COMPARISON OF GEOGRAPHIC INFORMATION SYSTEM SOFTWARE (ARCGIS 9.0 AND GRASS 6.0): IMPLEMENTATION AND CASE STUDY

being

A Thesis Presented to the Graduate Faculty

of the Fort Hays State University in

Partial Fulfillment of the Requirements for

the Degree of Master of Science

by

Todd R. Buchanan

B.S., Hardin – Simmons University

Date ______________________________ Approved ____________________________

Major Professor

Approved ____________________________

Chair, Graduate Council

ABSTRACT

Beginning in the early 1960s, Geographic Information Systems (GIS) have evolved from mainframe computer programs written in FORTRAN into the highly complex desktop PC software that we have today. In the early 1980s, GIS development went in two directions. Commercial GIS companies such as ESRI began producing commercial GIS packages such as ArcView while the United States Army Construction Engineering Research Laboratory (CERL) began developing a "no fee" GIS package called Geographic Resource Analysis Support System (GRASS). ESRI's ArcView would ultimately evolve into present day ArcGIS 9.0, while the CERL version of GRASS would eventually become today's open source and freely available GRASS 6.0.

The objective of this study was to characterize and contrast the advantages and disadvantages of ArcGIS 9.0 and GRASS 6.0 for application involving land use classification analysis. The hypothesis of the study was that GRASS GIS implementation would be more involved and time consuming than ArcGIS, but would ultimately provide an effective alternative to ArcGIS.

A land use classification analysis case study of Eugene-Springfield, Oregon, was used to compare the two packages. Throughout the course of the study, a detailed journal was kept documenting the land use classification process in both programs. Categories ranging from installation to post-classification analysis and map layout capabilities were documented for comparison. Upon completion of the analysis, each category was scored to compare strengths and weaknesses of both packages.

The results of the study supported the hypothesis within the limitations of the case study. ArcGIS is indeed more straightforward than GRASS. However, GRASS proved itself fully capable of performing the operations required by the case study and outperformed ArcGIS in several categories. When used in a land use classification analysis, results showed GRASS to be an effective, low cost alternative to ArcGIS.

ACKNOWLEDGMENTS

I am in debt to a number of people who helped me throughout the course of this project. I would like to thank Dr. John Heinrichs, my thesis advisor, for all of the hard work and patience that he put into this project. Thank you for the times when you went out of your way so that I could stick to my crazy schedule. I have learned a lot from you over the last couple of years and I am very lucky to have had you as my thesis advisor.

I would also like to thank my other graduate committee members, Dr. Richard Lisichenko and Dr. Mark Bannister. I have learned a lot from both of you and I greatly appreciate all of your help throughout the course of my graduate education.

I also could not have made it through this without the constant encouragement and friendship of my fellow Geoscience graduate students Tim Mullin, John Rebar, and Mark Vishnefske. You definitely made life in Hays, Kansas somewhat tolerable!

I am also extremely grateful to my wife Hannah, who selflessly gave up her job in Tennessee to follow me to Hays to complete this degree. You have encouraged me throughout the course of this thesis, pushing me harder when I needed pushing and sympathizing with me when I needed a pat on the back.

Thank you to my parents who have supported me from the beginning. You are truly a blessing in my life and I would not be here without your constant love and support.

Finally, I would like to thank my savior Jesus Christ who has blessed me and remained faithful throughout.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDIXES

INTRODUCTION

A Geographic Information System (GIS) is a set of tools for analyzing spatial data (Clarke, 2001). It is a special type of information system where the database consists of observations on spatially distributed features, which are definable as points, lines, or areas. The GIS can manipulate data about these features in order to conduct analysis and extract information (Dueker, 1979).

GIS History

GIS development began with the increased use of thematic maps during the first half of the twentieth century (Clarke, 2001).

In the 1960s, early attempts to develop GIS were made on large computers using programming languages such as FORTRAN. This combination dominated until the early 1980s (Clarke, 2001). In 1982, Environmental Systems Research Institute (ESRI) launched ARC/INFO 1.0, the first commercially available GIS software for main frame computers (ESRI, 2005). The same year, International Business Machines (IBM) introduced the personal computer (PC) (Clarke, 2001).

In order to support land management at U.S. military installations, the United States Army Construction Engineering Research Laboratory (CERL) began development of Geographic Resource Analysis Support System (GRASS) in 1982. CERL decided that they could reduce costs of extremely expensive systems by developing their own “no fee” software in a UNIX environment (Neteler and Mitasova, 2004). This made GRASS the first GIS package to be available on a PC. It wasn’t until 1986 that some of the larger commercial GIS packages such as ARC/INFO had made the transition from large computers to the PC (Clarke, 2001).

The decade of the 1990s saw remarkable growth in the GIS world. The lower cost of PCs and their increased mobility played an important factor in the growth of GIS (Clarke, 2001). Desktop GIS products such as ArcView emerged, lowering the cost of previously expensive systems and improving usability (Clarke, 2001). In 1991, GRASS 4.0 was released to the public for the first time over the internet (Neteler, 2005a). An unprecedented 10,000 copies of ArcView shipped within the first half of 1992, expanding the global presence of ESRI (ESRI, 2005). During the 1990s GIS also became fully integrated with the Global Positioning System (GPS), and the development of the Internet increased availability and use of geographic data (Clarke, 2001).

Objective

The purpose of this study was to characterize and contrast the advantages and disadvantages of alternate types of Geographic Information Systems (GIS), specifically ArcGIS 9.0 and GRASS 6.0, for applications involving land use classification analysis.

Even with all of the improvements in GIS over the last decade, implementation of a GIS can be an extremely daunting task, requiring a substantial commitment in organizational money, staff, and effort (Longley et al., 1999). The value of adopting a new technology must be visible and concrete. According to Innes and Simpson (1993), "Those responsible for the adoption must know what they are getting and be able to assess what it is worth." Although the cost of software is not the only expense involved with successful implementation of a GIS, it is nevertheless a substantial contributor to the overall cost of a project. Elimination of software costs by using an open source GIS such as GRASS may provide justification for a company to move ahead with a proposed GIS implementation.

Most major GIS projects traditionally include the phases of choosing a GIS supplier and GIS implementation (Longley et al., 1999). This study specifically addresses these phases, potentially saving the user time and money. It will be of interest to companies, government agencies, and others interested in implementing a GIS. More importantly, it may help answer the question of which will be the more appropriate GIS to use in order to meet the users' needs.

ArcGIS Overview

One of the programs examined in this study, ArcGIS 9.0, is a licensed commercial GIS software package developed by ESRI. The evolution of ESRI ARC software began with the release of ARC/INFO in 1982. ARC/INFO was designed to run on larger computers such as mainframes and minicomputers. In 1986, PC ARC/INFO was released on a UNIX platform. The ability to implement a GIS on a PC lowered the cost of implementing a GIS and provided greater access to the emerging technology (Clarke, 2001).

In 1992, ESRI released ArcView. ArcView was the first commercial GIS to be released on a Windows platform and utilize a graphical user interface (GUI). ArcView went through updates and revisions until 2001, when ESRI began shipping ArcGIS 8.1. ArcGIS 8.1 combined the analytical power of ARC/INFO with the straightforward, easier-to-use GUI employed by ArcView. The 22-year evolution of ARC software has led to ESRI’s current release, ArcGIS 9.0, which began shipping in May of 2004.

ArcGIS 9.0 is available for Microsoft Windows NT/2000/XP and is composed of many different geospatial data analysis and display tools, including data management tools, 3D visualization tools, and data processing tools. Many different extensions are available for purchase for ArcGIS 9.0, which give the user further analysis, data, conversion, or display capabilities (ESRI, 2004).

GRASS Overview

The other program examined in this study, GRASS 6.0, is the latest open-source GIS available. GRASS was originally developed in 1982 by CERL, in response to the new responsibility for the environment encoded into the National Environmental Policy Act of the late 1970s (Neteler, 2005a). GRASS was developed to conduct environmental research, monitoring, and management of U.S. Department of Defense lands.

During the early 1980s, costs involved in implementing a GIS (acquiring data and hardware, learning GIS skills and computer maintenance skills) became so high, CERL decided that no-fee software could reduce the technology hurdle involved in implementing a GIS. This proved to be true (Neteler and Mitasova, 2004). By the mid-1990s many of the original government GRASS GIS users were switching to proprietary software such as ArcView. CERL discontinued development of GRASS in 1995 and turned over development to the first GRASS research group at Baylor University in 1997.

Currently, GRASS is revised and updated by a worldwide GRASS development team. It is released under the General Public License (GPL) and is available for free download online (Neteler, 2005b).

GRASS is currently available for UNIX, Linux, or a Windows port program called Cygwin, which emulates a Linux environment on a Windows Operating System. Like ArcGIS 9.0, GRASS 6.0 incorporates tools for geospatial data management and analysis, image processing, 3D visualization, spatial modeling, and display.

METHODOLOGY

Case Study: Characterization of Urban Sprawl for Eugene-Springfield, Oregon

The topic chosen for the case study was an analysis of urbanization of Eugene-Springfield, Oregon. This topic was chosen for several reasons. The power in GIS lies in its ability to manage spatial relationships over time (Morain, 1999). Many municipalities, consulting companies, and government agencies will be interested in this type of analysis because it demonstrates how the analytical capabilities of GIS can be used to manage natural resources over space and time. Secondly, this type of analysis utilizes a broad spectrum of GIS functionality that serves as a basis of comparison for the two GIS being used.

The Eugene-Springfield metropolitan area is located in west-central Oregon in fast-growing Lane County (Figure 1). Both the Willamette and the McKenzie Rivers run through the Eugene-Springfield area (Figure 2), which is surrounded by coniferous forest. Lane County currently has a population of 322,959 (U.S. Census Bureau, 2005) and has experienced a 61% growth since 1970 (U.S. Census Bureau, 2005). Population increase directly effects land use change (Hunter et al., 2003) and although the topic for the analysis portion of this project has significant real world relevance, final results were used only to compare and contrast each of the two types of GIS.

General Approach

This project started with an in depth literature review of GIS implementation and land use analysis techniques. The literature review produced no previous literature comparing GRASS and ArcGIS.

A land use classification analysis case study of Eugene-Springfield, Oregon, was used to compare the two packages. Both programs were used to analyze downloaded Landsat imagery and derive results.

During the course of the project, a detailed journal was kept to document the entire implementation and case study for each program. The journal included descriptions of processes, problems encountered, solutions for problems encountered, and any issues encountered during the course of the project.

Categories ranging from installation to post-classification analysis and map layout capabilities were documented for comparison. Upon completion of the analysis, each category was scored to compare strengths and weaknesses of both packages.

Installation

This study began with acquisition of each of the programs being examined. A student copy of ArcGIS 9.0 was obtained from ESRI. The installation software included ArcGIS 9.0 with an ArcEditor license, and Spatial Analyst, 3D Analyst, and Geostatistical Analyst extensions.

GRASS 6.0 was downloaded and installed with Cygwin from the Official GRASS GIS website. GRASS does not utilize the extension method for upgrade as ArcGIS does, so full functionality of the program was included with initial download.

Both programs were installed on a Systemax computer with an AMD Athlon 64 3200+ processor, 1024 MB of RAM, and a Microsoft Windows XP Operating System. The hardware configuration significantly surpasses the minimum requirements for each of the GIS programs being used (Table 1).

Database Creation

Due to the UNIX file structure of GRASS as opposed to the Windows file structure of ArcGIS, the methods used by the two programs to access and store spatial data are very different. ArcGIS 9.0 has two methods of storing and accessing spatial data. Spatial datasets can be stored in individual Windows files that can be accessed from any location, or in a geodatabase that organizes spatial data on an object relational basis. Geodatabases are useful for businesses and organizations that have multiple users accessing and editing large amounts of spatial data (ESRI, 2004). Traditional file based directories were used for this project. An ArcGIS data directory was created and used to store data specifically being used for the ArcGIS portion of this study.

Unlike ArcGIS, the user must set up a specified database when using GRASS. This database must be created when starting a new GRASS project. Within the database are Locations and Mapsets. Each Location is based on a specified projection and specified region, with individual Mapsets that can be set up for multiple users (Neteler and Mitasova, 2004).

A GRASS data directory was created to store data specifically being used for the GRASS portion of this study. This directory served as the GRASS database.

Since GRASS does not support on-the-fly projection, which automatically converts and overlays spatial data in different projections, each Location was created based on the projection of the downloaded data to be imported. For instance, if the projection of the initial data to be imported was Lambert Conformal Conic (LCC) with a NAD83 Datum and WGS84 ellipsoid, then a Location was created with these exact specifications in order to import the data into GRASS. These data files were then reprojected from the LCC Location into the Universal Transverse Mercator (UTM) Location that was used for final analysis.

Data

Digital geospatial data was obtained online from the Oregon Geospatial Data Clearinghouse (Oregon Geospatial Data Clearinghouse, 2005). Satellite imagery was obtained from the Global Land Cover Facility (University of Maryland, 2005), administered by the University of Maryland.

Initial file formats of the downloaded data included satellite imagery in TIF (Tagged Information File) format; vector data in Shapefile format; aerial photos in MrSID (Multi Resolutional Seamless Image Database) format; and Digital Elevation Models (DEM) in Spatial Data Transfer Standard (SDTS) format. All raster data were initially in UTM projection and all vector data were initially Lambert Conformal Conic projection.

Downloaded vector datasets included basemap data for cities, counties, roads, streams, and water bodies for the State of Oregon. City limits for 1996, 1999, and 2003, as well as forested lands, urban growth boundaries, and eco-regions for Oregon were also downloaded and used in the study. Environmental Protection Agency (EPA) Geographic Information Retrieval and Analysis System (GIRAS) land use – land cover datasets were also downloaded in order to compare final classification results.

The Landsat sensor system was developed primarily to acquire Earth resource information (Jensen, 2000). Landsat was chosen as the type of imagery to be used for this study for several reasons. The imagery exhibits the spectral, spatial, and temporal resolution (Table 2) necessary for Earth resource and urban landscape analysis (Jensen, 2000). Landsat has a greater number of bands than the other widely used Earth resource satellite system SPOT, which has only four (Jensen, 2000). Availability was also a factor since there were three Eugene-Springfield area Landsat images available online for free download.

Landsat Imagery from Multi Spectral Scanner (MSS), Thematic Mapper (TM), and Enhanced Thematic Mapper Plus (ETM+) was downloaded from the University of Maryland Global Land Cover Facility (GLCF) and used for the project. The earliest available Landsat data at the GLCF for Eugene/Springfield was a Landsat 2 MSS image from July 30, 1977. A Landsat 5 TM image from July 12, 1987 and a Landsat 7 ETM+ image from July 26, 2001 were also downloaded and used in the study. A full description of vector and raster data is included in Table 3, detailing projection, format, source, and compilation dates of each of the datasets.

In order to minimize error and promote a better comparison of the two GIS, all datasets were converted to UTM projection in both programs.

ArcGIS

To begin this study, all downloaded data were brought into ArcGIS. Downloaded vector datasets were already in shapefile format, which is ESRI's commercial file format. Data available for download over the Internet are commonly in this format, making data use easy when using ESRI products but somewhat more involved if using other software. Landsat images, vector datasets, Digital Elevation Models (DEM), and aerial photos were opened in ArcGIS and saved in map document (.mxd) files for later viewing and analysis.

Basemap data were loaded into ArcGIS and saved in map document files for eventual use in display and analysis of classified images. These base maps included a mosaic of aerial photographs of the Eugene/Springfield area. The aerial photographs being used were 1 meter resolution black and white aerial photographs. These images provided help with land cover classification.

Area boundaries were established in order to encompass the area of study and produce reliable statistics. The aerial imagery mosaic was used along with city limit shapefiles to identify the extent of urban sprawl. Area boundaries for the project were defined according to this extent. The upper left coordinates of the area boundary were set at UTM zone 10 4892328 meters N, 479633 meters E, and the lower right coordinates were set to 48964837 meters N, 512157 meters E. These coordinates translate to an area bounded by the coordinates 123.25 W longitude, 44.18 N latitude and 122.85 W longitude, 43.94 N latitude. Establishment of a consistent boundary throughout the study served as a means of determining a growth percentage and more effectively characterizing urban sprawl of the area. The area was held constant in both programs throughout the course of the study. All subsequent raster datasets were clipped to this extent.

In order to process Landsat Imagery and derive land cover classification datasets, ENVI Remote Sensing Software (Research Systems, Inc., 2001) was employed. Remote sensing software such as ENVI is generally used in conjunction with ArcGIS in order to conduct a study that requires image classification (Eiumnoh and Shrestha, 2000). ArcGIS can be manipulated to produce certain types of classified images through very lengthy processing but the program does not offer any built-in image classification or remote sensing tools.

Landsat images were opened in ENVI and saved as ENVI Standard image files. Supervised image classification was used in order to provide a greater level of user interactivity. As opposed to unsupervised classification, supervised image classification allows the user to define regions of interest (ROI) by selecting pure pixels for each region (Research Systems, Inc., 2001).

In order to define ROIs in ENVI, individual pixels were selected based on their spectral response. Selected pixels were considered pure pixels for the class being defined. For instance, a pixel determined to be completely urban would be selected for a ROI in the urban class. Other pixels with similar spectral readings were grouped in the same region. Each region was then defined as belonging to one of the four classes that were derived from each Landsat image. These classes were Water, Forest, Grass, and Urban and were applied to the images in both the ArcGIS and GRASS portions of the project. Standard deviations were checked for each region in order to produce pure ROIs and more accurately characterize the class. Based on interactive interpretation and ROI standard deviation, approximately 10 – 20 ROIs were defined for each image class. The number of regions defined in each class was based on interactive interpretation using the linking and dynamic overlay tools in ENVI.

Three individual bands from each Landsat image were assigned RGB values in ENVI in order to help determine which pixels were pure. A significantly greater amount of visual information can be extracted when individual image bands are combined to form RGB composites (Jensen, 2000). For Landsat MSS, bands 3, 2, and 1 were used for pixel selection, respectively. Landsat TM and ETM+ image bands 7, 4, and 2 were used for urban pixel selection and bands 4, 3, and 2 were used for grass, forest, and water pixel selection. These bands enhanced contrast between land cover classes and aided in pixel selection.

After all regions had been selected, a maximum likelihood classification was done based on the selected ROIs for that image. Maximum likelihood classification assumes that the statistics for each class in each band are normally distributed (Research Systems, Inc., 2001). The probability is calculated for each pixel being assigned to each class. That pixel is then assigned to the class with the highest probability. Unless a probability threshold is specified, all pixels in the image are assigned to one of the classes. All regions were assigned a threshold of 0.5, except for urban regions, which were given a threshold of 0.9.

Initial probability thresholds were set at 0.5 for all regions. The 0.5 threshold instructed ENVI to assign a pixel to a region only if there was at least a 50 percent certainty that the pixel belonged to that region. If the probability was less than 50 percent for every region, the pixel was left unassigned. This approach produced an incorrect urban classification, due to the fact that pixels found in border areas between classes exhibited a similar spectral response to pure urban pixels. For example, sometimes forest and grass classes were found within one pixel. Due to spatial resolution limitations, this caused a combined spectral response that may be similar to pixels in a completely different class, in this case the urban class.

After several attempts, it was found that if a higher threshold was assigned to urban regions, ENVI would produce a more conservative urban classification. With the urban probability threshold set to 0.9, ENVI would not assign a pixel to an urban class unless there was 90 percent certainty that the pixel belonged to that class. A probablity threshold of 0.9 reduced the amount of border area pixels that were incorrectly assigned to the urban class. The same process was completed for each Landsat image and resulted in classified raster datasets for each image.

ArcGIS Post-classification

After the classification process was complete for each of the three Landsat images, classification images were saved in the GeoTIFF format and opened in ArcGIS. The Reclassify tool in the Spatial Analyst Extension was used to combine the individual ROIs into five classes: unclassified, water, forest, grass, and urban.

Upon completion of reclassification, the reclassified raster was converted to a polygon vector dataset for statistical analysis. The Calculate Areas tool in ArcToolbox was then used to determine the area of each individual polygon in each class. When using the Calculate Area tool, ArcToolbox creates a new field in theattribute table of the dataset called F_area with the area of each polygon given in square meters. These records were queried based on class. After areas had been calculated, urban class records were queried and selected and then exported to a text file. In order to produce class statistics, the text file was then imported into Microsoft Excel where each class was summed and area values were converted to other units. Percentage of urban area for each class was also calculated. This process was done for each classification image.

After calculation of class statistics, an accuracy assessment was done in ENVI in order to determine accuracy and error statistics. The kappa value is a coefficient generally used to determine accuracy in land cover classification projects (Bektas, 2003). The kappa coefficient is a measure of agreement between two individual elements, in this case two raster datasets. The coefficient evaluates how well the classified images represent the real world. Kappa is always less than or equal to 1, with a value of 1 implying perfect agreement and a value less than 1 implying less than perfect agreement. ROIs that had been previously defined for each Landsat image were used to produce a training map. In order to accurately assess the amount of classification error, both the classified image and the training map were reclassified by combining individual ROIs into the appropriate class. The ENVI confusion matrix tool was then used to conduct an accuracy assessment by comparing the training map to the classified image.

Upon completion of the ArcGIS analysis, hardcopy maps were created. Map layouts included a legend, compass, scale, and a small statistics table for easier comparison of urban area statistics. The map layout was done entirely in ArcGIS, with the exception of the urban statistics table, which was added using Microsoft Word.

For another area of comparison to GRASS, an image displaying the Landsat ETM+ image draped over a shaded relief map was created. The shaded relief map image was created using the Spatial Analyst Extension in ArcGIS. Three dimensional visualizations based on DEMs were also produced for final display of the datasets. Three dimensional (3D) visualizations improve analysis of data by incorporating elevation and aspect as part of the study and providing a means for understanding complex spatial relationships (Neteler and Mitasova, 2004). The user can then determine whether or not these characteristics play a role in urban development. Creating 3D images also allowed another area of comparison for the two programs.

GRASS

In order to view and analyze data in GRASS, all data were imported into GRASS vector or raster format from their native formats. All Landsat data were imported into a GRASS Location set up in UTM projection. Each Landsat image was imported one band at a time. All vector data were imported into an alternate Location set up in Lambert Conformal Conic (LCC) projection. Region settings, which define the analysis extent, were set to the same coordinates used for ArcGIS.

Once all data had been imported into the proper GRASS Location, vector data in LCC projection were reprojected into UTM, using the r.proj module. R.proj, as well as a variety of other modules, can be started through the GRASS GUI or by the GRASS command line prompt known as the bash shell.

After all data were reprojected into UTM coordinates, the analysis portion of the study could begin. All Landsat images were combined into imagery groups using the i.group module. This tool combines individual bands of an image into one group and/or subgroups, each consisting of at least two bands. For this study, subgroups were created to include all bands that would be used for classification. Bands 1, 2, 3, and 4 were subgrouped for the MSS image and bands 1, 2, 3, 4, 5, and 7 were subgrouped for TM and ETM+ images.

After image groups were created, false color composite images were created using the r.composite module. This module creates a single composite map layer based on specified red, green, and blue layers. Composites were created for each Landsat image using bands 3, 2, and 1 and bands 4, 2, and 1 for MSS and bands 7, 4, and 2 and bands 4, 3, and 2 for TM and ETM+. Composite images were also created for TM and ETM+

based on outputs of the i.oif module. This module calculates an optimum index factor table and determines which three TM bands produce the greatest contrast between land cover classes. Based on output from this module, RGB composites were created with bands 3, 4, and 5 for the ETM+ image, and bands 1, 4, and 5 for the TM image. The module does not support MSS imagery.

Interactive Classification

Interactive supervised classification in GRASS is very different from the ENVI/ArcGIS method. All classification was done in GRASS itself using the i.class and i.maxlik modules. The i.class module was used first. The first image to be classified was the Landsat 2 MSS image from 1977. I.class was started and the image group for the Landsat image being classified was selected. The previously created composite image composed of bands 4, 2, and 1 was specified as the map layer to use to define training areas. GRASS training areas are equivalent to regions of interest in ENVI.

The i.class module produced a signature file, which is needed in order to conduct a maximum likelihood classification using i.maxlik.

Once i.class started, training areas were selected based on land cover. The way GRASS approaches classification is somewhat opposite of ENVI. In ENVI, pixels were selected for ROIs and ENVI calculated a standard deviation for the selected pixels. Pixel selection was then based on the standard deviation calculations and pixels were added or removed as needed.

In GRASS, training areas were defined by digitizing a polygon around the pixels to be selected. Once the training area was defined, i.class was used to calculate a mean spectral signature based on pixels within that training area. Based on a histogram of each image band on the screen, a standard deviation was specified. GRASS then displayed the pixels that fell within the specified standard deviation of the mean spectral signature.

For this study, standard deviations were set between 2.0 – 5.0 for TM and ETM+, and 4.0 – 8.0 for MSS. Due to the lower number of bands present in the MSS data, the lower standard deviation values were insufficient to produce an accurate MSS classification. Once the desired amount of pixels was selected, the training area was saved to the signature file and a new training area was digitized. This process was done repeatedly until a sufficient number of training areas were selected to produce an accurate classification. GRASS reports training area statistics at the time of interactive classification but does not have a tool that reports statistics for previously defined training areas. All standard deviations and size samples were recorded manually during interactive classification.

The next step in the GRASS image classification process was to run the i.maxlik module. This module produced the final classified raster for the Landsat MSS image. The signature file that was generated from the i.class module, as well as the image group and sub group, and a resultant output file were specified in i.maxlik. A maximum likelihood analysis was then run on the image and the classified image was generated. This process was completed for each of the Landsat images being used for this study. Since GRASS does not incorporate the same type of hard copy map design and layout tools as ArcGIS, map layouts were created outside of GRASS using GIMP open source image manipulation software in combination with Microsoft Paint.

For comparison with ArcGIS, the classification for the Landsat ETM+ image was draped over a hillshade of the region. Using the NVIZ module, 3D images displaying the

ETM+ classification as well as a real color RGB composite were also developed for comparison to ArcGIS. In addition, a composite dataset of each classification was produced for final display of results.

Non-interactive Classification

GRASS interactive classification using i.class allows pixel viewing by interactively manipulating standard deviations but does not allow these statistics to be saved to a file that can be worked with at a later time. Once a training area is saved to a signature file, the training area or the statistics associated with it can no longer be viewed. This limits the user's ability to assess accuracy.

In order for the user to complete an accuracy assessment, the user must digitize training areas into a separate raster or vector map. This training map is then input into the i.gensig module, which creates a signature file to be used in i.maxlik. Just as using the interactive classification method, the i.maxlik module uses the input signature file to create a classified output image.

The non-interactive classification method allows the user access to the original training map. This map can then be compared against the classified image in the r.kappa module, which produces an error matrix including percent omission, percent commission, number or pixels correctly classified, and overall kappa coefficient. Although the non-interactive method gives the user access to original training data and allows an accuracy assessment, it does not allow the user to manipulate standard deviations or view regional histogram data for each band.

To better compare GRASS with ArcGIS/ENVI, a non-interactive classification was completed in addition to the interactive classification. Training maps were digitized using the same Landsat RGB composites that were created during the interactive process. These maps were then used to generate signature files that were used by i.maxlik for classification.

After non-interactive classification for each Landsat image, r.kappa was used to produce an error matrix for each. R.kappa calculates percent omission, percent commission and an estimated kappa coefficient for each class. The r.kappa module also calculates overall kappa, kappa variance, number observed, number correct, and percent correct.

GRASS Post-classification

Initially, each training area generated its own subclass and was displayed separately. Once the classified images had been generated, the r.reclass.rules module was run in order to combine the individual training areas into appropriate classes.

The r.colors.rules module was used to create symbology similar to the symbology used for classified images in ArcGIS. Using like symbology promotes a more accurate comparison of images generated in ArcGIS/ENVI with those created in GRASS.

Once each image had been classified and displayed appropriately, statistics for each classified image were generated. Using the r.report module, statistics were calculated and displayed based on user specified units. Statistical reports were generated in square miles, square meters, acres, hectares, pixels, and percentage.

EPA GIRAS Data

EPA GIRAS land cover – land use data were used to compare both ArcGIS and GRASS class statistical results. EPA GIRAS data was collected by the United States Geological Survey (USGS) National Mapping Division between the mid 1970s and the early 1980s. The national GIRAS dataset was compiled and published in 1994 (U.S. Environmental Protection Agency, 1994). These statistics served as a general comparison for both ArcGIS and GRASS class statistical results. The GIRAS data sets were clipped to the same region boundaries used for both programs. A map layout was created for the EPA GIRAS dataset to serve as a visual comparison for both GRASS and ArcGIS classification images. The same methods used to determine class statistics for the GRASS portion of the study were implemented for this dataset. .

Overall Functionality and Ease-of-Use Scoring

At the end of the case study, results were evaluated and a performance level of excellent, good, fair, or poor was given to each program for each functionality category in order to provide an easy to-read multi-categorical comparison of both programs.
To determine which program scored higher overall, the performance levels were then assigned a numerical score. Numerical scores for performance levels were:
Excellent = 4; Good = 3; Fair = 2; Poor = 1; and No functionality present = 0.

RESULTS

ArcGIS Results

Cost

The initial financial cost of purchasing ArcGIS software varies greatly, depending on the users needs. ESRI offers three different license levels of ArcGIS 9.0. These licenses start with the ArcView license, which gives the user basic display and analysis capability. The intermediate license, ArcEditor, includes all the functionality of ArcView but also offers extended editing and analysis capabilities. The ArcInfo license gives the user even more options by including more advanced geo-processing and data conversion capabilities. Many different extensions are also available for purchase for ArcGIS 9.0, giving the user further analysis, data conversion, or display capabilities. A list of costs is included in Table 10.

Installation

The single user license installation of ArcGIS 9.0 used for this project was very straightforward. Total installation time took approximately 15 minutes. Extensions were installed at the same time as the software itself by entering key codes that were supplied by ESRI. ArcGIS installation supported registration of the product in several ways – by fax, mail, or internet. The online method was chosen for this study and ArcGIS automatically registered the product within a matter of seconds.

ArcGIS 9.0 installation requires any earlier version of ArcGIS (8.x) to be completely uninstalled, but does not require removal of ArcView 3.x. Anyone running versions of ArcGIS 8.x will not be able to run it on the same computer as ArcGIS 9.0.

Training and Technical Support

ESRI, as well as numerous third party vendors, offer an abundance of manuals and training courses for the ArcGIS software (ESRI, 2004). ESRI also offers technical support and an extensive online knowledge base for ArcGIS. Neither technical support nor the knowledge base was needed during the course of this project.

Database Creation, Data Import and Display

In general, accessing and storing spatial data with ArcGIS was straightforward. Spatial data could be saved in any Windows directory and accessed directly from that location. ArcGIS also offered a geodatabase option that was more structured and similar to the GRASS database. The geodatabase option is good in environments where multiple users are accessing and editing data (ESRI, 2004). Traditional file based directories were used for this project.

A variety of data formats were used during this study (Table 3). All spatial data that were downloaded for this project was already in a format that was readable by ArcGIS and quickly opened when selected.

In ArcGIS itself, there were very few problems encountered when displaying and working with the data downloaded for this project. However, when exporting data from ArcGIS to other file formats, the resultant export files could not be read in two other programs. When working with raster data in ArcGIS, the default file format was a GRID (an ESRI proprietary raster file format). GRIDs worked very well in ArcGIS, but the “raster to other format” tool did not export the GRID to alternate formats as it was designed. ArcGIS, in the configuration used for this project, offers the user three raster export options: .tif (TIFF), .img (ERDAS Imagine), or ASCII. Neither the exported

ERDAS Imagine file nor the TIFF could be read by other software. Both ENVI and Lizardtech GeoExpress, a program that writes raster data to MrSID format, failed when attempting to read the exported files. Each program indicated that the image header that was created by ArcGIS during export could not be read.

ESRI technical support was not contacted about this problem since converting these files was not necessary for analysis. The raster files being exported were aerial imagery DOQQs that were to be mosaicked and written to MrSID format. The individual files were used instead. It is unknown whether this problem was a limitation of the software or an interoperability conflict between the programs being used.

Although problems were encountered when exporting raster data from ArcGIS, the program was very useful when working with vector data. Underlying vector attribute data could be quickly viewed, queried, and or edited by selecting the "open attribute table" feature. Export of vector attribute data was quick and straightforward when exporting to both .txt (text files) and .dbf (dbase files). This feature was used to export the F_area field created during area calculation of classified images. Exported files were successfully opened in other programs such as Microsoft Access and Excel.

In addition to the straightforward interface (Figure 3), it was also found that a very useful feature incorporated in the ArcGIS software is on-the-fly projection. This feature added to both data import and display qualities of the software.

On-the-fly projection allowed quick display of data in any projection. When the first dataset was added to the map, all subsequently added data were automatically displayed in the projection of the initial dataset. The projection could then be re-defined as needed. This feature was employed to quickly display downloaded LCC data in UTM projection. On-the-fly projection was valuable when quickly displaying and querying data but did not reproject the underlying dataset. The data were only displayed in the defined projection. In order to conduct accurate analysis, the data were reprojected.

Image Classification

As stated in the Methodology chapter, ArcGIS does not offer image classification tools, a major limitation of the package. In order to complete a satellite image analysis in ArcGIS, a remote sensing package - in this case ENVI - was required for segmenting images into types.

Although ENVI was not directly tested in this study, the program was very straightforward and useful when conducting image classification. ENVI maximum likelihood classification employed a very straightforward interface (Figure 4) and gave the user interactive access to ROI statistics (Table 5) during classification. Being able to view these statistics during classification made pixel selection much easier and ultimately more accurate. Classification results for each Landsat image can be seen in Figures 5 – 7. Image classification results are comparable to those of the downloaded EPA GIRAS land use / land cover dataset for the Eugene-Springfield area (Figure 8).

Even with the efficiency and straightforwardness of ENVI, file conversion between ArcGIS and ENVI increases time involved and requires higher user experience level. It also adds a significant financial investment to purchase the remote sensing software (Table 4).

Post-classification Analysis

ArcGIS is limited in the amount of post classification analysis that it can perform on the data in this type of project. ArcGIS was useful in calculating areas of individual classes of the classified images. Classified images were converted to vector datasets in order to employ the calculate areas tool. When implemented, the tool created a new vector dataset with all of the attribute data of the old one, and added an F_area column that contained area data in square meters for individual polygons. The quickest way to calculate class area was to query the data based on each class, then export selected records to a text file. The text file was then imported into Microsoft Excel where it was summed and converted to other units. Conversion to other units was done manually by entering conversion formulas into the Excel spreadsheet.

The query and export strategy produced the desired class area statistics but took several hours to complete. Class statistics for the ArcGIS/ENVI portion of the study are displayed in Table 6, showing that using the process of query and export will produce the class area statistics. These results are comparable to class statistics calculated for the downloaded EPA GIRAS land cover classification dataset (Table 7).

In order to evaluate the error of the ArcGIS/ENVI classification, ENVI was used to conduct the accuracy assessment. ENVI offers specific built-in tools for the user to employ when conducting an accuracy assessment, ArcGIS does not.

ENVI was efficient when conducting an accuracy assessment. The software offers a confusion matrix tool that will analyze a training map created from ROIs and compare it to the classified image that was created from them. As seen in Table 8, results include the total number of pixels observed, total number of pixels observed that were classified correctly, percent correct, percent commission, percent omission, producer and user accuracy, and an overall kappa coefficient. These parameters were very useful when assessing classification accuracy. The ArcGIS/ENVI classification process was extremely accurate, exhibiting at least 99.69 percent of the pixels observed to be correct for each of the classified Landsat images.

The hillshade tool in the ArcGIS Spatial Analysis extension is very useful when displaying or interpreting imagery. The hillshade tool, which produces a two dimensional (2D) shaded relief map using an input DEM dataset, allows the user to overlay other raster and vector datasets (Figure 9). The result is a 2D map with an elevation effect.

The ArcGlobe tool in ArcGIS offers the user many 3D display options. Both Figures 10 and 11 were produced using ArcGlobe. Figure 10 displays a 3D view of a real-color RGB composite of the ETM+ image while Figure 11 displays a 3D view of the ETM+ classification image.

Map Layout Capabilities

ArcGIS offers many styles of built-in compass arrows, legends, scales, text, and borders. In addition, each built-in map element could be customized in a variety of ways. This functionality was used to create a template for multiple map print outs, seen in Figures 5 - 7. These figures not only illustrate the type of map layouts that can be created using ArcGIS, but they also display the classification results from each of the Landsat images. Although outside the scope of this study, Eugene-Springfield urban areas were 7.5 percent in 1977, 11.53 percent in 1987, and 16.97 percent in 2001. These figures show that ArcGIS, used in conjunction with ENVI, is a useful means of conducting a time series growth analysis.

GRASS Results

Cost

There is no financial cost associated with the GRASS software itself. It is available for free download and is released as open source software under the General Public License (GPL).

Installation

Although the installation of GRASS on Cygwin is somewhat different than installation of most Windows based applications, it is nevertheless fairly straightforward and easy to understand. Both Cygwin and GRASS can be installed at the same time. One thing that should be mentioned is that the installation wizard requires the user to select the GRASS database when installing Cygwin, entailing connection to one of several servers that are listed for Cygwin download as well as adding a specific server for GRASS download. If the user does not specify both, Cygwin will be installed without GRASS.

Total download and installation time took about 30 minutes on a T1 Local Area Network (LAN) connection. Overall, installation of GRASS was comparable in difficulty to installation of ArcGIS, although it took approximately twice as long.

Training and Technical Support

The amount of GRASS reference manuals and training available at the time of this study was very limited. During the course of this study, the GRASS worldwide development was in the process of developing an online GRASS 6.0 tutorial. The tutorial, which can be found on the GRASS website, is an outline of the modules incorporated into the software. Each module has a dedicated web page that explains how to use that module. In general, this tutorial explained GRASS well but assumes the user is already familiar with GIS concepts and procedure.

Database Creation, Data Import and Display

The GRASS database is much more structured than that of ArcGIS, which can be both positive and negative. In general, the GRASS database does not allow the user as much freedom in working with geo-spatial data. GRASS forces the user to develop a specific Location with a specific analysis area (region) in a specific projection. In order to work with data outside the region or in another projection, the region settings of the database were altered or the data were reprojected from another Location.

Once the GRASS database design concept was fully understood, import and export of data in several different formats was rapidly conducted. Import and export of numerous data formats were conducted during this study. The variety of import and export format options that GRASS provides was very valuable in that it offered the user greater facility with working with different data sources.

This study showed that display of data in GRASS is somewhat limited. One of the biggest limitations of the software is that it does not incorporate on-the-fly projection, forcing the user to reproject all datasets in order to overlay and view multiple layers. GRASS did provide the ability to open multiple display monitors at the same time, which proved to be a very helpful feature when working with the large number of map layers that were involved in this study. The GRASS GIS Manager (Figure 12) was also very straightforward and gave a GUI alternative to the command line bash shell.

Image Classification

Unlike ArcGIS, GRASS offers all the tools necessary for an image classification analysis. These classification tools were found to be somewhat lacking in their capabilities. Both interactive and non-interactive supervised maximum likelihood classifications (MLC) studies were conducted during this project in order to produce all of the necessary statistical results.

The first method, interactive MLC, gives the user the ability to select pixels interactively by defining training area polygons. As seen in Figure 13, GRASS reports training area statistics at the time of interactive classification. Interactive MLC was very useful in that the user could view the pixels that were within a defined standard deviation of the mean value of training area pixels. However, GRASS does not have a tool that reports statistics for previously defined training areas. In addition, lack of ability to view previously defined training areas during interactive classification is a major limitation of the software that will hopefully be addressed in the next version of GRASS. Without the ability to view previously defined training areas, an accuracy assessment cannot be conducted. In order to keep track of training area statistics during interactive classification, training area sample sizes and standard deviations were manually recorded and listed in Table 9.

Another limitation of interactive MLC that should be mentioned was the "choppiness" of the i.class module interface. Running the i.class module for interactive classification used almost 100% of the system processor, resulting in a slow computer response time. Whether this is a bug in GRASS itself, or the result of running GRASS using Cygwin is unknown.

The lack of a way to view previously defined training area data resulted in the inability to conduct an accuracy assessment. It was found that if non-interactive classification was used, a training map could be created that could later be used for an accuracy assessment. This method forces the user to select all training areas at one time and does not give the user the ability to control standard deviations or view pixels selected for a given training area. A greater level of user image interpretation experience was needed in order to accurately define training areas.

The non-interactive method ultimately proved to be the best option for this type of classification. Data represented in Tables 10 and 11 show that class statistics for both interactive and non-interactive MLC were very similar. An accuracy assessment could not be performed using interactive MLC, but was conducted from the training map during non-interactive MLC.

The fact that GRASS offers the ability to conduct image classification is very impressive. By comparing classification images to aerial imagery, both methods of classification proved to produce classified images that were visually accurate. Both methods also produced images comparable to the EPA GIRAS dataset (Figure 8). It was found that non-interactive classification was the only useful method if quantitative analysis of accuracy is needed.

Post Classification Analysis

One of the most useful features of GRASS is its post-classification analysis. Other than the lack of ability to generate training area statistics, which was addressed in the image classification section, GRASS offered all of the tools needed to generate statistical reports. As seen in Tables 10 and 11, the r.report module gives the user the ability to generate class statistics in miles, meters, kilometers, acres, hectares, cell count, or percent cover. The r.report module was very useful and displayed the statistical results in an easy-to-read list.

When compared to ArcGIS, class statistical results were very similar. Interactive classification image results can be seen in Figures 14 -16. Urban class percentages for GRASS interactive classification were 5.68 percent in 1977, 11.11 percent in 1987, and 15.98 in 2001. GRASS non-interactive classification image results exhibit similar numbers, showing 5.01 percent in 1977, 8.57 percent in 1987, and 16.44 percent in 2001 (Figures 17 – 19). These are compared to the ArcGIS urban class statistics of 7.5 percent in 1977, 11.53 percent in 1987, and 16.97 percent in 2001. As in the case of ArcGIS and ENVI, these figures illustrate that GRASS is a useful means of conducting a time series growth analysis. Class results were also in line with EPA GIRAS dataset statistics (Table 7).

The r.kappa was also found to be very useful by reporting accuracy parameters in an easy-to-read error matrix. This matrix was straightforward and provided a means of assessment of classification accuracy. Results of the non-interactive MLC accuracy assessment are displayed in Table 12. Accuracy parameters calculated by the r.kappa module included percent omission, percent commission, estimated kappa for each class, percent observed correct, overall kappa, and kappa variance. Percentages observed correct were 98.35 percent or higher for each Landsat image, showing GRASS classification accuracy to be comparable to ArcGIS/ENVI classifications.

To further display the capabilities of GRASS for a land use classification analysis, an urban boundaries composite was created from interactively classified images (Figure

20). This image illustrates urban growth from 1977 through 2001. GRASS is very useful when developing images such as that seen in Figure 20.

GRASS also offers shaded relief and 3D capabilities comparable to ArcGIS. Figure 21 displays a shaded relief map created in GRASS with the ETM+ interactive classification image draped on top. The GRASS shaded relief module is very straightforward and produced an image that was visually comparable to the ArcGIS shaded relief map (Figure 9).

GRASS 3D capabilities are also very useful. Figure 22 displays a 3D view of a real-color RGB composite of the ETM+ image and Figure 23 displays a 3D view of the ETM+ classification image. The NVIZ module that was used to create these images is very straightforward and offers functionality that is very useful when conducting a land use classification analysis.

Overall, the modules incorporated in GRASS for post-classification analysis are very useful features that can be quickly employed. Post-classification analysis should be considered a major selling point of GRASS.

Map Layout Capabilities

GRASS does not offer useful tools for map design and layout. Tools for creating hard copy maps are limited in GRASS because of its focus on modeling and spatial analysis (Neteler and Mitasova, 2004). GRASS gives the user the ability to add only a very simple and standardized legend, compass, and scale to the display monitor and then export the display to an external image file such as a .png or a .tif. These map features are very basic and can not be customized.

Although GRASS does not incorporate many cartography tools, map layout and design was easily accomplished in an external image manipulation program such as open source GIMP or Adobe Photoshop. This project used a combination of GIMP and Microsoft Paint to produce the map layouts that are seen in Figures 14 – 17 and Figures 20 - 23. These figures illustrate the classified images produced in GRASS using both interactive and non-interactive MLC. Using GRASS with an external image manipulation program will produce map layouts comparable to those designed in ArcGIS (Figures 5 – 7).

Scores

Overall, both programs proved to be very useful tools. Total performance scores were very close. GRASS scored 31 points, slightly higher than the 30 points scored by ArcGIS. Table 13 provides a performance level for each program for each category. A brief description of the program's performance for each category is included. In order to easily compare and contrast the two programs, categorical and total scores associated with performance levels are listed in Table 14.

CONCLUSIONS

The results of this study support the hypothesis. Initially, GRASS GIS implementation is more involved and time consuming than ArcGIS, mainly due to the fact that there is very little training or tutorial material available for GRASS. The fact that GRASS is a UNIX based program and requires Cygwin to run on a Windows operating system can be a source of further complication.

Although GRASS implementation is more involved, it does produce accurate results and it is an effective alternative to ArcGIS. Once the initial problem of installing and learning how to use GRASS had been solved, the open source software provided analytical capability and function comparable to ArcGIS. In some areas, GRASS capability and function was found to be superior to that of ArcGIS. In other areas, GRASS lags behind. Overall, results of this study show that GRASS is an effective, low cost alternative to ArcGIS that will give the user comparable analysis and display capability for a fraction of the cost.

Each of the programs that were tested exhibit a variety of advantages and disadvantages. This section lists categorical recommendations based on the results of this study and categorical performance level scores listed in Tables 13 and 14. Each recommendation includes a brief explanation on why the recommendation was made for that particular category.

Cost

GRASS is undeniably the best program in this category. A complete cost comparison can be seen in Table 4. The comparison does not include costs associated with computer hardware, data, or user training.

Although GRASS does not require an initial financial investment, consideration should be given to the fact that ArcGIS is more straightforward and easier to learn for inexperienced GIS users. In the business world, time often translates into money. Depending on the user's needs and the project that the GIS is being implemented for, the amount of time it takes to train others on GRASS may be a source of increased financial investment.

Installation

Installation of both of these programs is very rapid. Installation of GRASS 6.0 is slightly more involved, since it requires the installation of Cygwin.

Training and Technical Support

If the user is unfamiliar with either ArcGIS or GRASS, then the user may want to consider purchasing an ArcGIS license. ArcGIS definitely has the advantage in this category. At this point in time, and depending on the potential user's knowledge and experience with GIS, learning the GRASS software will require a much greater investment of time than ArcGIS. The amount of training materials and courses related to ArcGIS greatly exceeds the amount of training materials available for GRASS.

ESRI offers an extensive online knowledge base for ArcGIS and the GRASS website also provides a search engine for user mailing list archives. If technical support is needed then the user will need to purchase ArcGIS. Currently, there is no technical support for GRASS. In order to get professional help, the GRASS user would have to contact a commercial GRASS consultant. There are several listed on the GRASS website, but compared to ArcGIS there are very few options.

Database Creation, Data Import and Display

ArcGIS and GRASS use different methods to access and store spatial data. The type (raster or vector) and file format of the data being used must be taken into account in order to fully compare the two.

Raster data is generally available in standardized image formats such as GeoTIFF or Spatial Data Transfer Format (SDTS). ArcGIS supports viewing, analysis, and editing of raster data in TIFF, GRID, or ERDAS Imagine file format. In order to manipulate the raster in ArcGIS, the data must be in one of these three formats. The ability to use any of these three formats is valuable when using multiple programs for analysis. However, as stated in the Results section, there were problems when reading these file formats in other software.

GRASS gives the user a wide variety of file format import options, but the datasets must always be imported (converted into a format that can be read by the program) into GRASS raster format before viewing or manipulation can take place.

A large majority of vector data online is already in ESRI's proprietary file format. Shapefile data can be directly loaded into ArcGIS, but must be imported into GRASS. There is very little, if any, data that are in GRASS vector format available for download. This translates into an increased amount of time and effort when working with GRASS.

The database structure of GRASS requires more user experience with GIS. The user must know in exactly which projection the initial data is and exactly which projection is ultimately needed. Bounding coordinates of the region must also be known. This requires a greater time investment and greater user experience level than that of ArcGIS.

There are benefits to this structured database environment. The user is forced to know the parameters of the data, including metadata such as projection, scale, datum, ellipsoid, and resolution, in turn promoting a better user understanding of the data as well as a better understanding of analysis results. A structured database environment also forces the user to store data in an organized manner, based on projection and location.

In general, ArcGIS is more straightforward and easier to understand than GRASS when storing and working with geo-spatial data. A novice GIS user can sit down and display data within a matter of minutes. It requires less time than GRASS from the time that data are downloaded until the time that data are displayed. This is partially due to the on-the-fly projection that is incorporated into ArcGIS. On-the-fly projection is a valuable feature that allows automatic overlay of data, no matter what the initial projection but it does not transform the underlying data, requiring that the initial data be reprojected when performing analysis. The difference between on-the-fly projection and actual re-projection can be confusing to an inexperienced GIS user.

Although not as easy to use, GRASS does offer similar data storage and access. Import and re-projection of data into GRASS works quickly. Once the data are imported and re-projected into the proper Location, they are viewed, edited, and analyzed in a manner comparable to ArcGIS. The database structure of GRASS promotes a data analysis environment but lags behind ArcGIS when it comes to user friendliness.

A major benefit of GRASS is the increased ability for import and export. GRASS offers a much greater variety of import and export options than does ArcGIS. GRASS can be used to import and export almost any type of vector or raster dataset.

Comparable data import and export is also available in ArcGIS, but is only offered as an extension.

The best program in this category will depend on the user's specific needs. If the user wants to quickly display and query data, ArcGIS will be the best choice. If the user wants to ultimately focus on analysis and has a variety of data formats to work with, GRASS will be the best option.

Image Classification

The fact that ArcGIS does not offer image classification tools severely limits its usefulness in conducting a land use classification study such as this. However, when ArcGIS is used in conjunction with ENVI, the user is given the ability to perform all the necessary classification steps.

When choosing between GRASS and ArcGIS, GRASS is the obvious choice for image classification. GRASS offers the tools necessary to successfully classify imagery and ArcGIS does not. When compared to ENVI, GRASS is slightly harder to use and doesn't offer as many options. ENVI gives the user slightly more classification functionality than GRASS, such as the types of classifications that can be conducted. ENVI offers Parallelepiped, Minimum distance, Mahalanobis distance, maximum likelihood, spectral angle mapping, and binary encoding supervised classification. GRASS offers only maximum likelihood and sequential maximum *a posteriori* (SMAP) classification. ENVI also gives the user access to more functionality when generating statistical data during interactive supervised classification. Although ENVI offers somewhat more functionality than GRASS in this category, it is a specialized remote sensing package and it does not possess the same data editing ability that GRASS exhibits.

Post Classification Analysis

ArcGIS does not incorporate the necessary tools in this category. The query and export strategy used to obtain class statistics in ArcGIS was a much more cumbersome method than using the straightforward, specific tools that are included in GRASS. ENVI, is similar to GRASS in post-classification analysis capabilities. A comparison of ENVI and GRASS in this category would probably be more appropriate.

Map Layout Capabilities

ArcGIS is, without a doubt, the best choice when it comes to map layout capabilities. Creating a map layout in ArcGIS, for either digital or hard copy use, is straightforward and efficient. The map layout capabilities greatly exceed those of GRASS and should be considered a major point in favor of ArcGIS.

Overall Functionality

Results show that both programs exhibit advantages over the other in specific areas. In general, ArcGIS is the better program when it comes to cartographic design. If the user wants a program that makes it easy for inexperienced GIS users to create maps, ArcGIS is best. It is also much easier to use than GRASS. However, ArcGIS is inferior to GRASS when it comes to analytical capabilities.

GRASS exhibited exceptional performance in numerous categories. It does need improvement in certain areas. It is hard to learn and harder to use than ArcGIS. It is also inferior to ArcGIS in cartographic design and map layout capability.

GRASS not only provides an effective alternative to ArcGIS, but its performance is superior to that of ArcGIS in many ways. It is much better than ArcGIS in statistical analysis. It also gives the user the ability to conduct image classification, something ArcGIS does not offer. Combine these advantages with the financial savings associated with using GRASS and it is hard to see why more people do not choose this option.

DISCUSSION

Results of this study supported the hypothesis and showed that GRASS is an effective alternative to ArcGIS within the limits of the land use classification case study. There are still numerous areas that remain to be compared, including several key functions of GIS. Further study needs to be conducted in order to completely assess the advantages and disadvantages of both programs.

These areas include GIS functions such as data creation and editing, which are common GIS applications. Future studies should be conducted to determine if GRASS data creation and editing capabilities are as effective as those of ArcGIS.

Studies involving alternative scoring methods that weight categorical raw scores should also be conducted. This type of study would be useful when evaluating software based on categorical needs. Although scores may change in a weighted study, the results of this study prove that GRASS is a useful alternative to ArcGIS.

The question of whether GRASS is an effective program when employed in an enterprise GIS setting (a setting in which multiple users access and use data simultaneously) is also unanswered. If it can be implemented, what are the constraints and limitations of GRASS in this type of GIS architecture? If these questions are addressed, GRASS could one day become as common in the business world as ArcGIS.

There are undoubtedly many more areas in which GRASS can be compared to other GIS packages. Future projects should include comparison of GRASS to other GIS software, such as Intergraph GeoMedia or MapInfo.

Continued research comparing GRASS to other GIS packages supports the open source concept. Not only can results prove the ability of GRASS to skeptics, but it can also aid GRASS developers in improving the software where it is insufficient. These improvements should bring more GIS users and professionals into the GRASS community, in turn leading to further advances in GRASS and GIS in general and ultimately benefiting the public.

LITERATURE CITED

Bektas, F. 2003. *Remote sensing and geographic information integration: A case study; Bozcaada & Gokceada Island*, Msc Thesis, Institution of Science and Technology, Istanbul Technical University.

Clarke, K. C. 2001. *Getting started with geographic information systems.* Upper Saddle River, New Jersey: Prentice-Hall. 352 pp.

Dueker, K. J. 1979. Land resource information systems: A review of fifteen years experience. *Geo-Processing*, vol.1, no.2, pgs.105-128.

Eiumnoh, A. and Shrestha, R. P. 2000. Application of DEM data to Landsat image classification: Elevation in a tropical wet-dry landscape of Thailand. *Photogrammetric Engineering and Remote Sensing*, vol.66, no. 3, pgs. 297-304.

ESRI. 2004. *ArcGIS 9: What is ArcGIS?* Redlands, California: ESRI Press. 119 pp.

ESRI. 2005. *ESRI History*. Retrieved June 21, 2005, from ESRI corporate website: http://www.esri.com/company/about/history.html.

Hunter, L. M., Gonzalez, M. G., Stevenson, M., Karish, K., Toth, R., Edwards, T. C., Lilieholm, R.J. and Cablik, M. 2003. Population and land use change in the California Mojave: Natural habitat implications of alternative futures. *Population Research and Policy Review*, vol. 22, pgs. 373-397.

Innes, J. E. and Simpson D. M. 1993. Implementing GIS for planning: lessons from the history of technological innovation. *Journal of American Planning Association* vol.59, no.2, 230 pp.

Jensen, J. R. 2000. *Remote sensing of the environment: An Earth resource perspective.* Upper Saddle River, New Jersey: Prentice-Hall. 544 pp.

Longley, P. A., Goodchild, M. F., Maguire, D. J. and Rhind, D. W. 1999. *Geographical information systems: Management issues and applications, Volume 2.* New Jersey: John Wiley and Sons, Inc.. 1296 pp.

Morain, S. 1999. *GIS solutions in natural resource management: Balancing the technical-political equation.* Santa Fe, New Mexico: Onword Press. 364 pp.

Neteler, M. 2005a. *GRASS history*. Retrieved June 28, 2005, from the official GRASS GIS homepage website: http://grass.itc.it/.

Neteler, M. 2005b. *Introduction to GRASS*. Retrieved Feb. 08, 2005, from official GRASS GIS homepage website: http://grass.itc.it/.

Neteler, M. and Mitasova, H. 2004. *Open source GIS: A GRASS approach, second edition*. Norwell, Massachusetts: Kluwer Academic Publishers. 387 pp.

Oregon Geospatial Data Clearinghouse. 2005. Spatial data library download page. Retrieved Feb. 10, 2005 from Oregon Geospatial Data Clearinghouse website: http://www.gis.state.or.us/data/alphalist.html.

Research Systems, Inc. 2001. *ENVI tutorial*. Research Systems, Inc. 620 pp.

University of Maryland. 2005. Earth science data interface download page. Retrieved Feb. 7, 2005, from Global Land Cover Facility website, http://glcf.umiacs.umd.edu/index.shtml.

U.S. Census Bureau. 2005. *U.S. Census Bureau quick facts: State and county quick facts.* Retrieved Feb. 15, 2005, from U.S. Census Bureau website: http://quickfacts.census.gov/qfd/.

U.S. Environmental Protection Agency. 1994. *1:250,000 scale quadrangles of land use/land cover GIRAS spatial data in the conterminous U.S. metadata.* Retrieved June 19, 2005, from U.S. EPA website: http://www.epa.gov/nsdi/projects/giras.htm.

TABLE 1
System requirements

ArcGIS system requirements	GRASS system requirements
Windows NT, 2000, or XP	Windows 95 or higher, Linux or Mac OS
256 MB of RAM	32 MB RAM
Pentium 800 MHz processor	No recommendation for processor

TABLE 2
General Landsat sensor characteristics

LANDSAT #	Dates of operation	Sensor	Swath	# of bands	orbit	temporal	inclination
1	72 - 78	MSS	115 m	4	917 km	Covers Earth in 18 days	99
2	75 - 83	MSS	115 m	4	917 km	Covers Earth in 18 days	99
3	78 - 83	MSS	115 m	4	917 km	Covers Earth in 18 days	99
4	82 - 87	MSS / TM	115 m	7	703 mi	Covers Earth in 16 days	98.2
5	84 - 99	MSS / TM	115 m	7	703 mi	Covers Earth in 16 days	98.2
6	93	ETM	115 m	8	703 mi	Covers Earth in 16 days	98.2
7	99 - present	ETM+	115 m	8	703 mi	Covers Earth in 16 days	98.2

m = meters, km = kilometers
MSS – Multi-Spectral Scanner; TM – Thematic Mapper; ETM – Enhanced Thematic Mapper

TABLE 2 (continued)
General Landsat sensor characteristics

LANDSAT MSS data			
Band #	**Wavelength range**	**spectral**	**Resolution**
1	0.5 – 0.6 µm	green	79 m
2	0.6 – 0.7 µm	red	79 m
3	0.7 – 0.8 µm	near infrared	79 m
4	0.8 – 1.1 µm	near infrared	79 m

LANDSAT TM data			
Band #	**Wavelength range**	**spectral**	**Resolution**
1	0.45 – 0.52 µm	blue	30 m
2	0.52 – 0.6 µm	green	30 m
3	0.63 – 0.69 µm	red	30 m
4	0.76 – 0.9 µm	near infrared	30 m
5	1.55 – 1.75 µm	middle infrared	30 m
6	10.4 – 12.5 µm	thermal infrared	120 m
7	2.08 – 2.35 µm	middle infrared	30 m

LANDSAT ETM+ data			
Band #	**Wavelength range**	**spectral**	**Resolution**
1	0.45 – 0.515 µm	blue	30 m
2	0.525 – 0.605 µm	green	30 m
3	0.63 – 0.69 µm	red	30 m
4	0.75 – 0.9 µm	near infrared	30 m
5	1.55 – 1.75 µm	middle infrared	30 m
6	10.4 – 12.5 µm	thermal infrared	60 m
7	2.08 – 2.35 µm	middle infrared	30 m
8	0.52 – 0.90 µm	panchromatic	15 m

m = meters; µm = micrometers

TABLE 3
Vector data used in case study

Data elements	Source (Scale)	Projection (Format)	Compilation date
Oregon city limits and city annexations	Oregon Dept. of Transportation (scale1:24,000)	LCC (shapefile)	1996, 1999, 2003
Oregon eco-regions	Oregon Natural Heritage Program (scale1:250,000)	LCC (shapefile)	2000
30 minute Quad Index	USGS (scale1:100,000)	LCC (shapefile)	1989
1 x 2 degree Quad Index	USGS (scale1:250,000)	LCC (shapefile)	1991
Oregon cities, Point locations	USGS (scale1:2,000,000)	LCC (shapefile)	1980
County Boundaries	USGS (scale1:500,000)	LCC (shapefile)	unknown
EPA GIRAS land use / land cover	EPA (scale 1:2,000,000)	LCC (shapefile)	1998
Oregon Rivers	EPA (scale 1:250,000)	LCC (shapefile)	1988
Oregon State Boundary	BLM (scale 1:24,000)	LCC (shapefile)	2001

*all vector data downloaded from the Oregon Geospatial Data Clearinghouse (http://www.gis.state.or.us)
LCC – Lambert Conformal Conic

TABLE 3 (continued)
Raster data used in case study

Data elements	Source (Resolution)	Projection (Format)	Compilation date
* Regional Digital Elevation Models (DEM) for Oregon	USGS (10 meter)	UTM (SDTS)	1999
* Digital Orthophoto Quads (DOQ) for Lane County	USGS (1 meter)	LCC (MrSID)	1994-1996, 2000-2001
** Landsat MSS, TM, and ETM+ imagery	EarthSat (MSS - 79 meter; TM and ETM - 30 meter)	UTM (GeoTIFF)	MSS - 07/30/1977; TM - 07/12/1987; ETM - 07/26/2001

* downloaded from Oregon Geospatial Data Clearinghouse (http://www.gis.state.or.us)
** downloaded from Global Land Cover Facility (http://glcf.umiacs.umd.edu/)
MSS – Multi-Spectral Scanner; TM – Thematic Mapper; ETM – Enhanced Thematic Mapper
UTM – Universal Transverse Mercator

TABLE 4
Cost comparison

Software Package	Description	Retail Price
ArcGIS 9.0 (ArcView)	Desktop GIS for mapping, data integration, and analysis.	$1,500.00
ArcGIS 9.0 (ArcEditor)	Desktop system for editing and managing geographic data.	$7,000.00
ArcGIS 9.0 (ArcInfo)	Includes all the functionality of ArcView and ArcEditor and adds advanced geoprocessing and data conversion capabilities. Used for all aspects of data building, modeling, analysis, and map display for screen and output.	$14,000.00
ArcGIS 3D Analyst Extension	Three-dimensional visualization and analysis.	$2,500.00
ArcGIS Spatial Analyst Extension	Advanced raster GIS spatial analysis.	$2,500.00
ArcGIS Interoperability Extension	Enables ArcGIS Desktop users to directly access and use most data formats without the need to convert between formats.	$2,500.00
All other extensions	Extend ArcGIS capability	$2,500.00
ENVI Node locked single user license	Remote sensing software	$4,500
GRASS 6.0	Open-Source Desktop GIS	Free

TABLE 5
ENVI ROI statistics

Region of Interest	7/26/2001 Landsat ETM+ Image Band					
	Band 1	**Band 2**	**Band 3**	**Band 4**	**Band 5**	**Band 7**
water2001-1	83.1 (1.86)	69.6 (1.56)	60.3 (2.91)	15.5 (0.92)	12.4 (1.24)	11.5 (1.57)
water2001-2	71.0 (1.92)	43.1 (1.48)	33.3 (2.05)	13.3 (0.88)	15.0 (1.29)	14.0 (1.69)
water2001-3	67.7 (1.92)	44.7 (1.76)	32.3 (2.97)	13.0 (3.51)	12.9 (3.98)	11.9 (2.30)
water2001-4	67.1 (1.77)	44.7 (1.39)	32.3 (1.92)	14.4 (0.80)	14.5 (1.31)	13.2 (1.66)
water2001-5	84.1 (2.70)	64.9 (2.11)	50.1 (2.60)	21.0 (0.94)	24.8 (1.86)	22.0 (2.07)
forest2001-1	60.9 (1.44)	52.0 (1.30)	34.2 (1.51)	130.2 (5.43)	65.5 (4.36)	28.3 (2.59)
forest2001-2	57.9 (1.71)	43.2 (1.53)	29.1 (1.70)	96.8 (5.28)	45.3 (4.07)	20.6 (2.50)
forest2001-3	54.0 (3.07)	37.0 (1.77)	25.3 (1.72)	48.0 (6.14)	31.3 (5.25)	17.8 (3.27)
forest2001-4	65.4 (1.81)	57.0 (1.21)	45.9 (2.60)	107.3 (7.00)	92.5 (5.55)	46.3 (3.99)
soil2001-1 (grass)	137.0 (3.49)	151.1 (3.83)	224.4 (5.41)	144.8 (2.98)	239.0 (3.13)	147.7 (3.30)
grass2001-1	108.9 (2.76)	113.6 (3.07)	163.7 (5.41)	111.2 (2.89)	205.7 (4.95)	129.0 (4.27)
grass2001-2	107.0 (2.38)	109.4 (3.54)	154.9 (4.36)	107.3 (2.05)	202.9 (4.22)	125.3 (3.06)
grass2001-3	100.4 (1.87)	99.5 (2.17)	132.5 (3.36)	105.0 (1.58)	182.0 (3.36)	113.7 (3.42)
grass2001-4	98.4 (1.77)	95.8 (2.26)	133.2 (4.99)	92.7 (2.39)	173.6 (4.03)	114.0 (2.78)
grass2001-5	109.1 (2.93)	102.2 (3.28)	133.0 (4.09)	79.0 (2.20)	159.2 (4.46)	147.7 (4.15)
grass2001-6	98.3 (2.04)	94.0 (1.78)	129.3 (3.68)	88.0 (2.07)	119.0 (3.22)	70.2 (2.01)
grass2001-7	94.2 (2.85)	80.9 (2.58)	98.4 (5.74)	58.9 (3.32)	117.8 (4.21)	107.2 (3.27)
grass2001-8	102.1 (1.88)	98.6 (1.69)	124.1 (3.61)	107.8 (1.73)	165.7 (2.42)	106.6 (2.27)
grass2001-9	108.1 (3.50)	113.8 (3.46)	156.5 (6.41)	134.4 (2.98)	179.2 (3.67)	100.3 (3.60)
grass2001-10	79.9 (2.57)	71.0 (2.69)	87.3 (3.60)	101.3 (5.95)	103.7 (3.47)	54.6 (3.87)
urban2001-1	85.0 (2.50)	64.4 (2.76)	66.3 (4.03)	35.4 (3.96)	85.9 (8.65)	75.6 (4.68)
urban2001-2	114.2 (5.17)	99.5 (4.44)	110.3 (5.81)	54.3 (3.85)	82.4 (5.12)	69.1 (4.16)

* mean (standard deviation)

TABLE 5 (continued)
ENVI ROI statistics

Region of Interest	7/12/1987 Landsat TM Image Band					
	Band 1	Band 2	Band 3	Band 4	Band 5	Band 7
water1987-1	100.3 (1.37)	46.2 (0.89)	44.3 (1.00)	11.3 (0.68)	4.1 (0.93)	1.6 (0.76)
water1987-2	88.6 (1.56)	26.4 (0.94)	19.0 (1.12)	8.8 (1.15)	5.8 (1.06)	2.7 (0.96)
water1987-3	79.9 (1.80)	27.0 (1.55)	22.6 (1.45)	9.72 (2.56)	8.25 (2.26)	3.23 (1.28)
water1987-4	86.6 (2.23)	30.4 (1.08)	24.7 (1.07)	13.9 (0.65)	13.7 (1.12)	7.6 (1.14)
water1987-5	78.4 (1.66)	26.2 (1.01)	18.6 (1.32)	9.4 (0.83)	6.5 (1.43)	3.0 (1.05)
Forest1987-1	79.7 (1.38)	31.6 (0.97)	24.1 (1.07)	160.9 (6.03)	68.6 (5.39)	16.3 (2.17)
Forest1987-2	68.2 (1.52)	26.2 (0.79)	19.7 (1.18)	108.6 (4.21)	51.6 (4.97)	12.4 (2.14)
Forest1987-3	61.9 (1.22)	22.4 (0.89)	15.9 (0.88)	64.9 (3.88)	30.4 (1.80)	7.7 (0.85)
Forest1987-4	77.4 (2.15)	36.6 (1.22)	32.1 (1.86)	122.0 (5.84)	93.9 (6.29)	31.1 (3.36)
soil1987-1 (grass)	165.5 (2.29)	95.8 (1.49)	150.8 (2.61)	172.5 (2.34)	218.8 (2.70)	90.1 (1.36)
Grass1987-1	124.1 (3.16)	64.7 (2.40)	99.8 (4.12)	128.2 (1.80)	177.3 (3.12)	78.4 (2.07)
Grass1987-2	121.3 (1.43)	62.6 (0.91)	89.2 (1.71)	129.5 (1.29)	175.2 (2.90)	77.5 (1.83)
Grass1987-3	124.5 (2.47)	64.4 (1.88)	104.2 (2.51)	110.1 (2.11)	182.7 (2.66)	85.1 (2.21)
Grass1987-4	138.0 (3.68)	70.2 (2.10)	99.0 (2.86)	94.1 (2.64)	165.1 (4.26)	103.5 (4.24)
Grass1987-5	107.1 (2.01)	48.1 (0.95)	65.5 (1.50)	60.6 (2.67)	85.5 (5.85)	51.5 (3.73)
Grass1987-6	83.5 (1.08)	27.9 (0.89)	27.8 (0.99)	20.0 (0.66)	44.0 (3.65)	35.8 (3.11)
Grass1987-7	102.0 (1.93)	49.1 (1.06)	71.9 (2.49)	118.2 (2.16)	157.8 (4.09)	61.5 (2.26)
Grass1987-8	115.8 (2.36)	57.0 (1.34)	79.8 (2.06)	109.0 (1.43)	171.4 (2.29)	81.3 (1.83)
Grass1987-9	93.6 (1.37)	48.2 (0.77)	47.4 (0.78)	112.9 (1.83)	62.5 (1.20)	25.4 (1.06)
Grass1987-10	93.8 (1.27)	43.1 (0.78)	64.8 (1.67)	70.8 (1.27)	86.3 (1.82)	37.7 (1.59)
Urban1987-1	117.3 (2.41)	52.3 (1.14)	66.5 (1.71)	50.1 (1.36)	65.9 (1.94)	40.1 (1.33)
Urban1987-2	149.4 (6.97)	71.6 (2.74)	92.8 (4.38)	73.3 (4.46)	96.8 (4.65)	59.8 (2.56)

* mean (standard deviation)

TABLE 5 (continued)
ENVI ROI statistics

Region of Interest	7/30/1977 Landsat MSS Image Band					
	Band 1	Band 2	Band 3	Band 4	-	-
Water1977-1	20.2 (1.46)	21.8 (1.24)	11.0 (1.16)	1.2 (0.60)		
Water1977-2	11.6 (1.36)	10.0 (1.05)	6.9 (3.94)	3.4 (3.36)		
Water1977-3	11.5 (0.78)	7.8 (0.83)	2.6 (1.18)	1.0 (0.00)		
Water1977-4	13.3 (1.51)	9.3 (0.82)	3.2 (1.87)	1.2 (0.94)		
Forest1977-1	10.1 (0.45)	8.2 (0.53)	34.1 (1.96)	36.9 (2.03)		
Forest1977-2	9.4 (0.61)	6.5 (0.57)	16.4 (2.36)	17.3 (2.28)		
Forest1977-3	9.8 (0.48)	7.7 (0.76)	30.6 (1.98)	32.8 (2.28)		
Forest1977-4	11.3 (0.90)	10.8 (1.22)	51.8 (3.75)	58.5 (3.65)		
Forest1977-5	9.5 (0.51)	6.7 (0.81)	17.1 (3.01)	17.4 (3.41)		
Forest1977-6	10.3 (0.61)	9.9 (1.17)	45.6 (3.15)	50.0 (3.83)		
Grass1977-1	43.7 (2.26)	74.5 (3.97)	88.0 (4.62)	80.0 (2.47)		
Grass1977-2	30.3 (1.47)	52.9 (1.78)	63.7 (1.40)	60.8 (1.37)		
Grass1977-3	29.4 (1.50)	49.4 (2.47)	59.4 (1.71)	55.6 (1.84)		
Grass1977-4	10.4 (0.50)	10.9 (0.70)	7.8 (0.90)	3.0 (1.75)		
Grass1977-5	11.7 (0.62)	11.5 (0.77)	84.9 (4.99)	97.9 (5.55)		
Grass1977-6	10.9 (0.42)	11.7 (0.48)	8.7 (1.03)	3.8 (0.96)		
Grass1977-7	15.8 (0.63)	19.3 (1.15)	57.8 (1.85)	65.1 (1.90)		
Grass1977-8	20.0 (1.49)	32.8 (0.97)	41.2 (1.63)	39.3 (1.43)		
Grass1977-9	22.5 (1.56)	30.0 (2.54)	67.6 (2.28)	70.9 (1.79)		
Grass1977-10	15.3 (0.74)	21.4 (1.43)	43.1 (3.16)	44.3 (2.91)		
Grass1977-11	19.7 (0.71)	25.4 (1.50)	43.4 (2.58)	42.1 (2.50)		
Urban1977-1	25.4 (0.79)	28.6 (1.62)	43.6 (2.30)	38.7 (2.42)		

* mean (standard deviation)

TABLE 6
ArcGIS / ENVI class statistics

Landsat MSS, 7/30/1977 - interactive						
unit	**water**	**forest**	**grass**	**urban**	**unclassified**	**total**
sq. miles	4.73	110.06	206.31	26.15	0.64	347.90
sq. meters	12,241,823	285,055,344	534,352,099	67,738,542	1,663,812	901,051,620
Acres	3,025.02	70,438.71	132,041.28	16,738.56	411.14	222,654.70
Hectares	1,224.18	28,505.53	53,435.21	6,773.85	166.38	90,105.16
percent cover	1.35	31.64	59.30	7.52	0.18	100

Landsat TM, 7/12/1987 - interactive						
unit	**water**	**forest**	**grass**	**urban**	**unclassified**	**total**
sq. miles	3.20	214.98	88.07	39.94	0.08	346.27
sq. meters	8,287,640	556,795,761	228,106,377	103,447,168	1,989,86	896,835,932
Acres	2,047.92	137,587.23	56,366.31	25,562.35	49.17	221,612.99
Hectares	828.76	55,679.58	22,810.64	10,344.72	19.90	89,683.59
percent cover	0.92	62.08	25.43	11.53	0.02	100

Landsat ETM+, 7/26/2001 - interactive						
unit	**water**	**forest**	**grass**	**urban**	**unclassified**	**total**
sq. miles	3.58	198.79	82.01	58.75	3.14	346.27
sq. meters	9,276,180	514,859,306	212,409,087	152,155,306	8,124,998	896,835,877
Acres	2,292.19	127,224.51	52,487.43	37,601.11	2,007.73	221,612.97
Hectares	927.62	51,485.93	21,240.91	15,216.63	812.50	89,683.59
percent cover	1.03	57.41	23.68	16.97	0.91	100

TABLE 7
EPA GIRAS data class statistics

EPA GIRAS Land Use Statistics - 1998						
unit	**water**	**forest**	**grass**	**urban**	**unclassified**	**total**
sq. miles	2.40	145.12	128.16	69.22	0.32	345.22
sq. meters	6,211,449	375,860,200	331,929,822	179,281,076	834,737	894,117,284
Acres	1534.85	92,875.06	82,019.86	44,300.35	206.26	220,936.38
Hectares	621.15	37,586.02	33,192.98	17,928.11	83.47	89411.73
percent cover	0.69	42.04	37.12	20.05	0.09	100

TABLE 8
ArcGIS/ENVI Classification accuracy assessment

Landsat MSS, 7/30/1977				
Class	**Percent commission**	**Percent omission**	**Producer accuracy**	**User accuracy**
Water	0.08	0.38	99.62	99.92
Forest	0.00	0.00	100.00	100.00
Grass	1.27	0.26	99.74	98.73
Urban	0.00	0.00	100.00	100.00

Total observed: 1963 **Overall kappa:** 0.9940

Observed correct: 1957 **Percent observed correct:** 99.69%

Landsat TM, 7/12/1987				
Class	**Percent commission**	**Percent omission**	**Producer accuracy**	**User accuracy**
Water	0.00	0.00	100.00	100.00
Forest	0.00	2.82	97.18	100.00
Grass	1.26	0.00	100.00	98.74
Urban	0.00	1.45	98.55	100.00

Total observed: 5969 **Overall kappa:** 0.9955

Observed correct: 5964 **Percent observed correct:** 99.92%

Landsat ETM+, 7/26/2001				
Class	**Percent commission**	**Percent omission**	**Producer accuracy**	**User accuracy**
Water	0.00	0.00	100.00	100.00
Forest	0.00	0.00	100.00	100.00
Grass	0.00	0.00	100.00	100.00
Urban	0.00	0.00	100.00	100.00

Total observed: 15534 **Overall kappa:** 1.0000

Observed correct: 15534 **Percent observed correct:** 100.00%

TABLE 9
GRASS interactive classification training area statistics

Landsat MSS, 7/30/1977		
Region of Interest	**Sample Size (pixels)**	**Standard Deviation**
Water1977-1	36	4.0
Water1977-2	41	4.0
Forest1977-1	502	6.0
Forest1977-2	407	6.0
Forest1977-3	286	6.0
Grass1977-1	48	7.0
Grass1977-2	117	6.5
Grass1977-3	128	6.5
Grass1977-4	209	7.0
Grass1977-5	91	8.0
Grass1977-6	154	8.0
Grass1977-7	46	6.0
Urban1977-1	41	4.0
Urban1977-2	54	4.0

Landsat TM, 7/12/1987		
Region of Interest	**Sample Size (pixels)**	**Standard Deviation**
Water1987-1	100	2.5
Forest1987-1	482	2.5
Forest1987-2	207	2.5
Forest1987-3	2014	2.5
Grass1987-1	1877	4.0
Grass1987-2	1049	4.0
Grass1987-3	601	4.0
Grass1987-4	603	5.0
Urban1987-1	97	2.0

Landsat ETM+, 7/26/2001		
Region of Interest	**Sample Size (pixels)**	**Standard Deviation**
Water2001-1	100	2.5
Forest2001-1	508	2.5
Forest2001-2	142	2.5
Forest2001-3	3245	2.5
Grass2001-1	1804	4
Grass2001-2	940	4
Grass2001-3	680	5
Grass2001-4	705	5
Grass2001-5	845	2
Urban2001-1	92	1.75

TABLE 10
GRASS interactive class statistics

Landsat MSS, 7/30/1977 - interactive						
unit	**water**	**forest**	**grass**	**urban**	**unclassified**	**total**
sq. miles	2.78	126.44	196.38	19.62	0	345.22
sq. meters	7,200,166	327,476,092	508,618,442	50,822,584	0	894,117,284
Acres	1,779.16	80,919.34	125,679.62	12,558.26	0	220936.38
Hectares	720.02	32,747.7	50,861.84	5,082.29	0	89411.73
percent cover	0.81	36.63	56.88	5.68	0	100

Landsat TM, 7/12/1987 - interactive						
unit	**water**	**forest**	**grass**	**urban**	**unclassified**	**total**
sq. miles	1.9	159.55	142.22	38.35	3.18	345.22
sq. meters	4,930,979	413,244,155	368,357,476	99,336,360	8,248,314	894,117,284
Acres	1,218.44	102,112.63	91,021.13	24,546.01	2,038.16	220936.38
Hectares	493.1	41,324.42	36,835.748	9,933.636	824.831	89411.73
percent cover	0.55	46.22	41.2	11.11	0.92	100

Landsat ETM+, 7/26/2001 - interactive						
unit	**water**	**forest**	**grass**	**urban**	**unclassified**	**total**
sq. miles	2.03	102.53	185.5	55.15	0	345.22
sq. meters	5,265,954	265,552,502	480,437,301	142,861,526	0	894,117,284
Acres	1,301.22	65,618.02	118,716.06	35,301.08	0	220936.38
Hectares	526.6	26,555.25	48,043.73	14,286.15	0	89411.73
percent cover	0.59	29.7	53.73	15.98	0	100

TABLE 11
GRASS non-interactive class statistics

Landsat MSS, 7/30/1977 – non-interactive						
unit	**water**	**forest**	**grass**	**urban**	**unclassified**	**total**
sq. miles	5.02	135.06	187.83	17.31	0	345.22
sq. meters	12,989,945	349,811,788	486,477,480	44,838,071	0	894,117,284
Acres	3,209.82	86,438.49	120,208.59	11,079.49	0	220936.38
Hectares	1,298.99	34,981.18	48,647.75	4,483.81	0	89411.73
percent cover	1.45	39.12	54.41	5.01	0	100

Landsat TM, 7/12/1987 – non-interactive						
unit	**water**	**forest**	**grass**	**urban**	**unclassified**	**total**
sq. miles	1.21	152.22	162.21	29.57	0	345.22
sq. meters	3,123,734	394,261,325	420,135,459	76,596,765	0	894,117,284
Acres	771.87	97,421.97	103,815.47	18,927.06	0	220936.38
Hectares	312.37	39,426.13	42,013.55	7,659.68	0	89411.73
percent cover	0.35	44.10	46.99	8.57	0	100

Landsat ETM+, 7/26/2001 – non-interactive						
unit	**water**	**forest**	**grass**	**urban**	**unclassified**	**total**
sq. miles	1.19	132.47	154.80	56.76	0	345.22
sq. meters	3,069,706	343,099,264	400,942,820	147,005,494	0	894,117,284
Acres	758.52	84,779.83	99,072.97	36,325.06	0	220936.38
Hectares	306.971	34,309.93	40,094.28	14,700.55	0	89411.73
percent cover	0.34	38.37	44.84	16.44	0	100

TABLE 12
GRASS non-interactive classification accuracy assessment

Landsat MSS, 7/30/1977			
Class	**Percent commission**	**Percent omission**	**Estimated kappa**
Water	0.00	0.00	1.00
Forest	1.16	0.00	0.96
Grass	2.64	3.57	0.96
Urban	3.38	0.99	0.96

Total observed: 4718
Observed correct: 4640
Percent observed correct: 98.35%
Overall kappa: 0.963643
Kappa variance: 0.000016

Landsat TM, 7/12/1987			
Class	**Percent commission**	**Percent omission**	**Estimated kappa**
Water	0.00	0.00	1.00
Forest	0.64	0.00	0.99
Grass	0.98	0.51	0.98
Urban	0.95	8.38	0.99

Total observed: 25533
Observed correct: 25333
Percent observed correct: 99.22%
Overall kappa: 0.985559
Kappa variance: 0.000001

Landsat ETM+, 7/26/2001			
Class	**Percent commission**	**Percent omission**	**Estimated kappa**
Water	0.00	0.00	1.00
Forest	0.92	0.03	0.98
Grass	1.15	0.66	0.98
Urban	0.46	8.38	0.99

Total observed: 22568
Observed correct: 22349
Percent observed correct: 99.03%
Overall kappa: 0.982770
Kappa variance: 0.000001

TABLE 13
ARCGIS vs. GRASS – Categorical summary

Category	ArcGIS	GRASS
Cost	Poor; can be extremely expensive, depending on the license and extensions needed.	Excellent; freely available to the public.
Installation	Excellent; very easy.	Good; very easy but slightly more involved due to Cygwin.
Training and Technical Support	Excellent; abundant resource material, training, and support.	Poor; training and support virtually non-existent.
Database	Excellent; easy to use and develop, several options.	Fair; harder to use and develop than ArcGIS, does not give the user as much freedom.
Data Import and Export	Fair; offers the options necessary to work with most downloadable data.	Excellent, offers a wide range of options.
Display	Excellent; on-the-fly projection makes display of data extremely easy.	Fair; offers the tools necessary to display data but does not offer on-the-fly projection.
Image Classification	Cannot perform image classification. Second program such as ENVI is needed.	Good; offers the tools necessary to conduct several types of classification but could definitely be improved.
Post Classification Analysis	Poor; offers a few options such as calculate areas but does not give the user many options. ENVI or other program must be used.	Excellent; offers an abundance of statistical tools that are easy to use.
Map Layout	Excellent; offers the user the ability to produce highly customizable, professional maps.	Poor; very few cartographic design or map layout tools. A second image manipulation program such as GIMP or Photoshop must be used.
Overall Functionality	Fair; offers a wide variety of tools but since it does not offer image classification or statistical tools it was given a lower score.	Excellent; offers a wide variety of analytical and statistical tools
Overall ease of use	Excellent; very straight forward and easy to use.	Fair; could be much more straightforward and user friendly. User must be much more experienced with GIS and GIS concepts.

TABLE 14
Score totals

Category	ArcGIS	GRASS
Cost	1	4
Installation	4	3
Training and Technical Support	4	1
Database	4	3
Data Import and Export	2	4
Display	4	2
Image Classification	0	3
Post Classification Analysis	1	4
Map Layout	4	1
Overall Functionality	2	4
Overall ease of use	4	2
Total	**30**	**31**

Figure 1 – Map of Eugene-Springfield, Oregon developed in ArcGIS.

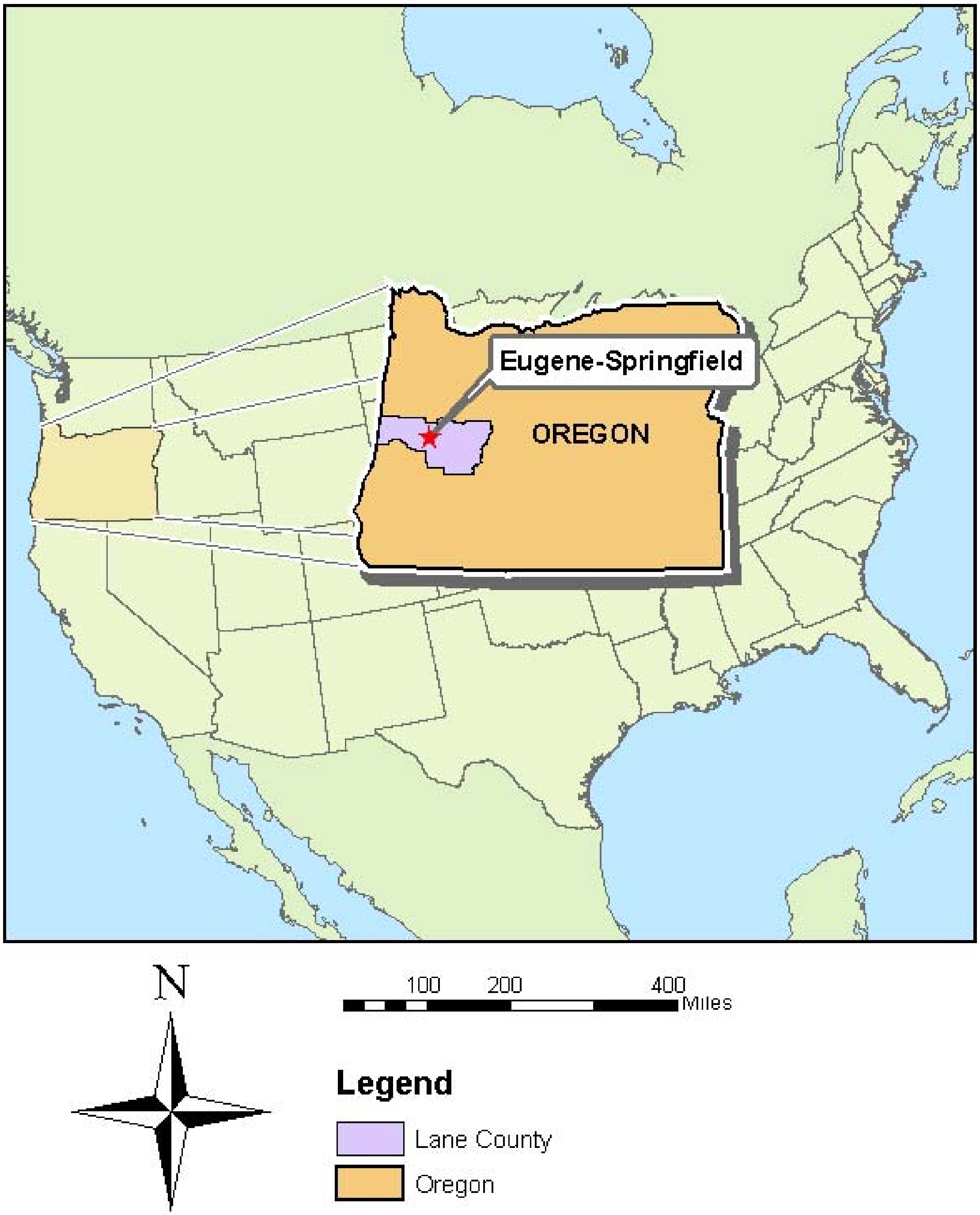


Figure 2 – Location map based on Landsat true-color RGB composite image.

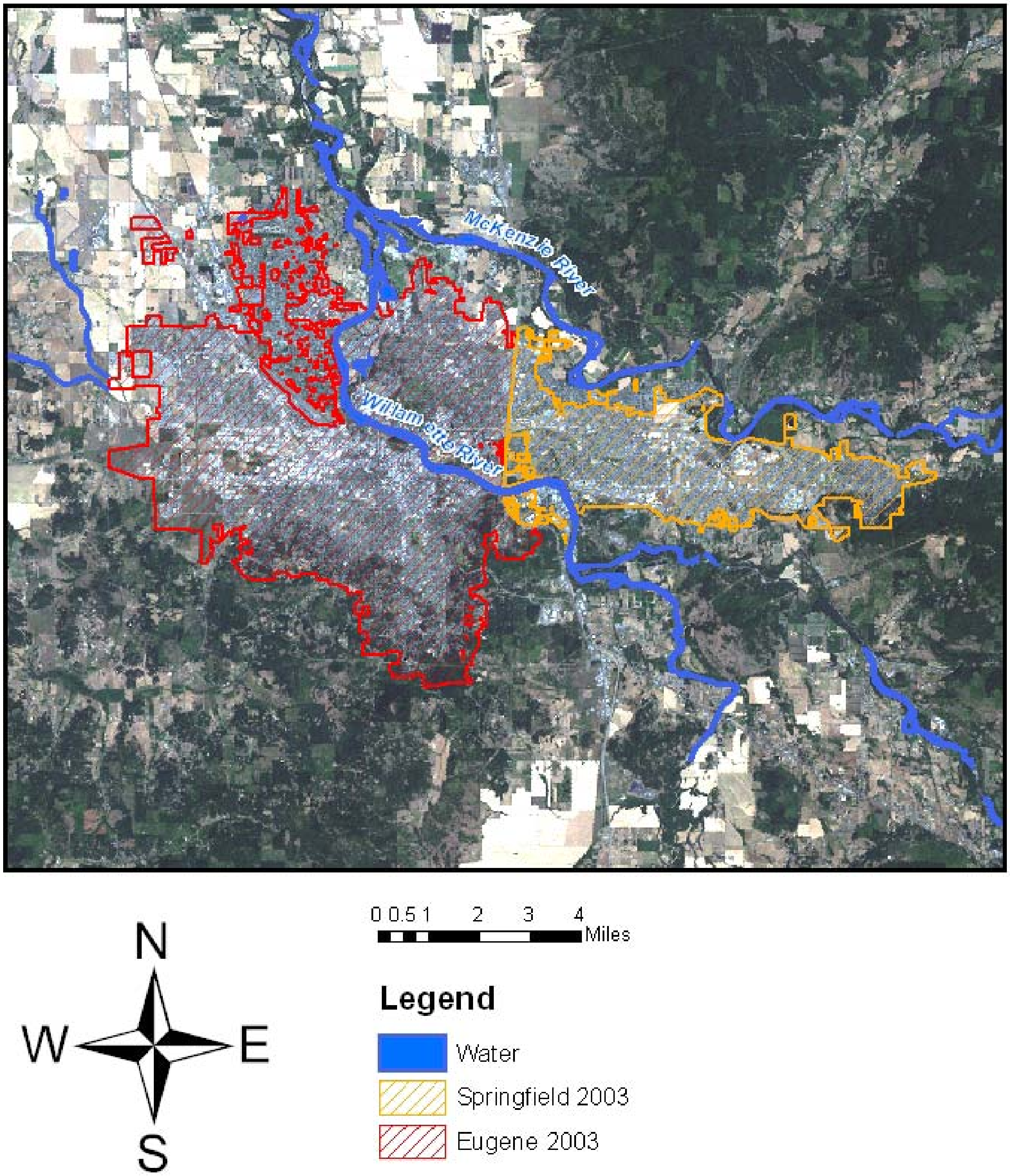


Figure 3 – Screenshot of ArcGIS ArcMap interface.

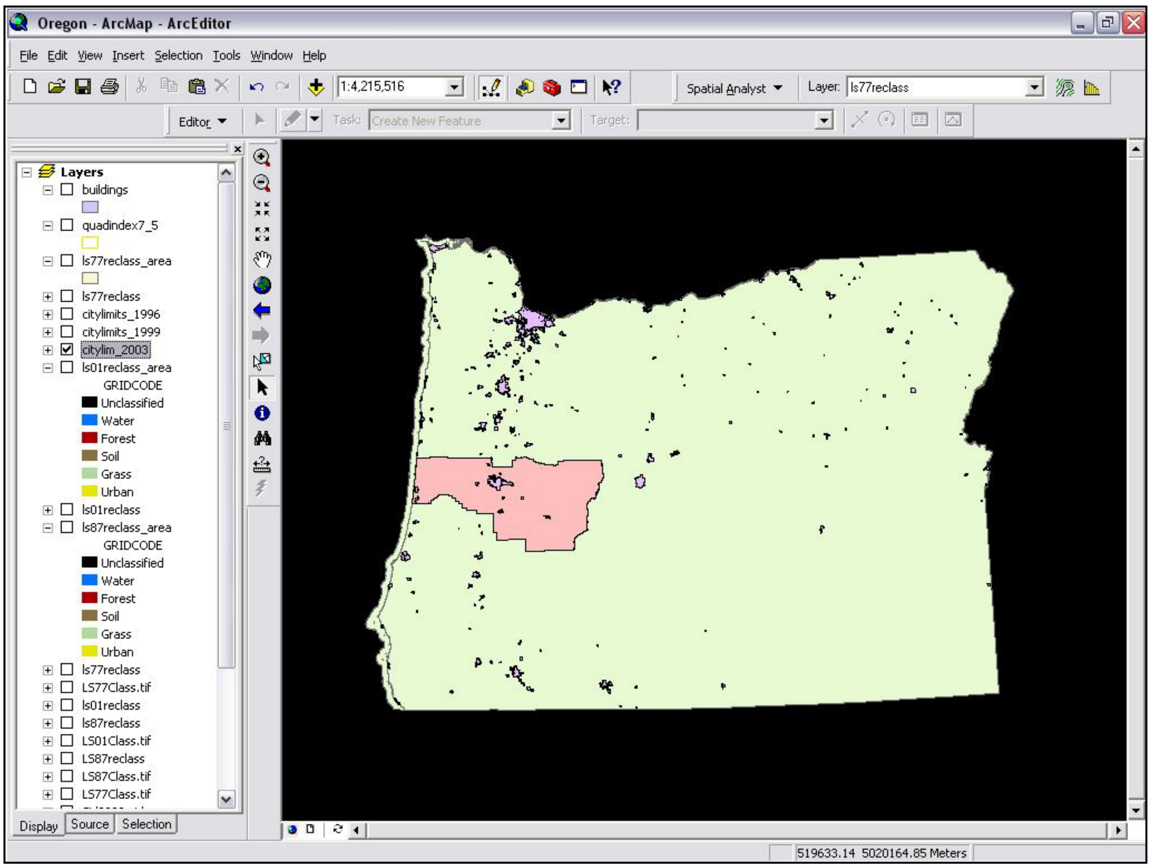


Figure 4 – Screenshot of ENVI supervised classification interface.

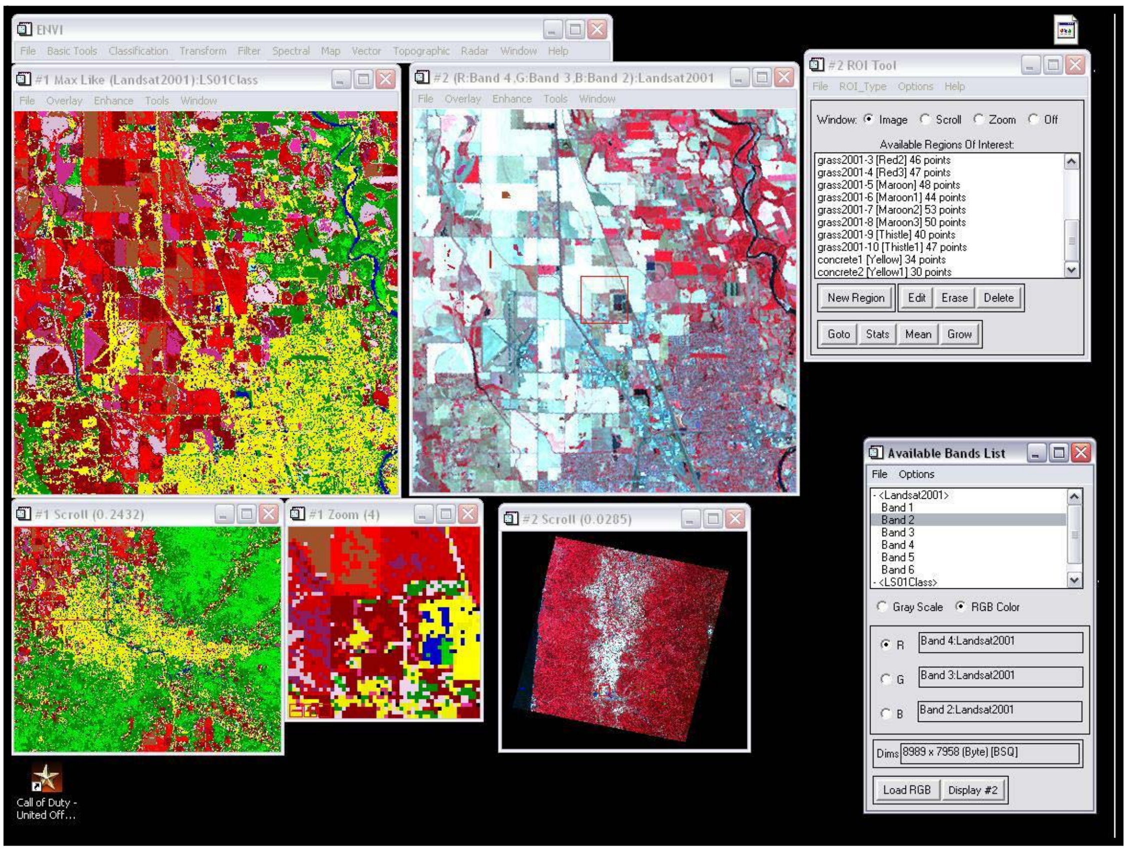


Figure 5 – ArcGIS / ENVI classification of Landsat MSS image, 07/30/1977.

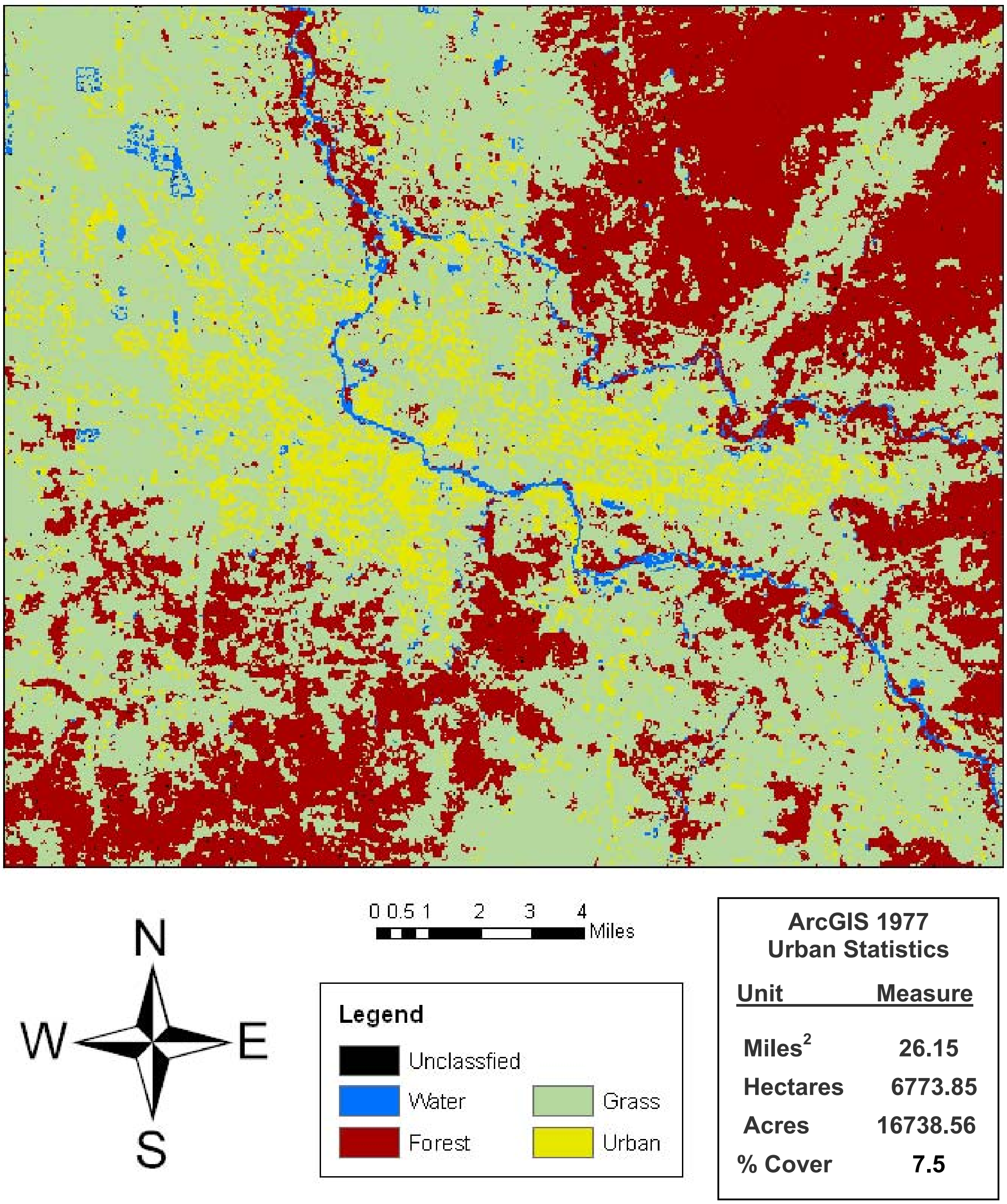


Figure 6 – ArcGIS / ENVI classification of Landsat TM Image, 07/12/1987.

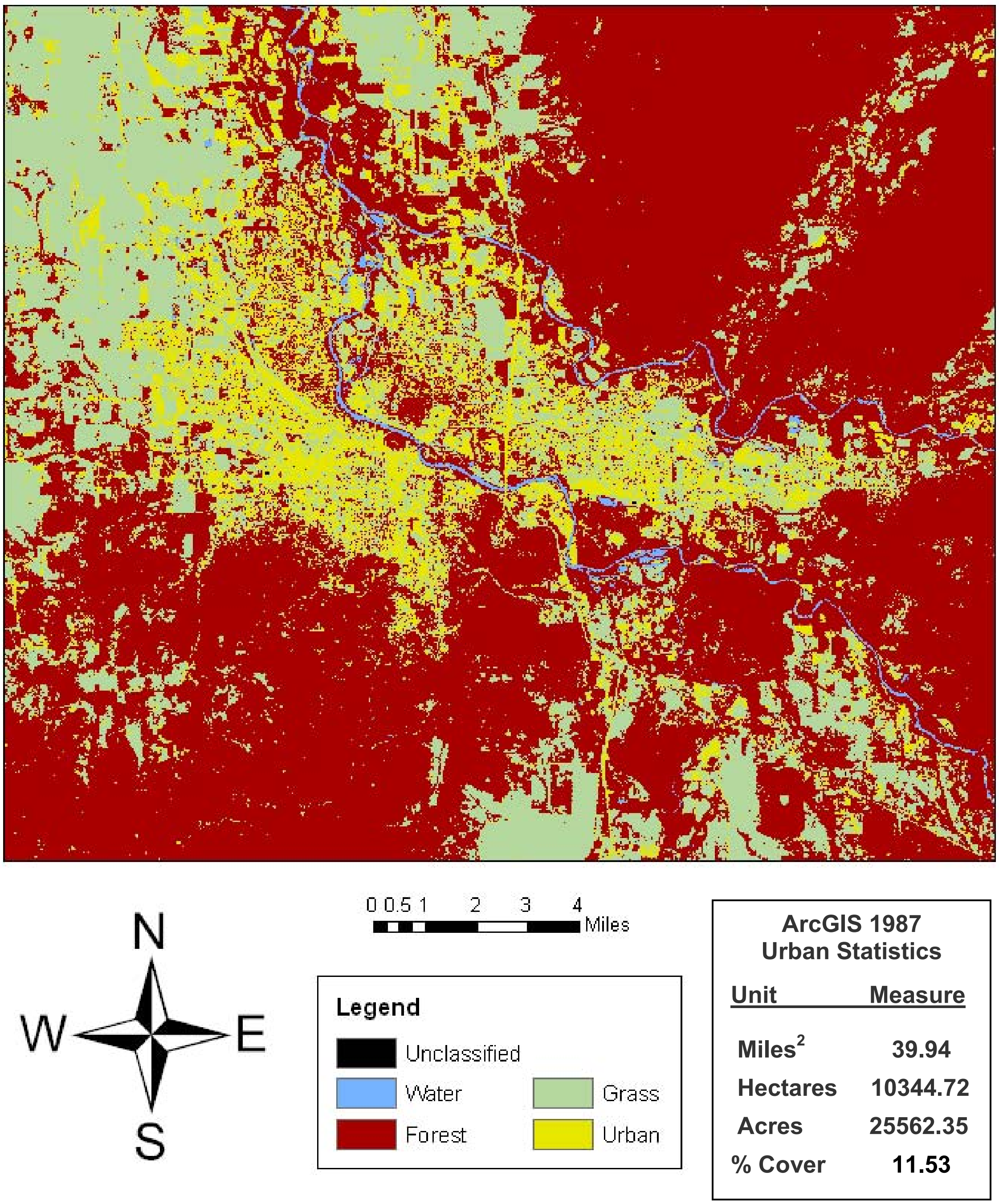


Figure 7 – ArcGIS / ENVI classification of Landsat ETM+ Image, 07/26/2001.

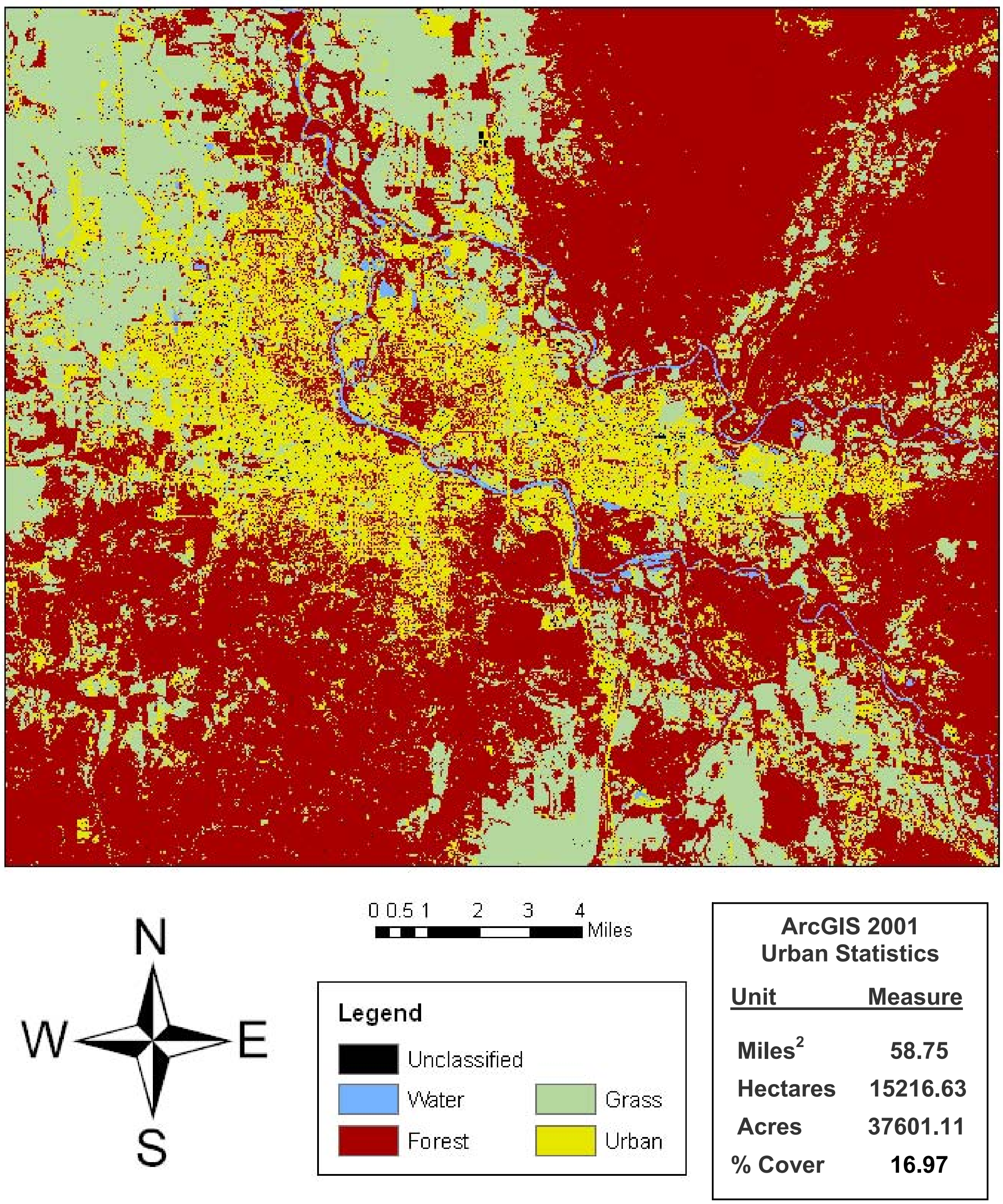


Figure 8 – EPA GIRAS land use classification image.

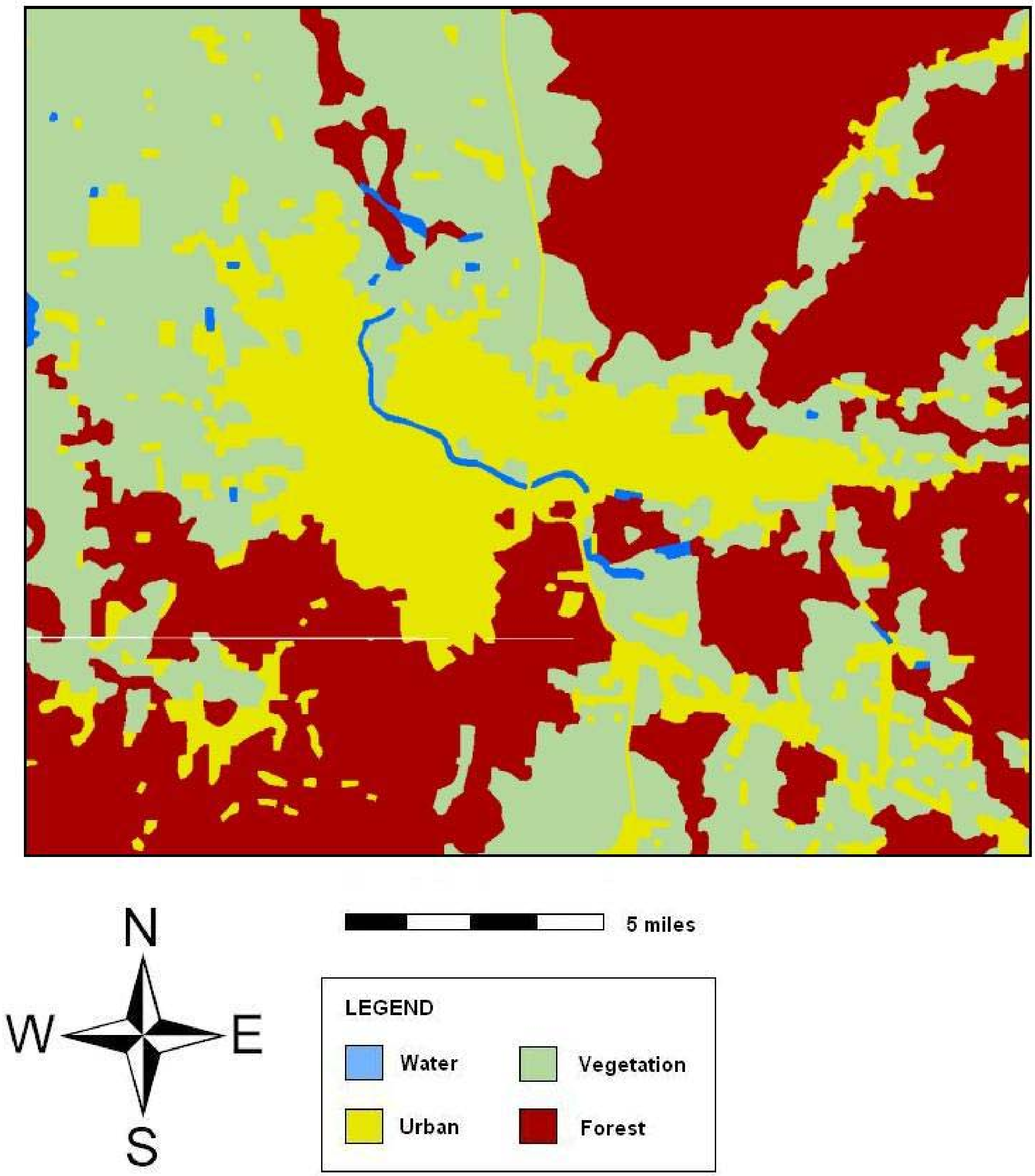


Figure 9 – ArcGIS Landsat ETM+ classification draped over shaded relief map.

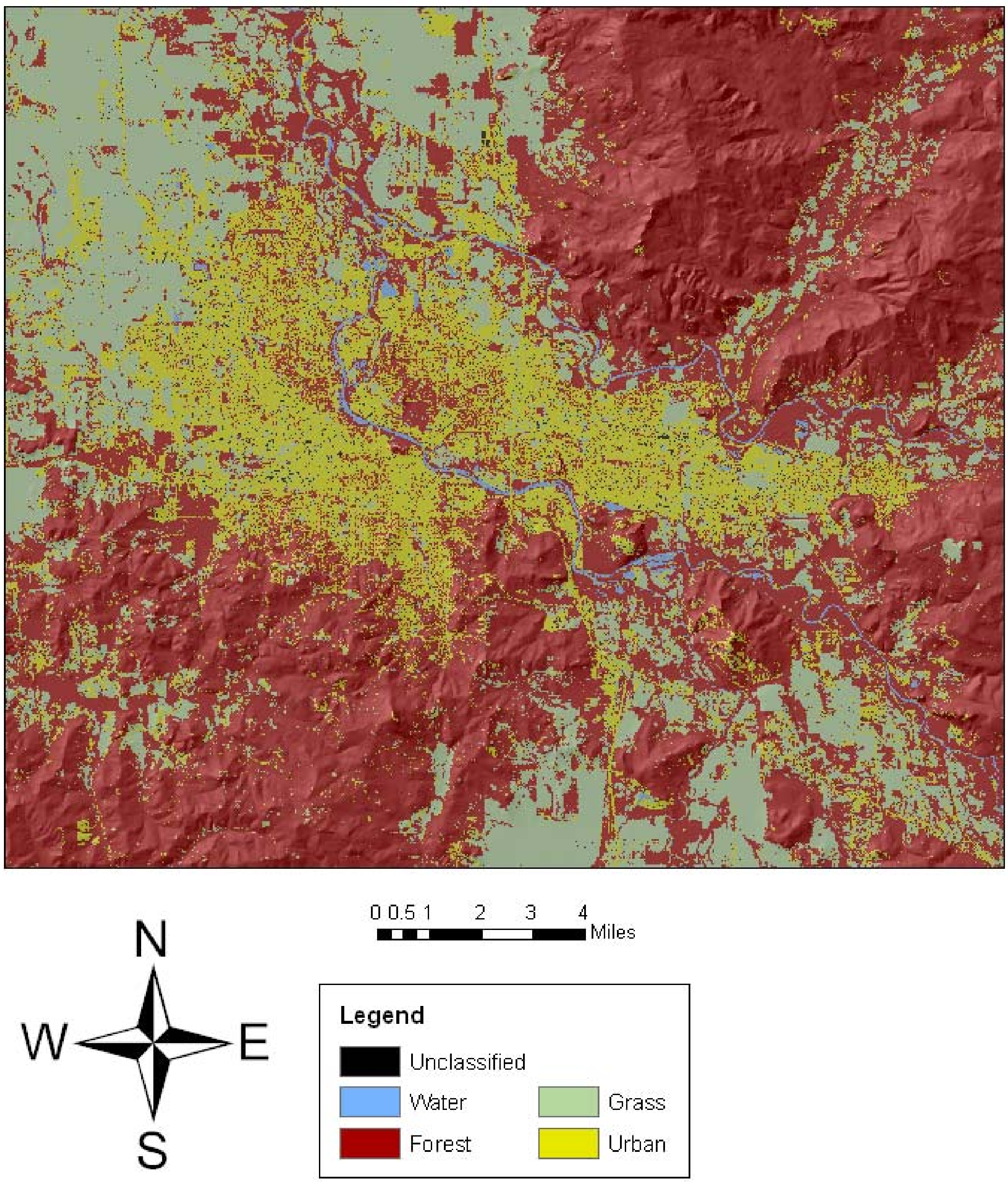


Figure 10 – ArcGIS 3D view of real-color RGB composite (bands 3, 2, and 1) of Landsat ETM+ image, 07/26/2001, draped on DEM of region.

Figure 11 – ArcGIS 3D view of classification of Landsat ETM+ image, 07/26/2001, draped on DEM of region.

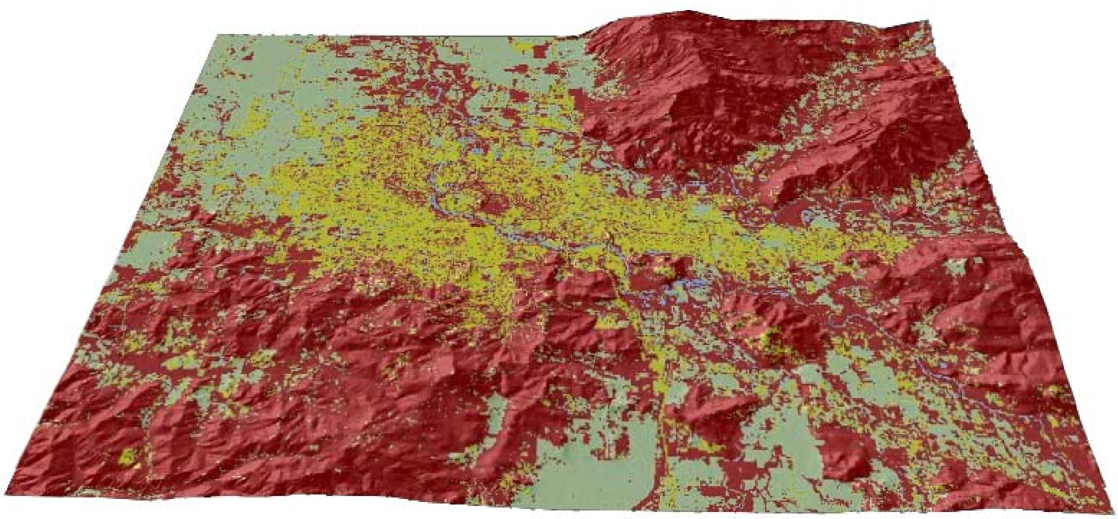


Figure 12 – Screenshot of GRASS GIS manager interface.

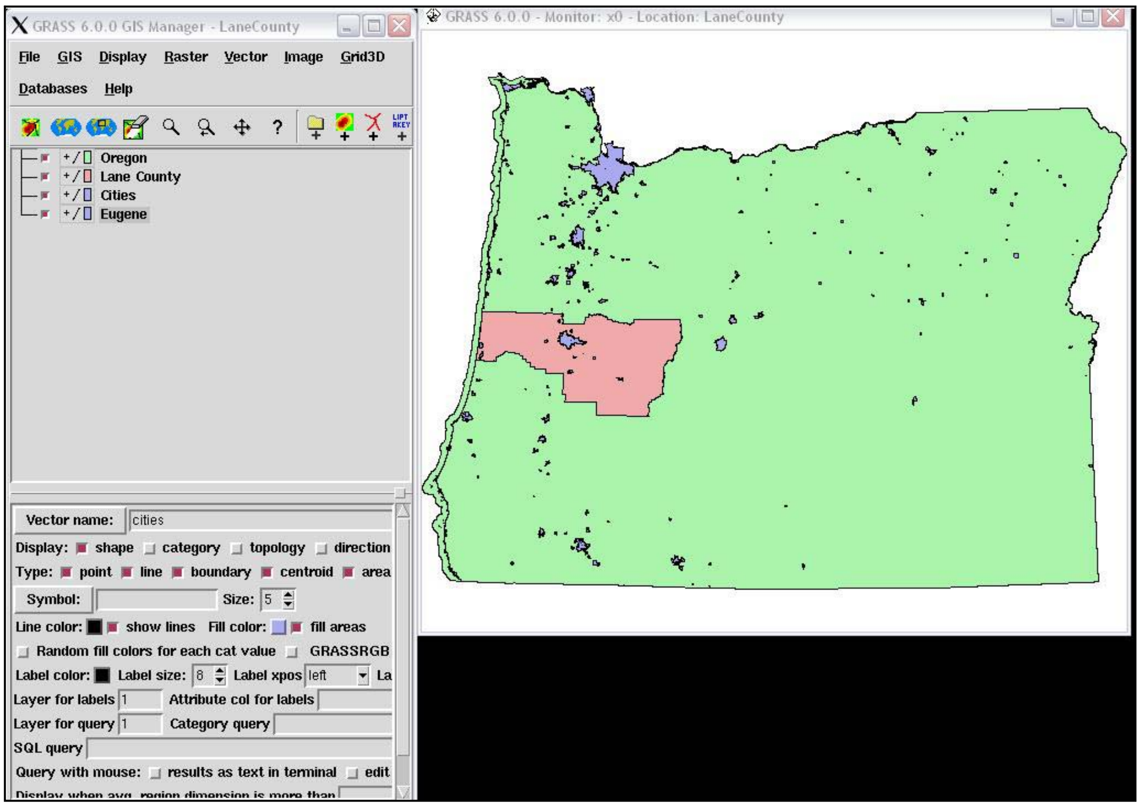


Figure 13 – Screenshot of GRASS interactive supervised classification (i.class module) .

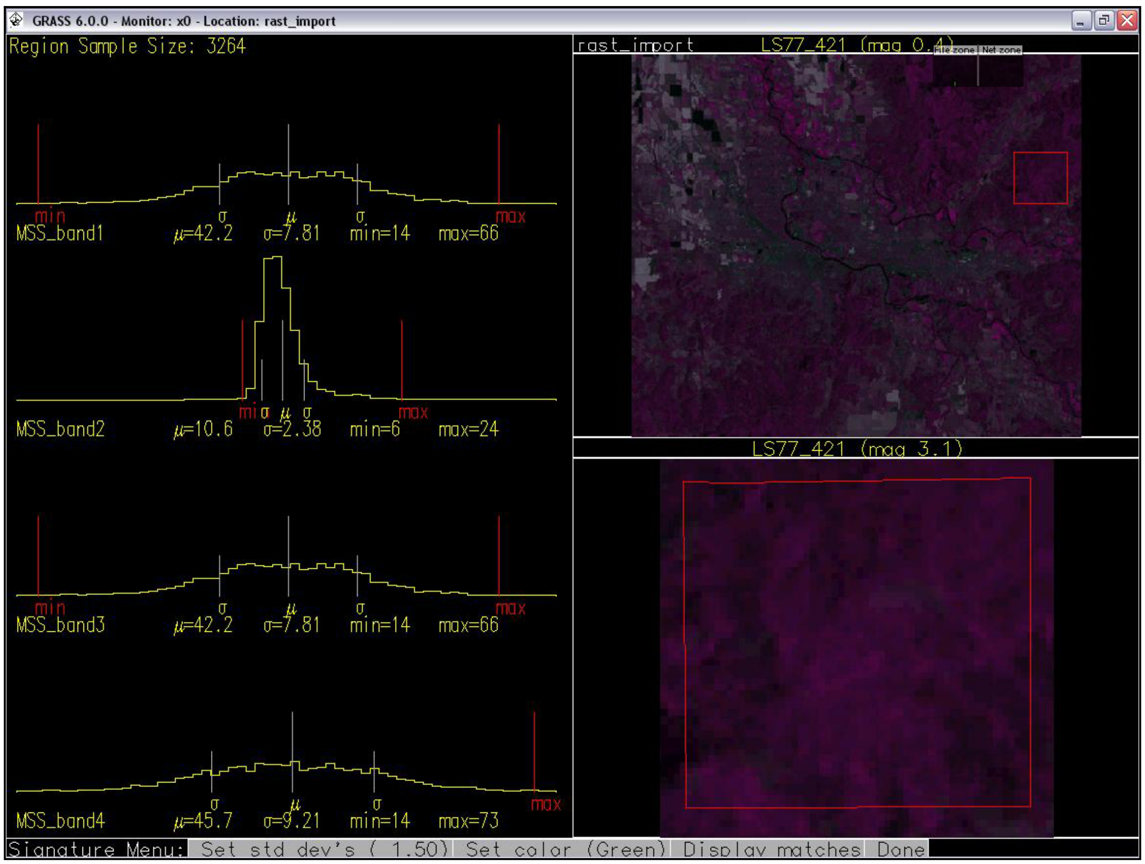


Figure 14 – GRASS interactively classified image of Landsat MSS image, 07/30/1977.

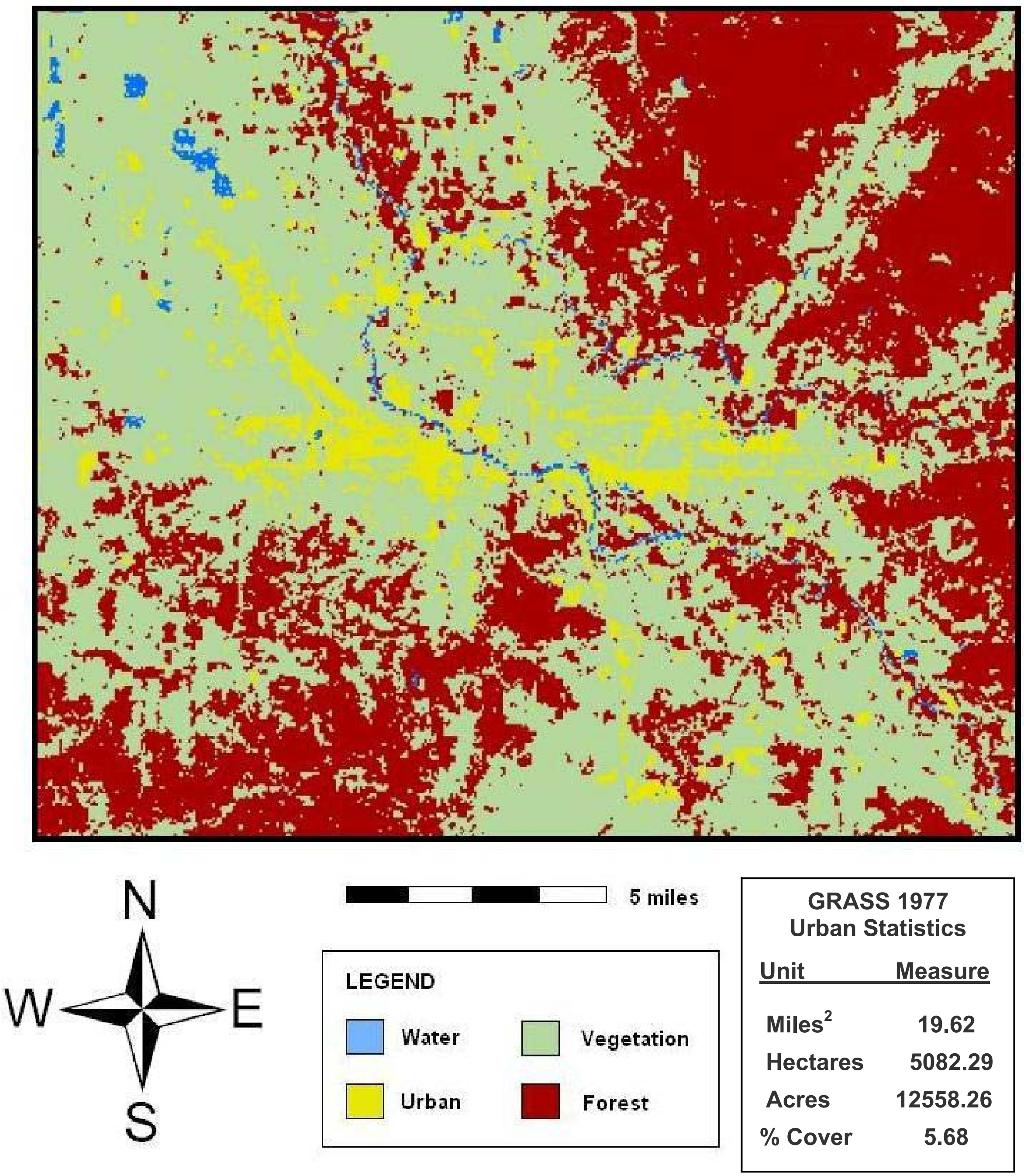


Figure 15 – GRASS interactively classified image of Landsat TM image, 07/12/1987.

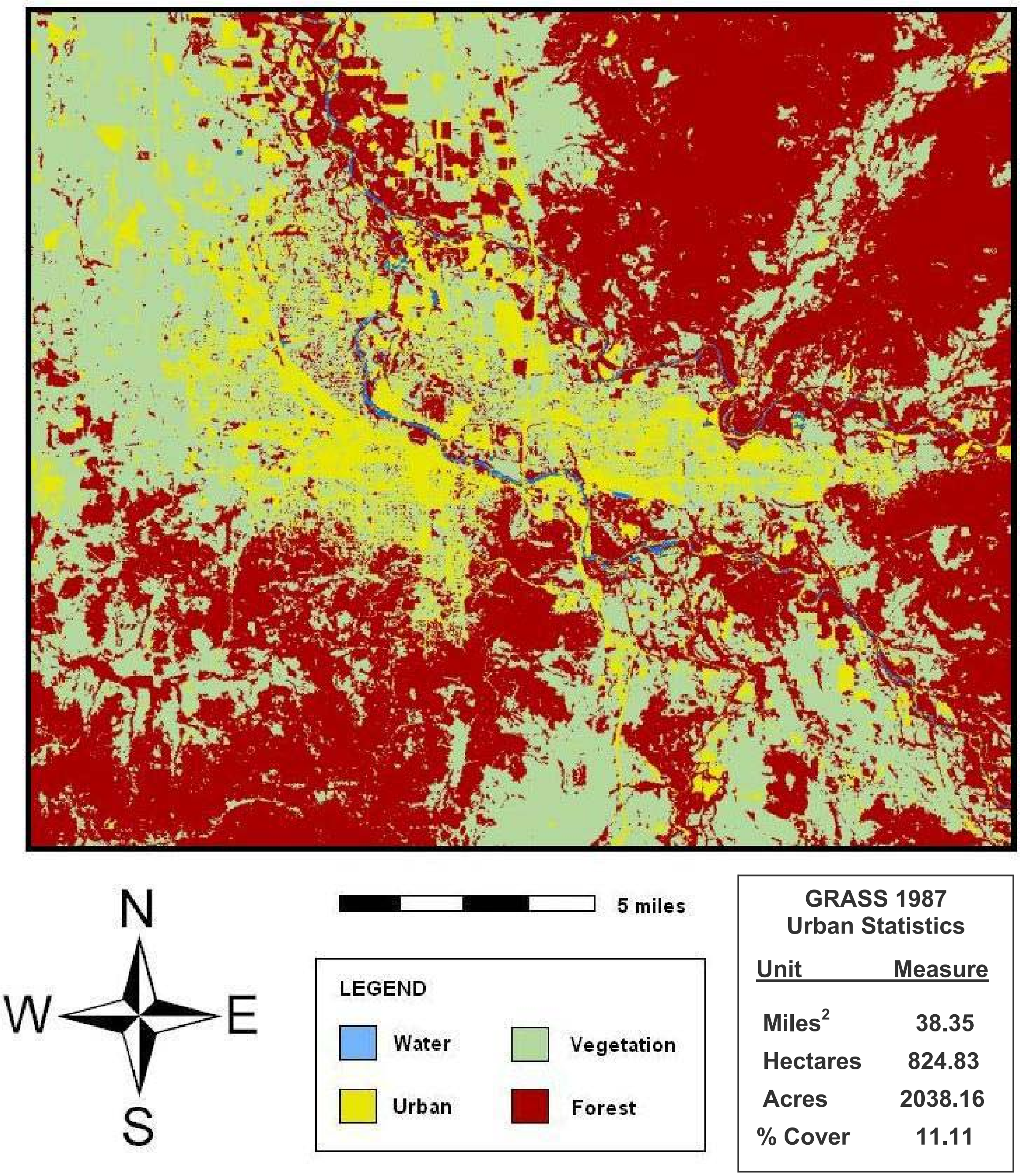


Figure 16 – GRASS interactively classified image of Landsat ETM+, 07/26/2001.

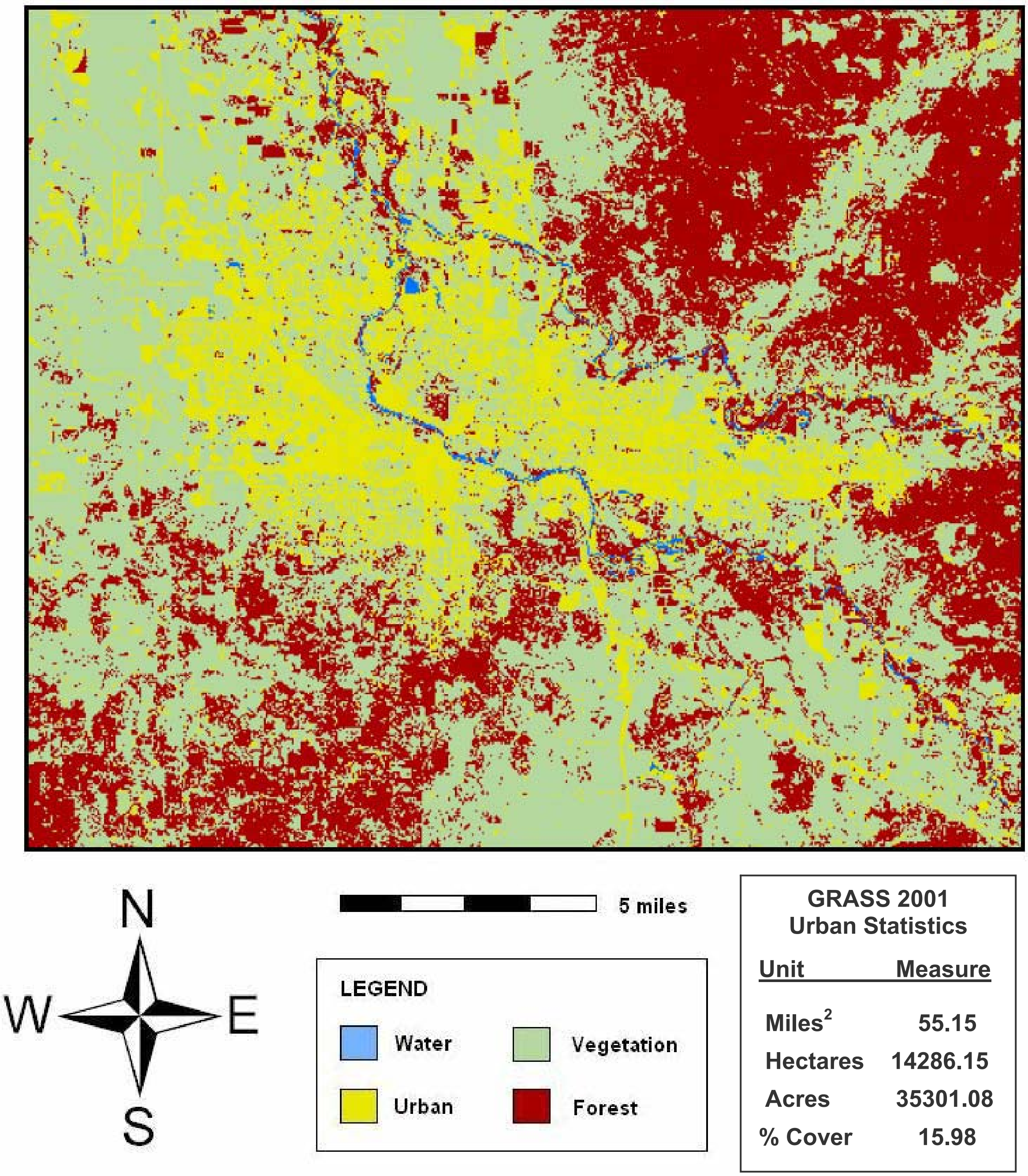
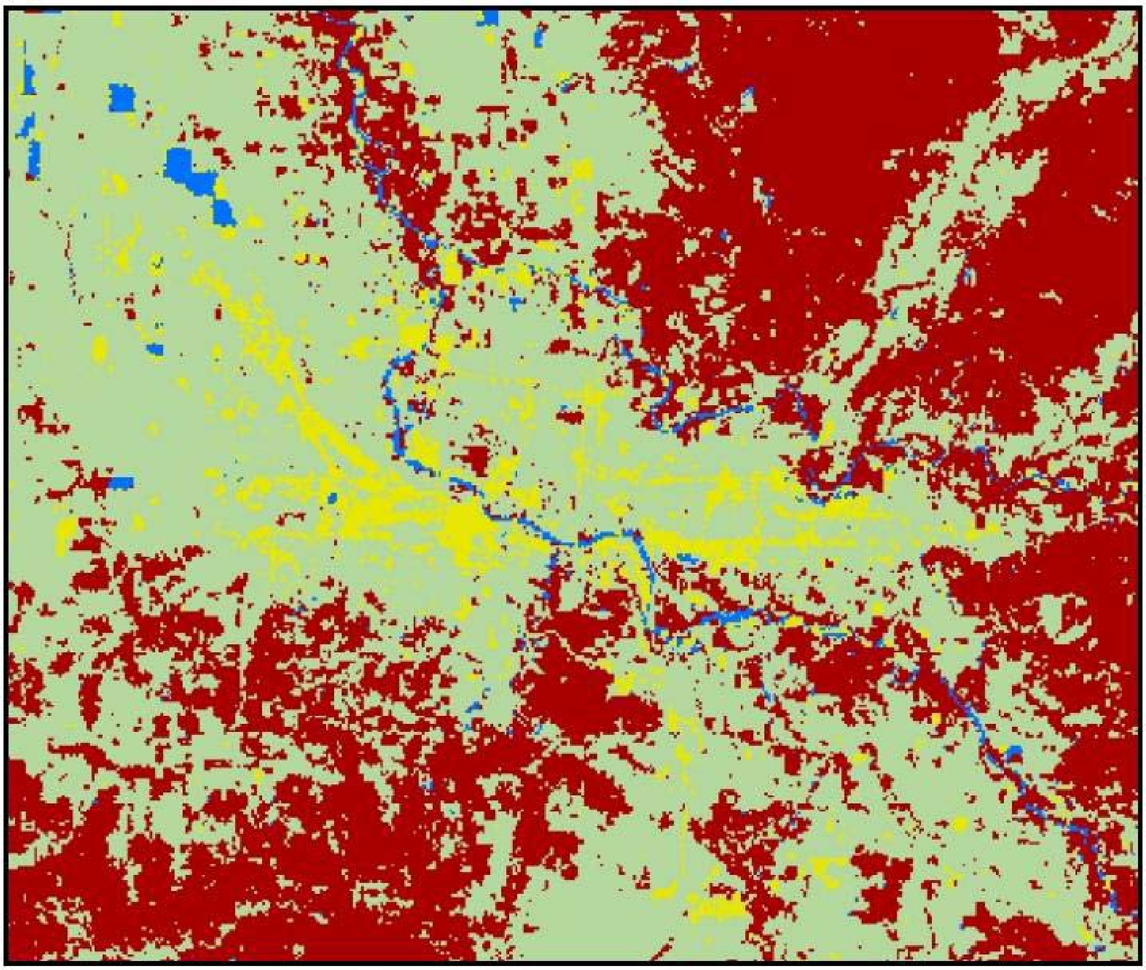
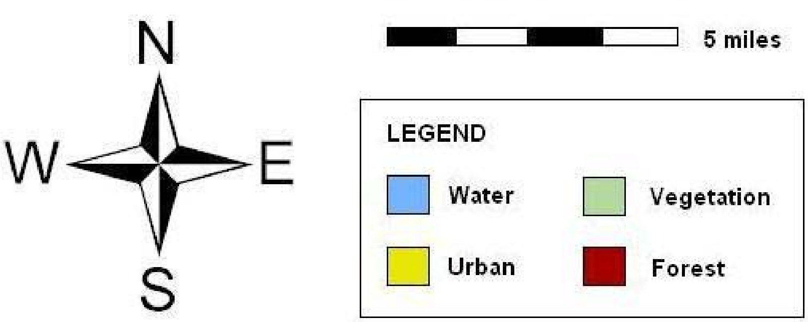


Figure 17 - GRASS non-interactively classified image of Landsat MSS, 07/30/1977.





GRASS 1977
Urban Statistics

Unit	Measure
Miles^2	17.31
Hectares	4483.81
Acres	11079.49
% Cover	5.01

Figure 18 - GRASS non-interactively classified image of Landsat TM, 07/12/1987.

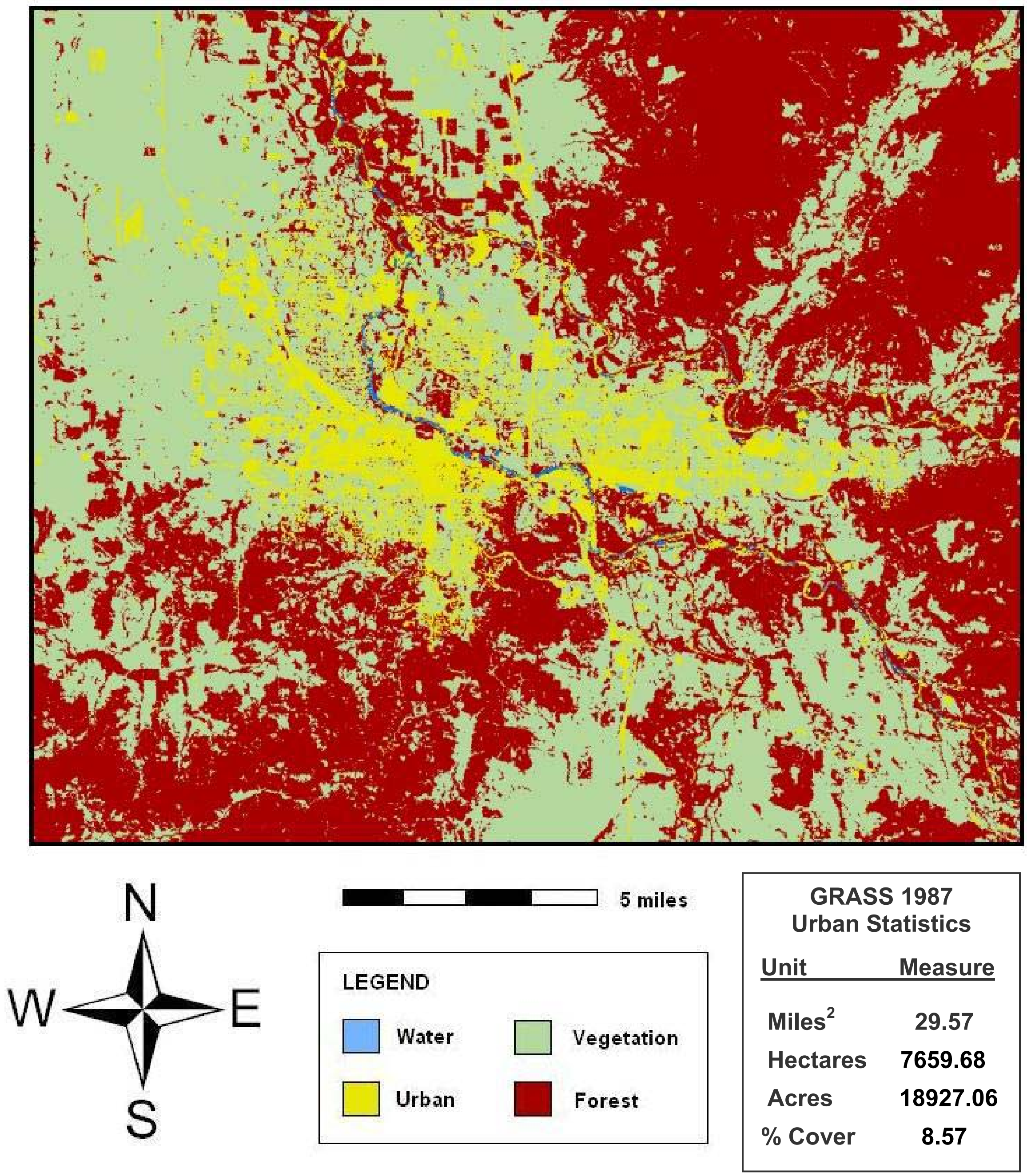


Figure 19 - GRASS non-interactively classified image of Landsat ETM+, 07/26/2001.

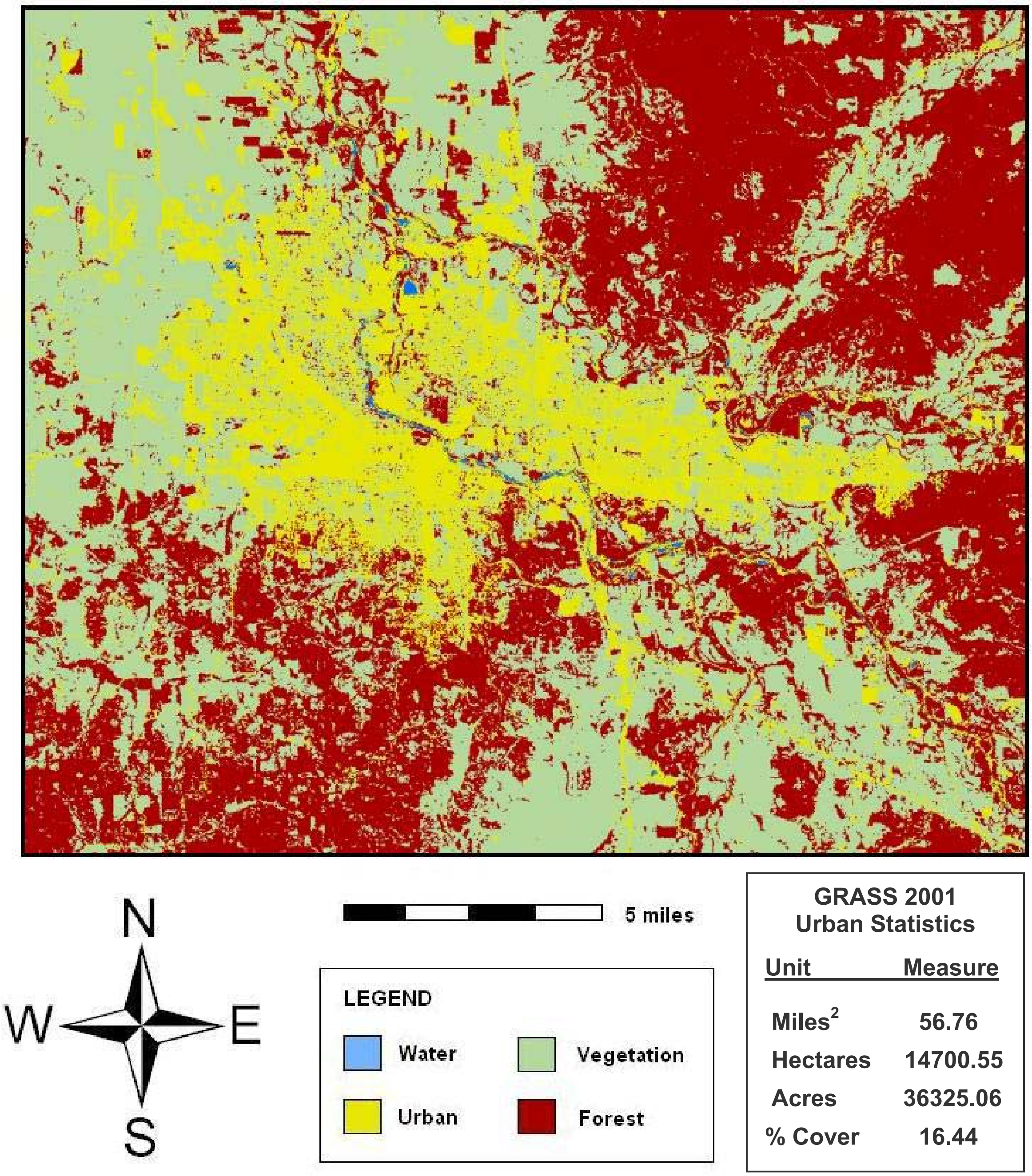


Figure 20 – GRASS urban boundaries by year created from interactively classified images.

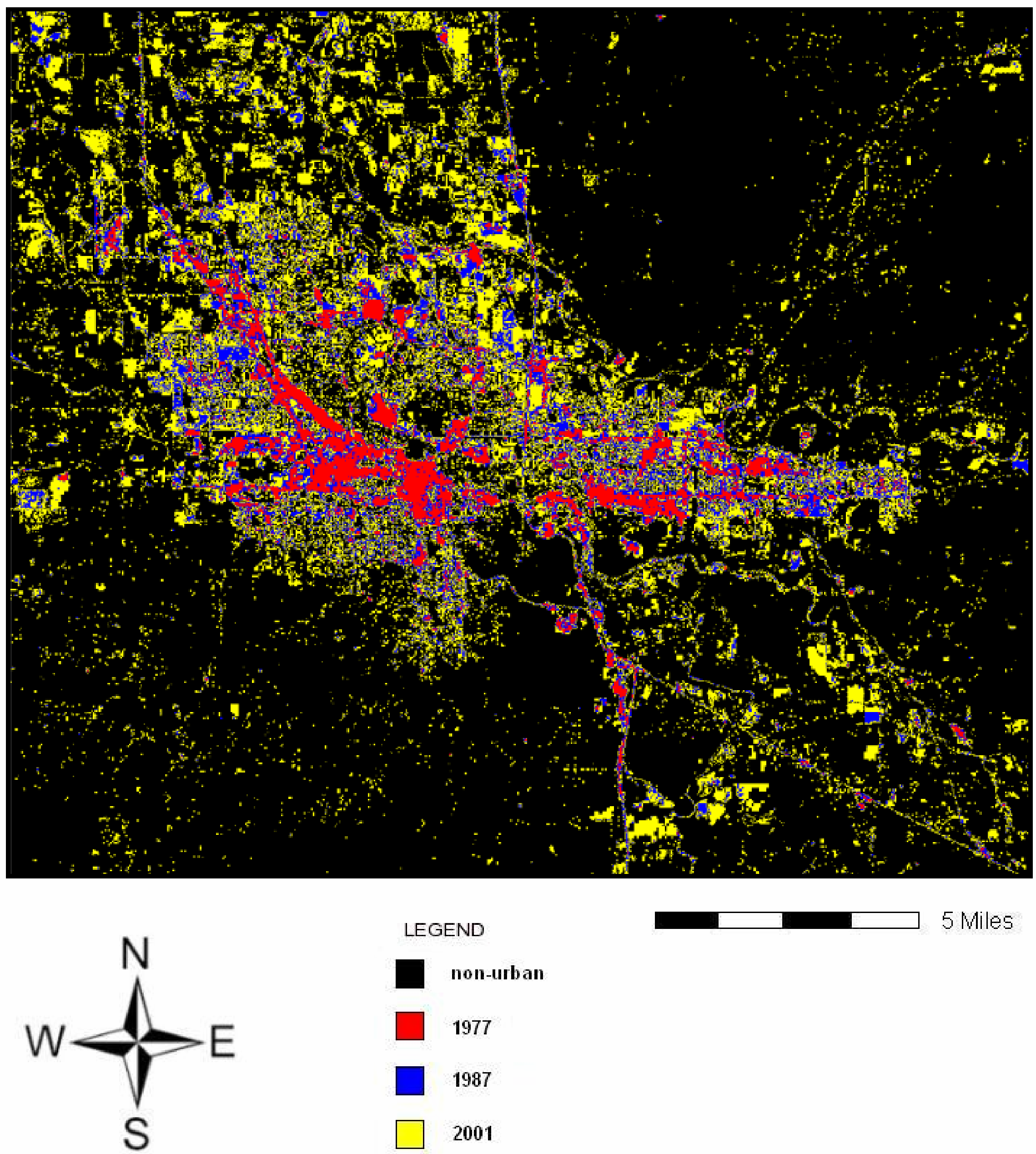


Figure 21 - GRASS interactively classified image of Landsat ETM+ draped over shaded relief map.

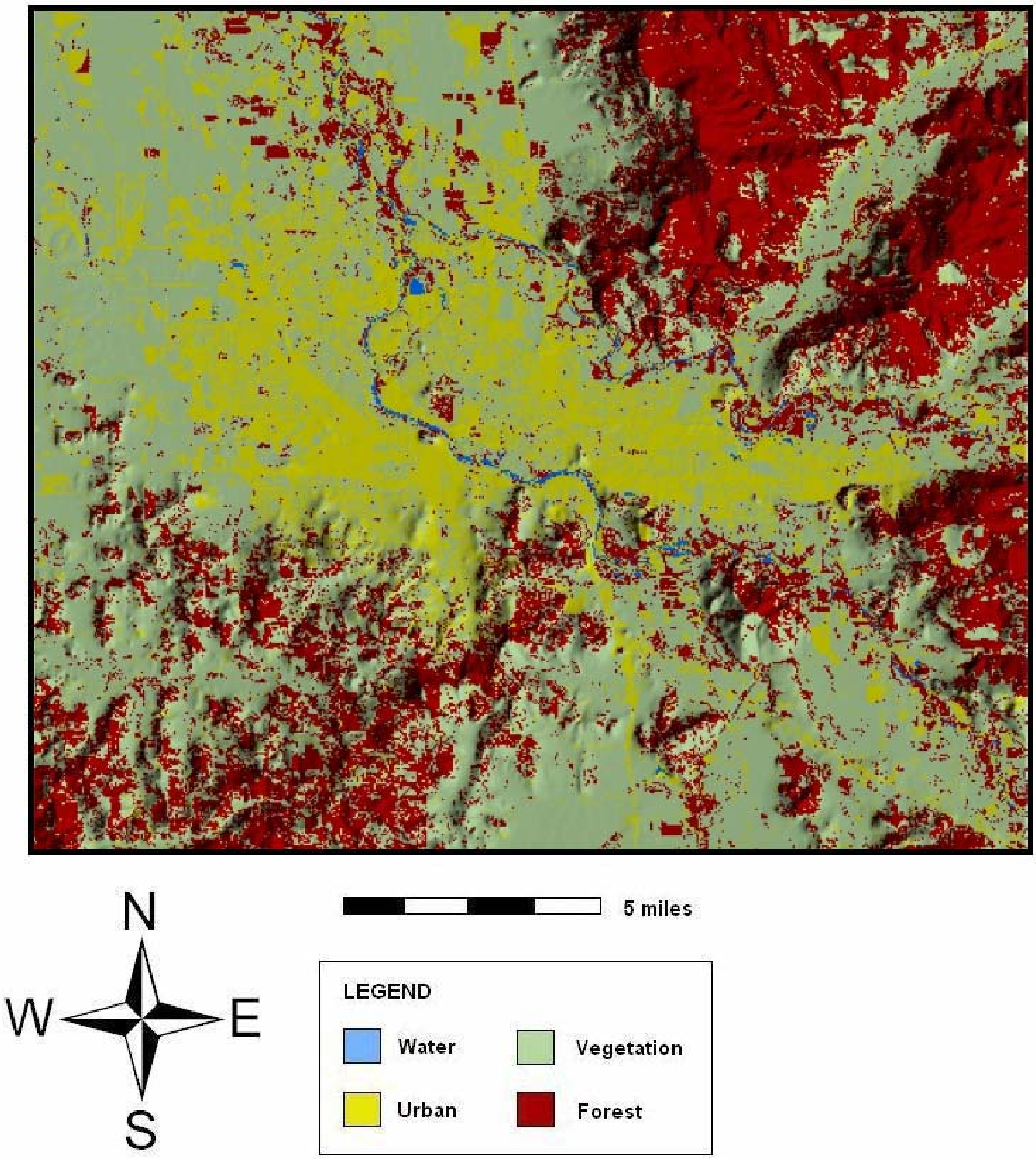


Figure 22 – GRASS NVIZ module 3D view of real-color RGB composite (bands 3, 2, and 1) of Landsat ETM+ image, 07/26/2001, draped on DEM of region.

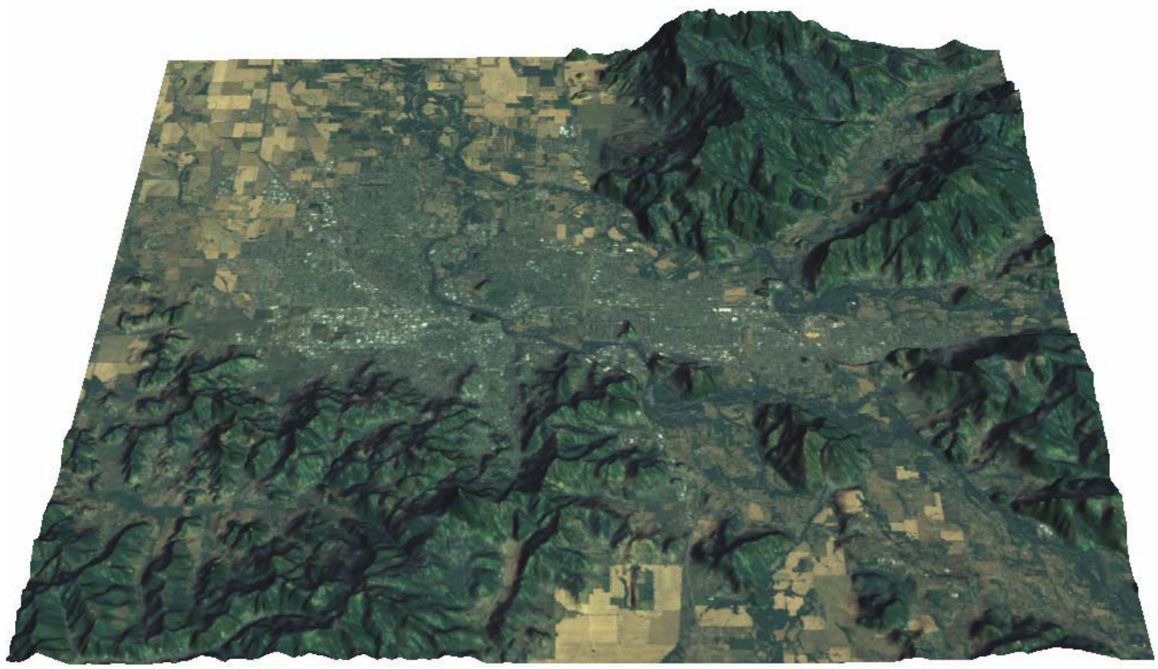
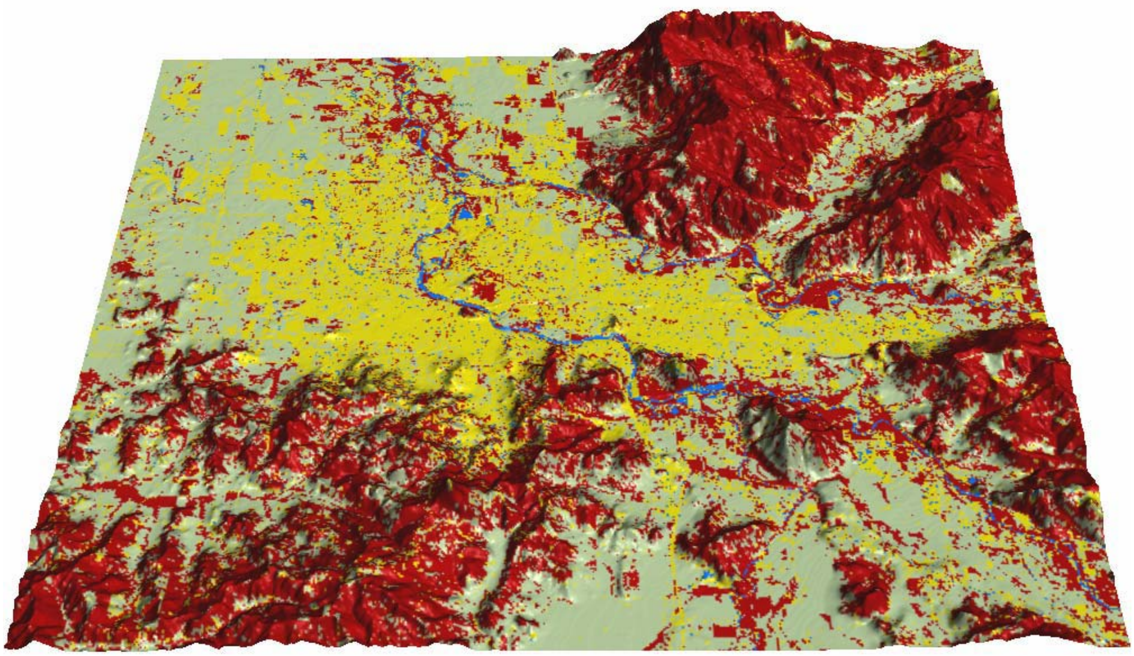


Figure 23 – GRASS NVIZ module 3D view of interactively classified Landsat ETM+ image, 07/26/2001, draped on DEM of region.



APPENDIX A
GLOSSARY

Bash shell – UNIX commend prompt; this is one type of user input used by GRASS on Cygwin.

Cygwin – A collection of tools which provide a Linux-like environment for a Windows operating system.

DEM – Digital Elevation Model; raster dataset representing elevation values.

DOQQ – Digital Orthophoto Quarter-Quad. A digital raster image of an aerial photo mapped by the USGS, specifically a "quarter-quad," or the extent of one fourth of a 7.5-minute quadrangle, often used in GIS applications.

ETM+ - Enhanced Thematic Mapper Plus; A sensor placed on Landsat 7 which was deployed in 1999 and operational today. ETM+ consists of 8 bands.

GeoTIFF – Georeferenced TIFF file which includes a separate file that describes the ground truth location of the data and ties the imagery to real world coordinates.

GPL – General Public License; A license with the purpose to grant any user the right to copy, modify and redistribute programs and source code from developers that have chosen to license their work under the GPL.

GUI – Graphical User Interface; A computer terminal interface, such as Windows, that is based on graphics instead of text.

Interactive classification – A type of image classification which requires the user to actively input classification criteria for each ROI or training area defined.

LCC – Lambert Conformal Conic; A map projection commonly used for areas that are longer in east-west extent than north-south extent.

Maximum likelihood classification – A type of classification method used in which a pixel with the maximum likelihood is classified into the corresponding class.

MrSID – Multi-resolutional Seamless Image Database; compressed image file format developed by Lizardtech.

MSS - Multi-Spectral Scanner; A sensor placed on Landsat 1 – 3 which were in service between 1972 and 1983. MSS consisted of 4 bands.

Non-interactive classification – A type of image classification which requires the user to digitize all ROIs or training areas at once and does not allow the user to define classification criteria such as standard deviation.

On-the-fly projection – ArcGIS feature which automatically overlays data in varying projection for quick display and query.

Raster - Common GIS data model that represents continuous data made up of grid cells with varying values. Satellite and aerial imagery are common raster datasets.

RGB – Red Green Blue; color model based on additive primary colors for computer video display.

Shaded relief map – Dataset produced by interpolation of a DEM which displays a representation of elevation.

Shapefile – A set of files developed by ESRI that contain a set of points, arcs, or polygons that hold tabular data and a spatial location.

SDTS - Spatial Data Transfer Standard; The formal standard specifying the organization and mechanism for the transfer of GIS data between dissimilar computer systems.

TIFF – Tagged Image File Format; a type of uncompressed image file.

TM – Thematic Mapper; A sensor placed on Landsat 4 and 5 which were in service between 1982 and 1999. TM consisted of 7 bands.

UTM – Universal Transverse Mercator; a coordinate system based on a grid that was developed for the U.S. Army by the U.S. National Imagery and Mapping Agency (NIMA). In this grid, the world is divided into 60 north-south zones, each covering a strip 6° wide in longitude. Within each zone, coordinates are measured in Northing and Easting values.

Vector – Common GIS data model that represents discrete data and is made up of points, lines, or polygons, with associated attributes.