2006

Landsat image classification using fuzzy sets rule base theory

Avivit Shani

San Jose State University

DOI: https://doi.org/10.31979/etd.99w3-6ftf

LANDSAT IMAGE CLASSIFICATION USING FUZZY SETS RULE BASE THEORY

The Faculty of the Department of Geography

San Jose State University

In Partial Fulfillment

of the Requirements for the Degree

Master of Arts

By

Avivit Shani

August 2006

APPROVED FOR THE DEPARTMENT OF GEOGRAPHY

Dr. Gary Pereira

Dr Richard Taketa

Cindy Schmidt, San Jose State University Foundation

APPROVED FOR THE UNIVERSITY

08/07/06

ABSTRACT

LANDSAT IMAGE CLASSIFICATION USING FUZZY SETS RULE BASE THEORY

By Avivit Shani

This thesis explores a new methodology for the classification of Earth surface land cover categories, using fuzzy sets and fuzzy logic. The region under consideration is the San Francisco Bay Area. Landsat Enhanced Thematic Mapper (ETM+) Multispectral data, as well as National Land Cover Database (NLCD) 2001 interim classifications, are used to derive fuzzy sets that represent the radiance characteristics of two broad classes, tree canopy and impervious surfaces. A new type of confidence-weighted histogram is used in this derivation. The fuzzy sets are used in combination to independently classify the region into these two classes, and the results are compared with the original NLCD classification maps and with aerial photographs. Formal evaluation of results of this methodology as well as the NLCD results is performed with the help of high-resolution color aerial orthophotographs. The methodology used here was shown to be as accurate and robust as the NLCD results for most of the data, while being more transparent and more easily applicable than the NLCD methodology.

TABLE OF CONTENTS

LIST OF FIGURES:

FIGURE

LIST OF TABLES

TABLE

Introduction

The need for measuring and monitoring biophysical characteristics and human activities on Earth as the human population grows is increasing. In the electronic information age, which includes spatial information, remotely sensed data has taken a major role in managing natural resources and planning future development. Advances in technology have led to rapid developments in methods of classification and the formulation of new and more sophisticated decision rules. Among these advanced technologies are methodologies using Fuzzy Logic and Neural Networks for the analysis of remotely sensed data.

Federal and local governments in the United States have been developing land-use/ land-cover maps based on traditional classification methods. Traditional classification assumes that image pixels are pure. A pixel contains one and only one class. In reality a large number of pixels may show affinity with several information classes. These pixels are called "mixed pixels" and can cause a misclassification. Traditional classification methods do not provide a good mechanism for dealing with such uncertainty and imprecision. Fuzzy set theory provides a new concept of classification problems in an ambiguous environment.

The United States Geological Survey (USGS) Multi-resolution Land Characterization Consortium (MRLC) is currently producing a set of 30-meter resolution land-use/ land-cover maps called NLCD 2001, over the conterminous United States, Alaska, Hawaii, and Puerto Rico using Landsat imagery and ancillary data.

This is a database approach to land cover, which was created to meet the necessity of standardized National Map (USGS 2001) currently being created by the USGS for the United States. The NLCD 2001 project includes the concept of partial membership and uses different methods of analysis in order to achieve this classification for the U.S. The study presented in this thesis will make use of the data and partial assessments currently available from the NLCD 2001 project in order to inform and evaluate an alternative fuzzy rule-based methodology classification methodology. This is an alternative method of arriving at the same sort of categorization as the USGS project. It uses Fuzzy Sets and Fuzzy Logic, and is implemented in the Idrisi32 Geographic Information System (GIS).

Land-Use Land-Cover Classification

Land cover refers to the type of material present on the landscape (e.g. water, sand, crops, forest, wetland, and human-made material such as asphalt). Land use refers to what people do on the land surface (e.g., agriculture, commerce, settlement) (Jensen, 2005, p.340). Land cover data have proved valuable for predicting distributions of phenomena and creating models of spatial patterns especially in broad areas that otherwise would not be examined. All classes of interest must be selected and defined carefully to classify remotely sensed data into land-use land-cover information. Certain classification schemes have been developed that incorporate land-use land-cover data obtained from remotely sensed data. Some examples are:

1. American Planning Association *land-based Classification System*. The LBCS requires input from *in situ* surveys, aerial photography, and satellite remote sensor data to obtain information at the parcel level on the activity, function, site development, structure, and ownership with a unique code for every commercial and industrial land-use activity.
2. United State Geological Survey *Land-Use/Land-Cover Classification System.* A resource-oriented land-cover classification system in contrast with land-use classification systems. The system is designed to interpret remote sensor data obtained at various scales.
3. U.S National Vegetation and Classification system of the Federal Geographic Data Committee. A uniform Vegetation resource data at a national level (FGDC vegetation Subcommittee, 1997; FGDC, 2004)
4. U.S Department of the Interior Fish & Wildlife Service, *Classification of Wetlands and Deepwater Habitats of the United States.* A wetland classification system that incorporates information from remote sensor data and *in situ* measurements in order to monitor the continuum loses of inland to agriculture.
5. International Geosphere-Biosphere Program IGBP *Land Cover Classification System* modified for the creation of MODIS land cover products. A land-cover classification system used mainly to depict broad scale land-cover changes at a regional, national, and global scale.

Image Classification Methods

The aim of classification is to establish a relationship between a pattern and a class label (Tso, 2001, p.54). Classification can be performed using either supervised or unsupervised algorithms. The supervised classification requires *a priori* knowledge of the type and number of information classes that are represented in the study area. The unsupervised classification method is less dependent on user interaction. This method 'learns' the characteristics of each class and the number of classes directly from the input data. This study will focus on a supervised classification method.

Supervised classification methods can be categorized into parametric or nonparametric. Parametric classification assumes that the observation measurements behave according to statistical distribution, such as Gaussian (That is a normal distribution). Nonparametric classification makes no such assumptions. Some of the most often used nonparametric methods are:

- Parallelepiped
- Minimum distance
- Nearest neighbor
- Neural networks
- Fuzzy logic

Though Parallelepiped and Minimum distance methods are considered nonparametric, they do use some degree of statistical parameters. The Parallelepiped algorithm uses mean and standard deviation statistics for each band to establish thresholds; Minimum distance classifier requires the mean for each training class for each band. Although these traditional approaches can perform well, their general ability for resolving inter-class confusion is limited. All but Neural Networks and Fuzzy logic use linear algorithms, which is based on location, distance, and probability analysis. In recent years, alternative strategies have been proposed such as Neural Networks, Fuzzy Sets Theory, Decision trees, and the incorporation of secondary information such as texture, context and terrain features (Tso, 2001, p.56). In this study, a nonparametric supervised method using Fuzzy Sets Theory will be examined for land cover classification. The results of this method will be compared to the results of the NLCD 2001 classification.

Introduction to Fuzzy Set Theory

The difference between crisp, or hard classification, and fuzzy sets can be characterized by the membership function. The membership function in a crisp set can only output two choices {yes, no} or [0,1] (Tso, 2001, p.150). An element of a crisp set can be a member of only one group with a grade of 1. The concept of the fuzzy set provides greater flexibility by allowing one data element to hold several non-zero membership grades for different groups. This approach can handle problems in indistinct boundaries that are common in the natural world (Burrough and McDonnell, 1998, p.268). Figure 1 illustrates the indistinct boundaries of a pixel and its classification in hard and fuzzy methods:

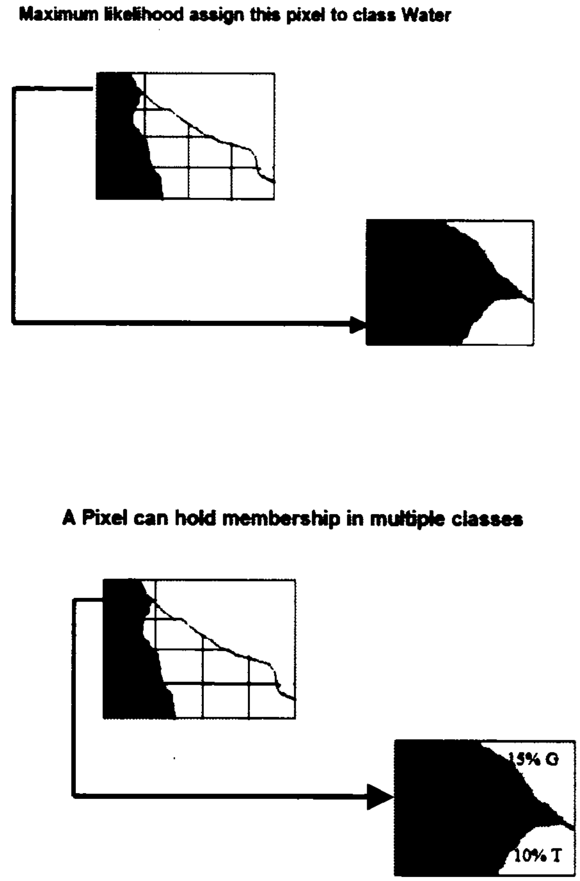


Figure 1. An illustration of a hard and fuzzy classification

In addition to the partial membership in different groups, each rule in a fuzzy rule base contains a strength or weighting or certainty parameter (Burrough and McDonnell, 1998, p.269). In a fuzzy rule base, rules can be triggered simultaneously, even though these triggered rules may act against each other. However since the level of strength or certainty triggers each rule, a decision can be made in favor of the rule or rules that contains the greatest strength (Tso, 2001, p.159).

Fuzzy Sets

A fuzzy set is determined by membership function to each element (Klir, 1995, p. 12). If we let S represent a universal set of generic elements, and let s represent a fuzzy set, the membership grade can range between zero and one, as expressed by:

$s: S \rightarrow [0,1]$

A fuzzy rule base generally comprises three principal steps, as shown Figure 2.

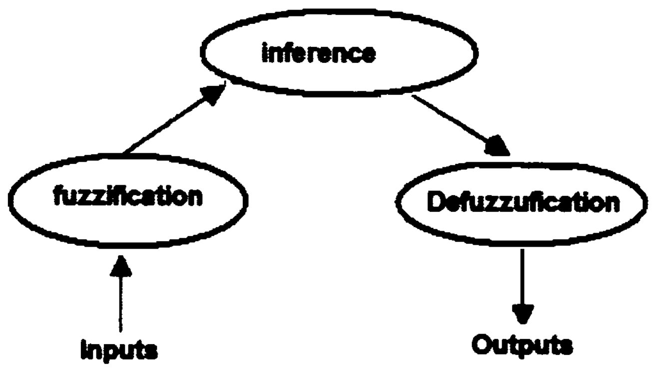


Figure 2. Three steps of Fuzzy rule base method

The first step is *Fuzzification*, which involves the division of the input into fuzzy sets, defined by fuzzy membership function. The second step, *Inference*, requires the combination of fuzzy sets in logical rules, to produce calculations of the strength of each rule being triggered. The final step, *Defuzzification*, combines all triggered rules and generates a non-fuzzy, crisp outcome. The study described here does not utilize this third step, since the classification is left in terms of partial membership values. This study will therefore describe common methods of Fuzzification and inference only.

Fuzzification

To carry out the process of Fuzzification, a membership function has to be defined in order to calculate the membership grade for the input pixel. Figure 3 illustrates six common membership functions that are most frequently used in fuzzy rule base practice: monotonic (linear), triangular, trapezoidal, bell shaped, sigmoidal and j-shaped functions.

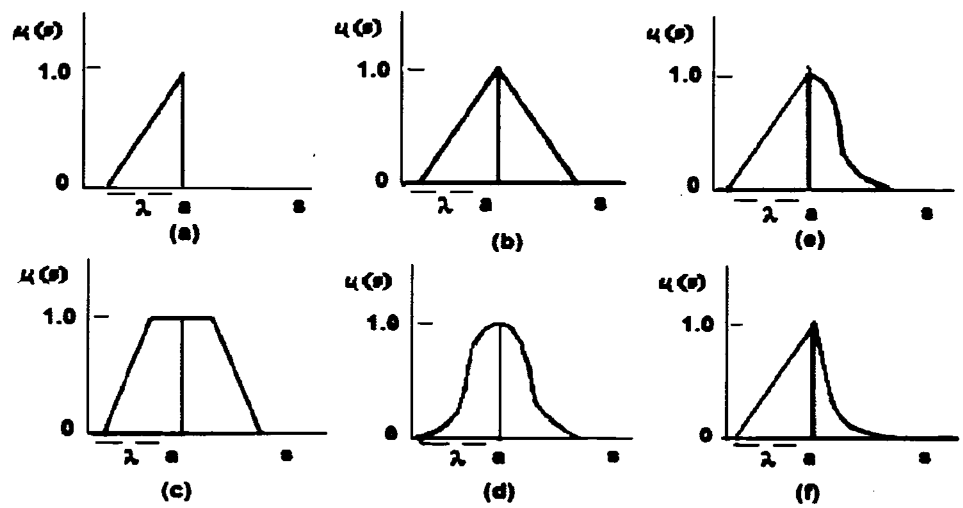


Figure 3. Six common types of fuzzy membership functions: (a) monotonic (linear), (b) Triangular, (c) trapezoidal, (d) bell-shaped, (e) sigmoidal, (f) j-shaped.

The mathematical descriptions of the most common functions are:

- Monotonic function: $\mu(s)=1-(a-s/\lambda)$, for $0 \le a-s \le \lambda$; $\mu(s)=0$
- Triangular function: $\mu(s)=1-(|s-a|/\lambda)$, for $0 \le |a-s| \le \lambda$; $\mu(s)=0$
- Trapezoidal function: $\mu(s)=\min\{2-(2(|s-a|/\lambda),1\}$, for $c-\lambda \le |s-a| \le a+\lambda$; $\mu(s)=0$

- Bell shaped function: $\mu(s)= 2\cdot(1-(|s-a| / \lambda))^2$, for $(\lambda/2) \leq |s-a| \leq \lambda$;

  $\mu(s)= 1 - 2\cdot(1-(|s-a| / \lambda))^2$, for $0 \leq |s-a| \leq (\lambda/2)$;

Fuzzy partitions can be performed using combinations of different, possibly overlapping membership function as illustrated in Figure 4.

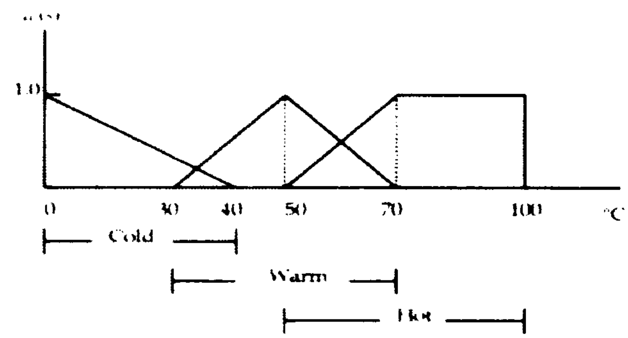


Figure 4. Fuzzy partitions establishing degrees of hot, warm and cold, (Tso, 2001, P. 163)

Inference

Fuzzy sets can be manipulated using logical query methods to select and combine data from several sets. The basic operations on fuzzy sets are similar and are a generalization of the AND/OR/NOT/XOR used for Boolean sets. The most used operations are union (maximize), intersection (minimize), negation (complement), and the convex combination (weighted sum). All these operations lead to the computation of a new Membership Function value, which is called the *Joint membership function* value or *JMF* (Burrough and McDonnell. 1998, p.272). This study will use the intersection (Minimum) operation.

Methodology

The theoretical basis for Fuzzy sets is well established and already used in different fields such as fuzzy logic control, fuzzy neural networks and fuzzy rule base. In remote sensing classification, fuzzy-based classifiers are becoming increasingly popular. Most of the research methods use fuzzy algorithms such as *c-mean clustering, fuzzy maximum likelihood,* and *fuzzy rule base.* The use of these algorithms is based on extracting statistics from the spectral distributions of each class, and forming fuzzy sets on the basis of class mean and variance. The sets are most often Gaussian distributions (e.g., Melgani et al. 2000). This study takes the view that the simplifying assumption of Gaussian distributions is unnecessary. It can result in a lack of flexibility in the methodology, and ultimately in greater error rates than would result from a less restrictive methodology in the formation of fuzzy sets. Literature also reveals that none of the researches derive the values of the pixel directly from histograms. This study extracts radiance values directly from histograms, to be used in the formation of fuzzy sets employed by the rule base of the classification methodology. Figure 5 illustrates extraction of radiance values directly from histograms and the threshold values used to describe trapezoidal fuzzy sets established for each band. Each histogram describes the radiance values of pixels previously classified as a particular land cover (either impervious surface or canopy) in the NLCD 2001 classification. In other words, the NLCD 2001 classification is used as a mask to extract histograms from the Landsat data, on a per-class basis. This is described in more detail below.

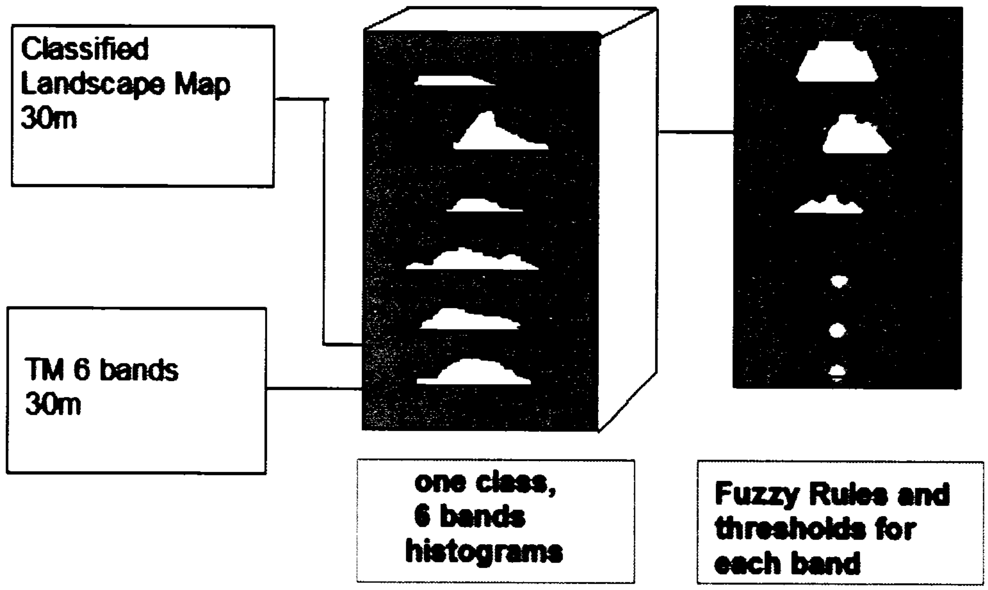


Figure 5. An overly of two images, the classified landscape and raw data of six bands, are used in order to establish the fuzzy rules.

Data Description

Data used for this research are obtained from the NLCD 2001, which is a Landsat, based land-cover database with several independent data layers, downloaded from the website http://seamless.usgs.gov. The layers used in this study are:

- Per-pixel estimates of percent imperviousness surface for the San Francisco Bay Area.
- Percent tree canopy for the San Francisco Bay Area.
- Landsat ETM+ full scene of the San Francisco Bay Area, encompassing six spectral bands.

Impervious surfaces refer to surfaces that are generally impenetrable to rainfall, such as rooftops, roads, or parking lots. Tree canopy refers the branches and foliage at the top or crown of a forest's trees. All layers were made geographically compatible by windowing each to precisely the same extent, and all projections were transformed to the longitude/ latitude coordinate system. All layers describe a rectangular region in the Bay Area, from 37.87 degrees latitude, -122.60 degrees longitude at the upper left corner, to 37.00 degrees latitude, -121.40 degrees longitude at the lower right corner.

Impervious Surface and Tree Canopy

Currently the NCLD 2000 classification project has not yielded a complete land cover land use classification. However, intermediate classifications have been performed and are available online, as described previously. These are Impervious Surface and Tree Canopy.

Imperviousness surfaces refer to impenetrable surfaces such as rooftops, roads, or parking lots (Homer *et al.* 2004). Quantification of imperviousness can be used as a measure of urban density. For NLCD 2001, Imperviousness was chosen as a classification for urban intensity in order to improve the precision of urban characterization used in the original NLCD 1992 (Homer *et al.* 2004). The method of classifying Landsat data into impervious class involved a regression tree technique. A one-meter resolution digital othophotos are used for each Landsat scene to derive reference to impervious data and Landsat spectral data. A regression tree algorithm was used and the model was applied to all pixels. This model produced a per-pixel estimate of imperviousness in urban areas. This 1 to 100 percent spatial distribution estimation layer was then masked to ensure only urban pixels are included.

The original NLCD 1992 classification provided four forest categories but made no distinction in forest canopy density (Homer *et al.* 2004). The NLCD 2001 developed a strategy for estimating tree canopy density at a spatial resolution of 30m using the same regression tree technique described above. The final product is a per-pixel 1 to 100 percent estimate of the canopy density.

Creation and Combination of Fuzzy Sets

A large region around San Francisco Bay was used in order to create the fuzzy sets and obtain rules. Results were then compared with other regions to the north and east. The impervious surface and canopy data were used to establish the fuzzy sets used here. These sets are generally derived from the shapes of confidence-weighted histograms, which are unique to this study. Each spectral band of the ETM+ data is used to create histograms associated with each of two classes: impervious surface and canopy. In order to extract radiance values from the radiance data associated with the impervious surface and canopy pixels, several steps were taken:

1. A mask was created Appendix A, canopyRadConfMacro). This mask is used to assign a value of -0- to original values to be ignored and the value of -1- to original values to be included. The mask is used to create ranges of 0-10, 10-20, 20-30, 30-40, 40-50, 50-60, 60-70, 70-80, 80-90, 90-100. These are confidence ranges.

2. The mask was applied to the impervious bands 1 to 6, creating intermediate layers of all ranges of class impervious in each band:

    a. Band 1, ranges 0-10, 10-20, 20-30 etc.

    b. Band 2, ranges 0-10, 10-20, 20-30, etc.

   At that point all ranges have equal weight. Overlaying the mask over each band created a count of pixels for each confidence range as explained in Appendix 1.

3. A numeric histogram was created from the actual values of the original layers of each ETM+ band, as masked in step 2 (Appendix B, A numeric histogram).

4. Out of the numeric histogram, only the frequencies of each range were imported to Microsoft Excel. The numeric histogram produced 256 classes of values. All the 0 value frequencies were eliminated in order to simplify the histogram by showing values with data only (Figure 6).

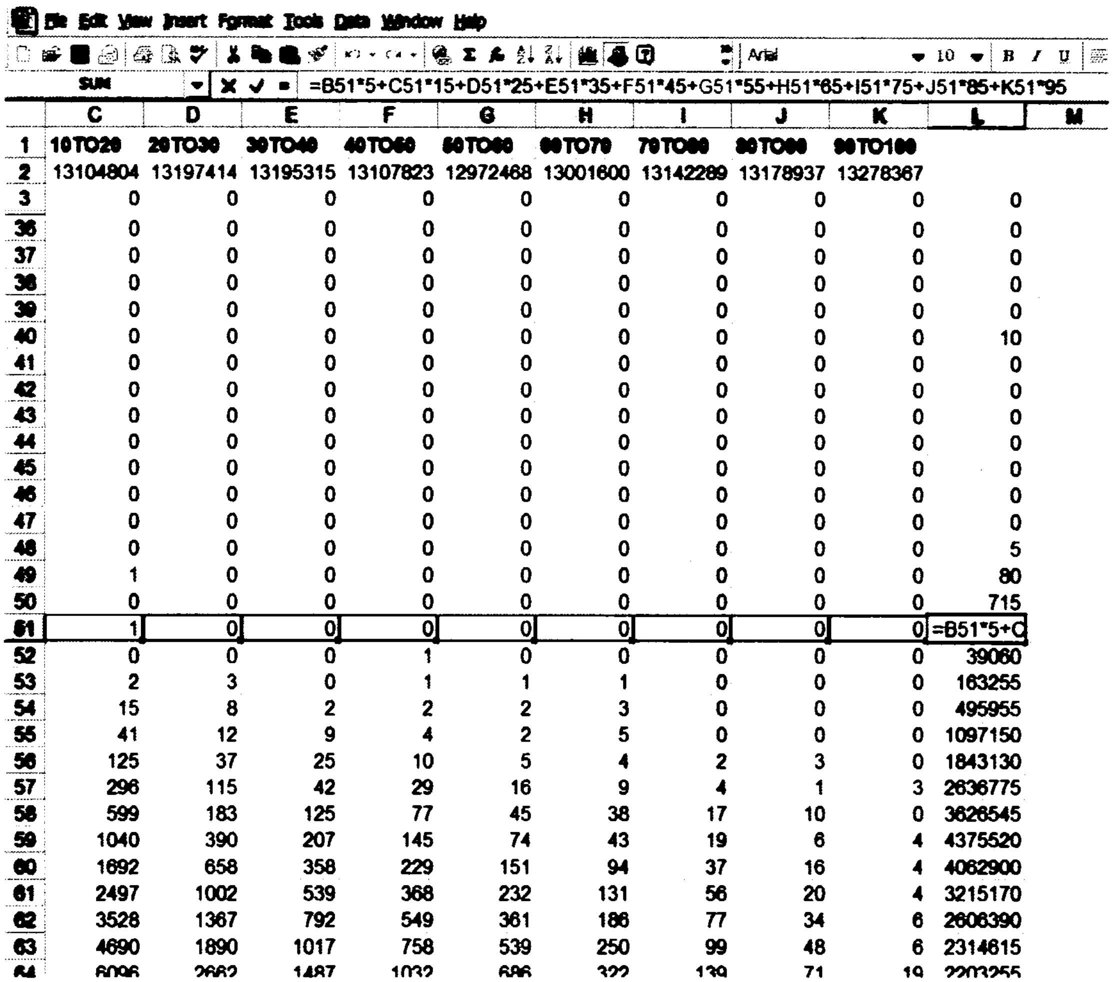
5. A weight was given to each range by multiplying the confidence values with the average of the range:

   Range 0-10 values were multiplied by 5

   Range 10-20 values were multiplied by 15 etc.

   Multiplying with the average range number created a *confidence weighed pixel count* for each range at each band. The use of weight creates a direct comparison with the NLCD 2001 results since their results are percent confidence to 100. Figure 6 shows the weight calculation for impervious layer band 1 that was created in Excel.

6. A histogram was created out of the confidence weighted pixel count for each band. This type of confidence-weighted histogram is unique to this study.



File Edit View Insert Format Tools Data Window Help

SUM =B51*5+C51*15+D51*25+E51*35+F51*45+G51*55+H51*65+I51*75+J51*85+K51*95

	C	D	E	F	G	H	I	J	K	L	M
1	10TO20	20TO30	30TO40	40TO50	50TO60	60TO70	70TO80	80TO90	90TO100		
2	13104804	13197414	13195315	13107823	12972468	13001600	13142289	13178937	13278367		
3	0	0	0	0	0	0	0	0	0	0	
36	0	0	0	0	0	0	0	0	0	0	
37	0	0	0	0	0	0	0	0	0	0	
38	0	0	0	0	0	0	0	0	0	0	
39	0	0	0	0	0	0	0	0	0	0	
40	0	0	0	0	0	0	0	0	0	10	
41	0	0	0	0	0	0	0	0	0	0	
42	0	0	0	0	0	0	0	0	0	0	
43	0	0	0	0	0	0	0	0	0	0	
44	0	0	0	0	0	0	0	0	0	0	
45	0	0	0	0	0	0	0	0	0	0	
46	0	0	0	0	0	0	0	0	0	0	
47	0	0	0	0	0	0	0	0	0	0	
48	0	0	0	0	0	0	0	0	0	5	
49	1	0	0	0	0	0	0	0	0	80	
50	0	0	0	0	0	0	0	0	0	715	
51	1	0	0	0	0	0	0	0	0	=B51*5+C	
52	0	0	0	1	0	0	0	0	0	39060	
53	2	3	0	1	1	1	0	0	0	163255	
54	15	8	2	2	2	3	0	0	0	495955	
55	41	12	9	4	2	5	0	0	0	1097150	
56	125	37	25	10	5	4	2	3	0	1843130	
57	296	115	42	29	16	9	4	1	3	2636775	
58	599	183	125	77	45	38	17	10	0	3626545	
59	1040	390	207	145	74	43	19	6	4	4375520	
60	1692	658	358	229	151	94	37	16	4	4062900	
61	2497	1002	539	368	232	131	56	20	4	3215170	
62	3528	1367	792	549	361	186	77	34	6	2606390	
63	4690	1890	1017	758	539	250	99	48	6	2314615	
64	6096	2682	1487	1032	686	322	139	71	19	2203255	

Figure 6. Impervious band 1, weight calculation.

7. Minimum and maximum values were established out of the histograms of each band in Idrisi32 macro commands (Appendix A: impervminrule). The AND

rule was used since the weakest membership grade of all bands for a specific class will determine the strength of the group as a whole. Figure 7 represents two classes with a signature over two bands. The weakest grade class urban, which is 0.2, determines the value of class urban as a whole; for class trees the 0.0 value determines the value of the class as a whole.

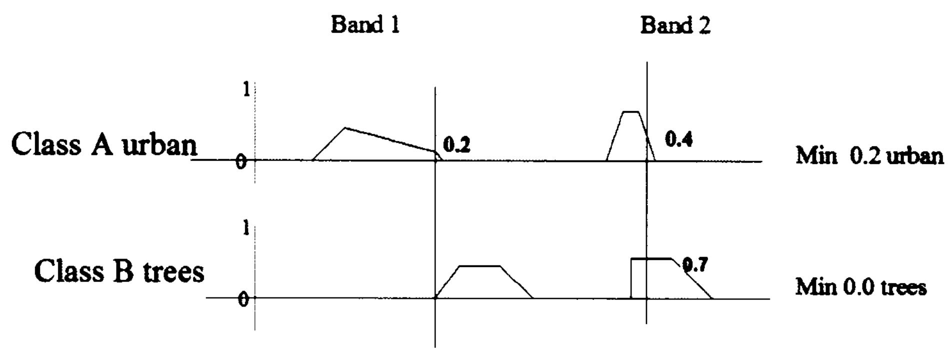


Figure 7. Class Urban min value = 0.2, Class Trees min value = 0.0

Figure 8 represents a two dimensional fuzzy rule for two bands, band 1 and band 2. If the MAX operation were to be performed (an OR rule), the highest value was to be taken, value of 1.0. However, using the MIN operation (an AND rule) the intersection of the two values occurs at a weaker point (yellow) because of the 0.5 value of band 1.

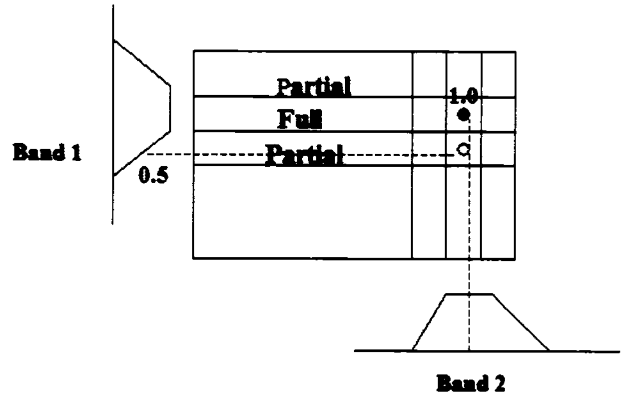


Figure 8. A two dimensional view of fuzzy sets.

Some histograms had two peaks, which required a performance of fuzzy partitions of two membership functions; Figure 9 represents impervious data on band six histogram with two peaks and its two fuzzy rules (all remaining histograms are shown in Figures 10 through 21). In addition, matching fuzzy sets to histograms requires decisions regarding the use of linear, j-shaped, or bell-shaped functions as segments of the 'trapezoidal' functions, as shown for example in Figure 9.

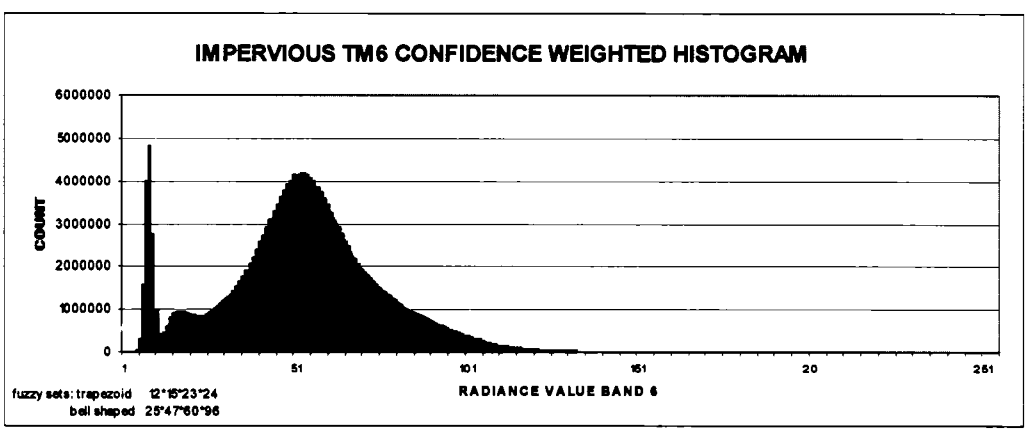


Figure 9. Two fuzzy rules threshold. First Lower thresholds at 12 and 24. Upper thresholds at 15, 23. Second threshold, lower at 25 and 96, upper at 38 and 70.

The combination of the membership functions of all bands was made by Idrisi32 macro command (Appendix: overlayminrule).

8. The final image was filtered with a mean 3x3 low pixel filter to clear the effect of salt and pepper (isolated single pixels of a class different from all its neighbors).
9. The final image was scaled to 100 so that NLCD results, which are percentage confidence to 100%, and this study's results are comparable.
10. The whole process was repeated with the NLCD canopy layer.
11. The resulting fuzzy classifications were visually compared with the NLCD classifications. Where there were significant discrepancies, the fuzzy sets were iteratively adjusted to yield better results. This process is described in more detail below.
12. After satisfactory performance was achieved with the training subscene, results were compared in a region of the North Bay that was not used in the above process.

Iterative Adjustment of Fuzzy Sets

A comparison with the NLCD results revealed that the initial classifications performed using this method included assignments to impervious and to canopy classes that obviously do not belong to those classes. An iterative adjustment to the fuzzy sets was made by narrowing the top plateau of the trapezoidal fuzzy sets that have high membership and by widening the base of the sets where values have a low membership.

In the final iterations adjustments were performed through comparison with both the USGS results and by direct interpretation of aerial photos. Composite images were made out of ETM+ bands 4, 5, and 6. A site surrounding the Mineta San Jose International Airport was chosen for the extraction of the Red Green Blue values of the aerial photo. According to those values the fuzzy sets were adjusted to include or exclude values from the membership function.

Histograms for Canopy bands 1-6

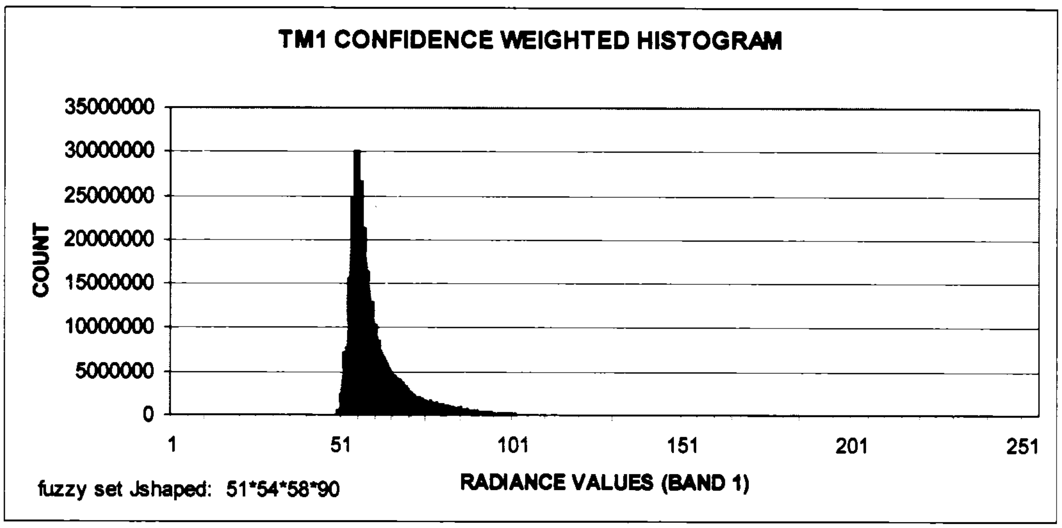


Figure 10. Canopy band-1 histogram of confidence weighted pixel count

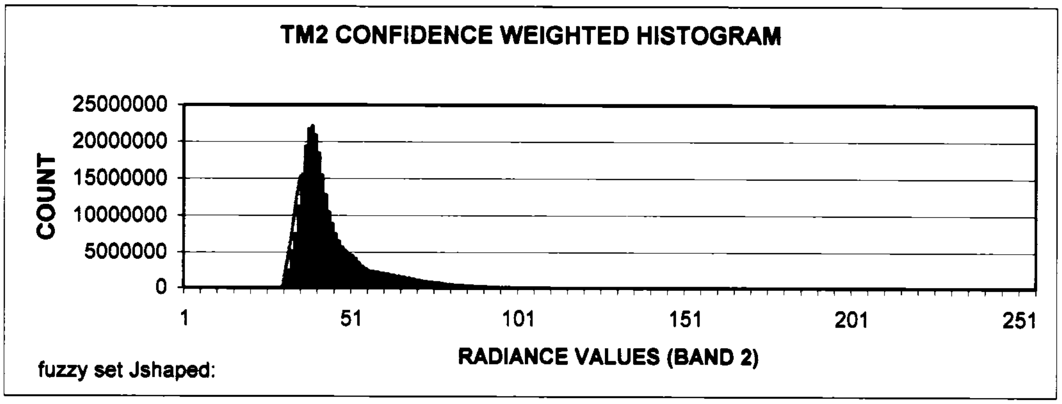


Figure 11. Canopy band-2 histogram of confidence weighted pixel count

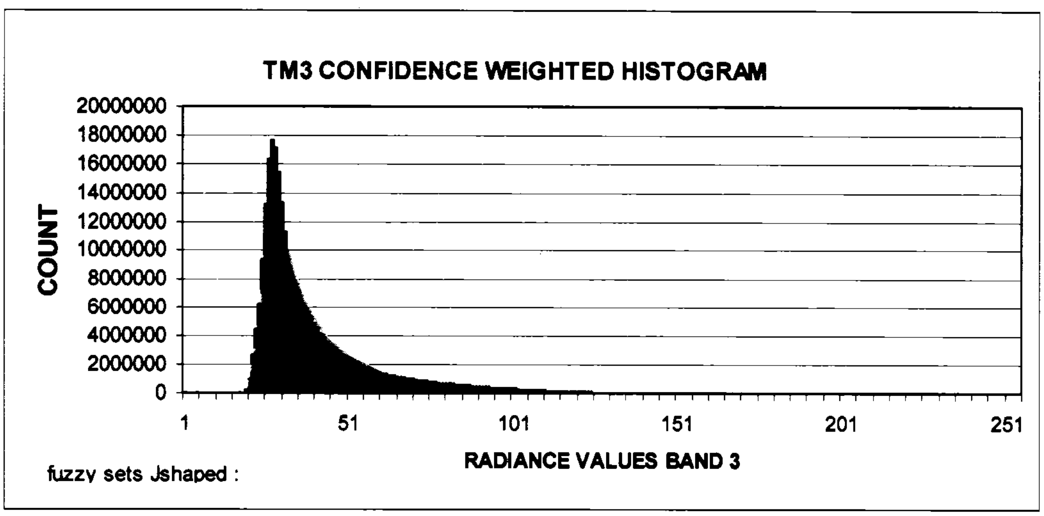


Figure 12. Canopy band-3 histogram of confidence weighted pixel count.

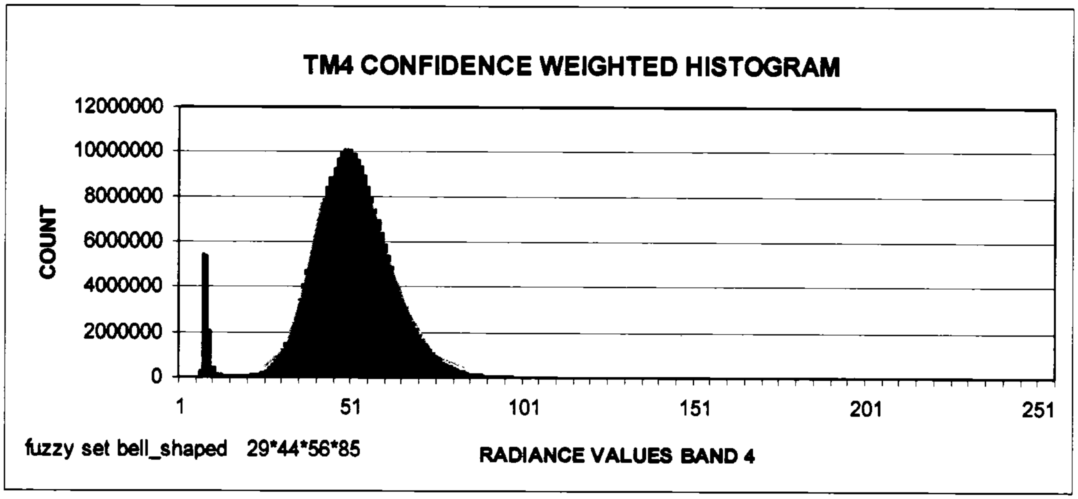


Figure 13. Canopy band-4 histogram of confidence weighted pixel count

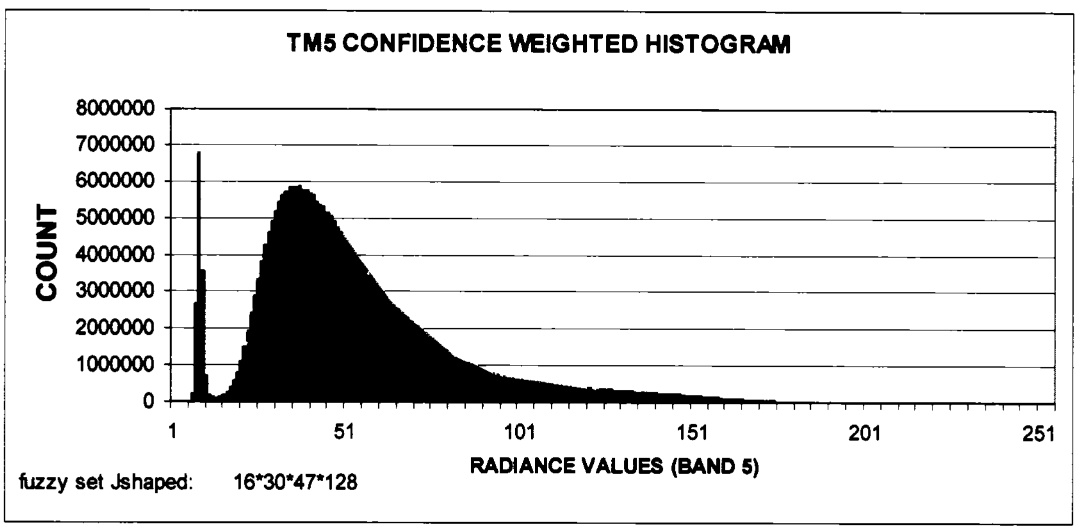


Figure 14. Canopy band-5 histogram of confidence weighted pixel count

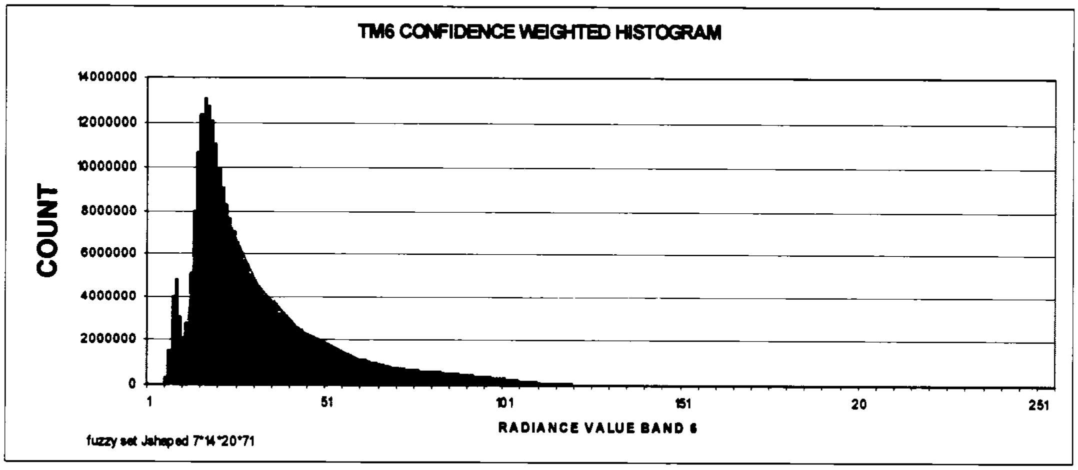


Figure 15. Canopy band-6 histogram of confidence weighted pixel count

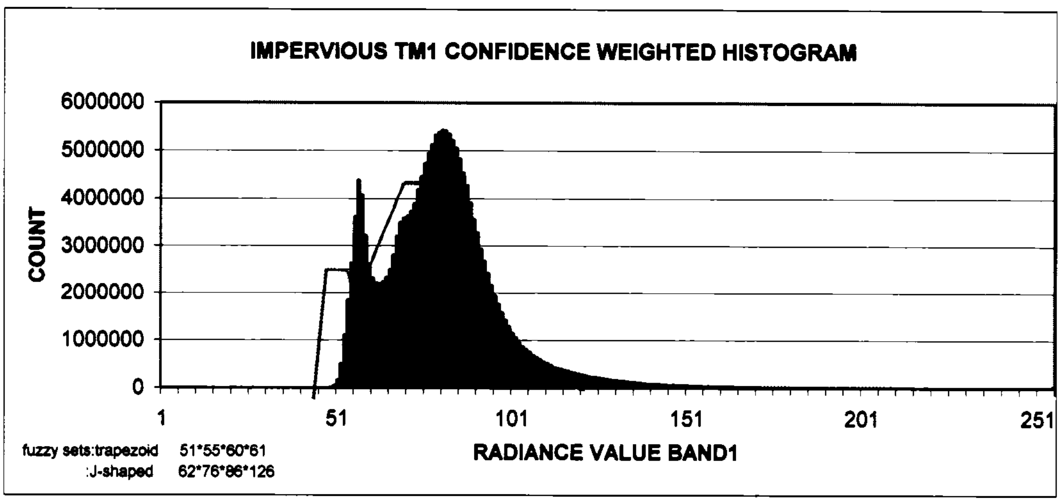


Figure 16. Impervious band-1 histogram of confidence weighted pixel count

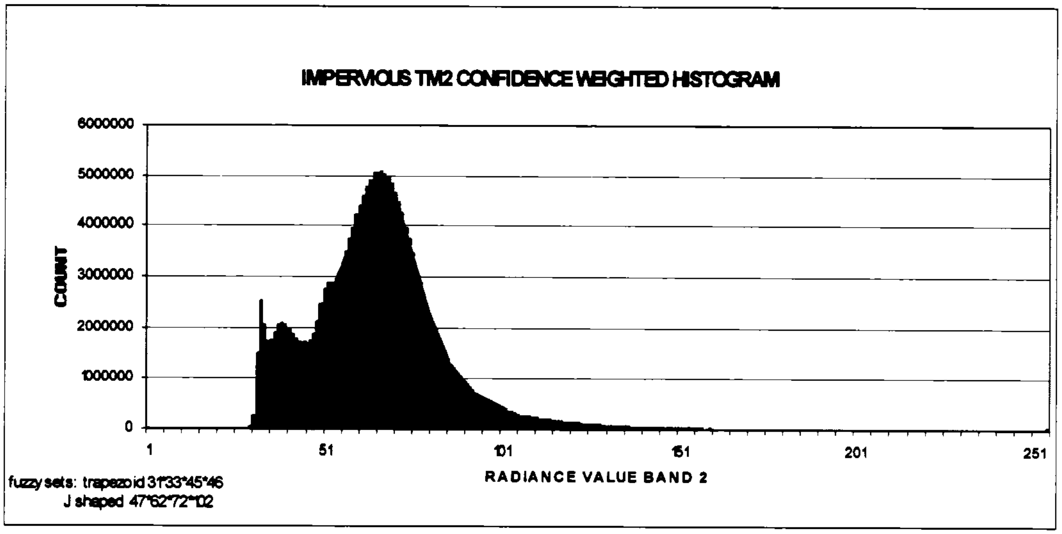


Figure 17. Impervious band-2 histogram of confidence weighted pixel count

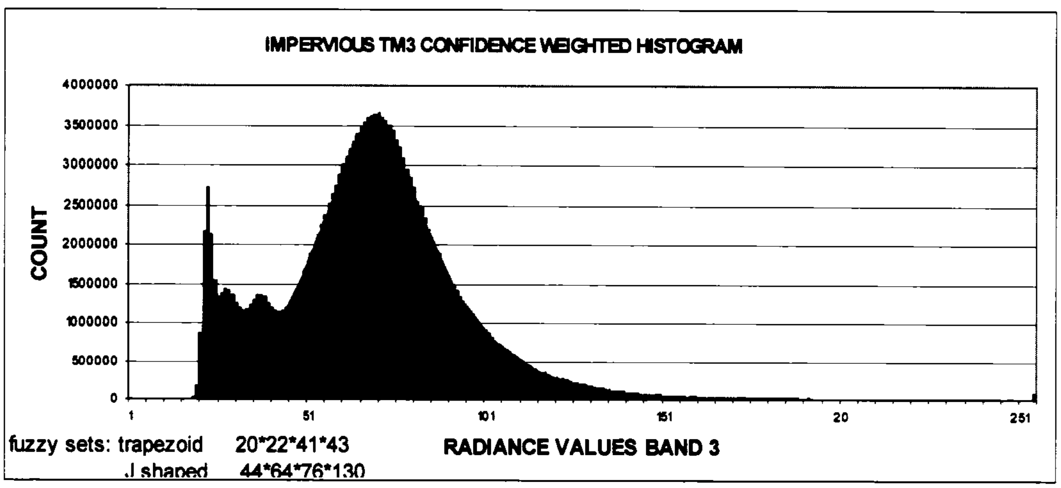


Figure 18. Impervious band-3 histogram of confidence weighted pixel count

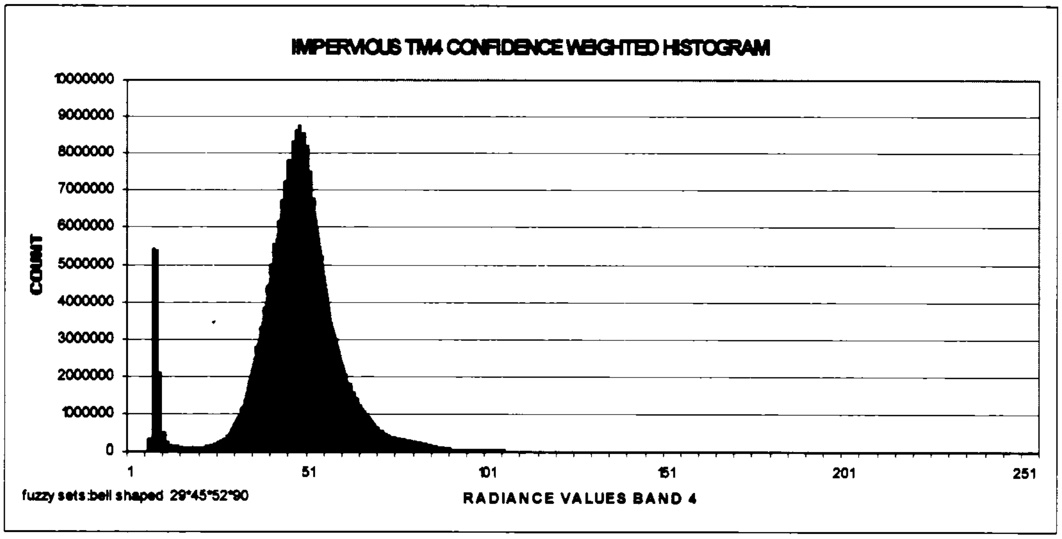


Figure 19. Impervious band-4 histogram of confidence weighted pixel count

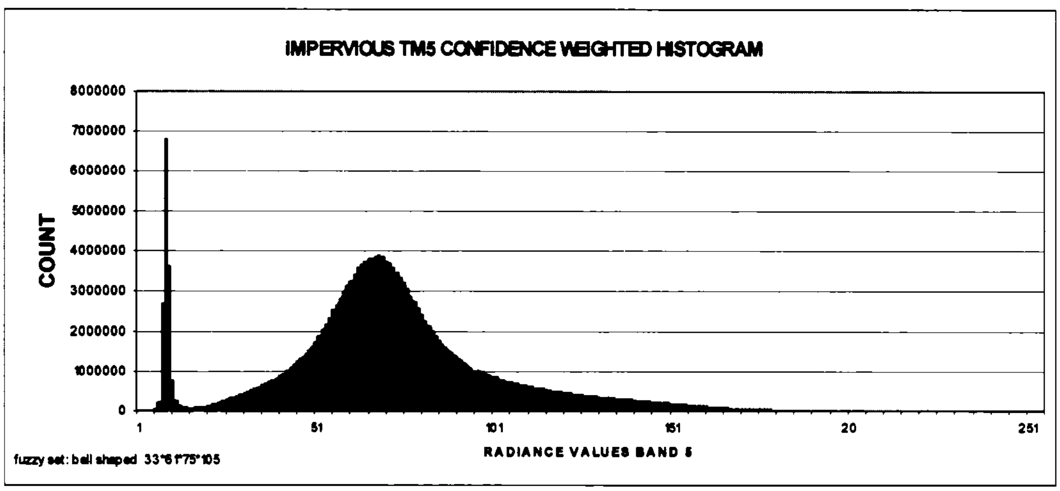


Figure 20. Impervious band-5 histogram of confidence weighted pixel count

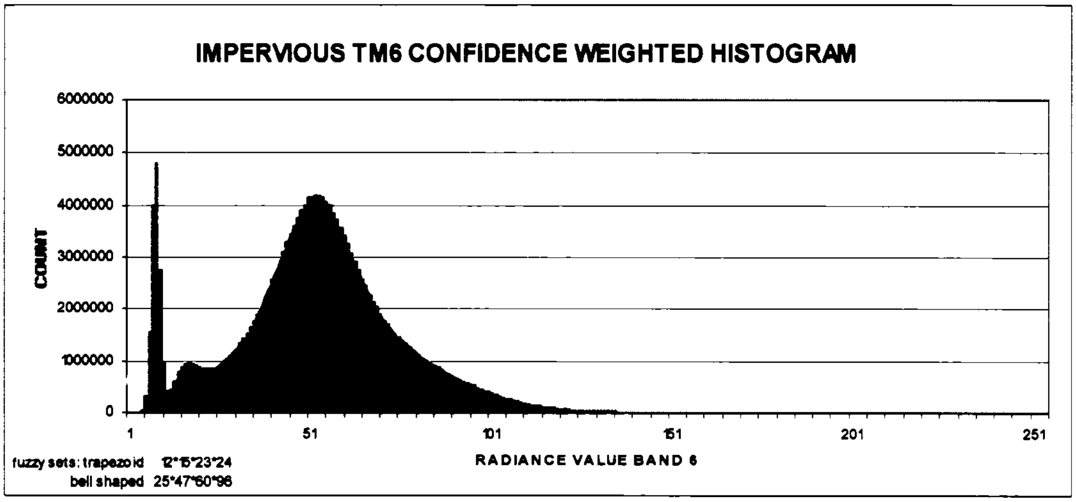


Figure 21. Impervious band-6 histogram of confidence weighted pixel count

Results of Training Data

A large sub scene centered on San Jose (Figure 22) was used to generate the confidence-weighted histograms and in creating the fuzzy sets used in this methodology. Although this subscene and the corresponding NLCD classifications may therefore be considered as "training" data, they are only so in a statistical sense. No per-pixel training was used. We can therefore expect to see some significant differences between the NLCD and the fuzzy classification even within this subscene.

Figure 22. Natural Color image of San Francisco Bay Area subscene used for training.

Figure 23 presents a broad scale comparison of the results of the NLCD and fuzzy classifications for the Canopy class, over the training set. Figure 24 presents the same comparison for the Impervious class. The confidence levels and fuzzy memberships in each of these images are expressed through a continuous gray shade palette. At this broad scale, we can view the results in general terms. Each classification strategy yields results that are broadly correct, although they differ statistically and morphologically. For the Canopy classification, the fuzzy method may be affected by terrain and shadowing, although it does seem to indicate the presence of forest in the region inland of Half Moon Bay better than the NLCD method. For the Impervious surface classification, the NLCD method seems to indicate the presence of roads and lots better than the fuzzy method. This may be due to an under-representation of fuzzy membership in spectral values returned by either asphalt or concrete, and may therefore be improved by further iteration. However, the city of Tracy seems to have been entirely unrepresented in the NLCD classification. The reason for this remains unknown; it may be due to errors in the creation of the data set that was downloaded. In general, the fuzzy method yields results that are more fragmented, with a "salt and pepper" appearance than the NCLD classifications, which are more homogeneous in appearance. Comparison of these broad scaled images cannot yield an absolute determination of which is more accurate; a closer examination of specific areas is required, as well as careful comparison with very high-resolution color orthophotographs at a number of locations. These results are discussed later.

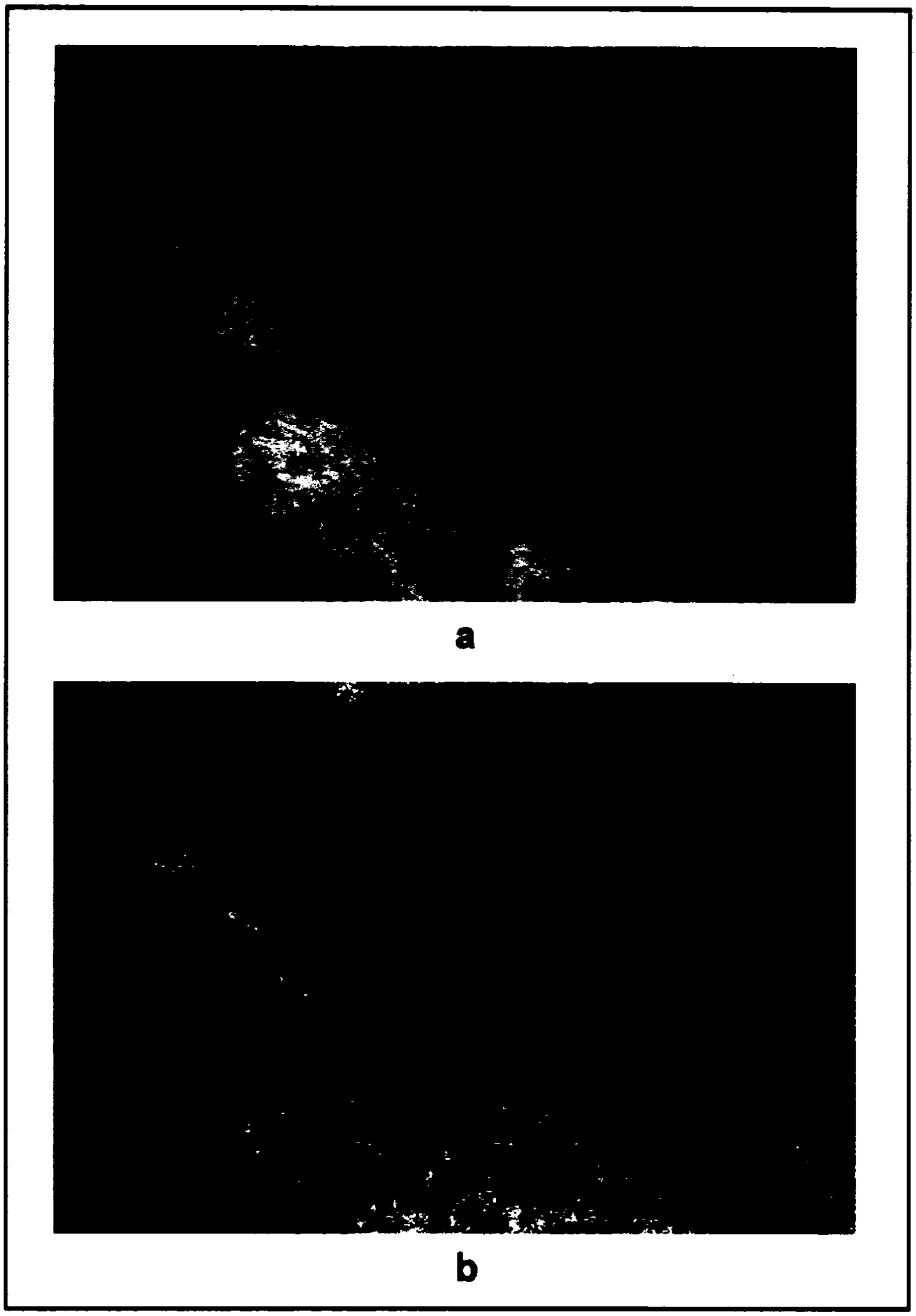


Figure 23. a) NLCD Canopy Classification; b) Fuzzy Canopy Classification. Brighter values indicate (a) greater confidence and (b) greater class membership. Lower left coordinates: 122.597461 W longitude, 37.831027 N latitude.

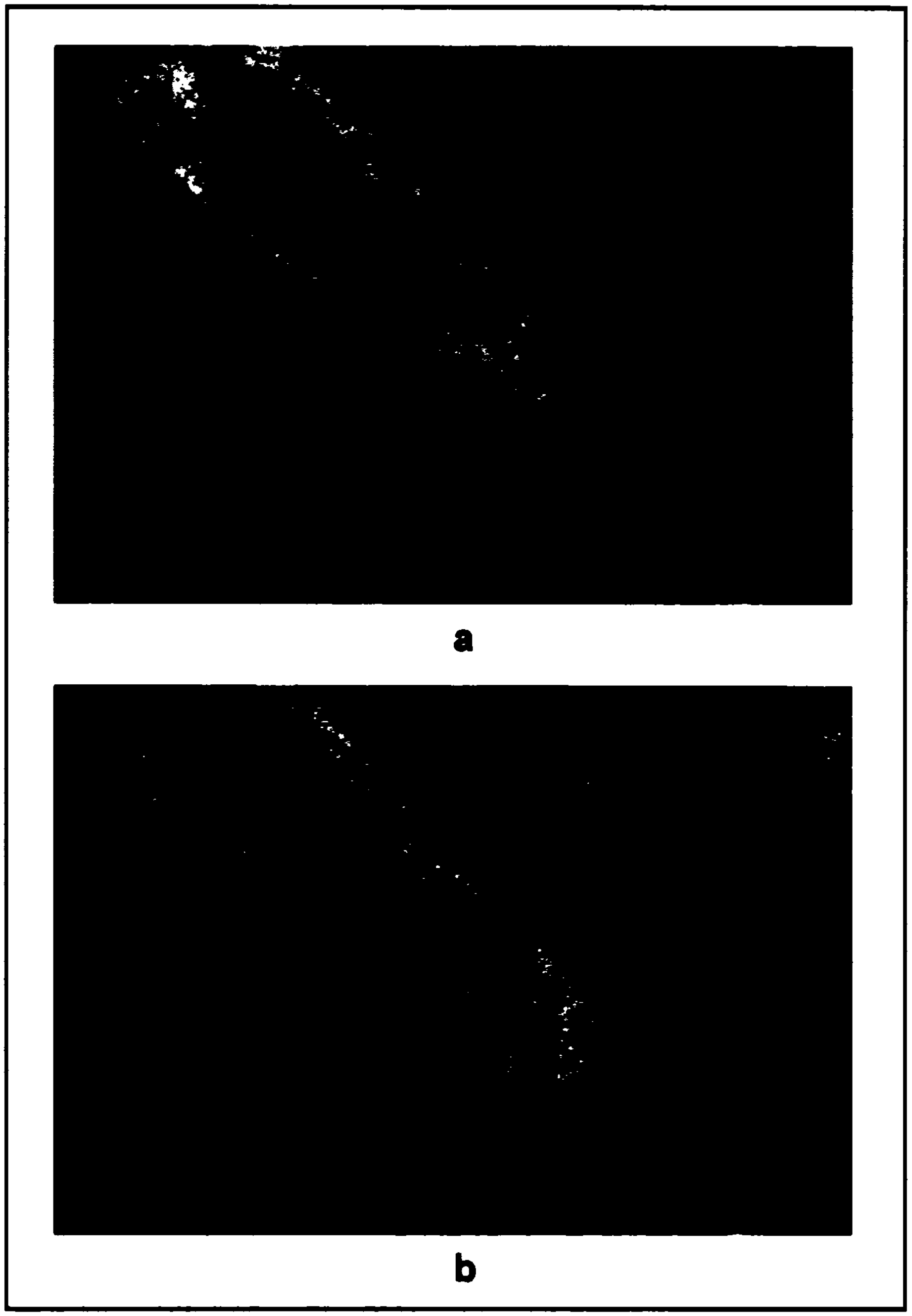


Figure 24. a) NLCD Impervious Classification; b) Fuzzy Impervious Classification. Brighter values indicate (1) greater confidence and (2) greater class membership. Lower left coordinates: 122.597461 W longitude, 37.831027 N latitude.

Looking at particular regions more closely, we can see that the two methods yield results that are more similar than anticipated. Figure 25 presents Canopy results in an area from San Jose southeastward. Note that the gray scales of the NCLD and fuzzy results do not match; this is a matter of normalization in palette assignment. It also remains difficult to detect and compare partial membership or confidence levels in either image. However, the detailed presence of Canopy is quite similar between methods. It even appears that the fuzzy method may have been affected less by terrain artifacts than the NLCD method, in the hills at the lower left.

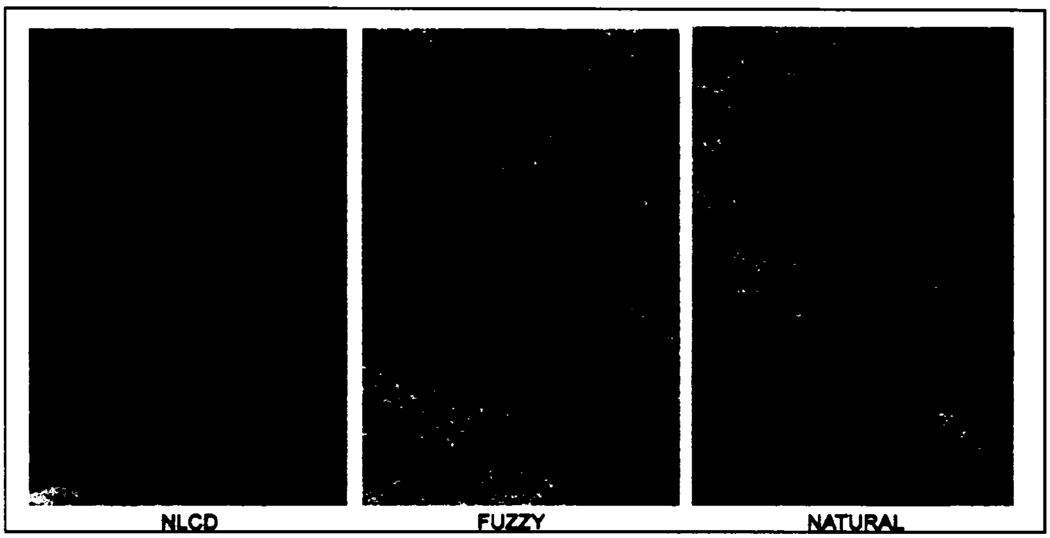


Figure 25. Canopy Classification of a region extending southeastward of San Jose. Lower left coordinates: 121.906072 W longitude, 37.358640 N latitude.

The fuzzy Impervious classification appears generally successful at this moderate scale. However, two items of concern become apparent when looking at the region from Livermore to Tracy (Figure 26). While the fuzzy method does successfully capture most of the main arteries, it is not as successful at capturing road surfaces as the NLCD method. In addition, it seems to misclassify some agricultural fields as impervious. The absence of Tracy in the NCLD image is probably the result of data drop somewhere in the file creation and transfer process.

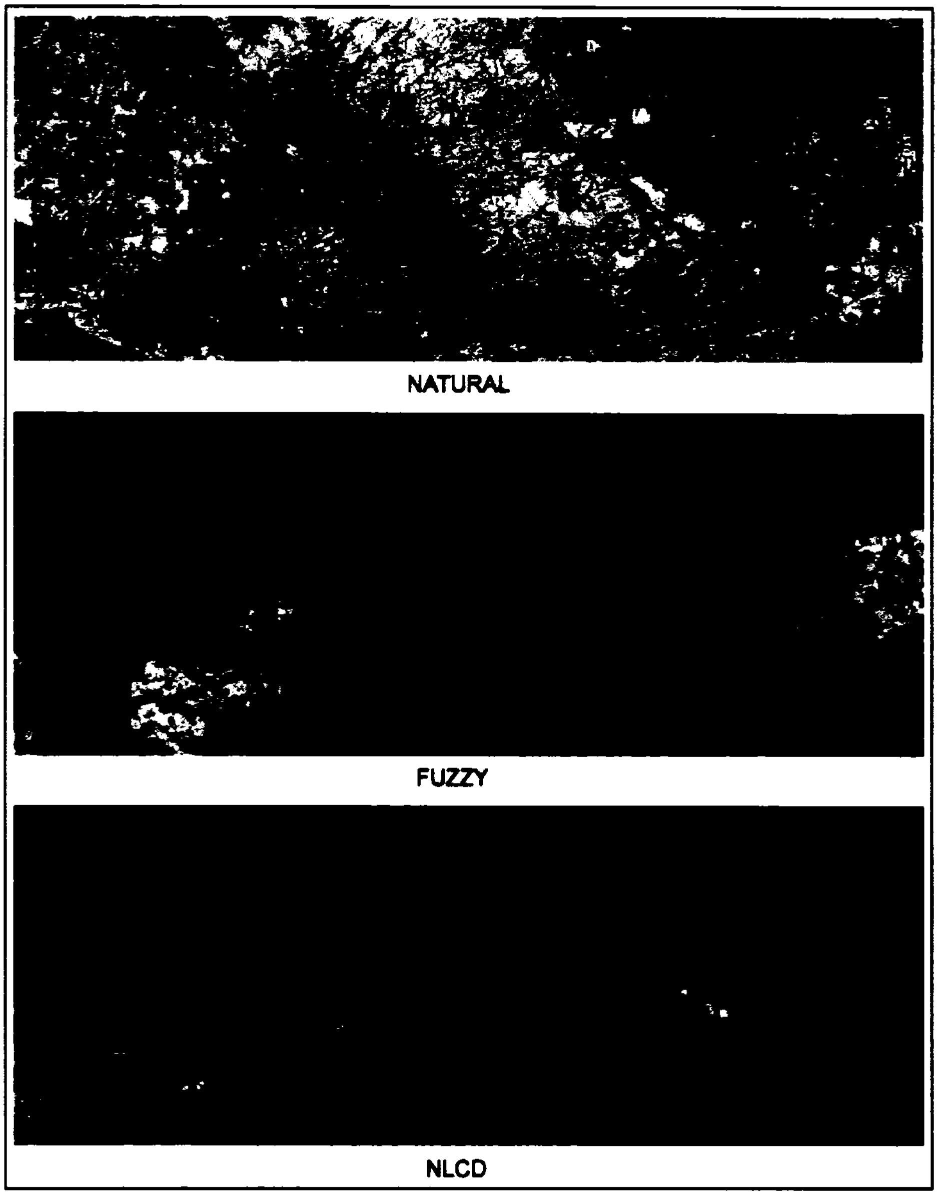


Figure 26. Impervious classification, from Livermore to Tracy.
Natural color image, Fuzzy, and NLCD. Note the failure of NLCD in classifying most of Tracy. Lower left coordinates: 121.814797 W longitude, 37.820639 N latitude.

Moving westward into Pleasanton (Figure 27), the fuzzy and NLCD results are more similar, although the misclassification of some agricultural locations in the fuzzy image is a matter of concern. The fuzzy classification within urban areas is also patchier than the NLCD classification. Their relative veracity can only be resolved through comparison with higher resolution data or ground surveys.

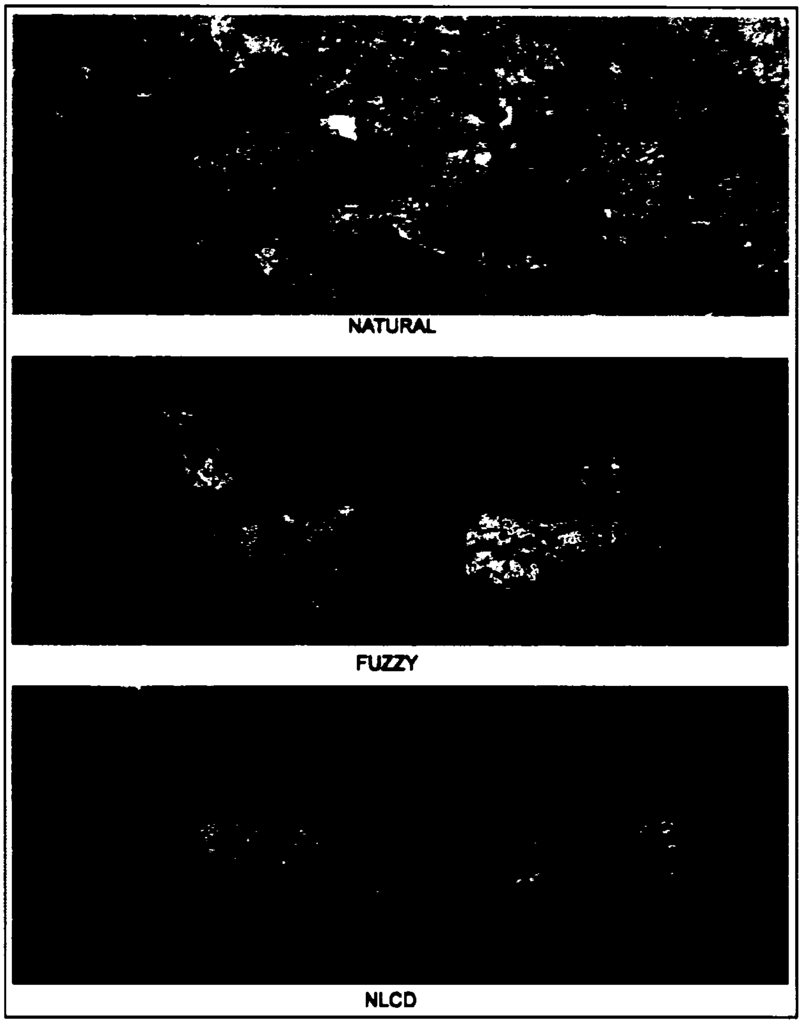


Figure 27: Impervious surface Lower left coordinates: 121.981936 W longitude, 37.770685 N latitude.in the vicinity of Pleasanton.

Figure 28 presents the respective classifications of impervious surface in the area of the San Jose International Airport. This discrepancy within urban areas becomes more apparent. The fuzzy results miss some obvious impervious details such as light industrial buildings in the area north and northwest of the airport, shown as bright structures in the natural color image, while the NLCD classification presents them more clearly. However, while the roadways are successfully captured, the airport tarmac is not captured. This confirms the earlier suspicion that either concrete or asphalt spectral returns were not used properly in the formation of fuzzy sets. Examination of the area around Milpitas and Fremont (Figure 29) provides further confirmation that light industrial buildings (bright returns in the natural color image) are not successfully captured, while asphalt roadways are in the fuzzy classification.

A couple of additional observations can be made here. The fuzzy method misclassifies a certain type of bare soil surface adjacent to the bay as impervious (but of low membership). This may be related to the misclassification of certain agricultural fields in earlier images. However, the NCLD classification may also have some problems here, indicating as impervious some of the dikes separating evaporation ponds. The materials there are actually earthen mounds topped with crushed stone (to be used as limited roadways), and it is probably not accurate to classify these surfaces as impervious.

Comparing roadway surfaces of large arteries, the concrete roadway of I 680 curving off into the hills at right is properly classified in both, while the wide asphalt surface of I 880 extending from lower right to upper left (and dark in the natural color image) is not successfully classified by the fuzzy method. This indicates that asphalt is the substance that is not properly represented in the fuzzy sets.

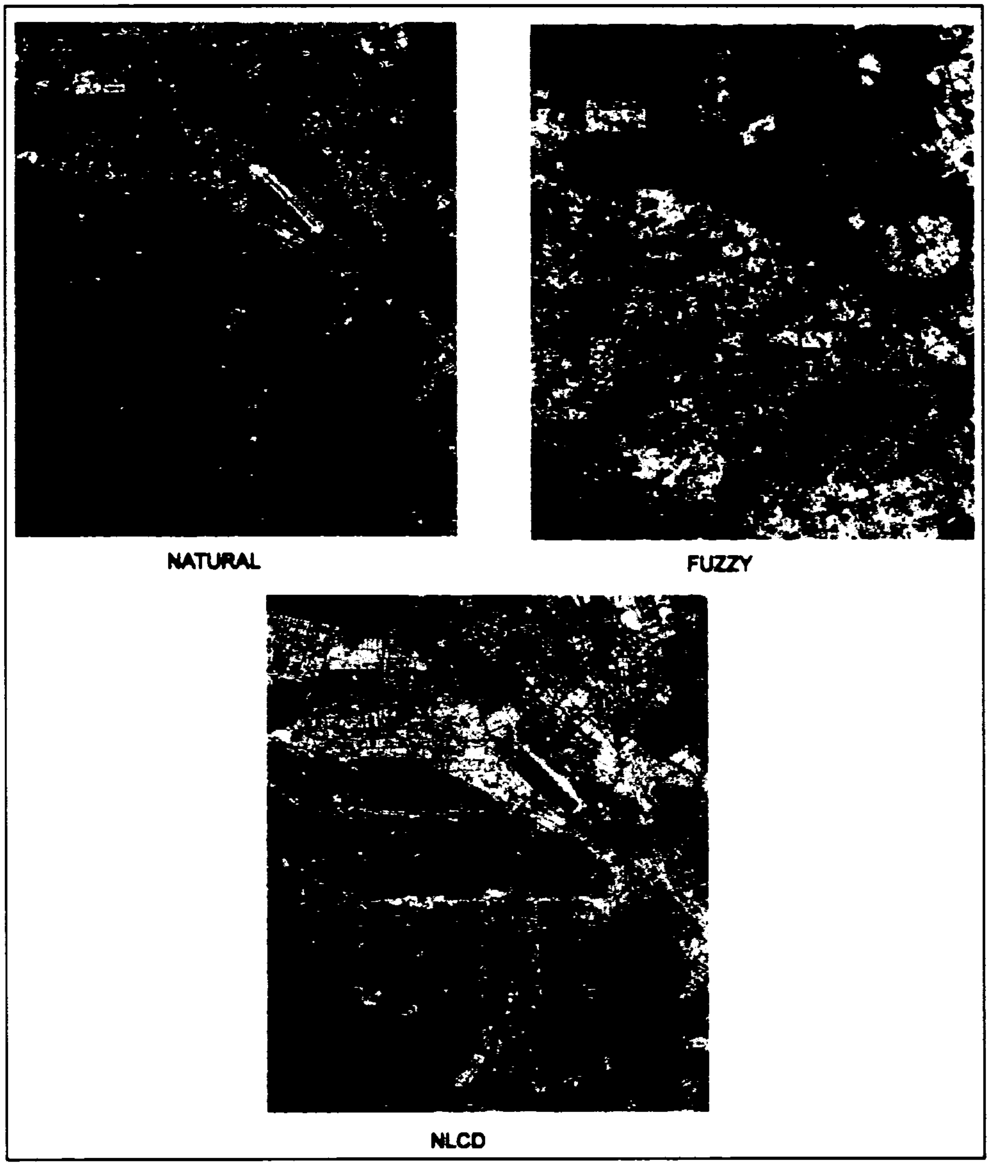


Figure 28. Impervious surface in the vicinity of Mineta San Jose International Airport. Lower left coordinates: 121.992247 W longitude, 37.408686 N latitude.

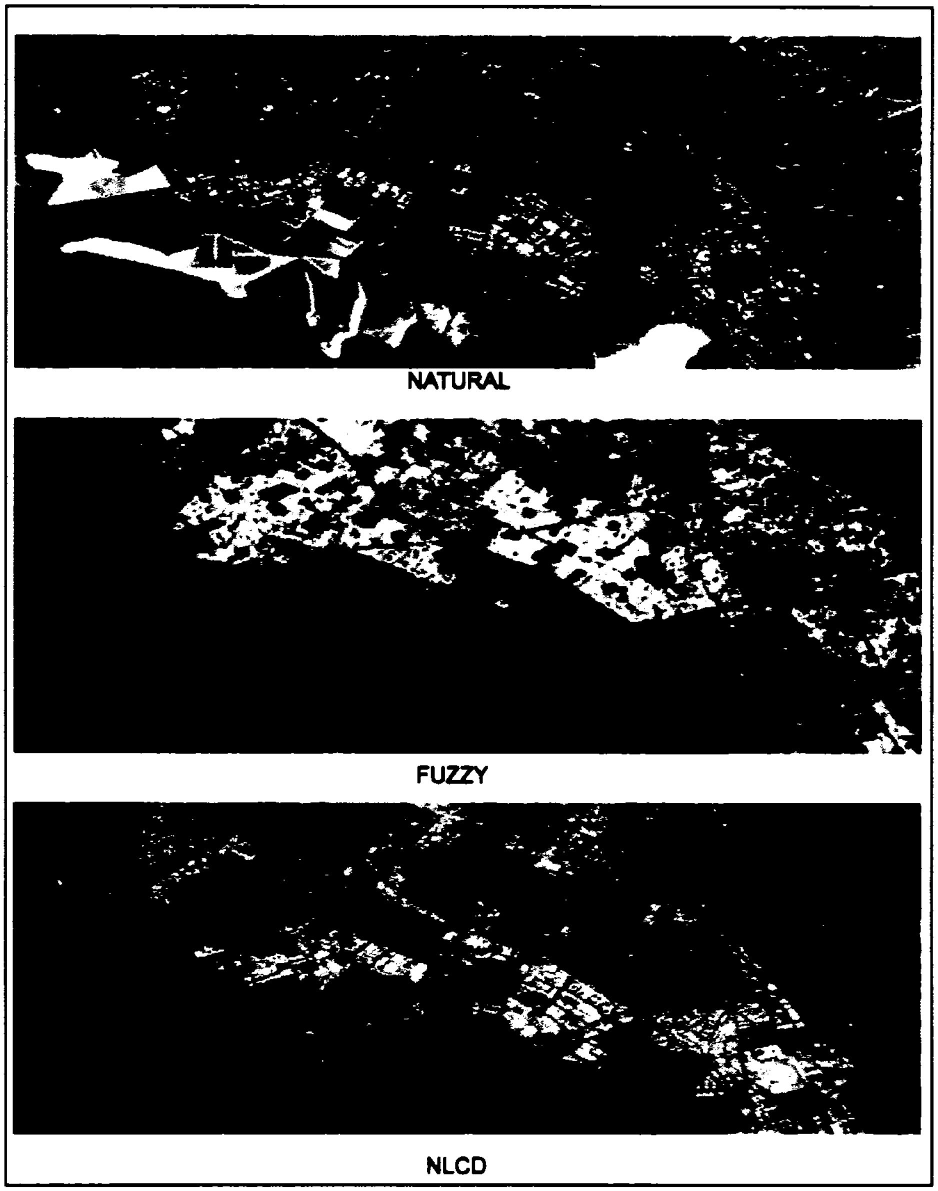


Figure 29. Impervious surface in the vicinity of Milpitas and Fremont, in the South Bay. Lower left coordinates: 122.122014 W longitude, 37.556818 N latitude.

Visual analysis at this intermediate scale is difficult to provide. Improvements in visual analysis as well as quantitative error assessments must be made at the per-pixel scale. This is done in a region that was not used in the creation of fuzzy sets. Fine-scaled assessments are therefore discussed below, with reference to the fuzzy testing data.

Results of Testing Data

In order to provide temporal and spectral consistency between our training and testing data, a different sub scene of the same Landsat ETM+ scene was selected for testing both the results of the fuzzy method used here as well as the results of the NCLD classifications.

Testing Area Data Preparation

The testing area is the northern Bay Area, extending from Marin County through Concord and Vallejo, into Stockton. NCLD classifications were downloaded from the USGS website at http://seamless.usgs.gov. This area was windowed out of the full scene, imported and reprojected from WGS84 to Lon/Lat in Imagine, exported to Idrisi32, where subsets bands 1-6 of the full scene were created.

- The same scripts were modified to contain the new subset files of bands 1-6 of the testing area.
- New masks were created with confidence ranges of 0-10, 10-20 ...90-100.
- The same fuzzy rules were experienced.
- The final results were scaled to 100 so that they are normalized to the NLCD results, which are expressed as percentage confidence to 100%.

- Several scenes of each class (canopy and impervious) were chosen for the error matrix evaluation. A pixel-by-pixel evaluation was made, comparing the scene to high-resolution aerial orthophotographs serving as ground truth data.

Figure 30 is a presentation of the respective Canopy classifications, and Figure 31 is a presentation of the respective Impervious classifications. A portion of the NCLD Canopy classification in the region of Stockton seems not to be well represented in the file provided (Figure 30c). Therefore, evaluation of canopy was performed mostly in Marin County.

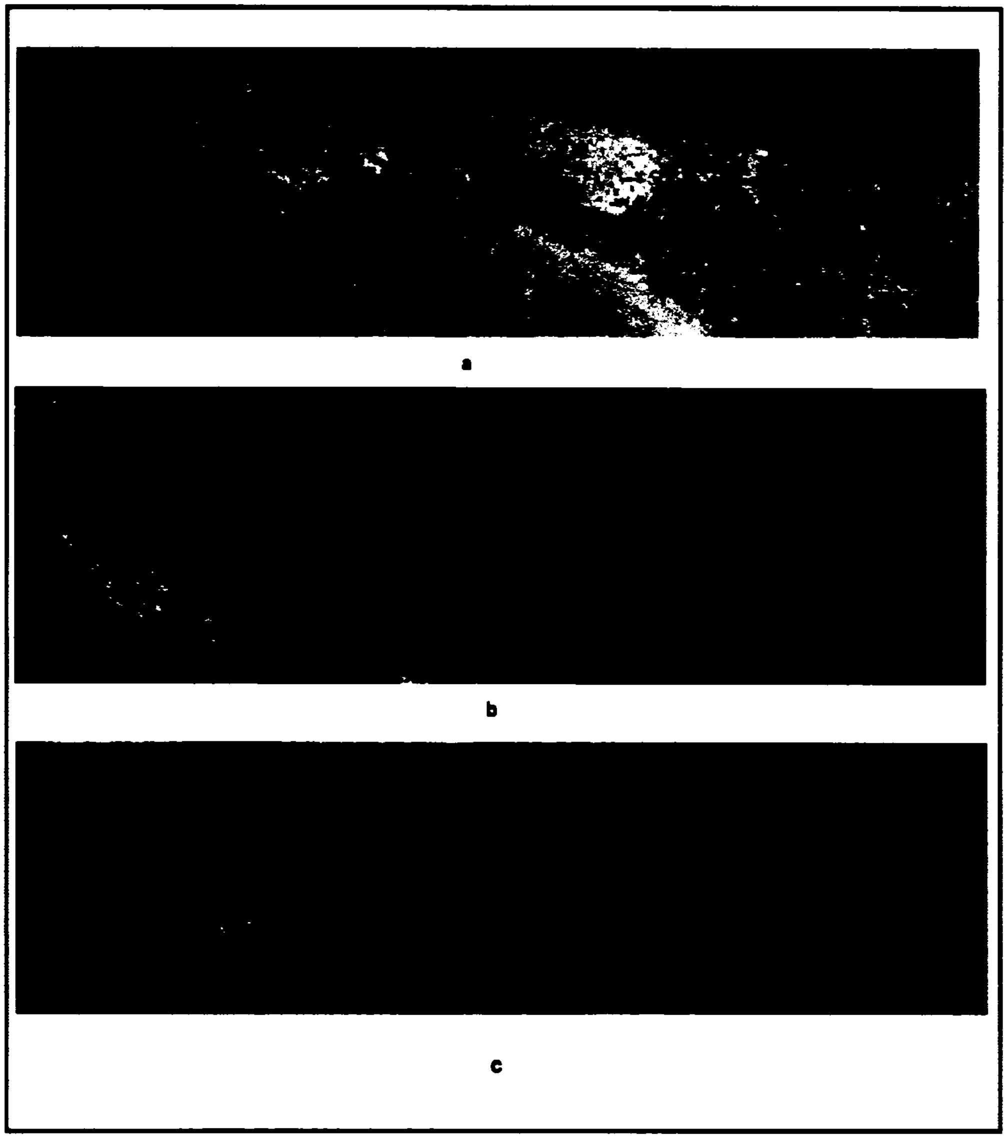


Figure 30. Testing region: a) Natural Color Natural Color orthophotograph; b) Fuzzy Canopy Classification; c) NLCD Canopy Classification.

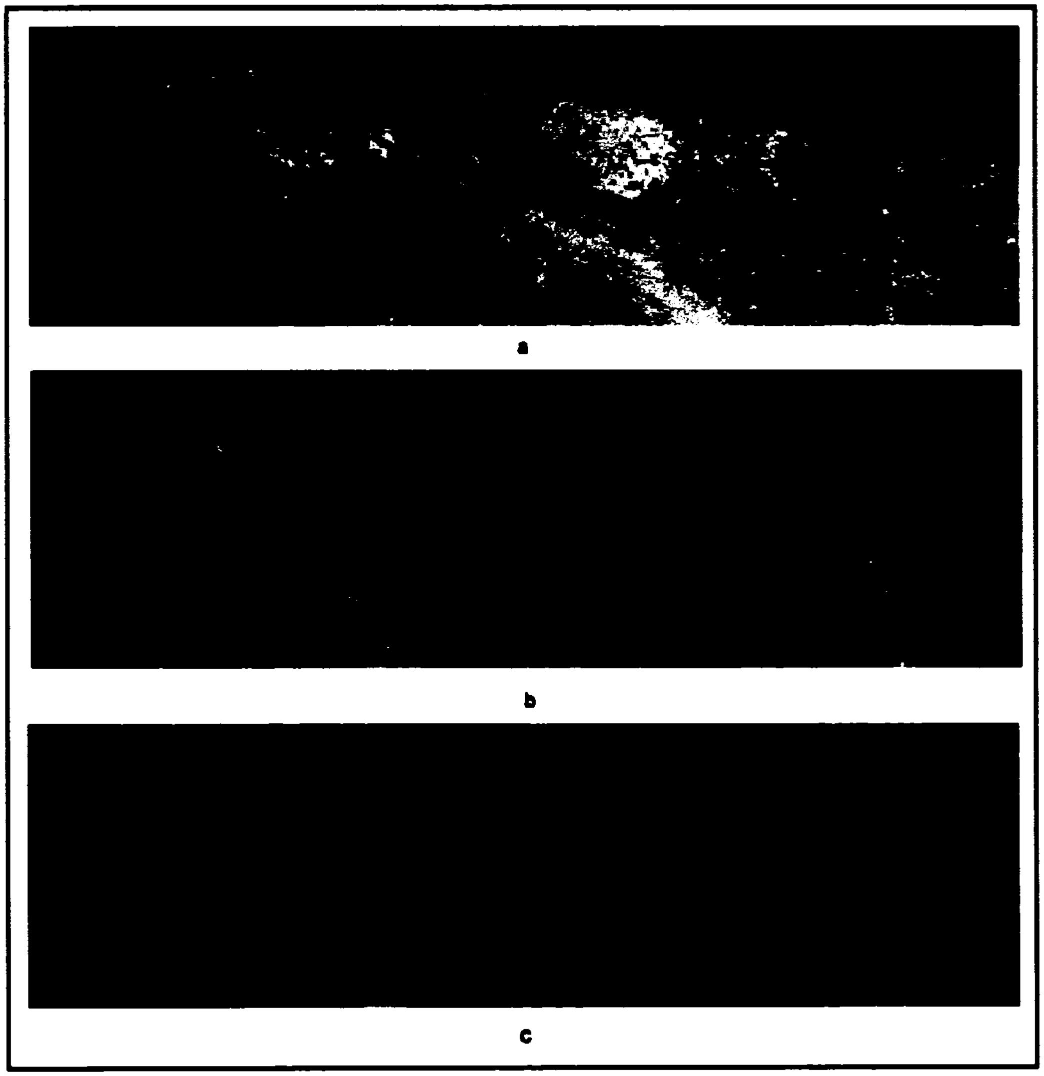


Figure 31. Testing region: a) Natural Color Natural Color orthophotograph;
b) Fuzzy Impervious Classification; c) NLCD Impervious Classification.

Examination of Error

Careful quantification of error and examination of both within-class confidence levels (NLCD) and membership values (fuzzy) on a per-pixel basis is a difficult process and requires perfect spatial rectification and co-registration. The source of "ground truth" for this process is the database of color orthophotography available for the USGS for urban regions. This photographic record is recent, and corresponds temporally with the 2000 data used in both the fuzzy and NCLD processes. In addition, it is of very fine sub-meter resolution, and visual inspection can yield an accurate estimate of class membership for each 30-meter pixel, if co-registration is accurate. Nevertheless, visual inspection is a subjective process (as is on-site inspection) and the precise meaning of "canopy" and "impervious surface" is a matter of interpretation.

The method used here involves the importation and registration of orthophotography for selected locations, and the creation of images with both the photographic and classified data side by side beneath an identical, indexed 30 meter grid, as shown in Figures 32 through 35. This is done on a high-resolution color monitor, so the interpretation of land cover on a cell-by-cell basis is actually easier than would appear from these figures.

As in the preceding discussion, consideration of Canopy and Impervious classes is done separately. This is done not only because of problems associated with visualizing more than one class simultaneously through a range of membership values, but because locations were chosen for the evaluation of each class on the basis of difficulty and the likely presence of pixels with partial memberships. The stratified random sampling method often used for evaluation of remote sensing classifications is neither necessary nor really appropriate, since classes are examined separately, and the examination of pixels chosen randomly over such a wide region would be very difficult, in terms of assuring proper registration. Therefore, the selection of several locations and the assessment of a number of contiguous pixels at each location yielded both a variety of terrain contexts as well as the assurance that pixel registration was accurate and uniform.

Figure 32 portrays one of the locations in Marin County that were used for the assessment of Canopy classification accuracy for the fuzzy method. Figure 33 portrays the same thing for the NCLD classification. Locations like this were chosen because of the dominance of canopy, as well as the heterogeneous presence of non-canopy land cover. This presents a challenging location for classification, far more than a continuous expanse of canopy that would likely yield excellent results for both methods. The 170 grid cells in this location were visually examined, and each was interpreted as being in one of three classes: 0% to 30% canopy; 30% to 60% canopy, and 60% to 100% canopy. This established the column location of each pixel in a modified error matrix (Tables 1 and *2*). The value assigned to each corresponding pixel in either the fuzzy or NCLD classification map determined its row location. This is a different sort of error matrix from that commonly used in such studies. Only one class is represented, but it is broken down into either membership or confidence levels. In each table, the totals along the diagonal represent successfully classified pixels, and are shown in blue. The total number of successfully classified pixels is shown as an extension to the diagonal (for example, 120 in Table 1). Non-diagonal totals are shown in black, and column and row totals are shown in red.

Indices can be derived from these tables: for example, total classification accuracy might be described as 120/170 for Table 1. However, this discussion will be based apon the general characteristics of the images and tables, rather than referring to specific numerical summaries. While providing specificity, they depend on many subjective factors, particularly the locations chosen and interpretation of boundaries between the three membership categories.

Figures 34 and 35 show results for one of the locations chosen for assessment of impervious surfaces, and the corresponding tables are compiled the same way. Discussion of these Canopy and Impervious results and others not portrayed follow the figures and tables.

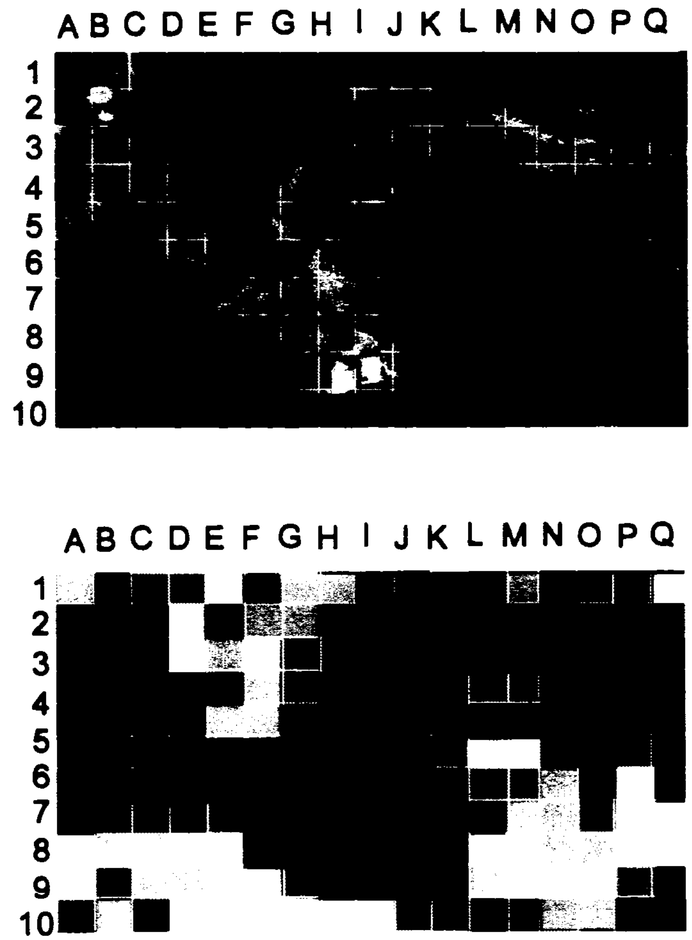


Figure 32. Color orthophotograph and image of results of fuzzy Canopy classification for a location in Marin County. Each pixel represents 30 m.

Table 1: An error matrix for the location shown in Figure 32

		Percent Canopy as interpreted from Orthophotographs			
		0-30%	30-60%	60-100%	Total
Percent Canopy From Fuzzy Classification	**0-30%**	**50**	**10**	**30**	**90**
	30-60%		**20**	**10**	**30**
	60-100%			**50**	**50**
	Total	**50**	**30**	**90**	**120**

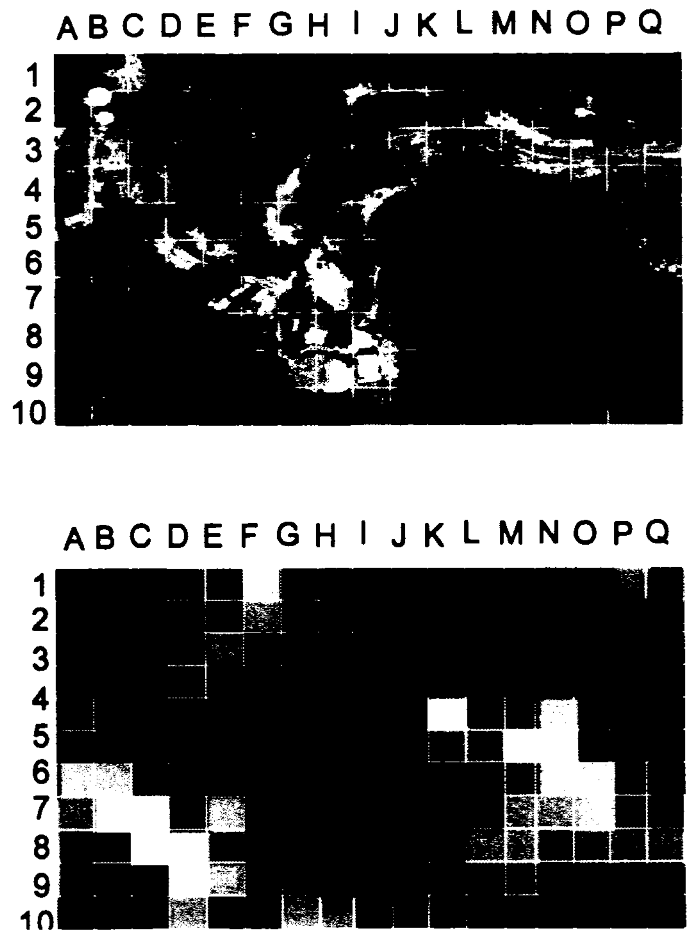


Figure 33. Color orthophotograph and image of results of **NLCD Canopy** classification for a location in Marin County. Each pixel represents 30 m.

Table 2: An error matrix for the location shown in Figure 33

	Percent Canopy as interpreted from Orthophotographs				
		0-30%	30-60%	60-100%	Total
Percent Canopy From NLCD Classification	0-30%	59	13	42	114
	30-60%	1	11	25	37
	60-100%			19	19
	Total	60	24	86	89

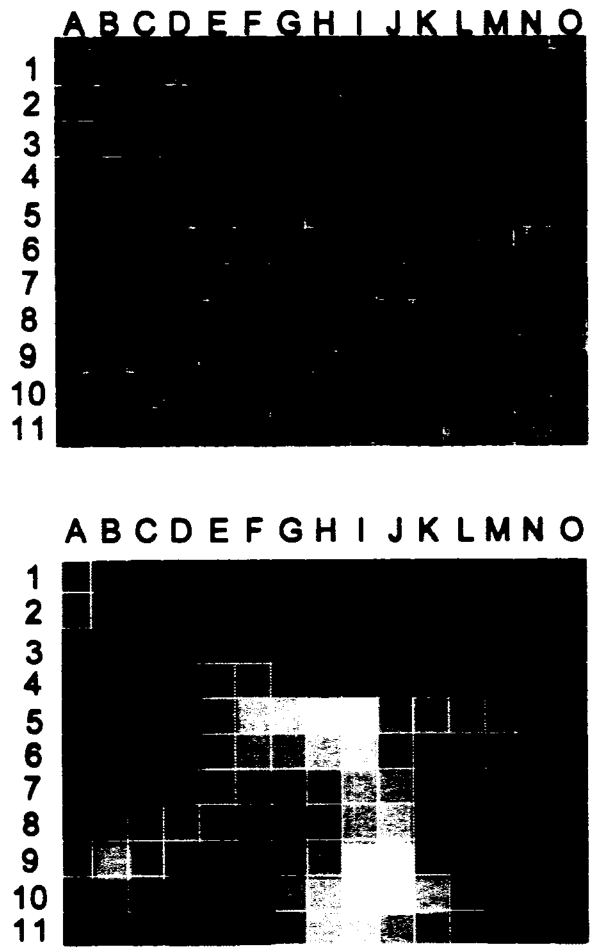


Figure 34. Color orthophotograph and image of results of **fuzzy Impervious surface** classification for a location in Marin County. Each pixel represents 30 m.

Table 3: An error matrix for the location shown in Figure 34.

		Percent Impervious Surface as interpreted from Orthophotographs			
		0-30%	**30-60%**	**60-100%**	**Total**
Percent Impervious Surface from Fuzzy Classification	**0-30**	**74**	**8**	**17**	**99**
	30-60	**9**	**25**	**3**	**37**
	60-100	**1**		**33**	**34**
	Total	**84**	**33**	**54**	**132**

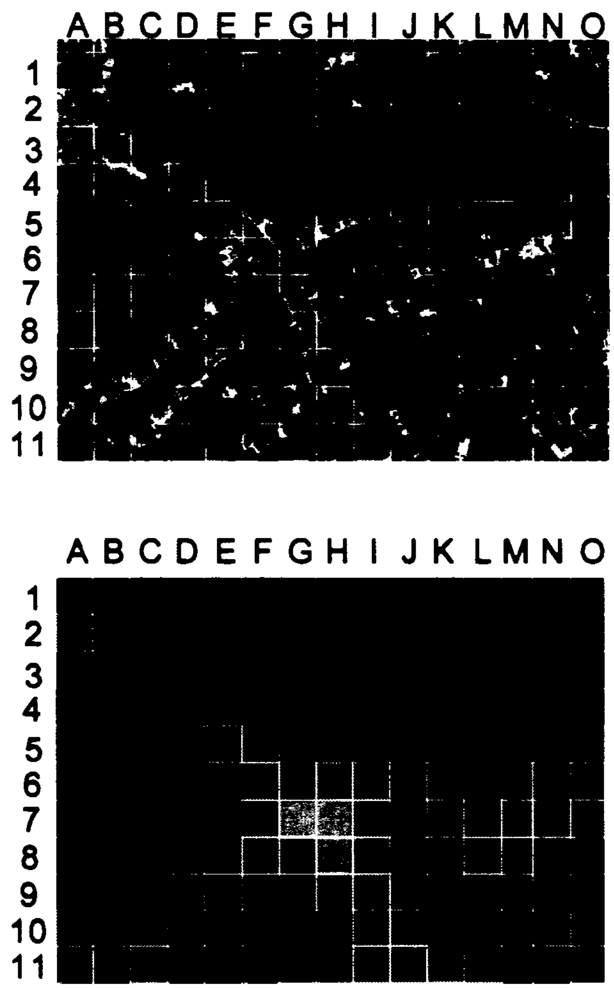


Figure 35. Color orthophotograph and image of results of **NLCD Impervious surface** classification for a location in Marin County. Each pixel represents 30 m.

Table 4: An error matrix for the location shown in Figure 35.

	Percent Impervious Surface as interpreted from Orthophotographs				
		0-30%	**30-60%**	**60-100%**	**Total**
Percent Impervious Surface from NLCD Classification	**0-30**	**91**	**2**	**11**	**104**
	30-60	**3**	**33**	**9**	**45**
	60-100		**1**	**20**	**21**
	Total	**94**	**36**	**40**	**144**

A few characteristics of both the fuzzy and NCLD results can be noticed in these tables. Other results not portrayed here are consistent with these observations, and are included in the error matrices over all locations examined (Tables 5 and 6). First, both the fuzzy and NCLD methods seem to have under-represented the amount of Canopy present at the location, as indicated by the numbers present above the diagonal. These errors of omission are more significant in the NCLD classification than in the fuzzy classification. Overall classification accuracy is somewhat better for the fuzzy method than for the NCLD method.

Results for impervious surfaces are somewhat different. While most errors for both methods are errors of omission, the fuzzy method also yields significant numbers of errors of commission (the entries below the diagonal in Tables 3 and 6). Overall accuracy is better for the NCLD method than for the fuzzy method.

To summarize, the limited number of locations and pixels used for assessment should caution against making categorical statements. Error assessment is a difficult, subjective process, and this study did not fully examine the sorts of errors that may occur in the assessment of errors. However, given the fuzzy sets chosen and portrayed in Appendix 1, the fuzzy method used here yields good results for the Canopy class, possibly better than the NCLD method. With regard to the Impervious class, the fuzzy method (again, given the fuzzy sets chosen) has some significant problems. At broader scales it misses some impervious surface materials, as shown by its failure to classify certain roads and industrial buildings. Examination at the pixel level indicates that even in a suburban environment classification accuracy is somewhat less than the NCLD method, both in terms of omissions and commissions. However, the NLCD method is also not perfectly accurate with regard to Impervious surface classification. Its success is only slightly better in the suburban locations chosen for evaluation.

Table 5: A Canopy class error matrix for the fuzzy method over all evaluated locations

	Percent Canopy as interpreted from Orthophotographs				
Percent Canopy From Fuzzy Classification		**0-30%**	**30-60%**	**60-100%**	**Total**
	0-30%	**153**	**23**	**62**	**238**
	30-60%	**3**	**62**	**34**	**99**
	60-100%		**5**	**134**	**139**
	Total	**156**	**80**	**230**	**449**

Table 6: An Impervious class error matrix for the fuzzy method over all locations.

		Percent Impervious Surface as interpreted from Orthophotographs			
Percent Impervious Surface from Fuzzy Classification		**0-30%**	**30-60%**	**60-100%**	**Total**
	0-30	**112**	**32**	**29**	**173**
	30-60	**19**	**93**	**10**	**112**
	60-100	**5**	**13**	**98**	**116**
	Total	**136**	**138**	**137**	**303**

Discussion

The fuzzy classification methodology described in this study was applied to the classification of land cover in a large region in the San Francisco Bay Area. Only two general land cover classes were investigated: tree canopy and impervious surfaces. The results were compared with those of the NCLD, and both were evaluated through comparison with high-resolution color orthophotography. The results of the fuzzy classification were as good or better than NCLD for the Canopy class, but fared somewhat less remarkably for the Impervious class. Impervious surfaces comprise a wider variety of materials than tree canopies, and therefore reflect a more heterogeneous variety of spectral signatures. If the method is applied to a disaggregation of several subclasses, it might yield better results.

The two classes of Canopy and Impervious were chosen because the ongoing NCLD project currently offers only these two interim classes. The idea behind this study was to use these results to provide a reliable way of masking out the spectral characteristics of these classes in order to build fuzzy sets that could be used in a logical way to duplicate the results of the complex NCLD regression tree approach. This was shown to be generally successful. It might be said that we have encoded knowledge derived from the results of the NCLD approach in the form of a simple fuzzy rule base, allowing us to duplicate their results anywhere for which Landsat TM or ETM+ data are available. Detailed evaluation of the performance of both approaches in specific locations indicate that both approaches err on the side of omission, but that some improvement might result from a more careful crafting of fuzzy sets and rules used in the classification of impervious surfaces.

Comparison with Similar Work

The use of rules involving fuzzy sets for various non-geographical applications is not new, and it would have been surprising if studies similar to this one had not been done elsewhere. In fact, a search of the literature did yield some similar work, but with some significant differences. This study is similar to work that has been done at the University of Baghdad, Iraq (Melgani *et al.* 2000), involving the use of a MIN fuzzy reasoning rule on pixels from different bands to be classified at several classes, as well as a fuzzy partition membership function.

A comparison between the methods used in this study and the method described below reveals three major differences. First, the Melgani study uses a prior assumption of the Gaussian distribution. This assumption is the basis for the establishment of their classes' extents and the extraction of statistics such as mean of class signature, and standard deviation of the class signature. According to their reasoning the mean represents the ideal pixel of the class. The standard deviation determines the width of the fuzzy subset (Melgani *et al.* 2000). In this thesis there are no assumptions of normal distribution of the classes. The lower ranges of confidence and higher ranges of confidence, and the pixel count of each range determine the width of the fuzzy set.

The second major difference between the methods is the pixel values extrapolation. In the Melgani study a matrix was created with training pixels and tested pixels based on the statistics mentioned above. This matrix then was analyzed with the MIN reasoning rule. In this study the values are extracted directly from the raw data of each band in each range of confidence, the MIN reasoning rule is established according to those raw values, no statistics are involved in determining the upper and lower boundaries and the fuzzy sets width.

The third difference is the comparison method. In the Melgani study a Defuzzification step was taken in order to be able to compare it to hard classification. The method to defuzzyfy was the MAX operation, which classifies only the absolute high confidence pixels to the examined class. In this study there is no need for Defuzzification as a comparison method since the ranges of confidence and the weighted pixel count of each range create a direct comparison.

Another similar research effort attempts to address accuracy and interpretability using fuzzy rule based classifiers from labeled data in an iterative approach (Roubos *at al.* 2001). This study creates a model of sampled pixels derived from clusters. A fuzzy rule base method is applied to this model. A Center and Covariance is calculated out of the fuzzy sets. A distance of elements to the class center is calculated. This distance is used as weight. Finally a real coded GA, genetic algorithm, is applied since many steps in the modeling process are sub-optimal (Roubos *et al.* 2001). Their study uses pragmatism to force parameters to respond to the desired classification results. In comparison, this study uses all raw data instead of a sample, simple fuzzy sets and a direct comparison method instead of a measure of distance from cluster. Statistics and algorithms are not involved in the process to force elements to certain classification.

Although others have used methods that are similar to those described in this thesis, the method used here is probably more robust for a wide variety of classification tasks. Landscapes to be classified must be similar to the ones that are used in creating the fuzzy sets, but once created they yield results that are comparable to officially sanctioned USGS results, and provide information about partial membership that would be useful in a variety of applications.

REFERENCES

Burrough, Peter A. and McDonnell, Rachel A., 1998, *Principles of Geographical Information Systems*, Oxford University Press, New York.

Console, Elena and Mouchot, Marie C., 1996, *Fuzzy Classification Techniques in the Urban Area Recognition,* 1996, IEEE Transactions On Geoscience and Remote Sensing, 0-7803-3068-4.

Foody, G. M. and cox d. O., 1994, *Sub-Pixel Land Cover Composition Estimation Using a Linear Mixture Model and Fuzzy Membership functions*, International Journal of Remote Sensing, 15, 619-631, part 3.

Homer, Collin, Huang, Chengquan, Yang Limin, Wylie Bruce and Coan Michael, 2004, *Development of a 2001 National Land-Cover Database for the United States,* [on-line]. Available: http://www.mrlc.gov

Jensen, John R., 2000. *Remote Sensing if the Environment An Earth Resource Perspective*, Prentice Hall, Upper Saddle River, New Jersey.

Jensen, John R., 1996. *Introductory Digital Image Processing*, Prentice Hall, Upper Saddle River, New Jersey.

Klir, George J. and Yuan, Bo, 1995. *Fuzzy Sets and Fuzzy Logic Theory and Applications*, Prentice Hall, Upper Saddle River, New Jersey.

Melgani, Farid Hashemy Al Bakir A. and taha Saleem M. R., 2000,*an Explicit Fuzzy Supervised Classification Method for Multispectral Remote Sensing Images*, IEEE Transactions On Geoscience and Remote Sensing, vol.38, no. 1.

Nedeljkovic, I., *Image Classification based on Fuzzy Logic,* The International archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. 34.

Roubos, Johannes A., Setnes Magne and Abonyi Janos, 2003, *Learning Fuzzy Classification Rules From Labeled* Data, Information Sciences, 150, 77-93.

Townshend, John R. G., 1981, *Terrain Analysis and Remote Sensing*, George Allen & Unwin LTD, London.

Tso, Brandt and Mather, Paul M., 2001. *Classification Methods For Remotely Sensed Data*, Taylor & Francis, London.

Vogelman, J.E, T.Soul, and S.M.Howard, 1988, Regional Characterization of Land cover Using Multiple Sources of Data, *Pohtogrammetric Engineering & Remote Sensing*, vol. 54, No. 1, January, pp. 45-57.

Zadeh, L. A. 1965, *Fuzzy Sets*, Information Country, vol. 8, 338-353.

APPENDIX: IDRISI 32 FUNCTIONS

RECLASS classifies or reclassifies the pixel values stored in images, the feature ID values of vector files or the second column values of attribute values files into new integer categories. Classification or reclassification is by equal intervals division of the data range, or by the application of user-defined limits.

OVERLAY produces a new image from the data of two input images. New values result from applying one of the nine possible operations to the two input images, referred to as the first and second images during program operation.

Assign a new value of:

To old values ranging from:

To those just less than:

HISTO produces a frequency histogram of cell values in an IDRISI image or a signature file. HISTO creates histograms by dividing the data range into classes of a user-specified width. The frequency within each class is tabulated. Both graphic and numeric output options are available.

FUZZY evaluates the possibility that each pixel belongs to a fuzzy set by evaluating any of a series of fuzzy set membership functions. The first point marks the location where the membership function begins to rise above 0. The second point indicates where it reaches 1. The third point indicates the location where the membership grade begins to drop again below 1, while the fourth point marks where it returns to 0.

CanopyRadConfMacro:

```
reclass¹ x i*canopy*canopy_90to100*2*0*0*91*1*91*101*-9999
reclass x i*canopy*canopy_80to90*2*0*0*80*1*80*90*0*90*101*-9999
reclass x i*canopy*canopy_70to80*2*0*0*70*1*70*80*0*80*101*-9999
reclass x i*canopy*canopy_60to70*2*0*0*60*1*60*70*0*70*101*-9999
reclass x i*canopy*canopy_50to60*2*0*0*50*1*50*60*0*60*101*-9999
reclass x i*canopy*canopy_40to50*2*0*0*40*1*40*50*0*50*101*-9999
reclass x i*canopy*canopy_30to40*2*0*0*30*1*30*40*0*40*101*-9999
reclass x i*canopy*canopy_20to30*2*0*0*20*1*20*30*0*30*101*-9999
reclass x i*canopy*canopy_10to20*2*0*0*10*1*10*20*0*20*101*-9999
reclass x i*canopy*canopy_0to10*2*1*0*10*0*10*101*-9999

overlay² x 3*canopy_90to100*tmplus_subset1*tm1_canopy_90to100
overlay x 3*canopy_80to90*tmplus_subset1*tm1_canopy_80to90
overlay x 3*canopy_70to80*tmplus_subset1*tm1_canopy_70to80
overlay x 3*canopy_60to70*tmplus_subset1*tm1_canopy_60to70
overlay x 3*canopy_50to60*tmplus_subset1*tm1_canopy_50to60
overlay x 3*canopy_40to50*tmplus_subset1*tm1_canopy_40to50
overlay x 3*canopy_30to40*tmplus_subset1*tm1_canopy_30to40
overlay x 3*canopy_20to30*tmplus_subset1*tm1_canopy_20to30
overlay x 3*canopy_10to20*tmplus_subset1*tm1_canopy_10to20
overlay x 3*canopy_0to10*tmplus_subset1*tm1_canopy_0to10

histo³ x 1*tm1_canopy_90to100*none*2*1*1*0*255*2*hist90to100
histo x 1*tm1_canopy_80to90*none*2*1*1*0*255*2*hist80to90
histo x 1*tm1_canopy_70to80*none*2*1*1*0*255*2*hist70to80
```

---

[1] reclass x i*canopy*canopy_90to100*2*0*0*91*1*91*101*-9999

The module "reclass" creates a mask to layer "canopy". This file will be called:"canopy_90to100".The format is: assign 0 to all values from 0 to 91; assign 1 to all values from 91 to 101.

[2] overlay x 3*canopy_90to100*tmplus_subset1*tm1_canopy_90to100. : File canopy_90to100 is overlaid on tmplus_subset1 and calls it tm1_canopy_90to100.This is an intermediate file for bringing out the actual values of each range only, from each band of canopy classification layer.

[3]Histo creates a numeric histogram out of the overlaid files for each range of values. See Appendix 2

```
histo x 1*tm1_canopy_60to70*none*2*1*1*0*255*2*hist60to70
histo x 1*tm1_canopy_50to60*none*2*1*1*0*255*2*hist50to60
histo x 1*tm1_canopy_40to50*none*2*1*1*0*255*2*hist40to50
histo x 1*tm1_canopy_30to40*none*2*1*1*0*255*2*hist30to40
histo x 1*tm1_canopy_20to30*none*2*1*1*0*255*2*hist20to30
histo x 1*tm1_canopy_10to20*none*2*1*1*0*255*2*hist10to20
histo x 1*tm1_canopy_0to10*none*2*1*1*0*255*2*hist0to10
```

CanopyMinRule:

```
FUZZY X 2*TMPLUS_SUBSET1*2*CANOPY1RULE*51*54*58*90
FUZZY X 2*TMPLUS_SUBSET2*2*CANOPY2RULE*31*37*41*73
FUZZY X 2*TMPLUS_SUBSET3*2*CANOPY3RULE*20*26*31*60
FUZZY X 1*TMPLUS_SUBSET4*2*CANOPY4RULE*29*44*56*85
FUZZY X 2*TMPLUS_SUBSET5*2*CANOPY5RULE*16*30*47*128
FUZZY X 2*TMPLUS_SUBSET6*2*CANOPY6RULE*7*14*20*71
```

ImpervMinRule:

```
FUZZY X 1*TMPLUS_SUBSET1*2*imperv1RULEa*62*76*86*126
FUZZY X 3*TMPLUS_SUBSET1*2*imperv1RULEb*51*55*59*61
FUZZY X 1*TMPLUS_SUBSET2*2*imperv2RULEa*47*62*72*102
FUZZY X 3*TMPLUS_SUBSET2*2*imperv2RULEb*31*33*45*46
FUZZY X 1*TMPLUS_SUBSET3*2*imperv3RULEa*44*64*76*130
FUZZY X 3*TMPLUS_SUBSET3*2*imperv3RULEb*20*22*41*43
FUZZY X 1*TMPLUS_SUBSET4*2*imperv4RULE*29*45*52*90
FUZZY X 1*TMPLUS_SUBSET5*2*imperv5RULE*33*61*75*105
FUZZY X 1*TMPLUS_SUBSET6*2*imperv6RULEa*25*47*60*96
FUZZY X 3*TMPLUS_SUBSET6*2*imperv6RULEb*12*15*23*24
```

OverlayMinRules:[4]

```
overlay x 9*imperv1RULEa*imperv1RULEb*imperv1RULE
overlay x 9*imperv2RULEa*imperv2RULEb*imperv2RULE
overlay x 9*imperv3RULEa*imperv3RULEb*imperv3RULE
overlay x 9*imperv6RULEa*imperv6RULEb*imperv6RULE
overlay x 8*imperv1RULE*imperv2RULE*step1
overlay x 8*step1*imperv3RULE*step2
```

[4] this macro overlays all impervious rules of all bands to create a final file called impervOverlayMinRule . for canopy : CanopyOverlayMinRule.

```
overlay x 8*step2*imperv4RULE*step3
overlay x 8*step3*imperv5RULE*step4
overlay x 8*step4*imperv6RULE*impervOverlayMinRule

overlay x 8*CANOPY1RULE*CANOPY2RULE*stepa
overlay x 8*stepa*CANOPY3RULE*stepb
overlay x 8*stepb*CANOPY4RULE*stepc
overlay x 8*stepc*CANOPY5RULE*stepd
overlay x 8*stepd*CANOPY6RULE*CanopyOverlayMinRule
```