISI32 software (Clark Labs). The scenes were radiometrically and geometrically corrected by USGS (Level 1G systematically corrected) and rotated and aligned to the UTM coordinate system using the WGS84 datum and a 30 meter resolution. Two Landsat 5 scenes were purchased due to the lack of cloud free ETM data during the time period of interest.

3. Image Preprocessing

Initial examination of the positional accuracy of the geometrically corrected scenes indicated a ground positional error of 60 meters or less for Landsat 7 ETM scenes and a positional error of 100 meters or less for the two Landsat 5 TM scenes. This level of positional accuracy was determined to be inadequate for classification and accuracy assessment purposes. Thus, the four 2002 winter scenes were registered to the TX DOT vector road layer using a linear first order transformation with nearest neighbor resampling. Image to image registration was applied to register the four 2001 scenes to the corresponding 2002 map registered scenes. This registration procedure also used a first order transformation and nearest neighbor resampling. All scenes were registered to within ½ a pixel (i.e., RMS error < 15 meters), according to national mapping standards. The number of control points selected for registration was dependent on the area that was to be extracted from the original scene and ranged from a minimum of 16 control points for smaller areas to 30 points for the larger areas extracted. The registered Landsat scenes were concatenated based on Path/Row position and date of acquisition, forming two eastern images (i.e., summer 2001 and winter 2002 images) and two western images of the study area. The study area was then masked from these images using a 6km buffer of the 15 counties encompassing the study area.

4. Classification Scheme

The target land cover classes were limited to 9 land cover classes that could be readily identified through visual observation of the satellite imagery and aerial photography. The target land cover classes were derived from level 1 of the NOAA Coastal Change Analysis Program (C-CAP) classification scheme. Please refer to the NOAA C-CAP: Guidance for Regional Implementation, which has more thorough explanations of the land cover classes (Dobson et. al., 1995). The following is a list of shortened land cover class descriptions extracted from the C-CAP classification scheme.

Low Intensity Developed

"Low Intensity, Developed Land includes areas with a mixture of constructed materials (e.g., roofing, metal, concrete, asphalt) and vegetation or other cover. Constructed materials account for 50 to 79% of total area. These areas commonly include single-family housing areas, especially in suburban neighborhoods, but may include scattered surfaces associated with all types of land use. As the percentage of constructed material cover decreases, this category grades into Cultivated, Grassland, Woody, and other land cover classes. A large building surrounded by several acres of grass, for example, might appear as one or more pixels of High Intensity Developed Land, one or more pixels of Low Intensity Developed Land and many pixels of Grassland."

High Intensity Developed

"High Intensity, Developed Land includes heavily built-up urban centers and large constructed surfaces in suburban and rural areas with a variety of different land uses. The High Intensity category contains areas in which a significant land area is covered by concrete and asphalt or other constructed materials. Vegetation, if present, occupies < 20% of the landscape. Examples of such areas include apartment buildings, skyscrapers, shopping centers, factories, industrial complexes, large barns, airport runways, and interstate highways."

Cultivated Land

"This category contains areas that have been planted, tilled, or harvested. Pastures and hayfields that are in a state of tilling or planting are also included. Otherwise, pasture or hayfield with well-established grasses are placed in the Grassland category."

Grassland

"The C-CAP category includes lands with herbaceous cover at the time of observation regardless of origin or potential. Pastures, hayfields, and natural rangelands are included. Also included are lawns and other managed grassy areas such as parks, cemeteries, golf courses, road rights-of-way, and other herbaceous-covered, landscaped areas."

Woody Land

"The Woody Land class includes any species with an aerial stem that persists, for more than one season." The Woody Land class includes shrub scrubland and the three C-CAP subclasses: Deciduous, Evergreen, and Mixed.

Open Water

"All areas of open water with < 30% cover of trees, shrubs, persistent emergent plants, emergent mosses, lichens, or other land cover are grouped under the heading, Water and Submerged Land, regardless of whether the area is considered wetland or deepwater habitat under the Cowardin et al. (1979) classification system."

Wetlands (separated into Woody and Non-Woody Wetland)

"Cowardin et al. (1979) define wetlands as lands where saturation with water is the dominant factor determining soil development and the types of plant and animal communities living in the soil and on its surface. The single feature that all wetlands share is soil or substrate that is at least periodically saturated with or covered by water. The upland limit for vegetated wetlands with soil is 1) the boundary between land with predominantly hydrophytic cover and land with predominantly mesophytic or xerophytic cover; 2) for non-vegetated wetlands with soil the boundary between soil that is predominantly hydric and soil that is predominantly nonhydric; or 3) in the case of wetlands without vegetation or soil, the boundary between land that is flooded or saturated at some time during the growing season each year and land that is not. The majority of all wetlands are vegetated and are found on soil… In the C-CAP Coastal Land Cover

Classification System, "Wetland" includes all areas considered wetland by Cowardin et al. (1979) except for Wetland Bottoms, Aquatic Beds, and Nonpersistent Emergent Wetlands."

This C-CAP class was separated into two classes, woody wetland and wetland. The woody wetland class includes all wetlands, both palustrine and emergent, whose spectral signature indicated the presence of persistent woody vegetation. All other palustrine and estuarine wetlands that did contain persistent woody vegetation were allocated to the wetland category. A classified image with 10 classes was also created, which further separated these two wetland classes into salinity regimes using the NWI data (i.e., Palustrine Emergent, Palustrine Woody, and Estuarine Wetlands)

Bare or Transitional Land

"Bare Land, modified from "Barren Land" in Anderson et al., 1976, is composed of bare rock, sand, silt, gravel, or other earthen material with little or no vegetation regardless of its inherent ability to support life. Vegetation, if present, is more widely spaced and scrubby than that in the vegetated categories. Unusual conditions, such as a heavy rainfall, occasionally may result in a short-lived, luxuriant plant cover. Wet, nonvegetated exposed lands are included in the wetland categories."

"Transitional Areas are dynamically changing from one land cover to another, often due to land use activities. This transitional phase occurs when, for example, forest lands are cleared for agriculture and wetlands are drained for development. Often land becomes temporarily bare as construction initiates the transition from Woody Land or Grassland to a future cover associated with residential, commercial, or other intensive land use. Lands, such as spoil banks and sanitary landfills, temporarily altered by grading and filling are considered transitional." Transitional areas also include woody transitional areas that have been clear cut and replanted in the past 2-3 years, and are transitioning to the woodland class.

5. Classification Methodology

Our classification methodology relied on an unsupervised classification algorithm to capture the inherent spectral signatures of the target land cover classes within the TM and ETM image bands. The unsupervised classification routine in IDRISI utilizes an iterative, self-organizing routine similar to the well-known ISODATA routine and cluster routines such as the H-means and K-means procedures. The algorithm also utilizes an 8bit composite image for initial seeding of clusters and a full maximum-likelihood procedure for cluster assignment. The overall approach to the classification process was to utilize image stratification to reduce spectral mixing of our target land cover classes over the entire region, and iterative cluster busting of the multi-temporal image bands to assign land cover labels.

5.1 Image Stratification

As described by Liliesand et.al, (1998) the eastern and western images were manually separated into spectrally consistent classification units (SCCU) using the Natural Regions of Texas data layer (Texas Parks and Wildlife Dept., 1978) and visual inspection of the general photomorphic characteristics of the satellite imagery. Each SCCU was further stratified into wetland and upland areas using NWI data (USFWS, 1992), urban areas using the TX DOT Road layer (Texas Dept. of Transportation, 2000) and Texas Urban Area map layer (Texas General Land Office, 1999), and cloud covered areas using on screen digitizing. The objective in dividing the imagery into spectrally consistent strata was to limit the amount of spectral mixing of our target land cover classes that would occur during the classification process. Although the method increased the amount of processing time required it greatly improved the accuracy of the classification.

5.2 Classification Process

The first step in the classification process was to separate each SCCU into upland, wetland, urban, and cloud covered stratum. Principal components analysis (PCA) was then applied to extract the first 3 components for both dates of imagery, forming 6 bands for input into the classification algorithm. Principal components were extracted separately for the upland, wetland and urban stratum. The extracted urban areas for each SCCU were then classified into low intensity and high intensity developed using the unsupervised clustering of the 6 PCA bands, and an iterative cluster busting process (Jensen, 1987). The cluster busting process requires the interpreter to assign clusters to the target land cover classes. Clusters that can be confidently assigned to target classes are then omitted from the next iteration of the clustering algorithm. Another iteration of the clustering algorithm is then performed on the remaining clusters that cannot be identified as to target class. This iterative process is continued until the interpreter can no longer confidently assign clusters to the target classes. Those clusters that remained spectrally mixed were manually stratified into the target classes using on screen digitizing and image interpretation of the satellite imagery and aerial photography. The classified urban areas for the SCCU were then removed from the wetland, upland and cloud covered areas, and these areas were in turn classified using the iterative cluster busting process.

Cloud covered areas were manually delineated using on screen digitizing and subsequently stratified into upland, wetland, and urban areas. These areas were then classified separately using a single date of cloud free imagery. The same cluster busting process was utilized for the classification with the exception of using the original bands for the single date of cloud free imagery (i.e., 6 bands with thermal omitted) as input in the classification algorithm. The classified wetland, upland, urban and cloud stratum for each SCCU were recombined to form a single classified image of the SCCU. All classified SCCUs were then aggregated to form a final classified eastern and western image of the study area. Figure 2 provides an overview of the classification procedure.

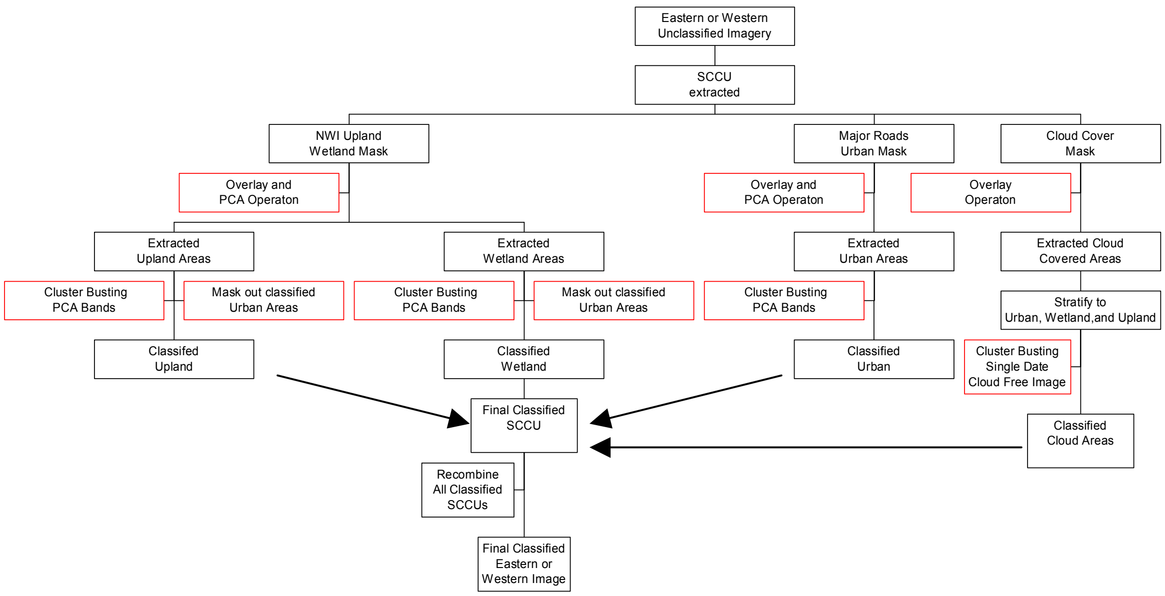

Figure 2: Flow chart of classification methodology

5.3 Rules for Wetland Classification

Accurate classification of wetland habitats is difficult due to spectral mixing, especially when classifying a large region. Consequently, we utilized NWI data (1992) as the primary source for determining the potential spatial location of wetlands during classification. Rather than rasterize and embed the data into our classification, we utilized attributes of the NWI data to stratify the imagery into wetland and upland areas. Within the wetland strata, principal components were extracted from the multi-temporal image bands and input into the clustering algorithm. The primary rules for interpreting the resulting clusters were as follows:

- The NWI data was considered to be accurate unless examination of the aerial photography and satellite imagery indicated the change had occurred (e.g., wetland has been converted to developed land or cultivated land, or woody wetland has been clear cut and is now in a transitional stage). If change was verified then the cluster was assigned to the suggested land cover class.

- If no change was indicated and the cluster consisted primarily of persistent woody vegetation then the cluster was assigned to the woody wetland category.

- If no change was indicated and the wetland cluster did not indicate the presence of woody vegetation than the cluster was assigned to the wetland category (non-woody).

- If no change was indicated and the cluster consisted primarily of water reflectance then the cluster was assigned to the open water class.

This method allowed us to limit spectral mixing associated with classifying wetlands over a large region and to spectrally update the NWI data if major change had occurred.

5.4 Transitional Land Classification

Transitional Land consists of areas that are transitioning to another land cover class. They include construction sites where land has been cleared, and woody land that has been harvested within the past two to three years and is currently in the process of regenerating. Spectral signatures of woody transitional areas were difficult to separate due to the mixture of spectral reflectance resulting from woody vegetation, grassland, and exposed soil typically found in harvested areas. Consequently, woody transitional areas were manually recoded to the bare/transitional class using on screen digitizing and image interpretation. The larger woody transitional areas were delineated using the satellite imagery and historical aerial photography to identify changes in forested areas and the geometric patterns indicative of harvesting activity. As a consequence of manual recoding, smaller woody transitional areas may have been omitted from the manual recoding process.

6. Post Classification Processing

The final eastern and western image classifications were projected to a common coordinate system so that they could be concatenated and edge matched to form the final classification of the study area. Because the study area contains both UTM zone 14 and 15, a customized UTM projection was created, which shifted the central meridian to -95.5 degrees. The customized projection reduces scale factor error on the margins of the study area to within the national mapping standard of 4 parts per 10,000. It was important to account for scale factor error when conducting the accuracy assessment of the final classified image.

The classification process can produce areas of isolated pixels that differ from the majority class. These isolated pixels are a result of the complexity of separating land cover signatures in a satellite image, or can reflect the actual heterogeneity of land cover. For mapping purposes it is a common procedure to generalize an image to reduce the occurrence of these areas of isolated pixels. After the projection and edge matching process, a 3x3 mode filter was applied to the unprocessed classification to improve homogeneity of the land cover classes and reduce the occurrence of stray pixels. The mode filter determines the most frequently occurring land cover class in a 3x3 pixel block and assigns the center pixel of the block to this land cover class. Applying the mode filter to the entire classified image essentially generalizes the entire classification, but may produce a more visually appealing result for general mapping purposes.

Examples of the final images produced in post classification processing can be found in Appendix B. The final images produced in post processing were the unprocessed classified image, the mode filtered classification, and a merged class image. NWI data was used to further stratify the wetland categories for the unprocessed classification into salinity regimes (i.e., Palustrine Woody, Palustrine Emergent, and Estuarine Wetlands), forming the classified image with 10 classes.

7. Preliminary Accuracy Assessment

This project called for an accuracy assessment methodology, which required limited investment in time and monetary resources. As a consequence the accuracy assessment relied primarily on current (2002) and historical (1995-1996 and 1999-2000) aerial photography for field verification of accuracy. The aerial photography consisted of 2.5 meter, 1 meter, and sub-meter digital orthophotos. The most current aerial photography with the highest resolution was selected whenever possible. Our target land cover classes were general enough to confidently identify on the aerial photography, allowing a larger sample size relative to ground-based sampling. The primary disadvantage occurred when cross-examination of historical photography and the satellite imagery indicated that change in land cover had occurred. In this case the sample point was either eliminated or an effort was made to visit the point on the ground. Ground-based sampling was conducted using a TRIMBLE PRO-XRS global positioning system integrated with IDRISI software running on a laptop. This method allowed the investigator to locate a position on the classified imagery in real time when out in the field.

The objective for the accuracy assessment was to obtain a minimum of 75 sample points per land cover class using a stratified random sample. Both conservative and optimistic accuracy assessments were conducted based on pixel-to-pixel comparisons and samples restricted to homogenous areas. Assessments were also conducted on the mode filtered image and a classified image with aggregation of problem land cover classes. Accuracy assessments using these different definitions of agreement between the ground truth data and the classified data will assist users in determining whether the data's accuracy will be suitable for their needs.

7.1 Sampling Strategy

A stratified random sampling methodology was selected for the accuracy assessment. The stratified random sampling scheme provides the strong geographic coverage of systematic sampling and the low potential for bias of random sampling. The "SAMPLE" module in IDRISI divides the classified image into a rectangular matrix of cells and randomly selects a sample point within each cell. Congalton (1991) recommends a minimum of 75 – 100 sample points per land cover class for classification of areas in excess of 1million acres or with classifications using more than 12 categories. The goal was to obtain a minimum of 75 sample points per land cover class. Anticipating that some of the sample points would be eliminated due to restricted access or the inability to interpret the point using historical aerial photography, additional points were randomly added to the initial sample to ensure that classes occupying a low proportion of the total area were adequately sampled (see Table 2). The grassland category was more heavily sampled due to the higher proportional area occupied by this class. It should be noted that the woody wetland category was not adequately sampled due to accessibility issues and difficulty identifying these areas on aerial photography. No statistical inference regarding the accuracy of the woody wetland category can be made due to the small sample size.

Table 2: Final sampling distribution per land cover class

Land Cover Class	Area (acres)	% of Area	Target Number of Sample Points	Points Sampled to Date
Low Intensity Developed	330,251	2.9	123	117
High Intensity Developed	381,322	3.3	96	93
Cultivated Land	1,066,233	9.4	96	81
Grassland	4,124,034	36.2	329	253
Woody Land	2,668,735	23.4	190	139
Open Water	1,729,878	15.2	127	125
Woody Wetland	467,277	4.1	89	13
Non-Woody Wetland	491,264	4.3	105	73
Bare or Transitional Land	141,905	1.2	77	67
Total	**11,400,899**	**100.0**	**1232**	**961**

7.2 Overview of Accuracy Assessment Statistics

The contingency matrix evaluates the correspondence of the ground reference data found in the columns of the table to the classified map data in rows (contingency matrices for each accuracy assessment are provided in Appendix A). The number of correct pixels per class can be found along the diagonal of the table. Values outside the diagonal of the table represent misclassified samples due to omission and commission errors. Statistical summary tables for each of the accuracy assessments, include: overall accuracy, overall Kappa, per category producer and user accuracies, calculated 95% confidence intervals for overall accuracy and per category producer and user accuracies, and per category Kappa statistics using both the ground data and classification data as the referent.

Producer's accuracy is the measure of omission error and is derived by dividing the number of ground truth samples that were correctly classified in the map by the total number of ground samples for that class. Thus, producer's accuracy is the probability that a pixel observed in the field is correctly depicted on the map. User's accuracy accounts for commission error, which is derived by dividing the number correct for a class by the total number of pixels classified as that class in the map data. User's accuracy is therefore a measure of the reliability of the map because it identifies the proportion of pixels in the classified map that may be committed to the wrong class. User's accuracy is the probability that a pixel on the map correctly identifies land cover categories, as they exist in the field.

Overall accuracy is a simple measure of accuracy and is derived by dividing the number of correct correctly classified samples along the diagonal of the contingency matrix by the total number of accuracy assessment samples. The overall Kappa statistic provides a more robust indicator of overall accuracy because it accounts for the agreement that may result due to chance. A Kappa of 0 (0%) would represent a classification that is no better than randomly assigning pixels to the land cover classes, whereas a Kappa of 1 (100%) represents perfect agreement and a classification that is 100% better than random in assigning pixels to the correct land cover classes. Congalton (1996) classifies Kappa values into three categories: a value of .8 (80%) or higher represents strong agreement, a value between .4 and .8 (40-80%) represents moderate agreement, and a value below .4 (40%) represents poor agreement.

7.3 Preliminary Accuracy Assessment Results

The 961 sample points were interpreted using the most current aerial photography available and a limited number of ground visits. Figure 3 below shows the geographical distribution of sample points that have been accumulated thus far. These interpreted points were input into a contingency matrix, which determines the number of correctly classified samples. The accuracy assessments were done using different levels of post-processing, sampling restrictions, and class aggregation to better enable users to determine the efficacy of the data set.

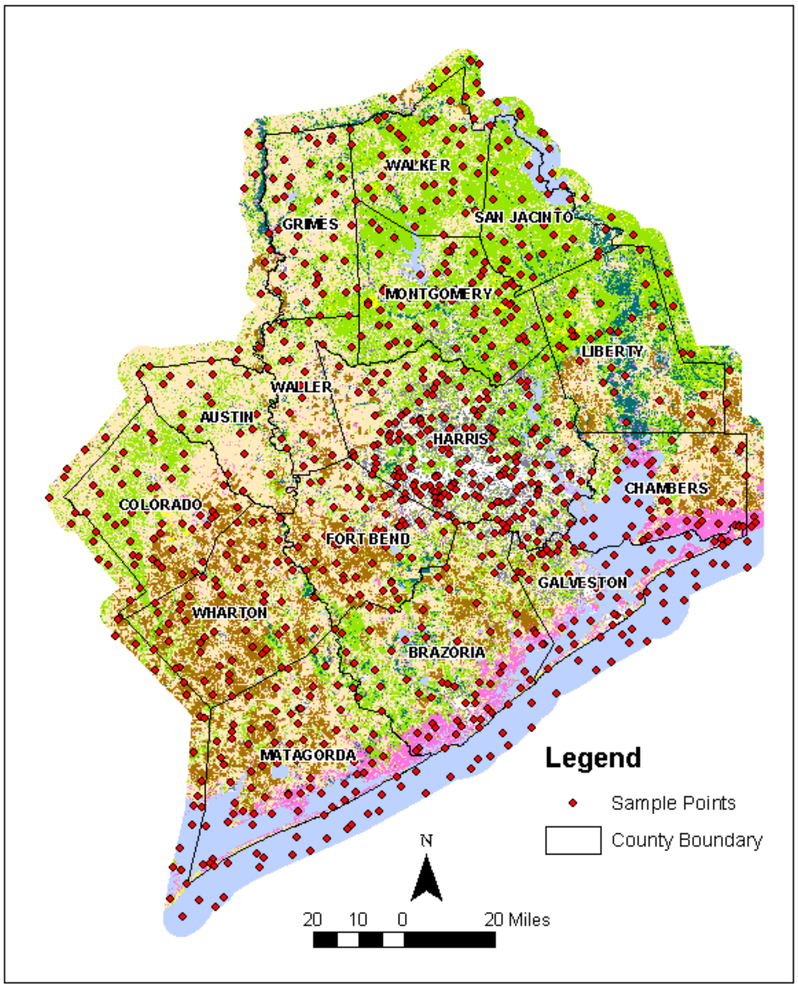

Figure 3: Geographical distribution of accuracy assessment sample points on the classified map

7.3.1 Conservative Accuracy Assessment

A pixel-to-pixel assessment was first conducted on the unprocessed, raw classified image. This type of accuracy assessment is considered to incorporate conservative bias

(Verbyla and Hammond, 1995) because of the difficulty in separating true classification error from error that may result due to misregistration of the satellite imagery (perfect registration is impossible to obtain), temporal differences between the collection of the ground truth data and satellite imagery, or the inability of the examiner to confidently interpret individual pixels on aerial photography. Thus the conservative assessment may actually underestimate the true accuracy of the classification.

A pixel-to-pixel assessment of the raw classified image with our ground truth data yielded an overall accuracy of 75% and a moderate overall Kappa of 71% (Table 3). Calculation of the 95% confidence interval for overall accuracy indicates that with a

Table 3: Statistical summary of conservative accuracy assessment

Category	User's Accuracy	User's 95% CI	Kappa Map Referent	Map Samples	Producer's Accuracy	Producer's 95% CI	Kappa Ground Referent	Ground Samples
Low Intensity Developed	73%	64.1% - 81.2%	68%	117	66%	57.8% - 75.0%	62%	128
High Intensity Developed	62%	52.0% - 72.7%	59%	93	75%	65.0% - 85.6%	73%	77
Agriculture	84%	75.3% - 92.6%	82%	81	75%	65.2% - 84.2%	72%	91
Grassland	67%	60.8% -72.8%	57%	253	77%	71.4% - 83.0%	69%	219
Woodland	82%	75.3% - 88.8%	78%	139	65%	57.4% - 72.1%	59%	176
Open Water	91%	85.8% - 96.6%	90%	125	91%	85.8% - 96.6%	90%	125
Woody Wetland	31%	1.8% - 59.7%	30%	13	50%	9.1% - 90.9%	49%	8
Wetland	75%	64.8% - 85.9%	73%	73	76%	65.9% - 86.9%	74%	72
Bare or Transitional	79%	68.6% - 89.6%	78%	67	82%	71.3% - 91.7%	80%	65

Overall Accuracy = 75% (95% CI: 72.1% - 77.7%) Total Observations = 961 Overall Kappa = 71%

sample size of 961 and repeated sampling, we can be 95% confident the true overall accuracy will lie between 72.1% and 77.7%. The lowest user's accuracies were for the grassland (62%) and high intensity developed (67%) classes. Examination of the contingency matrix indicates that the lower accuracy of the grassland class can be attributed to commission errors resulting from transitional zones with woodlands and low intensity developed, as well as spectral mixing with cultivated areas. The lower accuracy of the developed classes is closely associated with assigning clusters to either the high or low intensity developed category, and spectral mixing in transitional areas with grassland and woodland. The woodland class has a low producer's accuracy 59% when compared to the user's accuracy of 82%. The disparity of these results indicates that although the classified map is reliable in identifying woodland areas on the ground, there are woodland areas on the ground that are omitted from the map. This is primarily a result of transitional woodland and scrubland zones that have been omitted and classified as grasslands.

7.3.2 Optimistic Accuracy Assessment

The optimistic assessment of the unprocessed classification restricted sampling to areas in the image, which contained only one land cover class in a 3x3 pixel neighborhood. This type of assessment incorporates optimistic bias (Hammond and Verbyla, 1996) because sampling is limited to homogenous areas where land cover is more easily identified by the examiner. The optimistic assessment yielded a higher overall accuracy of 94% and a strong overall Kappa of 92% (Table 4). Kappa values per category were all greater than 80%. However, the restriction of sampling to homogenous areas reduced the

Table 4: Statistical summary of optimistic accuracy assessment

Category	User's Accuracy	User's 95% CI	Kappa Map Referent	Map Samples	Producer's Accuracy	Producer's 95% CI	Kappa Ground Referent	Ground Samples
Low Intensity Developed	93%	82.2% - 100%	93%	29	93%	82.2% - 100%	93%	29
High Intensity Developed	90%	76.8% - 100%	89%	29	93%	81.5% - 100%	92%	28
Agriculture	92%	83.5% - 100%	91%	50	82%	71.2% - 93.1%	80%	56
Grassland	85%	77.7% - 91.6%	80%	117	93%	88.2% - 98.6%	91%	106
Woodland	97%	93.2% - 100%	97%	77	91%	84.8% - 98.1%	90%	82
Open Water	100%	99.5% - 100%	100%	108	100%	99.5% - 100%	100%	108
Woody Wetland	100%	75.0% - 100%	100%	2	100%	75.0% - 100%	100%	2
Wetland	97%	89.3% - 100%	97%	32	97%	89.3% - 100%	97%	32
Bare or Transitional	97%	89.6% - 100%	97%	33	94%	84.7% - 100%	94%	34

Overall Accuracy = 94% (95% CI: 91.2% - 95.8%) Total Observations = 477 Overall Kappa = 92%

total number of sample points from 961 to 477, and also reduced the number of sample points for the low intensity and high intensity classes to n < 30. The lower sample size does not allow rigorous statistical inferences to be made for these two classes. However, comparison of overall accuracy of the optimistic assessment with overall accuracy of the conservative assessment indicates that the classification had greater success in classifying relatively homogenous areas of land cover and more difficulty in classifying transitional zones or areas that exhibit actual heterogeneity in land cover. This confirms visual observations made when comparing the classification to aerial photography. Many of the observed classification errors occurred in these transitional areas between land cover types, while areas of homogenous land cover appeared to be accurately classified.

7.3.3 Mode Filtered Accuracy Assessment

Mode filtering of the classified image increased both the overall accuracy (77%) and the overall Kappa (73%) several percentage points when compared to the unprocessed classification (Table 5). Per category user and producer accuracies were also improved with the exception of the high intensity developed class. A two-tailed hypothesis test at

Table 5: Statistical summary of mode filtered accuracy assessment

Category	User's Accuracy	User's 95% CI	Kappa Map Referent	Map Samples	Producer's Accuracy	Producer's 95% CI	Kappa Ground Referent	Ground Samples
Low Intensity Developed	76%	67.9% - 84.2%	72%	117	70%	61.2% - 77.9%	65%	128
High Intensity Developed	59%	48.9% - 69.1%	55%	100	77%	66.5% - 86.7%	74%	77
Agriculture	85%	76.5% - 92.9%	83%	85	79%	70.2% - 88.0%	77%	91
Grassland	69%	63.0% - 74.8%	60%	251	79%	73.4% - 84.6%	72%	219
Woodland	84%	78.1% - 90.6%	81%	147	70%	63.4% - 77.5%	65%	176
Open Water	94%	89.5% - 98.8%	93%	119	90%	83.8% - 95.4%	88%	125
Woody Wetland	44%	6.4% - 82.5%	44%	9	50%	9.1% - 90.9%	50%	8
Wetland	79%	68.7% - 89.1%	77%	71	78%	67.5% - 88.1%	76%	72
Bare or Transitional	85%	75.9% - 95.1%	84%	62	82%	71.3% - 91.7%	80%	65

Overall Accuracy = 77% (95% CI: 74.5% - 79.9%) Total Observations = 961 Overall Kappa = 73%

the 95% confidence level was conducted to determine whether overall accuracy of the mode-filtered image was significantly different than the accuracy of the unprocessed classified image. The test statistic did not exceed the critical *t* value of 1.96, so at this confidence level and sample size there is not a significant difference between the two overall accuracy results. However, mode filtering does appear to have positive rather than negative effects on classification accuracy and may be utilized for applications, which require some level of generalization of the classified map. Filtering with a 5x5 mode filter had negative effects on classification accuracy, decreasing per category and overall accuracy.

7.3.4 Merged Class Accuracy Assessment

Examination of the previous accuracy assessments indicated that classification errors could be attributed to confusion between individual land cover classes, particularly low intensity and high intensity developed, and cultivated and grassland. These classes and the wetland classes were merged to determine the effect on classification accuracy. As shown in Table 6, overall accuracy increased to 82% and the overall Kappa increased to 77% when compared to the unprocessed classification map. This analysis confirmed that the primary source of confusion in the classification is due to interpreter error in

Table 6: Statistical summary of merged classes accuracy assessment

Category	User's Accuracy	User's 95% CI	Kappa Map Referent	Map Samples	Producer's Accuracy	Producer's 95% CI	Kappa Ground Referent	Ground Samples
Developed	82%	76.5% - 87.3%	77%	210	84%	78.6% - 89.2%	79%	205
Grass/Agriculture	80%	75.5% - 84.4%	70%	334	86%	82.1% - 90.1%	79%	310
Woodland	82%	75.3% - 88.8%	78%	139	65%	57.4% - 72.1%	59%	176
Open Water	91%	85.8% - 96.6%	90%	125	91%	85.8% - 96.6%	90%	125
Wetlands	76%	65.9% - 85.2%	73%	86	81%	72.1% - 90.4%	79%	80
Bare or Transitional	79%	68.6% - 89.6%	78%	67	82%	71.3% - 91.7%	80%	65

Overall Accuracy = 82% (95% CI: 79.2% - 84.2%) Total Observations = 961 Overall Kappa = 77%

assigning clusters to the low intensity and high intensity developed classes, and spectral mixing of the grassland and agriculture classes. If a map with higher accuracy is needed for analysis, the merged class map may offer a better alternative than the unprocessed classification.

7.4 Comprehensive Kappa Analysis

One of limitations of the calculation of overall accuracy and the Kappa index is that these two measures do not provide information on spatial proximity and fail to distinguish between quantification error and location error (Pontius, 2000). Quantification error results when the quantity of pixels for a land cover class in a reference map differs from the classified map. Location error occurs when location of pixels for a land cover class in a reference map differ from the classified map. An error of location is disagreement, which can be corrected by moving the location of pixels in the classified map to improve the agreement with the reference map. The VALIDATE module in IDRISI was used to calculate more comprehensive Kappa statistics that distinguish between quantification and location error. The more comprehensive analysis enables one to break down the sources of classification success and error, and evaluate the ability of the classification process to spatially allocate individual land cover classes. Further explanation of the calculation of these statistics can be found in Pontius (2000).

Results of the analysis provided an overall Klocation of 76%, an overall Kquantity of 75%, and an overall Value of Perfect Information (VPIL) of 19%. Klocation and Kquantity values indicate that the classification was 76% better than random in specifying location of land cover in the classified map given the specified quantities, and the classified map was 75% better than random in specifying quantities of the land cover classes given the specified locations of the sample data. The overall VPIL indicates that the overall classification accuracy could be improved by 19% if the classification had perfect information on location of the land cover classes, given no change in quantity. Table 7 provides the Klocation and VPIL value for each of the land cover classes. The

Table 7: Per category results of KLocation and VPIL calculation

Category	RefPerc	SimPerc	Klocation	VPIL
Low Intensity Developed	13.3%	12.2%	68.5%	3.3%
High Intensity Developed	8.0%	9.7%	72.7%	2.0%
Agriculture	9.5%	8.4%	82.3%	1.4%
Grassland	22.8%	26.3%	69.0%	5.2%
Woodland	18.3%	14.5%	78.0%	2.6%
Open Water	13.0%	13.0%	89.9%	1.1%
Woody Wetland	0.8%	1.4%	49.3%	0.4%
Wetland	7.5%	7.6%	74.5%	1.8%
Bare/Transitional	6.8%	7.0%	80.2%	1.3%
Total	**100%**	**100%**		**19.0%**

Note: RefPerc: percentage of the ground reference sample.
SimPerc: percentage of the classified map sample.

most important information extracted from this analysis is that we can improve the accuracy of the classification significantly (i.e., in theory 19%) if we can improve the classification's ability to specify location. The lower Klocation values of grassland (69%) and low intensity developed (69%) relative to the other classes indicates that 9% (sum of VPIL for these two classes) of our classification error can be associated with spatial allocation of these two classes. Figure 4 shows the sources of classification error and success for the classification and illustrates that by correcting for the error due to location we can significantly improve classification accuracy. Improving the specification of location for the developed, grassland, cultivated and woodland classes can improve the classification accuracy 14% (sum of VPIL for these classes is 14.5%). Any efforts to improve the classification should focus on these five land cover classes.

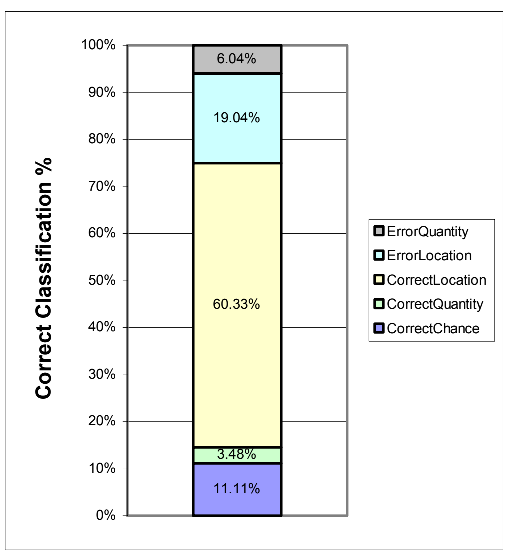

Figure 4: Sources of classification error and successes assuming fixed specification of quantity

8. Summary and Recommendations

Interpretation of the accuracy assessment results, indicate that the unprocessed pixel classification achieved moderate to good agreement with our ground truth data. Although the accuracy of the unprocessed image was lower than desired (i.e., it would have been nice to have Kappa and overall accuracy of .8 or greater), comparison of the optimistic and conservative accuracy assessments indicated that the classification is highly accurate (i.e., Kappa > .8 for all classes) in specifying areas of homogenous land cover, but this accuracy decreases when moving to transitional zones between land cover types and other areas that exhibit heterogeneity of land cover types. All classification efforts can expect classification errors in these land cover transitional zones, given that there is no defined physical boundary between land cover types in actuality. However, this disparity in accuracy should be kept in mind when utilizing the data. Filtering of the classification with a 3x3 mode filter had positive effects on the classification and slightly improved the overall classification accuracy and per category accuracy for most of the land cover classes. The filtered map will be useful for applications, which require generalization of the map data or when a more visually appealing map is needed for display purposes. If a map with higher accuracy is required it may be beneficial to aggregate the grassland and agriculture classes, and the developed classes. Results of the accuracy assessment of merged classes significantly improved overall accuracy and the overall Kappa for the classification. Comprehensive analysis of Kappa statistics indicated that the classification accuracy can be significantly improved by focusing on the ability of the classification process to specify location of the land cover classes, particularly the grassland, cultivated, woodland and developed classes.

The accuracy of the low intensity and high intensity developed classes may be improved through the more diligent assignment of land cover labels during the cluster busting process. It may be possible to improve the accuracy of these two categories by isolating the developed areas and reclassifying them using the iterative cluster busting process. Improving the accuracy of the cultivated class would require that training data and signatures for individual crop types be developed from additional temporal or hyperspectral satellite imagery. Selection of temporal satellite imagery should correspond to seasonal changes in individual crop phenologies. Using a supervised classification algorithm with spectral signatures extracted from the hyperspectral and/or additional temporal satellite imagery will be more effective in identifying cultivated land. Developing training data for woodland transition zones and scrub shrub areas, and inputting these signatures into a supervised classification algorithm may also reduce woodland omission. Because of spectral mixing with all of the before-mentioned classes, accuracy of the grassland class should improve if these modifications are implemented. Finally, the 2002 aerial photography for 6 of the study area counties (i.e., Waller, Fort Bend, Montgomery, Harris, Liberty, and Chambers) can be used to identify and recode additional classification errors.

This accuracy assessment of the land cover data set is considered to be preliminary. An effort should be made to complete the remaining sample points, especially for the woody wetland category so some statistical inference can be made regarding its accuracy. Increasing the sample size will also ensure that the stratified random sampling scheme is maintained and allow more robust statistical inferences to be made regarding overall accuracy and per category accuracies.

9. Appropriate Use of the Data

It is important for individual users of the data to determine what is an appropriate use of the H-GAC land cover data set. Users of the data are encouraged to review and comprehend the image processing and accuracy assessments protocols used in the creation of the data set in order to determine if the data would be suitable for their needs. The data set is not intended to serve as the primary tool for regulatory or jurisdictional decision-making. Regulatory applications using the land cover data should involve rigorous field verification before any decisions or conclusions are made. Specifically, the data set was created for broad-scale planning and research applications at the county and regional level. Some general examples of appropriate and inappropriate uses would include:

Appropriate Uses

- Regional and county planning.
- Large area resource management planning.
- Educational purposes for students and citizens.
- Regional or county level water quality and watershed analysis.
- Basic research on county or regional distribution of land cover to determine specific areas for monitoring or management focus.
- Broad-scale evaluation of the environmental impact or benefits of a major project.
- Change detection and time series analysis.

Inappropriate Uses

- Determining the accuracy of other data using the H-GAC data set.
- Determining the location of jurisdictional wetlands.
- Determining exact area coverage of land cover without consideration of the overall accuracy and per category accuracy of the data.
- Establishing exact boundaries for regulatory enforcement.
- Mapping areas finer than the original resolution of the data.
- Combining or altering the data set and redistributing them.
- Establishing definite occurrence or nonoccurrence of a feature without consideration of probabilities determined by the accuracy assessments.

References

Congalton, R.G. 1996. Accuracy Assessment: A Critical Component of Land Cover Mapping. Gap Analysis. ISBN-1-57083-03603 *American Society for Photogrammetry and Remote Sensing.* 1996. p. 119 – 131.

Congalton, R. G. (1991) A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment*. 37 pp 35-46.

Dobson, J.E., Bright, E.A., Ferguson, R.L., Field, D.W., Wood, L.L., Haddad, K.D., Ireland III, H., Jensen, J.R., Klemas, V.V., Orth, R.J. and Thomas, J.P. (1995) *NOAA* Coastal change analysis program (C-CAP): guidance for regional implementation, National Marine Fisheries Service, Seattle, Washington, NOAA Technical Report NMFS 123

Hammond, T.O. and D.L. Verbyla. 1996. Optimistic bias in classification accuracy assessment. *International Journal of Remote Sensing*, 17: 1261-1266.

Jensen, J.R., Ramsey, E.W., Mackey, H.E., Christensen, E.J., and Sharitz, R.R. 1987. Inland Wetland Change Detection Using Aircraft MSS Data. *Photogrammetric Engineering and Remote Sensing* 53(5): 521-529.

Lillesand, T., J. Chipman, D. Nagel, H. Reese, M. Bobo, and R. Goldmann.1998. Upper Midwest GAP analysis program image processing protocol. Report prepared for USGS, Environmental Management Technical Center, Onalaska, WI, June, 1998. EMTC 98-G001. 25pp. + Appendixes A-C.

Pontius Jr, R.G., 2000. "Quantification error versus location error in comparison of categorical maps." Photogrammetric Engineering and Remote Sensing. 66(8) 1011-1016.

Verbyla, D.L. and T.O. Hammond. 1995. Conservative bias in classification accuracy assessment due to pixel-by-pixel comparison of classified images with reference grids. *International Journal of Remote Sensing*, 16: 581-587.

Appendix A. Contingency Matrix Results

Table A. 1 Contingency matrix for unprocessed classification (conservative assessment)

Lab\Field	Low Intensity	High Intensity	Agriculture	Grass	Woody	Water	Woody Wet	Wetland	Bare/Trans	**Totals**	**Commission**
Low Intensity	**85**	11	1	10	9			1		**117**	**73%**
High Intensity	18	**58**		7	3	1		1	5	**93**	**62%**
Agriculture		1	**68**	8	2	1			1	**81**	**84%**
Grass	15	4	22	**169**	35	1		5	2	**253**	**67%**
Woody	9	1		12	**114**		1	2		**139**	**82%**
Water				2	2	**114**	1	3	3	**125**	**91%**
Woody Wet				1	3	1	**4**	4		**13**	**31%**
Wetland	1			4	3	7	2	**55**	1	**73**	**75%**
Bare/Trans		2		6	5			1	**53**	**67**	**79%**
Totals	**128**	**77**	**91**	**219**	**176**	**125**	**8**	**72**	**65**	**74.9%**	
Omission	**66%**	**75%**	**75%**	**77%**	**65%**	**91%**	**50%**	**76%**	**82%**		

KAPPA = 70.6% **Total Observations 961**

Table A. 2 Contingency matrix for optimistic assessment (limited to 3x3 homogenous areas)

Lab\Field	Low Intensity	High Intensity	Agriculture	Grass	Woody	Water	Woody Wet	Wetland	Bare\Trans	**Totals**	**Commission**
Low Intensity	**27**	2								**29**	**93%**
High Intensity		**26**			1				2	**29**	**90%**
Agriculture			**46**	4						**50**	**92%**
Grass	2		10	**99**	5			1		**117**	**85%**
Woody				2	**75**					**77**	**97%**
Water						**108**				**108**	**100%**
Woody Wet							**2**			**2**	**100%**
Wetland				1				**31**		**32**	**97%**
Bare\Trans					1				**32**	**33**	**97%**
Totals	**29**	**28**	**56**	**106**	**82**	**108**	**2**	**32**	**34**	**93.5%**	
Omission	**93%**	**93%**	**82%**	**93%**	**91%**	**100%**	**100%**	**97%**	**94%**		

KAPPA = 92.2% **Total Observations 477**

Table A. 3 Contingency matrix for 3x3 mode filtered classification

Lab\Field	Low Intensity	High Intensity	Agriculture	Grass	Woody	Water	Woody Wet	Wetland	Bare\Trans	Totals	Commission
Low Intensity	**89**	11	1	8	7			1		**117**	**76%**
High Intensity	20	**59**		10	3	2		1	5	**100**	**59%**
Agriculture		1	**72**	7	3	1			1	**85**	**85%**
Grass	14	3	18	**173**	33	1		7	2	**251**	**69%**
Woody	5	1		13	**124**		1	2	1	**147**	**84%**
Water				2		**112**	1	2	2	**119**	**94%**
Woody Wet					1	1	**4**	3		**9**	**44%**
Wetland				4		8	2	**56**	1	**71**	**79%**
Bare\Trans		2		2	5				**53**	**62**	**85%**
Totals	**128**	**77**	**91**	**219**	**176**	**125**	**8**	**72**	**65**	**77.2%**	
Omission	**70%**	**77%**	**79%**	**79%**	**70%**	**90%**	**50%**	**78%**	**82%**		

KAPPA = 73.3% **Total Observations 961**

Table A. 4 Contingency matrix for merged classes

Lab\Field	Developed	Grass/Ag	Woody	Water	Wetland	Bare/Trans	Totals	Commission
Developed	**172**	18	12	1	2	5	**210**	**82%**
Grass/Ag	20	**267**	37	2	5	3	**334**	**80%**
Woody	10	12	**114**		3		**139**	**82%**
Water		2	2	**114**	4	3	**125**	**91%**
Wetland	1	5	6	8	**65**	1	**86**	**76%**
Bare/Trans	2	6	5		1	**53**	**67**	**79%**
Totals	**205**	**310**	**176**	**125**	**80**	**65**	**81.7%**	
Omission	**84%**	**86%**	**65%**	**91%**	**81%**	**82%**		

KAPPA = 76.7% **Total Observations 961**

Appendix B. Final Classification Images

Figure B.1 Unprocessed final classification map

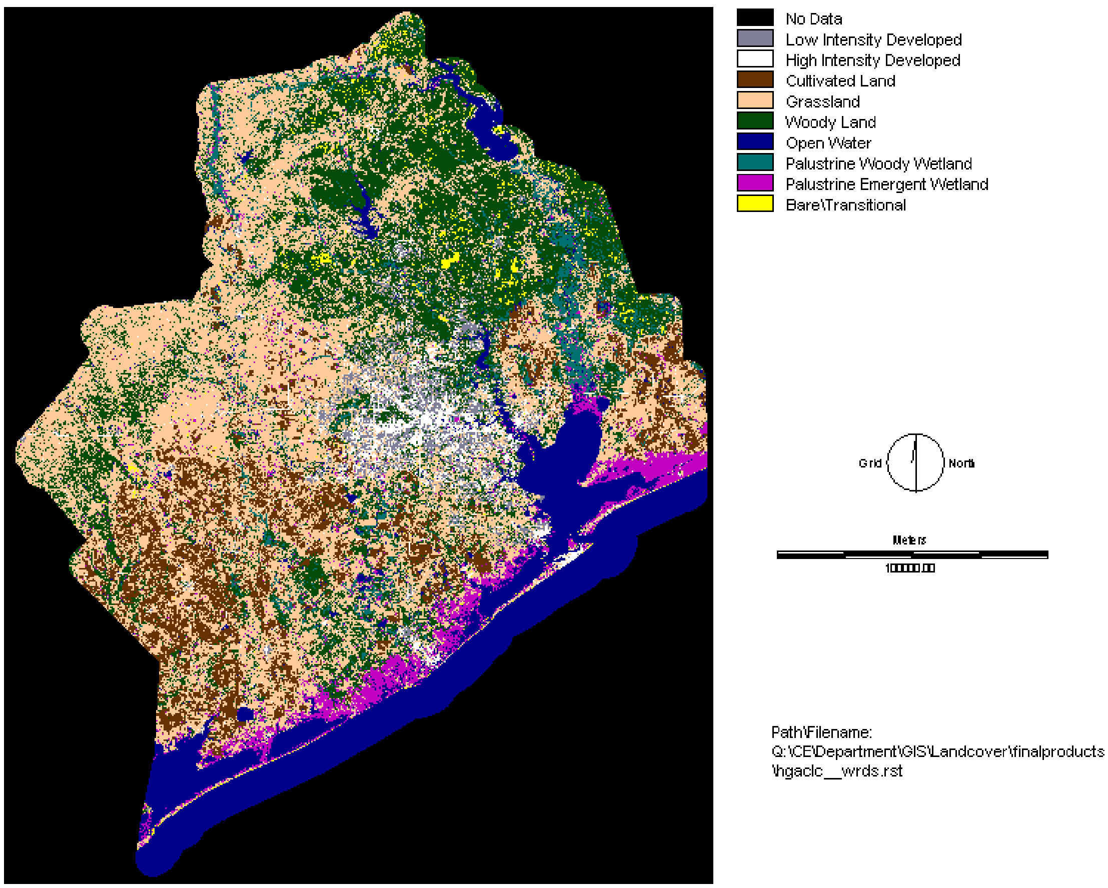

Figure B.2: 3x3 mode-filtered classification map

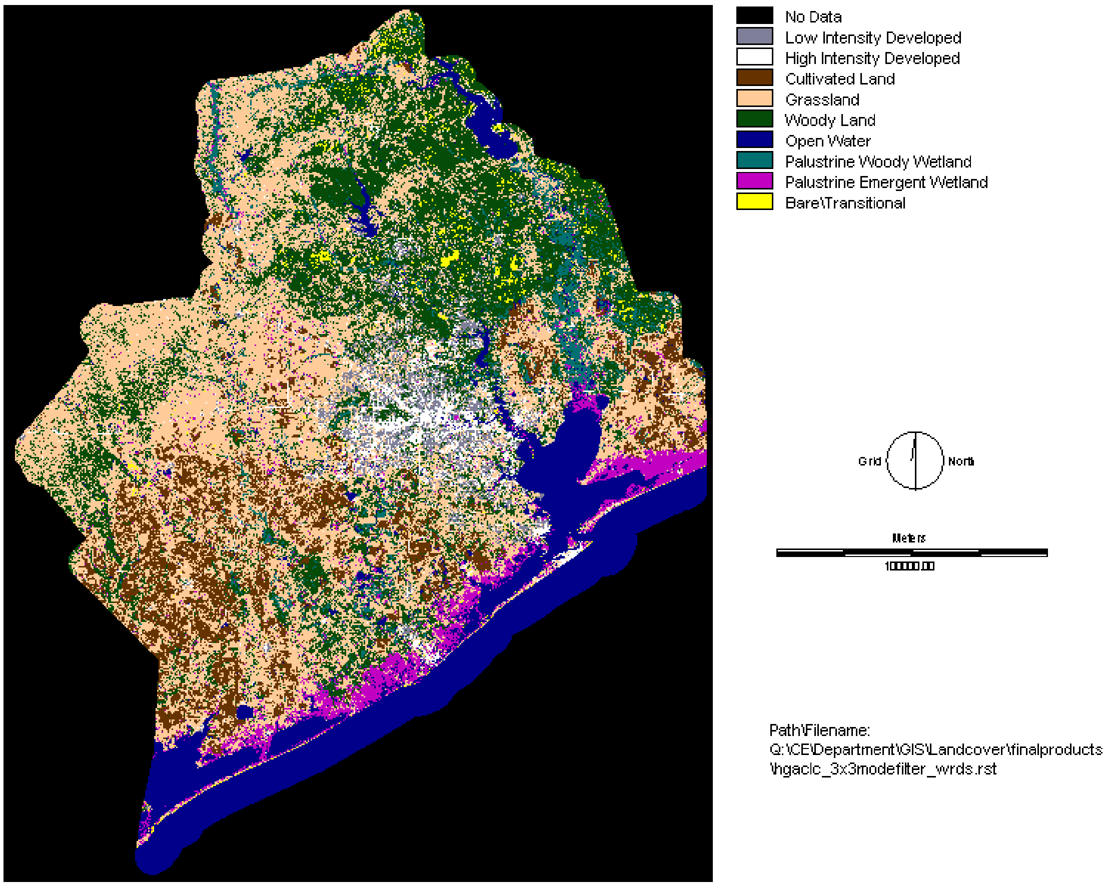

Figure B.3 Final classified map with merged classes

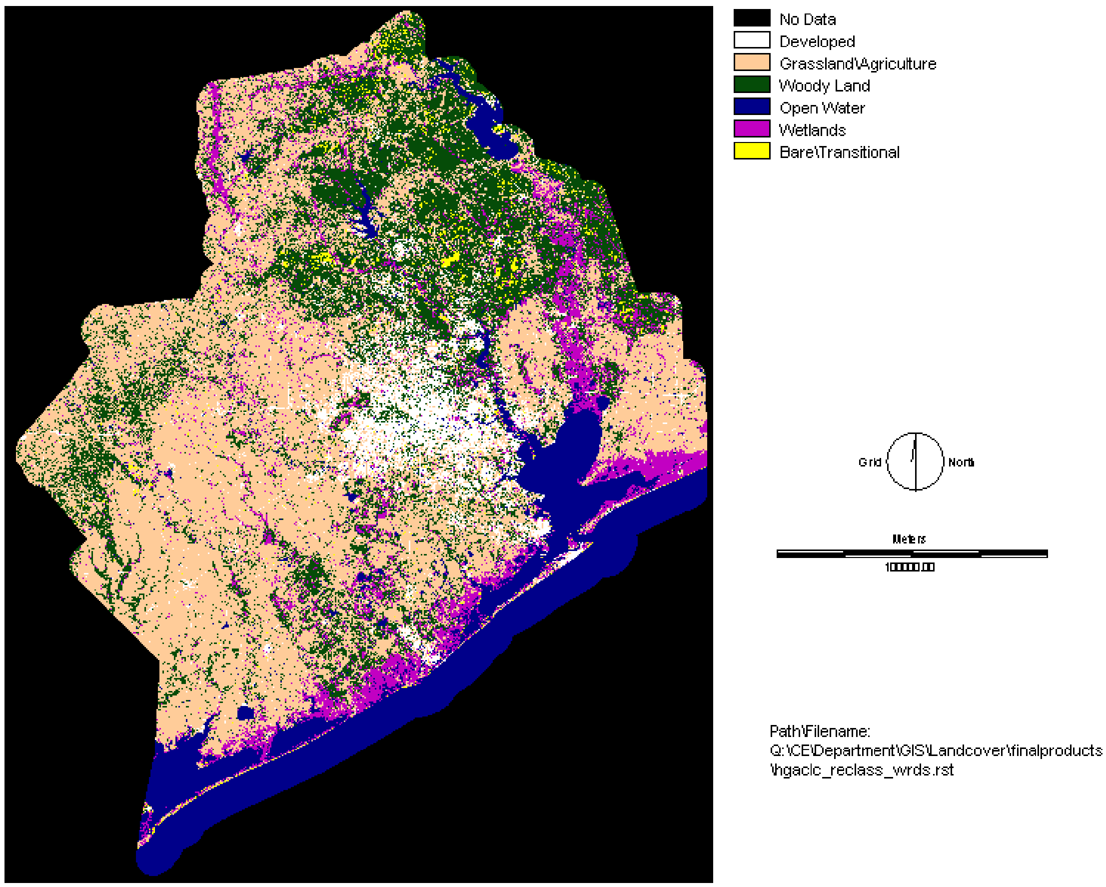

Figure B.4: Classified image with wetlands divided into salinity regimes using NWI data