2001

Measuring open space change : an analysis of change detection techniques

Keane T. Grivich
San Jose State University

MEASURING OPEN SPACE CHANGE: AN ANALYSIS OF CHANGE DETECTION TECHNIQUES

A Thesis
Presented to
The Faculty of the Department of Geography
San Jose State University

In Partial Fulfillment
of the Requirements for the Degree
Master of Arts

By
Keane T. Grivich
December, 2001

UMI Number: 1407291

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

APPROVED FOR THE DEPARTMENT OF GEOGRAPHY

Richard Taketa

Dr. Richard Taketa

David Schwarz

Dr. David Schwarz

Cindy Schmidt

Cindy Schmidt, M.A., CSU Monterey Bay

APPROVED FOR THE UNIVERSITY

ABSTRACT

MEASURING OPEN SPACE CHANGE: AN ANALYSIS OF CHANGE DETECTION TECHNIQUES

By Keane T. Grivich

This thesis compared two popular methods of change detection using image processing. The methods of change detection evaluated were *post classification change detection* and *image differencing change detection*. The primary evaluation criteria for the methods were accuracy and cost effectiveness. Imagery used for the analysis was Landsat Thematic Mapper data.

The study demonstrated that the image differencing change method could more accurately find change than the traditional post classification method. The overall accuracy for the image differencing method was 94 percent versus 84 percent for the post classification change method. This project did not find any meaningful cost difference between image differencing and post classification change detection. The image differencing method required more effort at the beginning of the change detection process but required much less time during the classification phase because the area to be classified was reduced through the differencing process.

ACKNOWLEDGMENTS

A special note of thanks to Professor Taketa, Cindy Schmidt and Professor Schwartz for their encouragement and for enduring my early drafts. Also, I would like to thank Christine whose support enabled me to complete this project. And finally, to my parents, thank you for encouraging my pursuit of more education.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF PLATES

Chapter 1

INTRODUCTION, STATEMENT OF PROBLEM AND GEOGRAPHY OF STUDY AREA

Digital remote sensing has given rise to many new ways to study our environment. One such topic of study is monitoring change in our environment. Measuring land use change is one of the most common applications of change monitoring. Measuring and inventorying change is important because it helps us understand our interaction with the environment.

Change detection is the process of measuring change on the ground using two or more dates of satellite imagery. Older methods to monitor change such as air photo interpretation and ground mapping are expensive and usually require long periods of time to complete. Monitoring change using digital satellite imagery can be accomplished faster and with the same accuracy as methods using air photographs because computer processing provides more consistent results and utilizes fewer people hours per project. Thus, change monitoring can be readily available to land managers and policy makers in a timely and cost effective manner (Green and Weinstein, 1996). Many different methods exist to monitor change using digital imagery. Deciding which method to use can be a difficult task. Which change detection methods are faster or more accurate?

Introduction to Change Detection

Most maps of change are made with one of two basic methods. The traditional method is *post classification change detection* or sometimes called "map-to-map" change detection. This type of change detection first classifies data from two different dates and then change data is extracted. The two classified images are overlaid and areas of change or no change are noted. An alternative change detection method is to process raw digital imagery first, finding spectral change areas, then only changed areas are classified.

"Image differencing" or "image to image" change comparison first identifies change areas and then performs a classification only on areas identified as change. Multispectral images reflect to some degree what is on the ground depending on the sensor, spatial resolution and atmospheric conditions. Thus, if multispectral images from two dates are compared, land cover differences can be identified and placed into a map of spectral change.

Image to image comparisons can overcome deficiencies of a map-to-map or post classification comparisons. Differences in classification systems and mapping techniques do not pose problems with the image differencing method. Inaccuracies introduced by differences in classification systems and mapping techniques do not affect image differencing because the change areas have already been extracted from the data. Conversely, post classification change detection relies on image classification to identify change areas. Accurate image classification is still an important part of the image differencing method, but image classification is not the basis for identifying change areas.

Statement of Problem

The primary objectives for this thesis are:

1. Which is more accurate: post classification change detection or image differencing change detection?
2. Which is more cost effective: post classification change detection or image differencing change detection?

The main steps of the thesis are:

1. Use two change detection methods to measure change from open space to developed land during the 10 year span from 1986 to 1996 in the Diablo Valley and surrounding hills.
2. Measure difference in accuracy and cost effectiveness between post classification change and image differencing methods.

Both post classification and image differencing change detection will yield a number of open space acres changed to other uses. Will the two methods be different in numbers of acres changed? Will the time to complete each change detection be different?

Evaluation Criteria

The core questions to be answered are which method will provide the more accurate measure of land use change and which method is more cost effective. To answer the questions, both methods have to be evaluated with the same criteria. Accuracy of each change method will be measured using a random sampling accuracy assessment technique. The accuracy assessment will provide the data necessary to evaluate which method was more accurate.

The two largest costs of change detection projects are labor and data. The cost of data is the same for both change detection methods: the same Landsat images will be used. Since the cost of data is the same, labor will determine which method is more cost effective. Labor costs are a function of time or effort expended on a project. Careful track will be kept of the time spent on each method, so that a cost effectiveness value can be placed on each change detection method.

Geography of Study Area

The Diablo Valley was chosen as the study area because of the urban growth during the last several decades. Central Contra Costa County, home to the Diablo Valley, has experienced a large increase in urban growth during the 1970's continuing through today (Figure 1.1). From 1970 to 1980, Contra Costa's population increased 18 percent (Hornbeck, 1983). The increase was even greater between 1980 and 1990, growing at a rate of 22 percent (U.S. Census Website). This trend appears to be continuing, with Contra Costa expecting more double-digit population growth in the 1990's and beyond. Urban development means that some other kind of land use is being changed or developed. In the case of the Diablo Valley, open spaces are being changed into housing developments.

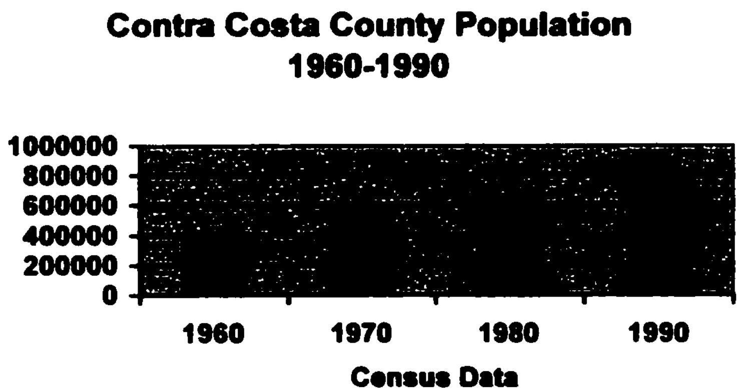

Figure 1.1

Contra Costa County is home to large areas of open space, which primarily consists of oak woodland and grasslands. Much of the development is suburban, occurring in areas previously occupied by open space. Open space provides important habitat to many species of birds, insects and mammals, some of which are now threatened.

A new threat to the ecology of the oak woodland and Contra Costa Counties open space is the increase of suburban development in outlying areas near existing development. These areas are sometimes referred to as interface zones. Suburban developments are springing up in areas that were previously oak woodland. When new housing is constructed, the land is graded with heavy machinery and most vegetation is removed, including trees.

Many local governments are enacting laws to protect open space and oak woodland. The City of Walnut Creek, located near the center of the Diablo Valley, has city ordinances to preserve oaks and woodland. Under Walnut Creek's ordinance, "A

tree can be removed only if the burden of saving it out weighs its public benefit."
("Valley Oaks Test Protection Policy," 2000).

Geomorphology

The study area includes the Diablo Valley and surrounding hills. This study area is 50 square miles, with a center point at latitude 37° 56' 29 N" and longitude at 121° 59' 09 W" (see plate 1). Diablo Valley is located near the center of Contra Costa County. Mt. Diablo dominates the geography of the valley at 3850 feet and is the valley's southern boundary. The valley floor averages less than 50 feet in elevation. Two ridges come out from the center of Mt. Diablo like arms surrounding the Diablo Valley. Lime Ridge is Diablo Valley's eastern boarder and Briones Park is the western border. The northern end of the valley opens into the Delta.

Vegetation

Dense stands of oak trees can easily be misinterpreted as residential neighborhoods during the image classification part of change detection. Oak woodland, shrub and grassland dominate open spaces in the Diablo Valley and surrounding hills. Oaks are more dominant on the north slopes while grasses and shrubs are dominant on the hotter and dryer south facing slopes. The two most common species of oak trees in the area are evergreen coast live oak (Quercus agrifolia) and the deciduous blue oak (Quercus douglasii).

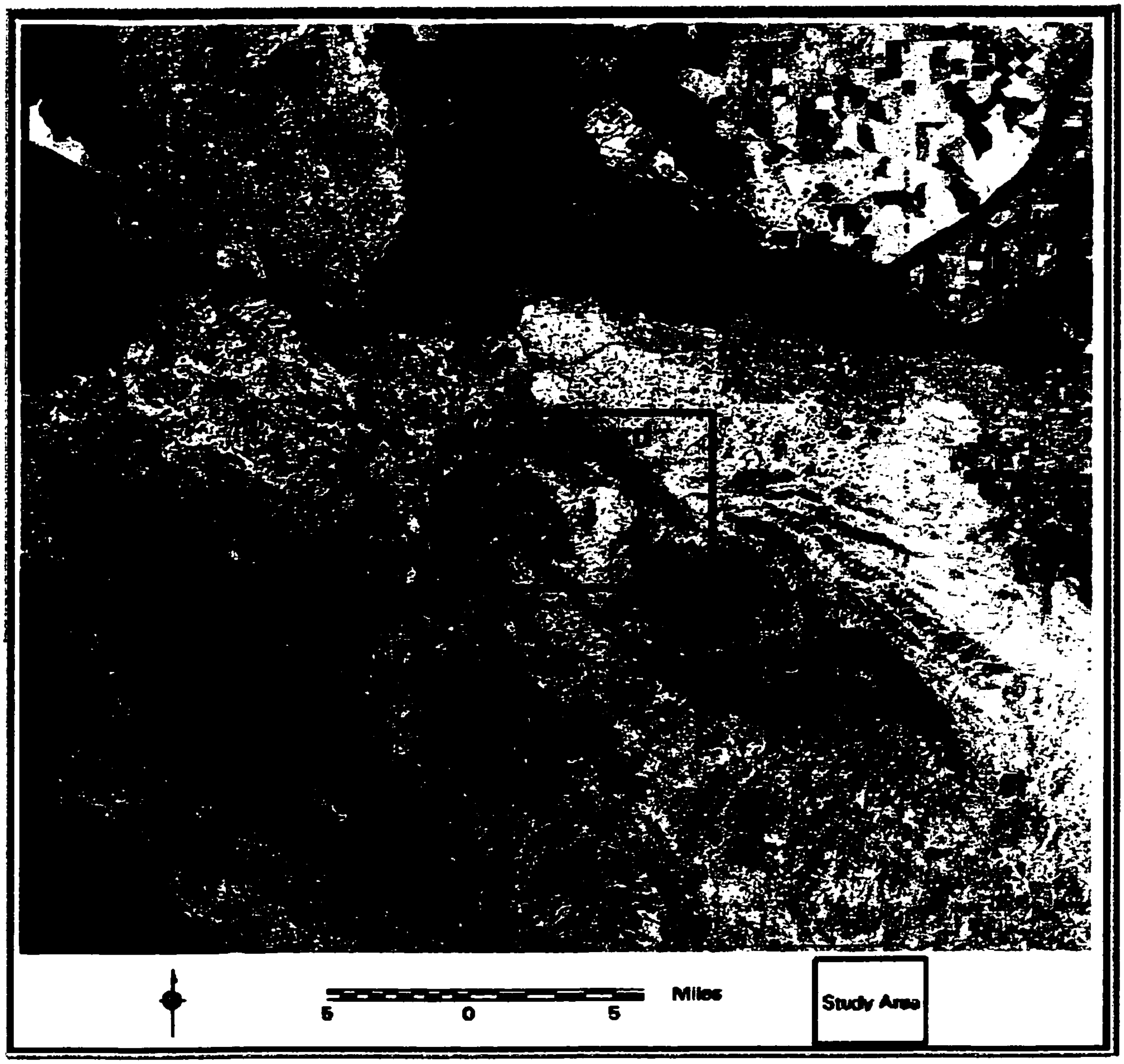

Plate 1 (Map depicting the study area)

Summary

The primary goal of this project is to compare the effectiveness of image differencing and post classification. Accuracy and cost effectiveness are the criteria for judging which change method is better. The Diablo Valley is a good location to test change detection methods because many land use changes occurred from 1986 to 1996. Application of post classification and image differencing to images of the Diablo Valley will provide a basis for the evaluation.

Chapter 2

CHANGE DETECTION METHODOLOGIES

Change Detection Overview

Choosing a change detection algorithm is one of the first tasks to be performed when starting a change detection project. Many different types of change detection methodologies exist. Listed below are some of the most widely used:

- Multi-date Composite Image Change Detection
- Manual On-screen Digitization of Change
- Post classification Change Detection
- Image Differencing Change Detection

Selection of the most suitable change detection algorithm is important (Jensen, 1996). Determining which change detection method to use depends on factors such as the type of data used, output needed, time restraints and budget.

Multi-Date composite image change detection uses bands from two images and combines them into one image to be classified. The classification result is a data set containing clusters of change and non-change pixels. Each cluster must then be labeled with the type of change (Jensen, 1996). This method can speed up the classification process because only one classification is needed. However, it is more difficult to extract *from-to* change data from the classification (Fung and LeDrew, 1987).

Manual on-screen digitization of change is the process of delineating change areas by hand, with the aid of a computer. The manual on-screen digitization method requires two dates of imagery to be displayed side by side on a computer screen. The two images are linked so the cursor can be seen in both images. Polygons are traced around change areas and the change information for each polygon is stored. One of the biggest advantages of manual on-screen digitization is that the output polygons will already be interpreted for change type. The length of time to produce manual on-screen change maps is the main limitation of manual on-screen digitization.

Post Classification Change Detection is the most commonly used change method (Jensen, 1996). The post classification method uses two rectified images for which land cover classes have already been identified. The classified images are overlaid and compared using a GIS algorithm. The result is a change map containing from-to change information, which is significant because the change information has already been extracted from the data. There are some drawbacks to this method of change detection. Non-change differences are sometimes incorrectly measured as change. Measuring non-change differences as change is the result of using different classification systems, mapping techniques, map accuracy, and registration systems. All these factors can introduce potential error into the final change map (Green and Weinstein, 1996).

Image differencing or image algebra change detection works by selecting a band or a band ratio from two different dates of imagery and subtracting the values of one image from the values of the other. The result is a change map containing positive and negative values. Values near zero indicate little or no change and values far from zero indicate change. The main drawback of using image differencing is that no from-to data is collected. After the change areas are discovered classifications need to be performed on the change areas.

Other change detection methods exist besides the ones discussed here. Many of these methods are variations on the types discussed earlier. For example, some types of change detection substitute a digitized map for one of the images. For dates that preceded satellites or for areas that do not have satellite image coverage, this often has to be done. Many different methods for change detection exist; this thesis will focus on image differencing and post classification change detection because they are two of the most widely used methods.

Image Differencing Methodology

As stated before, image differencing works by selecting a band or a band ratio from two different dates of imagery and subtracting the values of one image from the values of the other (see figure 2.1). In figure 2.1, change values of -2, -2, and -3 represent little change, but a change value of 50 may indicate significant change. The result of the image differencing is a map that contains positive and negative values. Pixels in the output map that have values near zero probably have not changed. In contrast, pixel values far from zero (negative or positive) denote pixels that are very different from one date of imagery to the other. For example, a grassland pixel in the 1986 image might reflect a value of 140 in band 4. In the 1996 image the pixel has been transformed into a parking lot. This pixel will now have a band 4 or infrared reflectance of 90 because of the removal of vegetation. Subtract 140 from 90 and that is a difference of 50; this represents a significant change from 1986 to 1996.

$$\begin{array}{|c|c|}\hline 140 & 140 \\ \hline 140 & 140 \\ \hline \end{array} \; - \; \begin{array}{|c|c|}\hline 138 & 138 \\ \hline 137 & 90 \\ \hline \end{array} \; = \; \begin{array}{|c|c|}\hline -2 & -2 \\ \hline -3 & -50 \\ \hline \end{array}$$

1986 Image Values **1996 Image Values** **Change Values**

Figure 2.1 (Image Differencing Process)

The output of image differencing is a map that contains areas of change and no change. The map does not contain information on what kind of change occurred. After the change map is produced, an image classification must be run on the 1st and 2nd dates of imagery. Usually, only areas that were identified as change are classified. Classifying only change areas can significantly reduce the number of pixels processed. The last step is to compare the two classified change maps digitally. Information describing what land use types have changed and what they have changed into can now be extracted. Extracting land use changes from the change areas is a similar process to the post classification change detection method.

Post Classification Change Detection Methodology

Post classification requires image classification of the entire study area of both scenes. Image classification is a way to transform the raw digital images into useful information (e.g., vegetation map or geologic map). The first step in classifying the image is to define a set of rules to follow during the classification process. Classification rules will include information like minimum mapping unit and definitions for land use classes. Detailed information about the classification rules for this project is discussed in chapter 3. The classification rules or methodologies should be the same for both dates of imagery. After the two dates of imagery are classified, the result is two thematic images with the same classes but from different dates. Next, the classified maps are combined digitally. Data values from the combined map depict areas of change or no change. Information describing what land use types have changed and what they have changed into has now been extracted.

Many different types of change detection methods exist. Image differencing and post classification change detection were chosen for analysis because they are two of the most common methods. Image differencing and post image classification are also different enough to provide for meaningful comparison. Close examination of the image differencing and post classification change methods will determine which is more accurate and cost effective.

Chapter 3

IMAGE PROCESSING AND CREATING CHANGE MAPS

Image Processing Overview

Many processing steps for image differencing and post classification change detection are the same. Image differencing and post classification change detection only differ in one major aspect: the manner by which change data is extracted. Post classification change detection first classifies data and then extracts change information. Image differencing first identifies change areas and then performs a classification only on areas identified as change.

Once change areas are identified, the image differencing method follows the same processing steps as the post classification change method. The one unique image differencing step is the creation of a difference image. The difference image contains only change areas, thus substantially reducing the area to be classified. Table 3.1 lists the principal steps for each change method.

Post classification change	**Image differencing**
Image Preparation	Image Preparation
----------------------------------	Create Difference Image
Image Classification	Image Classification
Extract Change Information	Extract Change Information

Table 3.1 (Image Processing Steps)

Image Preparation

Imagery selected for this project is Landsat Thematic Mapper (TM) imagery from two dates, June 1986 and June 1996. The pixel size of the imagery was sampled to 25 meters. EOSAT, the former reseller of Landsat TM data, acquired and fully rectified the imagery. Geometry of the images were corrected for x, y, and z values. The positional accuracy of the corrected imagery was verified by overlaying USGS 1:24000 Digital Line Graphs. Additionally, the two images were compared to make sure pixels for each band exactly registered to one another.

Landsat TM imagery is composed of 7 bands of data. The TM sensor collects electromagnetic data in 3 visible bands (blue, green, red), 3 infrared bands and a thermal band. Band 4, an infrared band, was chosen as the image difference band because of its sensitivity to vegetation reflectance. Band 4 covers the .76 to .90 micrometer wavelengths of the magnetic spectrum (see table 3.2). Band 4 was extracted from each image using Erdas Imagine. The image classifications used bands 1-5 and 7. Band 6 was excluded because the resolution of band 6 was much lower than the other 6 bands (120 meters versus 30 meters).

Band Number	Wavelength (Micrometers)	Wavelength (Color)	Pixel Resolution
1	0.45-0.52	Blue	30
2	0.52-0.60	Green	30
3	0.63-0.69	Red	30
4	**0.76-0.90**	**Infrared**	30
5	1.55-1.75	Mid-Infrared	30
6	10.4-12.5	Thermal-Infrared	120
7	2.08-2.35	Mid-Infrared	30

Table 3.2 (Thematic Mapper Bands and Wavelengths)

Erdas Imagine Professional 8.3 was used for all image processing. Data was processed on a Pentium 233 computer with 96 megabytes of memory. Image processing is processor and disk space intensive. Each TM image with 6 bands of data is approximately 400 megabytes. Five gigabytes of disk space were needed for processing and storage space of the TM images.

Create Difference Image (Image Differencing Method Only): Calibration and Normalization Process

Calibration or normalization is the process of minimizing non-change differences between two dates of imagery. Calibration is part of the image differencing change detection process because atmospheric differences between two scenes of imagery (i.e., smog, humidity, gases, etc.) affect the brightness values (BV) collected by the sensor. Atmospheric scattering affects how much electromagnetic radiation is reflected back to be collected by the orbiting sensor.

The two main types of atmospheric scattering are Rayleigh and Mie. Rayleigh scattering, sometimes called molecular scattering, primarily occurs at high altitudes (+4.5 km) and is caused by very small particles primarily composed of oxygen and nitrogen molecules (Avery and Berlin, 1985). Rayleigh scattering primarily affects the shorter wavelengths of the electromagnetic spectrum (i.e., UV and visible blue). Mie, or non-molecular, scattering occurs in lower altitudes (-4.5 km) and is caused by relatively larger particles composed of smoke, dust, salt and ash (Avery and Berlin, 1985). Mie scattering affects longer wavelengths of the electromagnetic spectrum (i.e., Red and Near IR).

Limiting the BV differences caused by atmospheric scattering can be accomplished with several different methods. The two most popular methods of atmospheric correction are *Absolute Radiometric Correction* and *Multiple-date Empirical Radiometric Normalization.* Absolute Radiometric Correction works by collecting atmospheric data for the time and day the area was imaged. The atmospheric data is run through a series of equations that are applied to each band of data, adjusting the BV's for the atmospheric conditions on that day. Absolute Radiometric Correction has two drawbacks: atmospheric data for that scene must be attained and the model assumes consistent atmospheric conditions throughout the entire scene (Jensen, 1996).

Calibration for this project was carried out using Multiple-date Empirical Radiometric Normalization because exact atmospheric conditions are unknown for the two dates of imagery. Multiple-date Empirical Radiometric Normalization works by applying a linear equation that transforms the BV's from a band in the target image to equivalent BV's in a reference image. The transformation equation requires collection of BV's from the same points for each date of imagery. The study area has many fixed targets that have no vegetation because of the urban environment.

The following standards were used to select targets:

- Targets should be fixed elements
- Targets should not contain vegetation
- A target's pattern or texture should be consistent between images
- Targets should have a similar elevation as the rest of the scene and should be in a reasonably flat area. (Eckhardt et al., 1990)

The 1986 image was used as the base image and the 1996 image as the target. Five targets were buildings and four targets were located in water. Table 3.3 lists the targets and the calculated Y-intercept and slope. After all 9 targets were selected, they were recorded in a spreadsheet and the Y-intercept and slope were calculated (see Table 3.3 and Figure 3.1).

1986	1996	1996 Normalized	UTM Coordinates
131	110	152	589031,4199982
150	115	159	590624,4200460
215	142	198	588740,4203484
210	141	198	584927,4199320
130	101	140	585910,4201246
20	16	16	584651,4202867
28	21	23	584935,4202690
20	17	17	584115,4197301
29	24	28	584174,4197129
		Slope	Y-intercept
		1.443989956	-6.557899995

Table 3.3 (Normalization Table- Target Points and Brightness Values)

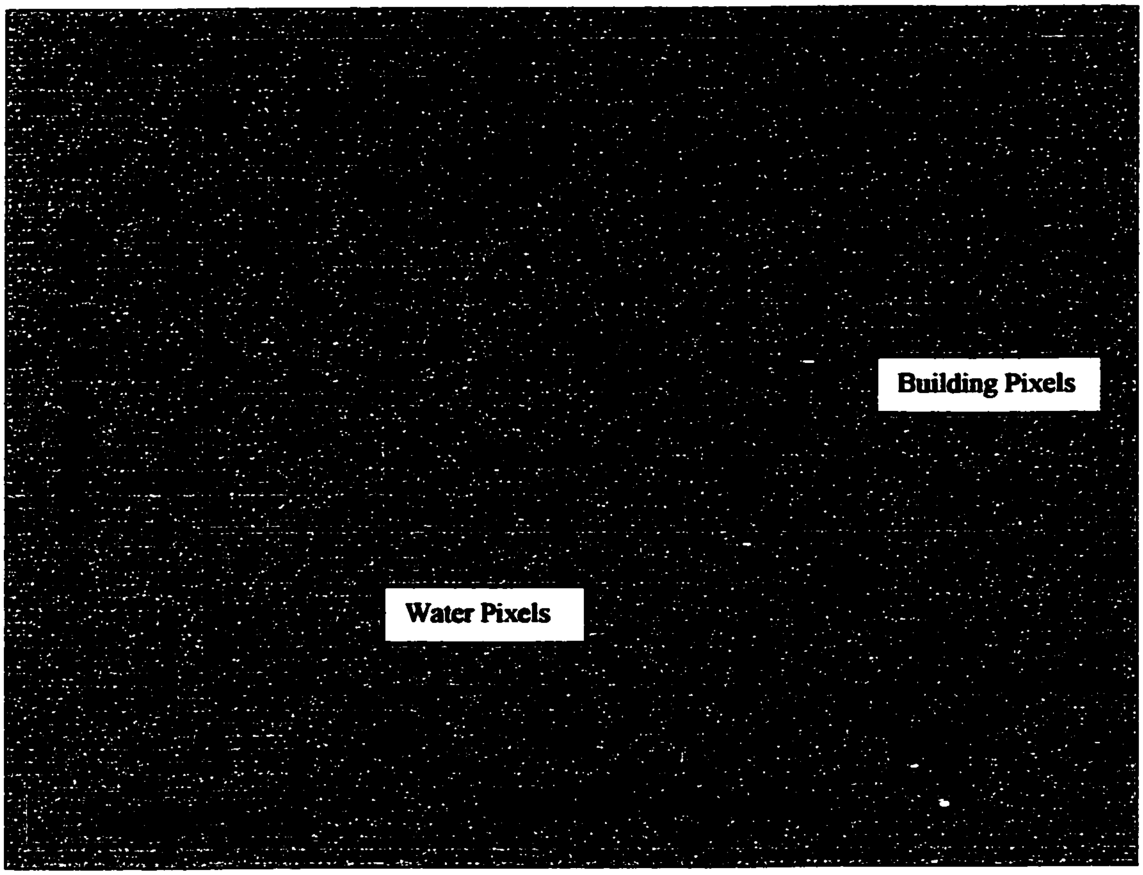

Figure 3.1 (Band 4 Target Points for Normalization)

Y-intercept and slope are calculated and used in the regression calculation as follows:

1996 Band 4 Pixel Values = slope * 1986 Band 4 Values + (Y-intercept)

The regression equation was applied to the 1996 band 4 values using Erdas Imagine. The output data is a 1996 band 4 image normalized to band 4 from 1986.

The difference image was then calculated from the normalized images. The image differencing calculation is a simple arithmetic subtraction:

$$D_{ijk} = BV_{ijk}(1) - BV_{ijk}(2) + c$$

where

D_{ijk} = change pixel value

$BV_{ijk}(1)$ = brightness value at time 1
$BV_{ijk}(2)$ = brightness value at time 2
c = a constant (e.g., 127)
i = line number
j = column number
k = a single band (e.g. TM Band 7)

(Jensen, 1996)

The imagery is 8-bit data, meaning it has 256 possible brightness values. Subtracting two images that have 0 to 255 values yields a possible range of difference values from –255 to +255 (Jensen, 1996). The actual difference range was from –107 to +92. To help make the data more readable, a constant of 127 was added to the difference image resulting in values from 20 to 219. The pixel distribution was normal, as expected (Figure 3.2).

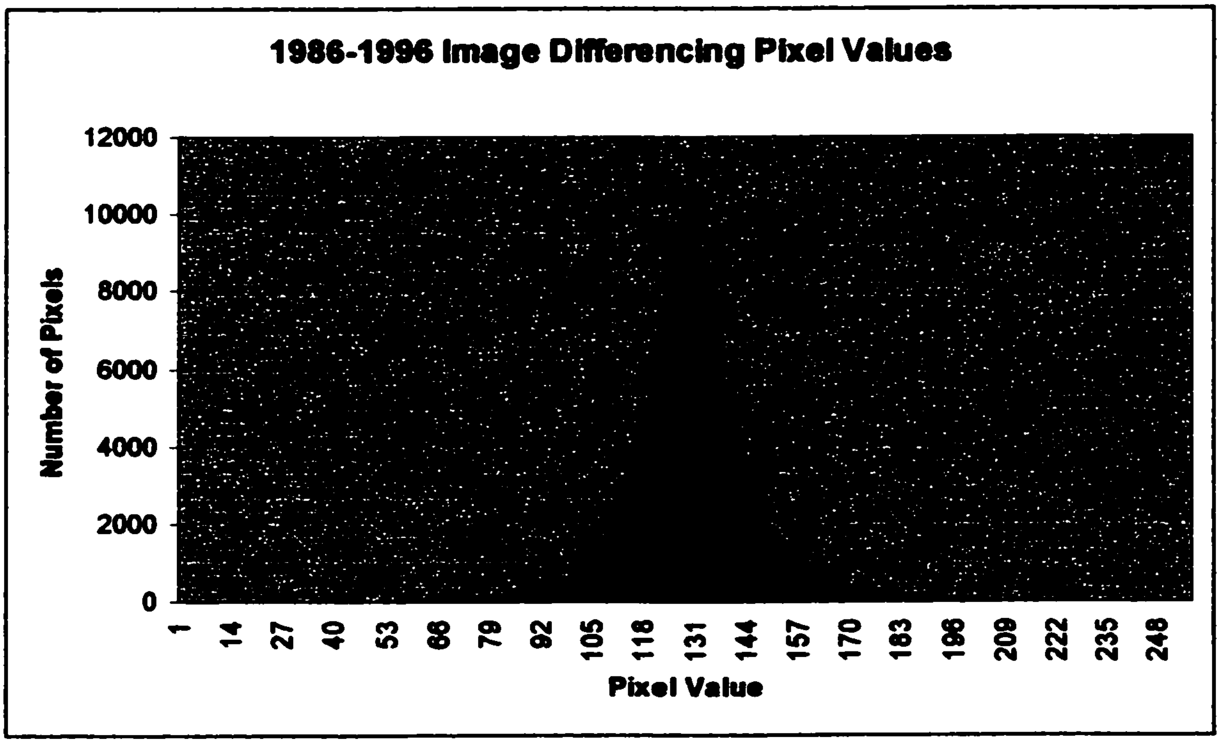

Figure 3.2 (Pixel Values for 1986-1996 Image Differencing)

The image differencing pixel value distribution was normal. Pixels with little or no change stay near the mean and pixels with high change are clustered near the ends of the distribution (Jensen 1996). The mean for the change image was near –6 after the image differencing. After a constant of 127 was added, the mean changed to 120 as depicted in figure 3.2. As noted earlier, the pixels near the mean have changed little in the 10 year span. In contrast, pixel values on the tails of the graph are the pixels that indicate change. The normal pixel distribution shows that the majority of pixels had little change.

The next part of the image differencing process is deciding where to place boundaries between areas of change and no change. Jensen suggests one standard deviation from the mean as a beginning threshold between change and non-change. However, most analysts use empirical testing at different thresholds until "realistic amounts of change" are found (Jensen, 1996). A realistic amount of change means that the analyst determines what is real change and what is not change. This is accomplished by comparing change pixels with the raw imagery.

Standard deviation for the 1986-1996 image differencing is 51.87, with a mean of 120. Using one standard deviation, the change boundaries were placed at 68 on the low end and 172 on the high end (see figure 3.3).

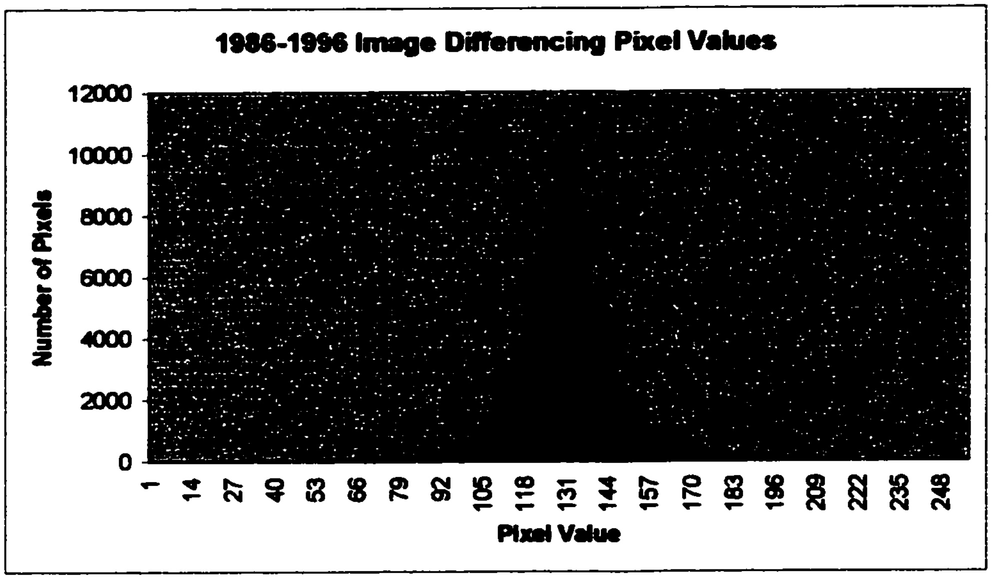

Figure 3.3 (Image Differencing Pixel Values with 1 Standard Deviation Boundary)

Interpretation of the standard deviation boundaries compared with the raw imagery found that they excluded too many areas that represented real change. Following the interpretation of the standard deviation boundaries, experimentation with empirical boundaries began. Using the standard deviation boundaries as a base, the boundaries were moved to include data closer to the mean. The boundaries were moved in increments of 1 unit while the results were being monitored with the 1986 and 1996 images open in linked viewers. The new empirical boundaries were placed at 80 on the low end and 145 on the high end (figure 3.4).

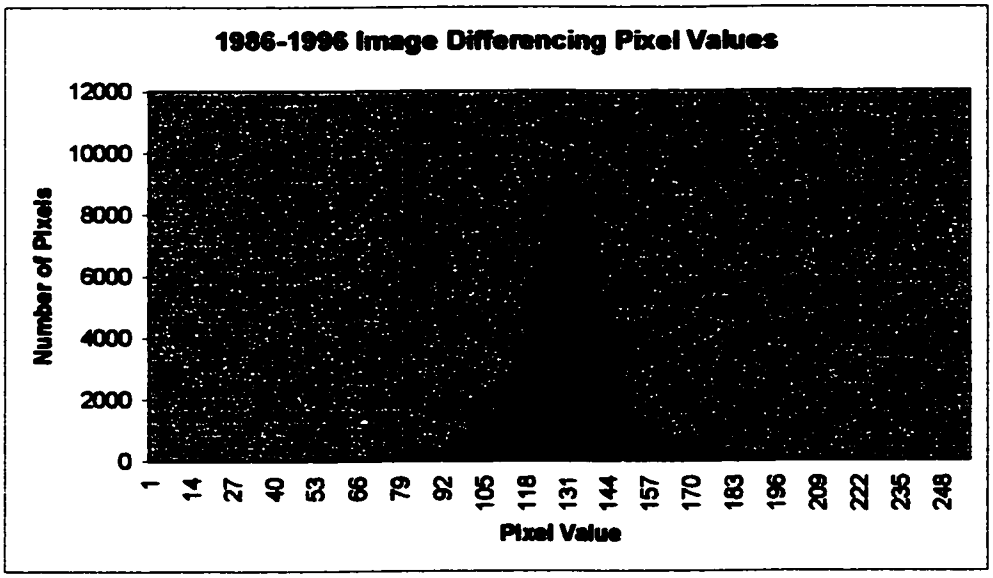

Figure 3.4 (Image Differencing Pixel Values, with Empirical Boundaries)

The image classification can begin now that change areas have been identified with the image differencing method. The image classification methodology is the same for both change detection methods. As noted earlier, only areas identified as change are classified using the image differencing method, while the entire study area is classified using the post classification change method. The results of the image classification for each change method are discussed in the following section.

Image Classification Methodology

The classification for both change methodologies was accomplished using multiple unsupervised (ISODATA) classifications. ISODATA is an acronym for Iterative Self-Organizing Data Analysis Technique (Tou and Gonzalez, 1974). ISODATA is an automated method for classifying pixels which requires the following user inputs:

(From ERDAS field Guide)

- The maximum number of clusters to be extracted, since each cluster is the basis for a class, this number becomes the maximum number of classes to be formed.
- A convergence threshold, which is the maximum proportion of pixels whose class values are allowed to remain unchanged between iterations.
- The maximum number of iterations to be performed

One of the first steps in image classification is defining a classification system. The focus of this change detection process was to determine the amount of land that has changed from open space to developed land. A classification system was designed with the focus of classifying open space and developed land. More specific classes (A-F) were broken out during the classification process to facilitate classification of the larger groups (1-3). For example, during the classification process, confusion between residential and some types of natural vegetation necessitated the division of larger clusters of pixels that included both land cover types. The larger cluster may have contained components of mostly residential with some components of shrub and hardwood. Without breaking out the more refined groups of shrub and hardwood, this large group of residential pixels would not have been correctly classified. At the end of the classification, the six refined classes were assigned into any of the three major classes: Open Space, Developed Land or Water.

1. **Open Space (Major Class)**
   A. Hardwood Trees (oak, cottonwood) –Informational Class
   B. Shrub –Informational Class
   C. Grassland –Informational Class

2. **Developed Land (Major Class)**
   D. Urban (Commercial, Industrial) –Informational Class
   E. Residential –Informational Class

3. **Water (Major Class)**
   F. Water –Informational Class

Interface zones containing oak woodland and residential are difficult to assign to either open space or residential, because they are both. To manage the mixing problem, a 75 percent rule was used to distinguish open space from developed land. This means that, in order for an area to be called open space, it must be covered by more than 75 percent open space pixels. This rule was made so that a small group of oak trees in a residential area would not be classified as oak woodland. The minimum-mapping unit for the classes is 5 acres or about 36 pixels squared.

Image Classification- Post Classification Change Detection

The ISODATA classifications used a 95 percent convergence factor with a maximum of eight iterations. All 6 bands of Thematic Mapper data were used for each date. Thirty clusters were used for the first round of ISODATA classification. Fieldwork and digital air photos indicated the 1986 image had 20 easily discernable clusters with minimal mixing. These groups were mostly made up of clusters of grass, urban land and water. As was expected, oak and residential clusters had considerable confusion. Established residential clusters were the most difficult to break out because of the heterogeneity of the pixels. The older, more established neighborhoods have large trees with wide canopies that are often confused with oak and shrub pixels.

The 1996 image was more difficult to classify due to a very wet winter and spring. Residential and open space vegetation had considerably more mixing compared with the 1986 image. A wet spring will cause normally brown (dry) wild land grasses to be green, much like an irrigated lawn. Enhanced image processing techniques were employed to help break out mixed clusters.

A "cluster busting" technique was employed to help minimize intra-cluster mixing (Jensen, 1996). The cluster busting technique is used on clusters that have a low correlation with what is actually on the ground. These are clusters containing fragments of many different classes (e.g., clusters that contain Oak, Residential, and Irrigated Lawns). First, high confidence clusters are separated from the low confidence classes from the first ISODATA. Next, a binary mask file is created from the low confidence clusters (which are assigned a value of one with everything else coded zero). The raw imagery is masked using the binary file and the resulting output is a raw image file that contains areas with considerable confusion during the first round classification. A second classification is then run on these remaining pixels.

The second unsupervised classification used 20 clusters. Three additional clusters were separated from the second classification. A third unsupervised classification was run on the remaining clusters and no more data could be extracted from further classifications. The remaining clusters of pixels, although made of different classes, had extremely similar properties. These clusters were assigned to the land cover class associated with the majority of pixels.

Cluster busting methodologies were also employed with the 1986 image. Many obviously residential areas were being confused with oak woodland. Again, the second classification was useful while the third classification attempt had very diminished returns. The final recoded classifications are depicted in plates 2 and 3 in the following pages. Plate 2 is from 1986 and plate 3 is from 1996.

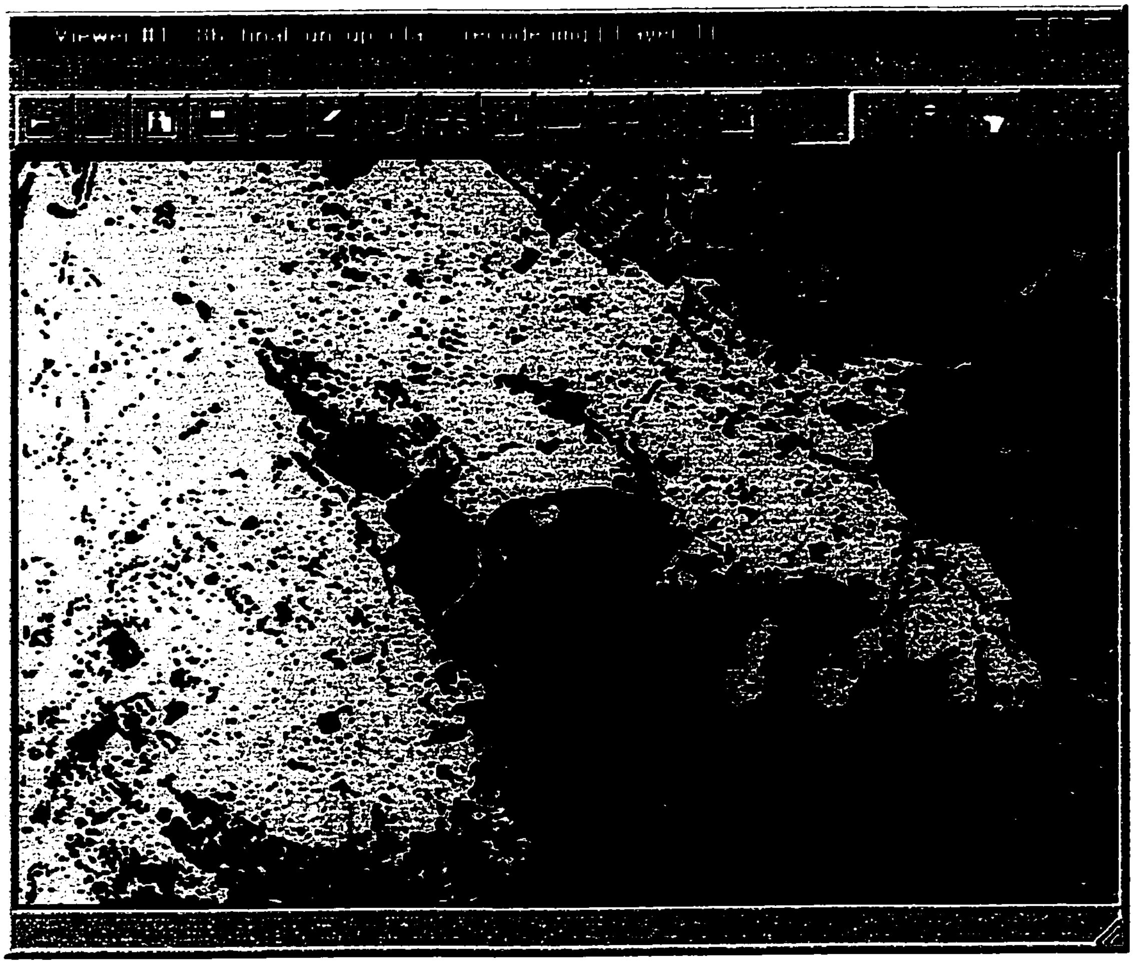

Plate 2 (1986 Image Classification)
(Plates 2 and 3 (Green =Open Space, Gray = Developed, Blue = Water)

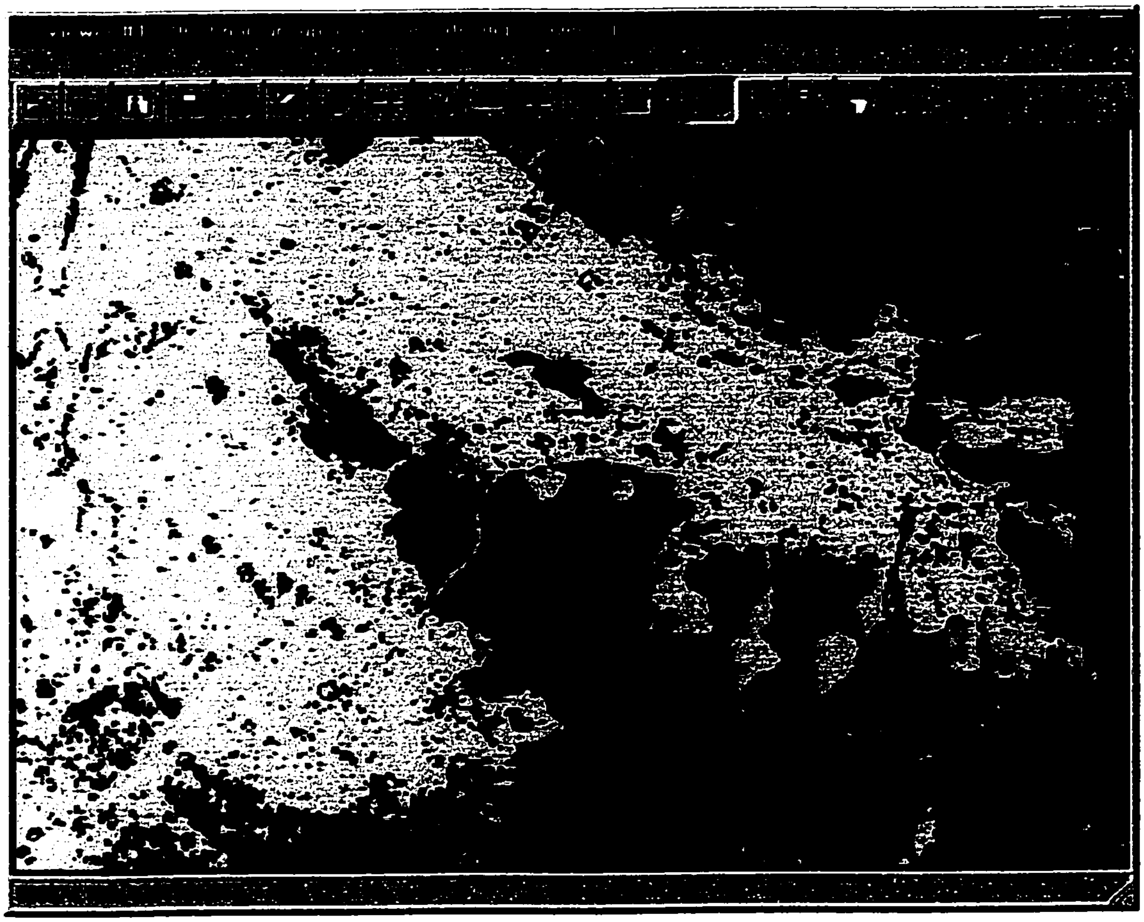

Plate 3 (1996 Image Classification)
(Plates 2 and 3 (Green =Open Space, Gray = Developed, Blue = Water)

Image Classification- Image Differencing Change Detection

Image classification of the image differencing results discovered change areas similar to the post classification change method. The main difference is that the area being classified is a great deal smaller. The total area of image classification is approximately 100 square miles for the post classification change versus 2.2 square miles for image differencing change areas. Classification of image differencing change areas was also made easier because only 2 classes were classified. The water class did not have to be classified in the image differencing method because the water areas did not change between 1986 and 1996.

After the change areas were identified using image differencing, a binary mask was built and applied to the raw data of the study area. The binary mask and raw data were combined to create an image that contains only change pixels. The reduction in area to be classified and number of classes was the result of the image differencing. Classifying only change areas greatly reduced the time to classify an area.

The ISODATA method of classifying the pixels was used. Thirty initial classes were extracted from the change areas. Only 1 ISODATA per scene needed to be run for the image difference data. In contrast, the classification of the whole study area for post classification change detection needed 3 ISODATA classifications run for each scene. After the thirty initial classes were assembled from the raw data, 5 classes were extracted, labeled A-E. Classes A, B and C were then collapsed into open space and classes D and E were folded into Developed Land. The results of the image differencing change classification are in plates 4 (1986) and 5 (1996). Areas classified as developed are color-coded yellow and areas classified as open space are coded cyan.

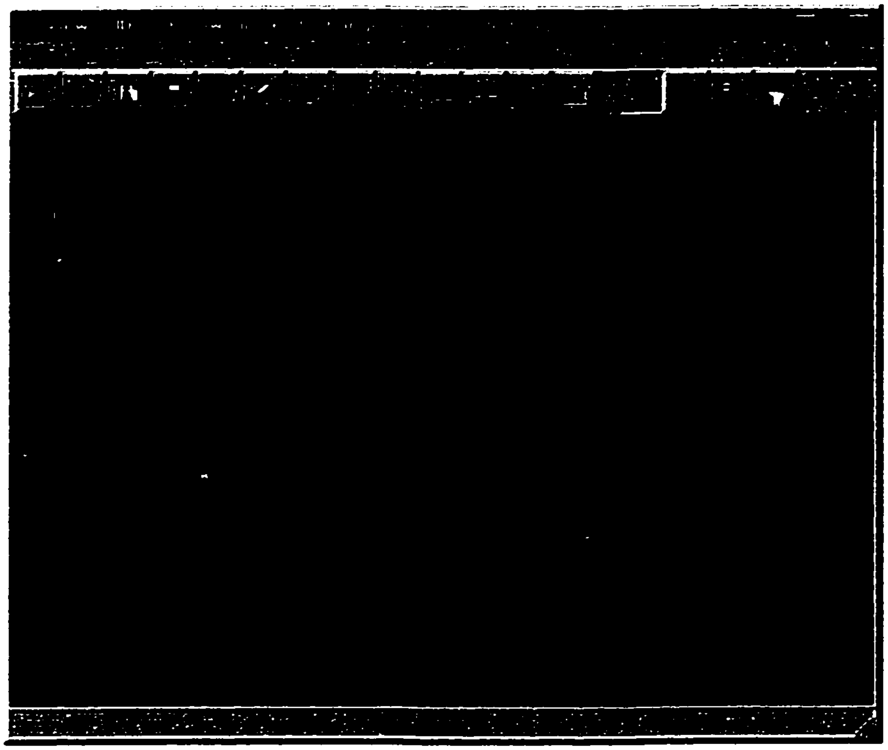

Plate 4 (1986 Image Differencing Classification Results)
(Plates 4 and 5 (cyan =Open Space, yellow = Developed)

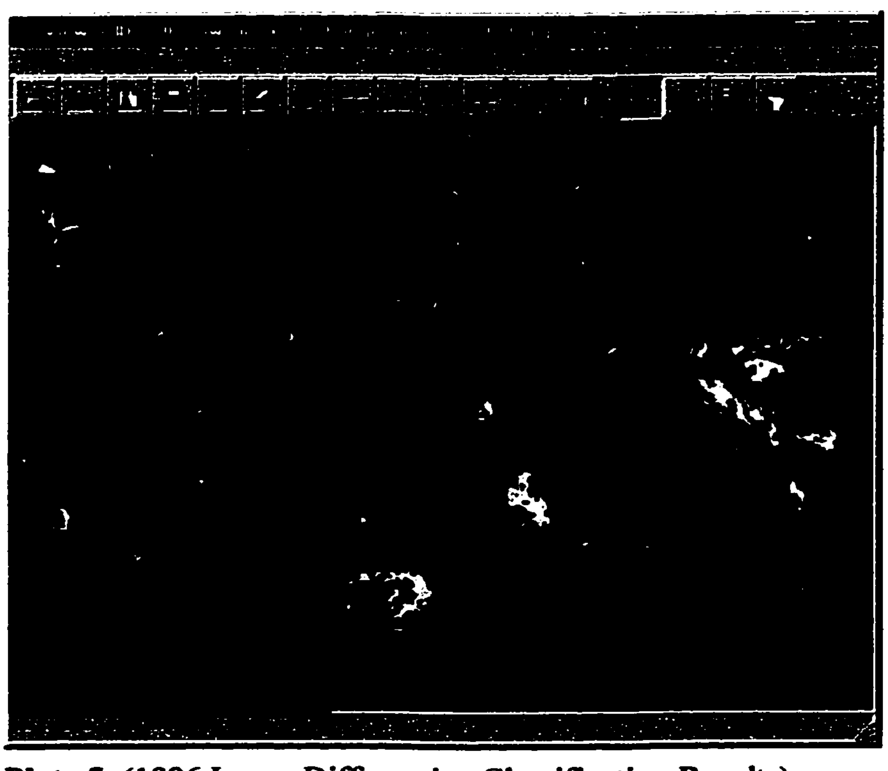

Plate 5 (1996 Image Differencing Classification Results)
(Plates 4 and 5 (cyan =Open Space, yellow = Developed)

Extracting Change Information

After the image classifications for both change detection methods were complete, the next step was to collect from-to change information. The images were recoded so they could be processed with a GIS operator function to extract change data (see table 3.4). In Erdas Imagine, operator functions are tools used to apply simple algebra or arithmetic to thematic layers. First, the 1986 land class values were coded to values 1-3 and 1996 land class values were recoded 10, 20 and 30. Next, the two thematic layers, were added together using the operator function. Table 3.5 is a matrix that lists all of the possible results from adding the two layers.

Final Recodes 1986	Final Recodes 1996
Value	Value
1 Wildland	10 Wildland
2 Urban Land	20 Urban Land
3 Water	30 Water

Table 3.4 (Recode Table)

1996 Data Values				
	30	31	32	33
	20	21	22	23
	10	11	12	13
		1	2	3

1986 Data Values

Table 3.5 (Recode Matrix, all possible values)

The resulting output from the combination of 1986 and 1996 classified images are depicted in plates 6 and 7. Plate 6 is the change image for post classification change and plate 7 is the output of the image differencing change method. In Plate 6, red areas indicate open space reduction. The post classification change comparison found a decrease of 1859 acres of open space. These are areas that were classified in 1986 as open space and in 1996 as developed. This represents a net decrease of 2.9 square miles of open space from 1986 to 1996. The change results were different using the image differencing method.

Image differencing found much less open space change than the post classification change method. The image differencing change map found a decrease in open space of 880 acres. Results of the image differencing are draped over a natural band combination of the study area to help reference the data (plate 7). Red areas were classified as open space in 1986 and as urban in 1996. The green areas were classified as open space in both 1986 and 1996. The green areas were classified because the image differencing process incorrectly detected change where there was none.

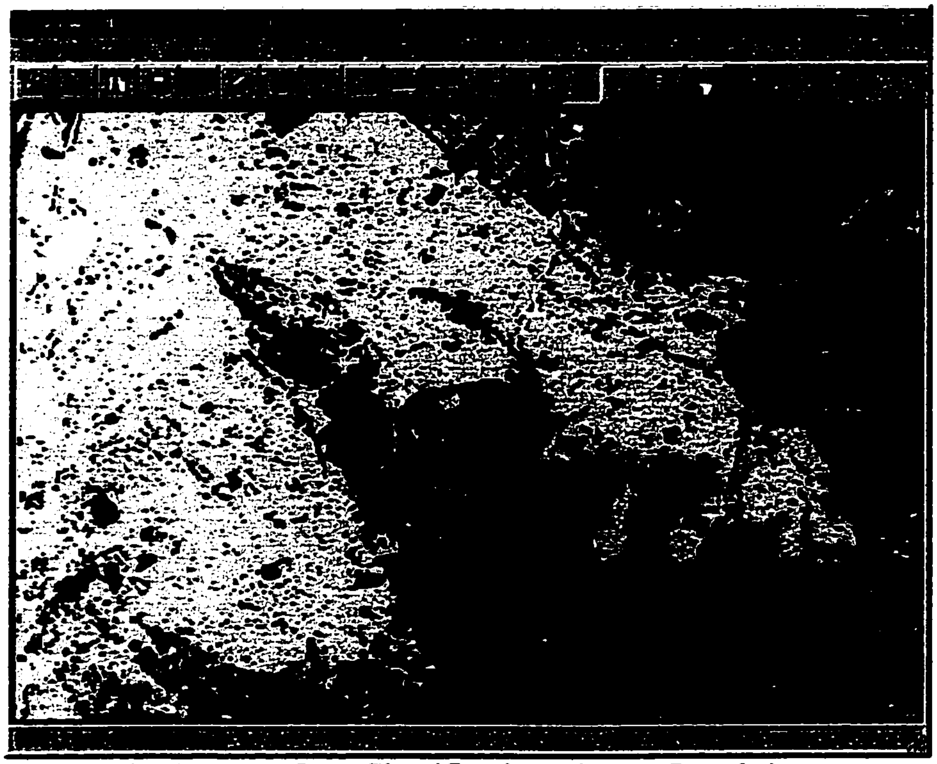

Plate 6 (1986-1996 Post Classification Change Results)

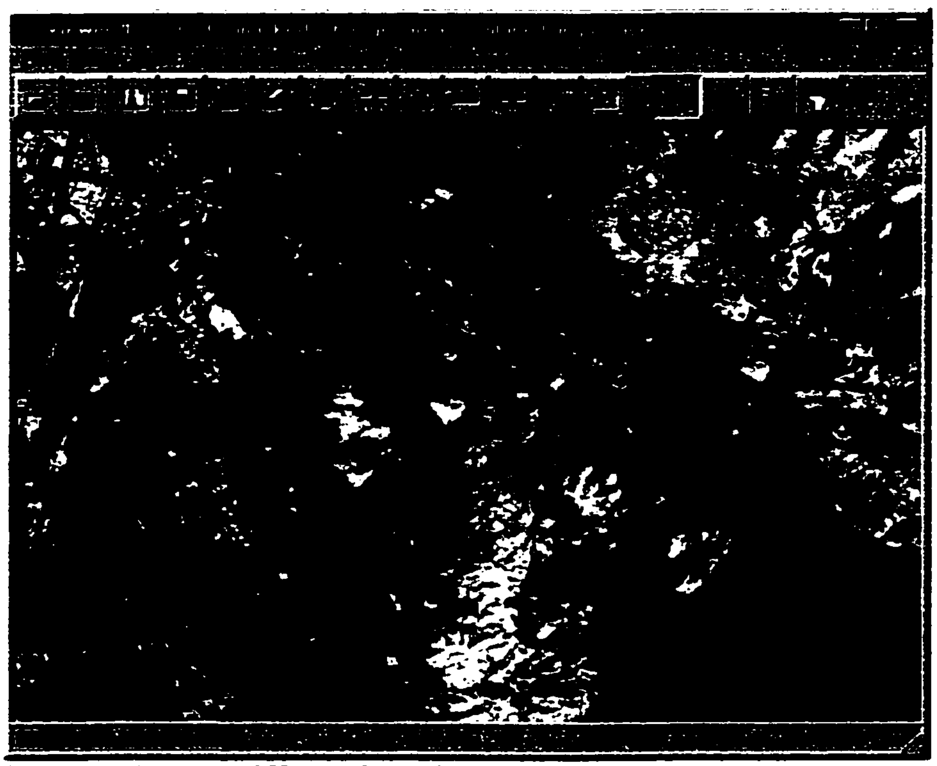

Plate 7 (Image Differencing Change Map)

Both change methods found a substantial decrease in acres of open space. The image differencing method found a decrease of 880 acres. The post classification change method found a decrease of 1859 acres. Which method is more accurate? The following chapter discusses the accuracy of each method as well as the cost efficiency of using image differencing versus post classification change detection.

Chapter 4

ANALYSIS OF POST CLASSIFICATION CHANGE DETECTION VERSUS IMAGE DIFFERENCING CHANGE DETECTION

Accuracy assessment was performed on each change map, since the primary objective of the thesis was to determine the better of two change detection techniques. Accuracy assessment is a way to measure the correctness of an image classification. When performing accuracy assessment, two sources of information are compared: the classified map and reference data (Jensen, 1996). Reference data is derived from collecting data from randomly selected sample points. Accuracy is measured by comparing data collected from sample points to the classified map.

Sampling Scheme

Seventy-five sample points per land cover class were collected for each classification using Congalton's guideline of at least 50 sample points per land cover class. Congalton suggests ramping up the number of sample points if the classification is over 1 million acres (Congalton, 1991). At approximately 32,000 acres, the study area falls well under this threshold.

The type of sampling used for this accuracy assessment was equalized random sampling. Equalized random sampling was used so the open space change and no-change classes had equal weight in the accuracy assessment. The most common method of sampling is a stratified random sampling method (Congalton, 1988). Stratifying the random sampling assures that smaller classes will get sampled during a regular random sampling of the data (Congalton, 1988). A stratified random sampling was not necessary for this study because the maps only have two sufficiently large classes.

The sample points were generated from each class using Erdas Imagine's accuracy assessment tool. Next, the UTM coordinates and land use class for each point were generated and saved to a spreadsheet. The sample points were then checked using either airphotos or field observations. Airphotos were used for areas that were inaccessible to fieldwork, such as those on private property or located too far from trails and roads.

After the ground truth data was collected, the data was input into Imagine's accuracy assessment tool to generate the error matrix. The results were written to error matrices in tables 4.1 and 4.2. Accuracy is depicted in three different ways: producer's accuracy, user's accuracy and overall accuracy. Producer's accuracy is "total number of pixels of that category as derived from the reference data (i.e., the column total)" or the "probability that a reference pixel is correctly classified" (Jensen, 1996). User's accuracy is the total number of correctly classified pixels in category "A" divided by the total number of pixels actually classified into category "A" (i.e., the row total) or the "probability that a pixel classified on the map actually represents that category on the ground" (Jensen, 1996). Overall accuracy is calculated by dividing the total number of correctly classified pixels by the total number of sample pixels. Jensen recommends reporting all three measures of accuracy because the numbers can have very different meanings depending on how the classification is to be used.

Producer's accuracy identifies the "probability of a reference pixel being correctly classified and is a measure of omission error" (Jensen, 1996). User's accuracy is the likelihood that a pixel classified in a particular class, matches what is really on the ground or a measure of commission error (Story and Congalton, 1986). For example, the post classification accuracy results reported a Producer's accuracy of 94 percent for identifying change areas. The 94 percent accuracy number was obtained by dividing the 54 correctly identified pixels by 57, the total number of change reference pixels. User's accuracy for the same class was 72 percent. The 72 percent accuracy number was obtained by dividing the total number of correctly classified change pixels (54), by the total number of classified change pixels (75). Ninety-four percent of change areas were correctly identified as change, however 72 percent of the time change areas identified on the map were really change areas in the field (Jensen, 1996). The 72 percent accuracy number is indicating that too many areas were identified as change that were not really change.

Post Classification Change Detection Error Matrix					
Classified Data	Non-Change	Change	Classified Totals		
Non-Change	72	3	75		
Change	21	54	75		
Reference Totals	93	57	150		
ACCURACY TOTALS					
Class Name	Reference Totals	Classified Total	Number Correct	Producer's Accuracy	User's Accuracy
Non-Change	93	75	72	77.42%	96.00%
Change	57	75	54	94.74%	72.00%
Totals	150	150	126		
Overall Classification Accuracy = 84.00%					

Table 4.1 (Post Classification Change Detection Accuracy Results)

Image Differencing Change Detection Error Matrix					
Classified Data	Non-Change	Change	Classified Totals		
Non-Change	71	4	75		
Change	5	70	75		
Reference Totals	76	74	150		
ACCURACY TOTALS					
Class Name	Reference Totals	Classified Totals	Number Correct	Producer's Accuracy	User's Accuracy
Non-Change	76	75	71	93.42%	94.67%
Change	74	75	70	94.59%	93.33%
Totals	150	150	141		
Overall Classification Accuracy = 94.00%					

Table 4.2 (Image Differencing Change Detection Accuracy Results)

Analysis: Post Classification Change Detection Accuracy Results

Producer's Accuracy

Table 4.1 describes the accuracy numbers for post classification change detection. Class 1 or non-change producer's accuracy is 77 percent. Analysis shows that 93 sample sites were found to be non-change but only 75 of the pixels were supposed to be non-change. Eighteen sample sites were incorrectly classified as change pixels when they should have been non-change. For class 2 or change pixels, producer's accuracy is 94 percent correct.

User's Accuracy

User's accuracy for Class 1 or non-change pixels is 96 percent. This means that of the 75 pixels that were supposed to be non-change, 72 of them turned out to be non-change. User's accuracy for class 2 or change pixels is 72 percent. Analysis shows that of 75 pixels originally selected as change pixels, only 54 turned out to actually be change pixels. Analysis shows there were many areas classified as change that were not really change.

Overall Accuracy

As stated earlier, overall accuracy is calculated by dividing the total number of correctly classified pixels by the total number of reference pixels. Overall accuracy for the post classification change detection is 84 percent. Of the 150 reference points data was collected for, post classification change detection correctly identified 126 of the areas.

Analysis: Image Differencing Change Detection Accuracy Results

Producer's Accuracy

Table 4.2 contains the accuracy numbers for Image Differencing Change Detection. Class 1 or non-change producer's accuracy is 93 percent. Analysis shows 76 reference sites were found to be non-change but 71 of the sites that were classified as non-change were actually non-change pixels. Similarly, the producer's accuracy for class

2 or change pixels is 94 percent. Seventy-four sample sites were found to be change but only 70 of the pixels that were supposed to be change were actually change pixels.

User's Accuracy

Class 1 or non-change user's accuracy is 94 percent. Seventy-one out of 75 accuracy assessment pixels were classified correctly. Class 2 or open space change user's accuracy is 93 percent. Class 2 was correctly identified in 70 out of 75 possible accuracy assessment pixels.

Overall Accuracy

Overall accuracy for Image Differencing Change Detection is 94 percent. Analysis shows image differencing change detection identified 141 out of 150 accuracy assessment pixels correctly.

Accuracy Assessment Discussion

The image differencing change detection's overall classification accuracy is 94 percent versus 84 percent for the post classification. Eighty-five percent or better is often a gauge used in the GIS/Remote Sensing industry to measure the usefulness of classified data. The post classification change map would not meet an 85 percent overall accuracy standard. At 94 percent overall classification accuracy, the image differencing change map would meet most standards for image classification accuracy.

Image differencing change detection is very accurate, with user's, producer's and overall accuracy all above 93 percent. The post classification change method accuracy problems were caused by overestimating the amount of change. Plate 8 shows an example of how the image differencing change detection was more precise than post classification change detection. The area is an established residential area that did not change from 1986 to 1996. Bright yellow pixels are areas the post classification method identified as change. Orange pixels are areas identified as change by the image differencing change method. Bright green pixels are areas identified as change by both methods.

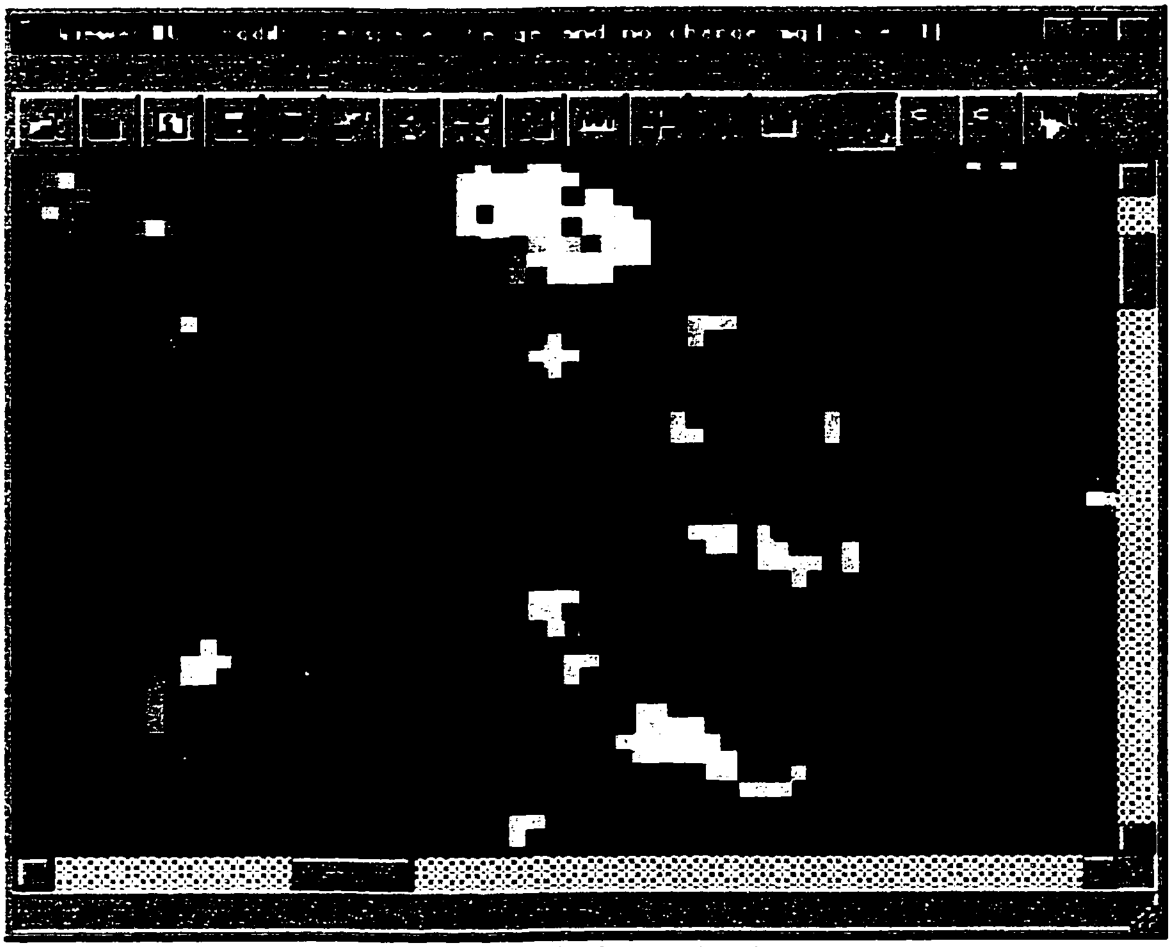

Plate 8 (Post Classification- Residential Change Area)

The post classification change method showed some areas with an increase in vegetation (e.g. areas changing from urban to open space). The box in plates 9 and 10 represent an area that gained vegetation from 1986 to 1996. Plate 9 depicts an urban/open space interface area that in 1986, when viewed in a false color band combination, looks to have little vegetation. Notice the blue/gray color in plate 10a that denotes a lack of vegetation. The color looks this way because of the lack of infrared reflectance. When viewed against plate 10b from 1996, one can see the more vibrant red color that denotes green vegetation on the ground.

Field data showed that this area is a fringe suburban area that is sometimes nearly bare soil in dry years and grass/weeds in wetter years. Spring of 1996 had more precipitation in the Diablo Valley than in 1986. The Diablo valley received 3.11 inches of rain in April and May of 1996 and only 1.16 inches during the same period of 1986 (National Climatic Data Center, Martinez Weather Station). Of the 3.11 inches of rain collected in the spring of 1996, 1.76 inches were collected in late May. The wetter spring in 1996 caused some post classification change areas to be labeled as open space in 1996 that were previously labeled as developed in 1986. The extra precipitation in the spring of 1996 did not cause problems with the image differencing method. Empirical change boundaries placed on the image differencing method were confining enough so that the area in plate 9 was not classified as change.

Zoom in of area changed by differences in rainfall

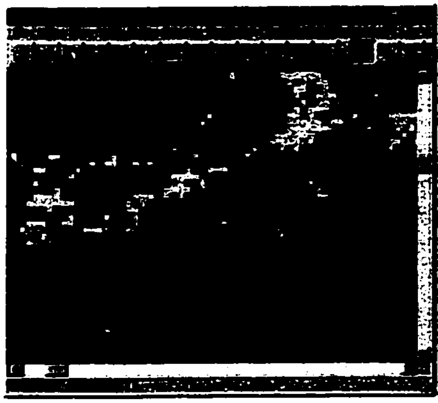

Plate 9 (Post Classification change method)
Grey = Urban\Developed
Green = Open Space
Orange = Veg Increase
Pink = Veg Decrease

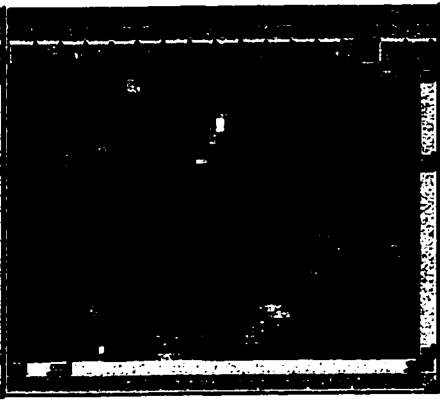

Plate 10a (1986 raw imagery)

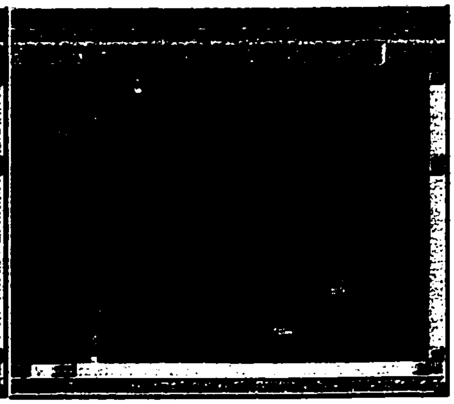

Plate 10b (1996 raw imagery)

The image differencing and post classification change methods correctly identified the areas shown on plates 11-14 as change areas. Both areas were converted from open space to residential developments. Plates 11 and 12 show an oak grassland area being converted to a residential development and a golf course. Notice the visual differences between the two dates of imagery in plates 11 and 12. In the 1996 image, the golf course is represented by bright red in the image, indicating high vegetation reflectance (e.g. irrigated grass).

The residential area in plate 12 shows some dark red in the imagery indicating there are some established oak trees still left in tact in the neighborhood. Ground truthing showed this to be correct. This neighborhood is approximately 10 years old, which means it was established shortly after the 1986 image was scanned. Plates 13 and 14 show a very new area being cleared for a residential sub division.

Plate 11 **Plate 12**

Plates 11 and 12 show an area converted from open space to developed land. The imagery is displayed in a 4, 5, 3 band false color image.

In plates 13 and 14, notice the change in reflectance between the 1986 and 1996 images. The 1986 image shows tan and red colors indicating dry grass and oak/shrub vegetation. The 1986 image clearly indicates some of the area has been prepared for development. In plate 14, 1996, you see that some of the neighborhood has been completed. Also, notice the newly cleared areas near the bottom of the box. The bluish/gray color in the 1996 image indicates little or no vegetation (i.e., cleared areas). This takes place when the hillsides are graded for development, thus removing the natural vegetation.

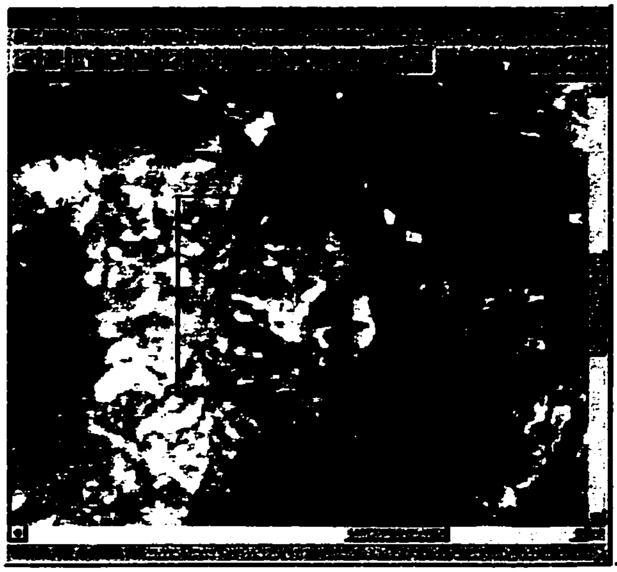

Plate 13 **Plate 14**

Plates 13 and 14 show an area converted from open space to developed land. The imagery is displayed in a 4, 5, 3 band false color image.

As shown in the description of plates 11-14, both change detection methods identified change areas successfully. The post classification change method identified many of the same change areas as image differencing. The primary problem with the post classification change method was that it classified some areas as change which were not change. The incorrectly classified areas were due to differences in the classifications. Residential areas were the most likely to be incorrectly classified as open space. As noted before, established residential areas have many of the same spectral characteristics as oak woodland. The similar characteristics are due the dense foliage and closed tree canopies of oak woodland and established residential areas. When the change map is produced from the two classifications, any differences in the classifications (right or wrong) appear as change.

The results of the accuracy assessment between the two methods of change demonstrate that image differencing change detection is more accurate. Both methods of change detection showed a reduction in the amount of open space from 1986 to 1996. Most of the open space developed was used for residential construction. Plates 15 and 16 are a comparison of the post classification method and image differencing method results. Notice both methods successfully identified many of the change areas. Plate 16 shows the results of the post classification change method. Observe some of the change areas in the developed and open space sections that were incorrectly identified as change by the post classification change method. Arrows mark some of these areas.

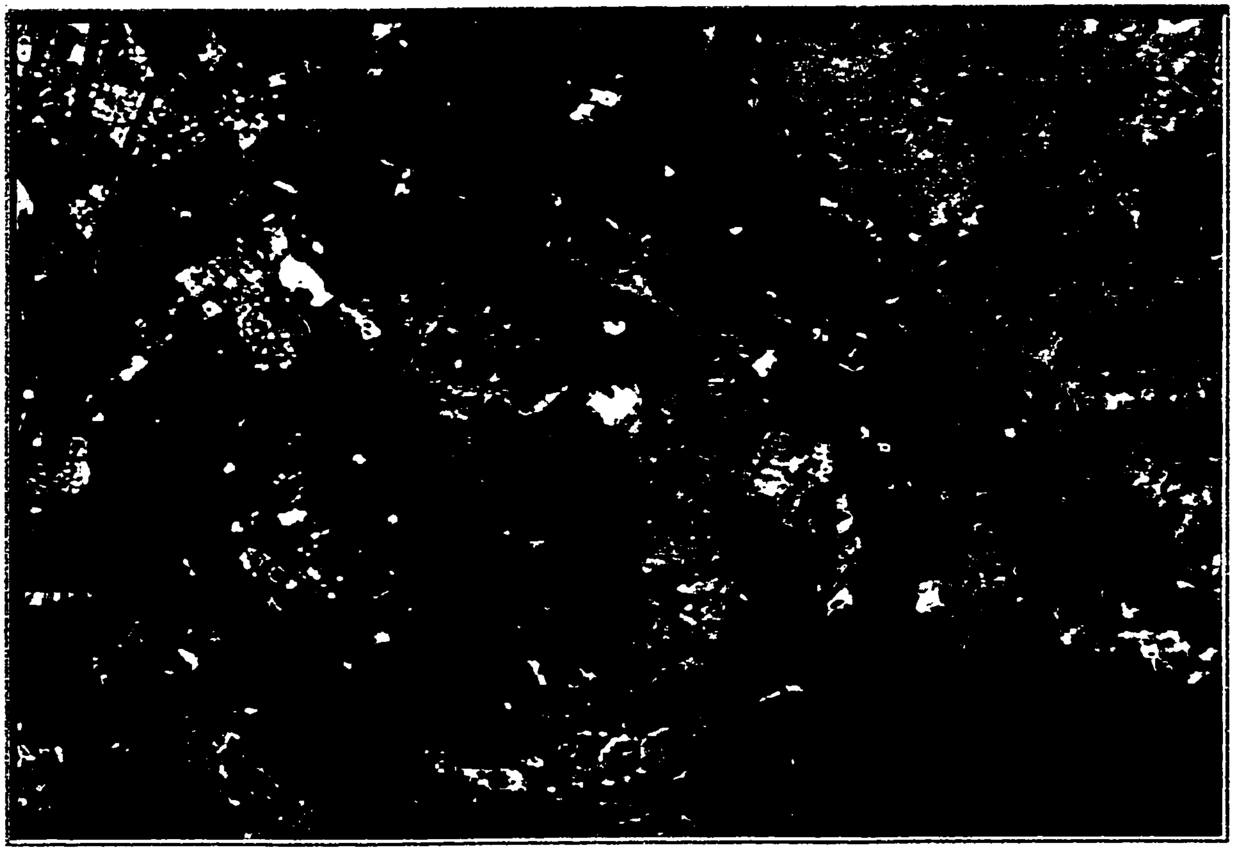

Plate 15 Image Differencing Change Results (overlaid on imagery).

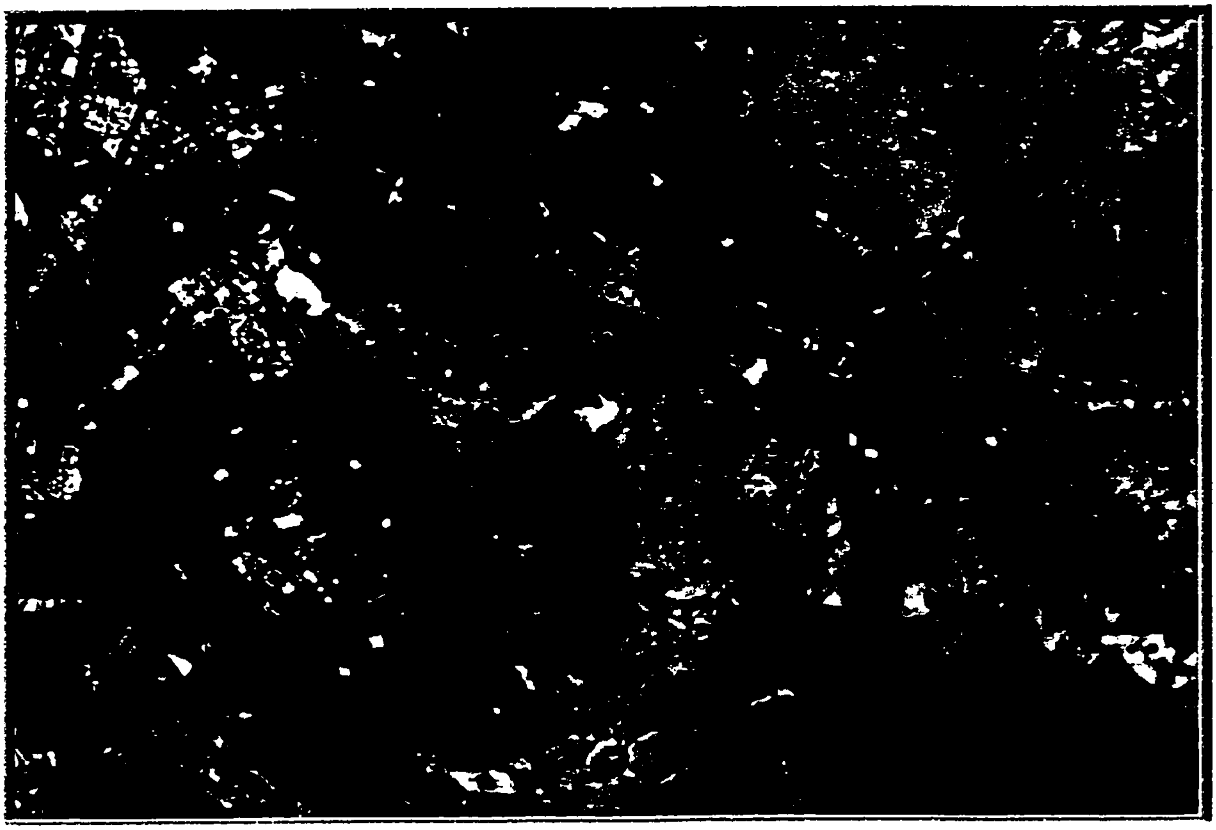

Plate 16 Post Classification Change Results (overlaid on imagery).

Accuracy assessment determined that the image differencing method was more precise. The second goal of the thesis was to determine which of the two methods was more cost effective. The following section discusses the cost effectiveness of each method of change detection.

Cost Efficiency

As stated in chapter 1, image classification and GIS projects are time intensive and can be very expensive. Cost is usually a primary factor when undertaking these projects. Budgets for GIS and image classification projects affect the type of processing performed, data used, and the size of the study area. Hours to complete each map of change were tracked since all other factors in producing the maps are the same. The static factors in producing maps of change were cost of data, hardware and software. Person hours are usually the first or second highest cost in producing maps of change. Table 4.3 identifies the time spent on each section of producing the post classification change map.

Image	Pre-Processing	Unsupervised	Change Map	Total
1986	1 hrs.	8 hrs.	N.A.	9 hrs.
1996	1 hrs.	10 hrs.	N.A.	11 hrs.
Both	N.-	N.A.	5 hrs.	5 hrs.
Total	2 hrs.	18 hrs.	5 hrs.	**25 hrs.**

Table 4.3 (Post Classification Change Number of Hours) N.A.= Not Applicable

Table 4.4 identifies the time spent on each section to produce the image differencing change map. Total time to complete the image differencing change detection map was 23 hours. The image differencing method required 2 less hours to complete than the post classification method. The image differencing method required less time because only the areas identified as change are classified. While the image differencing change method required 2 less hours to complete than the post classification change method, the difference is not significant enough to definitively state that image differencing is more cost effective.

Task	**Time**
Pre-Processing	2 hours
Calibrate/Normalize	5 hours
Image Differencing	6 hours
Classification	6 hours
Change Map	4 hours
Total	**23 hours**

Table 4.4 (Image Differencing Number of Hours)

Chapter 5

CHANGE DETECTION SUMMARY

This study started out with two main goals:

1. Which is more accurate: post classification change detection or image differencing change detection.
2. Which is more cost effective: post classification change detection or image differencing change detection.

The study demonstrated that the image differencing change method could more accurately find change than the traditional post classification change detection. The overall accuracy for the image differencing method was 94 percent versus 84 percent for the post classification change method. Many commercial remote sensing projects have a predefined accuracy target, which must be achieved to meet contractual obligations. A 10 percent increase in accuracy is significant when working with remotely sensed data.

This project did not find any meaningful cost difference between image differencing and post classification change detection. The post classification change process required 2 more hours to complete than the image differencing method. Two hours represents less than a 10 percent difference between the methods. The image differencing method required 6 hours for image classification versus 18 hours for the post classification change method. However, the 12 hours saved on the image differencing classification were offset by the time required for normalization and differencing.

Goals of the change detection project should be defined before selecting a methodology. Although image differencing yielded a more accurate result, note that in the process of producing the post classification change map, two valuable land use maps were created. With some post classification editing, the accuracy standards of those maps could be brought up to commercial standards. The image differencing change map only provides data for change areas.

Digital remote sensing systems and computer image processing have given us a powerful tool for monitoring change in our environment. Often change information is needed quickly because of natural disasters. For instance, if a wild fire strikes a wooded area, emergency personnel will need to know which areas received the most damage. Older methods of monitoring environmental change were expensive and required a long duration. Today, with digital image processing techniques, a change map of a study area similar in size to the Diablo Valley could be produced in 3 (8-hour) working days (not including accuracy assessment). Providing change monitoring to land managers and policy makers quickly and at a low cost can be essential in emergency situations.

This thesis analyzed two popular change detection methodologies. This study showed the image differencing method could produce a highly accurate change map in 23 hours. The study also demonstrated that image differencing may be more effective at finding change areas. Many future change detection projects exist that will require analysts to decide what kind of change detection methods to use. Hopefully, the knowledge gained by this study of image differencing and post classification change detection will help others in their choice of change detection methodologies.

REFERENCES

Avery, T.E., & Berlin, G.L. (1985). *Fundamentals of remote sensing and airphoto interpretation.* New York: Macmillan Publishing Company.

Congalton, R.G. (1988). Using spatial autocorrelation analysis to explore the errors in maps generated from remotely sensed data. *Photogrammetric Engineering and Remote Sensing*, 54(5), 587-592.

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

Eckhardt, D.W., Verdin, J.P., & Lyford, G. R. (1990). Automated update of an irrigated Lands: Gis using spot hrv imagery. *Photogrammetric Engineering and Remote Sensing*, 56(11), 1515-1522.

Fung, T., & LeDrew, E. (1987). Application of principal components analysis for change detection. *Photogrammetric Engineering and Remote Sensing*, 59(12), 1649-1658.

Gong, P., LeDrew, E.F., & Miller, J.R. (1992). Registration-noise reduction in difference images for change detection. *International Journal of Remote Sensing*, 13(4), 773-779.

Green, K., & Weinstein, D. (1996). Automated change detection using remotely sensed data. Proceedings, *GIS 96' Symposium*, Vancouver, B.C.

Hornbeck, D. (1983). *California patterns: A geographical and historical atlas*. Mountain View: Mayfield Publishing Company.

Jensen, J.R. (1996). *Introductory digital image processing: A remote sensing perspective*. Upper Saddle River: Prentice-Hall.

Story, M., & Congalton, R. (1986). Accuracy Assessment: A User's Perspective. *Photogrammetric Engineering and Remote Sensing*, 52(3), 397-399.

Tou, J., & Golzalez, R.C. (1974). *Pattern recognition principles*. Reading: Addison-Wesley Publishing Company.

Valley oaks test protection policy. (2000, August 20). *Contra Costa Times Newspaper*, p. A12)