

DETECTING THE PRESENCE OF DISEASE BY UNIFYING TWO METHODS
OF REMOTE SENSING

Steve Reames, B.F.A, M.S.

Dissertation Prepared for the Degree of
DOCTOR OF PHILOSOPHY

UNIVERSITY OF NORTH TEXAS

May 2002

APPROVED:

Ana D. Cleveland, Major Professor
Sam Atkinson, Committee Member
Cecil Hallum, Committee Member
Cathleen Norris, Committee Member
Jon Young, Committee Member and Chair
of the Department of Technology and
Cognition
Philip M. Turner Director Interdisciplinary
Ph.D. Program in Information Science
Neal Tate, Dean of the Robert B. Toulouse
School of Graduate Studies

Reames, Steve, Detecting the presence of disease by unifying two methods of remote sensing. Doctor of Philosophy (Information Science), May 2002, 174 pp., 4 tables, 48 illustrations, reference, 100 titles.

There is currently no effective tool available to quickly and economically measure a change in landmass in the setting of biomedical professionals and environmental specialists. The purpose of this study is to structure and demonstrate a statistical change-detection method using remotely sensed data that can detect the presence of an infectious land borne disease.

Data sources included the Texas Department of Health database, which provided the types of infectious land borne diseases and indicated the geographical area to study. Methods of data collection included the gathering of images produced by digital orthophoto quadrangle and aerial videography and Landsat. Also, a method was developed to identify statistically the severity of changes of the landmass over a three-year period.

Data analysis included using a unique statistical detection procedure to measure the severity of change in landmass when a disease was not present and when the disease was present. The statistical detection method was applied to two different remotely sensed platform types and again to two like remotely sensed platform types

The results indicated that when the statistical change detection method was used for two different types of remote sensing mediums (i.e.-digital orthophoto quadrangle and aerial videography), the results were negative due to skewed and unreliable data. However, when two like remote sensing mediums were used (i.e.- videography to videography and Landsat to Landsat) the results were positive and the data were reliable.

Copyright 2002

by

Steve Reames

ACKNOWLEDGMENTS

I want first to recognize my committee chair, Ana D. Cleveland for her unflagging support, expertise, and continuing regard throughout this process. I also want to say a heart-felt thank you to other committee members, Sam Atkinson, Cecil Hallum, Cathleen Norris, and Jon Young each of whom has given unfailingly of their time, energy, and expertise.

Secondly, I recognize and thank Sam Houston State University and The Texas Research Institute for Environmental Studies (TRIES), for grant monies dedicated to this research and the unrelenting assistance by the faculty and staff, including Randy Garner, Jamie Herbert, Richard Rush, and Gary Smith. I would be remiss in not thanking Julie Rawlins, with the Texas Department of Health; Steve Cerwin, of the Southwest Research Institute; Ray Taylor, of The American Radio Relay League; Intuitive Circuits, LLC; Charles Parks and James Jackson, of C.P. Electronics; programmer, Roberto Salazar; and Brian O'Connor of the University of North Texas. Without the assistance of all of these individuals this body of work would have been impossible.

Thanks as well go to my family and friends, including my wife, Tatiana, my daughters, Big Sasha and Little Sasha; my sister, Tammy Lackey; my mother, Nita Carter; and friends, Sol Schneider and Don and Sharon Booth. Without their support, love, and understanding I would not have had the courage to complete this degree.

TABLE OF CONTENTS

ACKNOWLEDGMENTS	iii
LIST OF TABLES	vii
LIST OF ILLUSTRATIONS	viii
Chapter	
1. INTRODUCTION	1
Statement of Problem.....	1
Purpose of Study	1
Significance of the Study	1
Background	2
Remote Sensing	2
Scale and Resolution of Image Data.....	4
Spatial	5
Spectral	5
Radiometric.....	6
Temporal.....	7
Data Quality	9
Data Analysis	10
Image Processing	15
Analog Image Processing	16
Digital Image Processing	17
Preprocessing	18
Enhancement.....	23
Classification.....	31
Remote Sensing Data Types	39
Aerial Videography.....	39
Digital Orthophoto Quadrangle (DOQ) Aerial Photography.....	40
Landsat	41
Definition of Terms.....	43
Research Question	44
Hypothesis.....	45
Assumption and Limitations	45
2. REVIEW OF THE LITERATURE	46

Aerial Videography	47
Aerial Videography Mapping-Wetland	47
Aerial Videography Mapping-Soil Sampling	48
Aerial Videography Mapping-Data Collection.....	49
Aerial Videography Mapping-Low Technology	51
Aerial Videography Mapping-Inexpensive.....	51
Digital Orthophoto Quadrangle (DOQ) Aerial Photography.....	53
Spatial Resolution of DOQ	53
Spectral Range of DOQ	54
Data Organization of DOQ	54
Accuracy of DOQ	55
Landsat	57
Wildlife and Resource Management Using A GIS Landsat GIS System.....	57
Landsat 7.....	61
Land Remote Sensing Policy Act of 1992.....	63
Remote Sensing and Health	64
Remote Sensing and Health Research	68
Remote Sensing and Health Education.....	69
Remote Sensing and Health Planning.....	70
Considerations in Landscape Epidemiology	71
Remotely Sensed Landscape Epidemiology and Tick-Borne Diseases	72
Rocky Mountain Spotted Fever (RMSF).....	74
Lyme Tick Disease	79
 3. METHODOLOGY	 86
Research Question	86
Research Design.....	86
Selection of Subject(s)	88
Disease Candidate Selection	89
Aerial Videography.....	94
Digital Orthophoto Quadrangle (DOQ) Aerial Photography.....	94
Subject Development	95
Tick-Borne Disease.....	96
Aerial Videography.....	96
Digital Orthophoto Quadrangle (DOQ) Aerial Photography.....	97
System Creation.....	97
Converter Box.....	101
System Test.....	103
Flight of Jewett Quad.....	104
Registration of Images	104
Detection.....	104

Change-detection-Differencing.....	106
Difference Characteristics-The Change Detector	107
Severity Values	107
Ranking Severities	108
Severity Data Set.....	108
Viewing Severity Frames.....	109
4. DATA COLLECTION AND ANALYSIS.....	110
DOQ and Videography Severity of Change Analysis	110
Data Gathering	110
Registration of Images	111
File Conversion	113
Data Review DOQ to Video Registration.....	114
Two-Band DOQ and Videography Severity of Change Analysis	116
Registration of Images	118
File Conversion.....	118
Data Review.....	119
Videography to Videography Severity of Change Analysis.....	119
Data Gathering.....	119
Registration of Images	120
File Conversion.....	120
Data Review.....	121
Landsat to Landsat Severity of Change Analysis	127
Data Gathering.....	127
Registration of Images	128
File Conversion.....	129
Data Review.....	129
5. FINDINGS and CONCLUSIONS	132
Findings.....	132
DOQ and Videography	132
Two-Band DOQ and Videography	133
Videography and Videography	133
Landsat to Landsat	134
Conclusions	135
Summary of Findings.....	136
Future Research	136
APPENDICES	139
REFERENCES	167

LIST OF TABLES

Table	Page
1. Categories and Dimensions of Data Quality.....	9
2. Elements of Image Interpretation.....	16
3. TDH Data 1995 to1999.....	91
4. Ratio of occurrence(s) of a Disease vs. Zip Codes Affected 1995-1997	92

LIST OF ILLUSTRATIONS

Figure	Page
1. Human transmission chain for RMSF tick disease	76
2. The two-year cycle of RMSF is based on the lifespan of the tick	77
3. The deductive process model.....	86
4. The relationship of model to research design	87
5. The research design flowchart	87
6. The selection of subject(s) flowchart.....	88
7. Sample query of the TDH database	90
8. The candidates for land base infectious diseases.....	90
9a. Occurrences by zip code of flea-borne diseases	92
9b. Occurrences by zip code of mosquito-borne diseases	93
9c. Occurrences by zip code of tick-borne diseases	93
10. The development of subject(s) flowchart	95
11. The system creation flowchart.....	98
12. The selected zip code 75856.....	99
13. The seven quads DOQ overlay of zip code 75856	100
14. DOQ Jewett Quad (2.5 meter depicted).....	100
15a. Wing camera	101
15b. Digital video camera.....	101
15c. Converter box.....	101

15d. GPS	101
16. The data stream flow to and from the unique converter box	102
17. The special built integrated circuit board.....	102
18. The aerial video frame with serial data overlay.....	103
19. The change-detection flowchart—a multivariate statistical procedure	105
20. Random video frames captured over Jewett DOQ Quad area	111
21. Video frame 14 selected to register with the Jewett DOQ.....	112
22. Registered video image.....	113
23. Registered DOQ image	113
24. Example of ASCII output by row-column-R-G-B value.....	114
25. DOQ and Video Color Sequencing.....	116
26. DOQ Green Data Band 1	117
27. DOQ Red Data Band 2	117
28. DOQ NIR Data Band 3	117
29. Video Blue Band 1.....	117
30. Video Green Data Band 2.....	117
31. Video Red Data Band 3	117

32. Registered Video Frame 5-24	120
33. Registered Video Band 7-22	120
34. Example of data polled	121
35. 80% Severity-Video.....	126
36. 90% Severity-Video	126
37. 95% Severity-Video.....	126
38. 99% Severity-Video.....	126
39. 99.9% Severity-Video.....	127
40. Nueces Landsat (10-20-1999).....	128
41. Nueces Landsat (4-6-2000).....	128
42. Nueces Registered Landsat	129
43. Nueces Registered Landsat	129
44. 80% Severity-Landsat.....	130
45. 90% Severity-Landsat.....	130
46. 95% Severity-Landsat.....	131
47. 99.9% Severity-Landsat.....	131
48. 99.99% Severity-Landsat.....	131

CHAPTER 1

INTRODUCTION

Statement of the Problem

A method for a quantitative assignment of pixel level severity of temporal change to the landscape from the use of aircraft-based, remotely sensed data does not currently exist. In particular, such a capability for reading a landscape change at the pixel level between a point in time when a disease was not present and a point in time when the disease was present is a critically needed tool.

Purpose of the Study

The purpose of the study was to structure and demonstrate a change-detection method using aircraft-based remotely sensed data that will assign quantitative severities at the pixel level. The study provides a capability to an analyst in answering the question: Can the detection and severity of an infectious disease(s) be determined by observing changes in landscape?

Significance of the Study

There is a need for a method for interpreting remotely sensed data that is accurate and timely. This method will assist an analyst in gaining further insight into the relationship or characteristics between the encroachment of a disease and changes in the landscape. Identifying a reliable method could have an impact on the health of a community by saving

time and predicting the possible spread of a land-based disease. Such a determination of the surrounding landscape can be made for the purpose of epidemic disease prevention and/or the pre-inoculation of a human population.

Background

Remote Sensing

Remote sensing includes information collected from a variety of platforms, such as satellite systems, fixed-wing aircraft, helicopters, and remotely piloted aircraft. Remote sensing techniques can be defined as a computer-assisted information management system of geo-referenced data. This system integrates the acquisition, storage, analysis, and display of geographic data. The application field and objectives of remote sensing can be varied, and they involve a great number of questions linking social and physical problems. Generally, the objectives of remote sensing techniques are the management (acquisition, storage, maintenance), analysis (statistical, spatial modeling), and display (graphics, mapping) of geographic data.

Remote sensing is the process of collecting data about objects or landscape features without coming into direct physical contact with them. The application of remotely sensed data to landscape epidemiology requires an understanding of several components: *the scale and resolution of image data, data quality, data analysis, and image processing*. Each component is important and requires consideration by the researcher.

Scale and Resolution of Image Data

The data characteristic of gray scale and resolution is the most important quality for data integrity. Remote sensing images acquired in various spectral bands are used to estimate certain geophysical parameters or to detect the presence or extent of geophysical phenomena. In a majority of cases, the raw image acquired by the sensor is processed using various operations such as filtering, compression, and enhancement. In performing these operations, the analyst is attempting to maximize the information content in the image to fulfill the end objective.

The information content in a remote sensing imagery for a specific application is greatly dependent on the gray-scale resolution of the image. It may seem obvious that, as the gray-scale resolution is decreased, the information content reduces; but in remote sensing images there may not be as great a loss as might be expected. This is because, although the value of a pixel may change as a result of the decrease in the resolution, the same pixel may, in most cases, be correctly classified. One of the measures to quantify information content is classification accuracy. The loss in information is exponential with respect to the number of gray levels. Using a mathematical model for the information content of images as a function of gray-scale resolution, one can accurately specify the "optimal" resolution of an image for a specific purpose. There are four types of resolution, and an understanding of each is necessary before conducting research. Each could affect the nature of the remotely sensed data. All remotely sensed data have each of these four resolution characteristics: *spatial*, *spectral*, *radiometric*, and *temporal*.

Spatial. Spatial resolution is a measurement of the minimum distance between two objects that will allow them to be differentiated from one another in an image (Jensen, 1996). This is a function of sensor altitude, detector size, focal length, and system configuration. The electromagnetic spectrum is the extent of energy propagated through space between electric and magnetic fields whose range includes (in increasing wavelength) gamma rays; X-rays; and ultraviolet, visible, infrared, microwave, and VHF radiation. For aerial photography, the spatial resolution is usually measured in resolvable line pairs per millimeter on the image (Jensen, 1996). It is important to understand that display scale and spatial resolution are not necessarily related. Remotely sensed data with any nominal spatial resolution may be displayed anywhere from 1:1200 to 1:1,000,000 and infinitum. The spatial resolution and display scale are independent. However, there are practical limitations—Advanced Very High Resolution Radiometer (AVHRR), for example. Images are generally scale independent. However, diseases may be studied at many different scales. The proper spatial resolution is the first requirement in data selection.

Spectral. The number and size of the electromagnetic bands recorded by a sensor determine the instrument's spectral resolution. Sensors are unique with regard to which portions of the electromagnetic spectrum they record. Different remote sensing instruments record different segments, or bands, of the electromagnetic spectrum. A sensor may be sensitive to a large portion of the electromagnetic spectrum but have poor spectral resolution because its sensitivity is contained within a single wide band. A hyperspectral sensor sensitive to the same portion of the electromagnetic spectrum, but with many small bands, would have greater spectral resolution. With higher spectral resolution both

computer and analyst are able to better distinguish among scene elements. When combined with better spatial resolution, higher spectral resolution promotes order of magnitude improvements in image interpretability. Historically, the spectral resolution of sensors was predetermined in order to best identify land surface features considered most important. Landsat satellite spectral thematic mapper (TM) data—particularly the red thematic mapper (TM3), near infra-red (TM4), and mid-infrared (TM7) regions—are sensitive to characteristics of the rice field canopy such as percent cover and leaf area, which can in turn be linked to plant filtering and mosquito larval habitat quality (Wood, Beck, Washino, et al., 1991). More detailed information (e.g., greater spectral resolution) about how individual elements in a scene reflect or emit electromagnetic energy increases the probability of finding unique characteristics for a given feature on the landscape, allowing it to be distinguished from other features in the scene. For instance, in identifying villages at high risk for malaria transmission, Beck et al. (1994) successfully used remote sensing reflectance measurements within a limited number of bands of Landsat satellite thematic mapper (TM) data to classify landscapes of high probability for malaria. While seeming ideal, hyperspectral data are so information rich that it is often difficult to decide which bands to use in a particular application or test. Furthermore, additional spectral bands mean additional data storage and manipulation requirements. It is often the case that the availability of data far exceeds the ability to manage it.

Radiometric. Radiometric resolution refers to the sensitivity of the sensor to incoming radiation. The number of different levels of radiance that a particular sensor can distinguish characterizes this form of resolution. It is easiest to think of this type of resolution with

respect to panchromatic data. Data sampled in 6 or 8 bits record 64 or 256 shades of gray. This sensitivity to different signal levels will determine the total number of values that can be generated by the sensor (Jensen, 1996). Wood (Wood, Washino, Beck, et al., 1991) used multispectral Daedalus (National Aeronautical and Space Administration [NASA] airborne scanner) data but found that the radiometric quality of the thermal band made that particular band unusable. However, special infrared glasses, crystalline, and hollow fibers, along with UV transmissive fibers, are being used in the radiometric measurement of temperature, evanescent-wave chemical sensing, and the delivery of UV and IR laser power for the ablation of tissue. As biomedical fiber sensors, the potential exists for rapid, cost-effective, point-of-care clinical diagnostics and invasive or noninvasive physiological monitoring of parameters such as blood pressure, blood chemistry, tissue typing, temperature, and the presence of pathological organism substances.

Temporal. Temporal resolution refers to the amount of time it takes for a sensor to return to a previously recorded location. Seasonally dependent research and change-detection require an understanding of and planning for temporally dependent data collections. Most orbital remote sensing platforms pass over the same spot at regular time intervals. These intervals were historically dependent strictly on the orbital parameters of the platform in question; today, however, many platforms carry sensors with pointing capabilities. This variable directional feature reduces apparent temporal resolution from weeks to days. Data collected on multiple dates allow the scientist to chart changes of phenomena through time. Examples include the growth of crops in various parts of the world, the expansion of urban areas, monitoring desertification, disease proliferation, and,

perhaps most common, the ever-changing weather. By improving one or any combination of these resolutions, a scientist can increase the chance of obtaining accurate and useful remotely sensed data (Jensen, 1996). Roberts et al. (1991) considered also the importance of the characterization of the preferred breeding sites by season of year as the first step in resolving the research problem with the quantification of the temporal availability of the breeding sites by category of habitat. Pope et al. (1992) were limited by the temporal resolution of the existing suite of satellite systems in the need for timely (in terms of seasonality) acquisitions.

In the vast majority of research using remote sensing as a data source, the scientists often express the need for better resolution. This general and inaccurate use of “resolution” invariably refers to spatial resolution. The negative aspect of increased resolution is the need for increased storage space and more powerful data-processing tools (Jensen, 1996). An increase in spatial resolution from 10 meters to 1 meter is a 100 times increase in data volume per band. For these reasons, it is important to determine the minimum resolution requirements needed to accomplish a given task from the outset. This avoids unnecessarily wasting time in processing more data than are needed. It will also help in avoiding the problem of too little data to allow completion of the task.

Data Quality

There is a long-standing belief that data quality has many attributes (Agmon & Ahituv, 1987; King & Epstein, 1983; Seddon, 1997). Wang and Strong (1996) proposed a data quality framework that includes the categories of accuracy, relevancy, representation, and accessibility. Wang and Strong proposed 20 data quality dimensions (see Table 1), which they later pared to 15, and which they assembled into four categories. They demonstrated that 15 dimensions were relevant to a generic population (Grendron & D’Onofrio, 2001).

Table 1

Categories and Dimensions of Data Quality

<u>Category</u>	<u>Dimension</u>
Accuracy of Data	Believability Accuracy Objectivity Reputation
Relevancy of Data	Value Added Relevancy Timeliness Completeness Appropriate Amount of Data
Representation of Data	Interpretability Ease of Understanding Representational Consistency Concise Representation
Accessibility of Data	Accessibility Access Security
Eliminated Dimensions	Traceability Variety of Data Sources Ease of Operation Flexibility Cost-Effectiveness

Ground truthing is one method used in remote sensing that can demonstrate by category and dimension a credible data quality measurement. Specifically, ground truthing is the acceptable method of surveying the ground to confirm the findings of an aerial survey or calibrate quantitative aerial observations. Often it is necessary to collect ground truth data specifically correlated with the image data. Spatial location or georeferencing is often dependent upon accurately collected ground control points or quantitative samples. These data are combined with the remotely sensed data to assign qualitative values to broad classes or provide statistical validation. Consequently, the collection of quality data may take many forms. From true synoptic field sampling to laboratory sampling, the range of acceptable scales of sampling are broad and are generally dependent upon the cost and inclinations of the researcher. With appropriate ground studies, remote sensing can identify and extensively map in a probabilistic manner the potential habitats of specific parasites and vectors (Hugh-Jones, 1989). Global positioning system (GPS) receivers are becoming commonplace, serviceable, and readily available. Differential GPS provides the precision to allow the accurate geolocation of the ground truth data.

Data Analysis

Data analysis is an analytic process designed to explore large amounts of data in search of consistent patterns and/or systematic relationships between variables and then to validate the findings by applying the detected patterns to new subsets of data. By definition, data analysis requires knowledge of data structure and semantics to analyze and quantify the data elements. However, remotely sensed data were first analyzed visually. Early aerial

photographs were used in military campaigns with the identification of known tactical features deemed most important. By using various image processing techniques and methods, an analyst may use the data to discern features invisible to the naked eye. These techniques include both visual processing techniques applied to hard copy data, such as photographs or printouts, and the application of digital image processing algorithms to the digital data (Jensen, 1996). The process of data visualization allows the analyst to examine data from all possible angles and to place entire images in context with their surroundings. Contextual organization in most cases helps to establish the link between abstract imagery and real places. This visual analysis by a remote sensing data analyst proved to be both time consuming, cumbersome, and outdated.

The demand for an effective method for the remote sensing data analyst grew as computer technology evolved. The rapid technology growth demanded a definitive data analysis system. The following outlines specific design goals implemented by remote sensing analysts as a result of this demand (Biehl & Landgrebe, 2001).

1. The implementation should be on a readily available computer platform that has adequate processing power, but is financially within the reach of a researcher (i.e., computer platforms).
2. The system should be easy to learn and easy to use, even for the infrequent user, using the most modern of software environments. Features of this system should include the following:

the ability to import data in either Binary or ASCII format with or without a header, and in Band Interleaved by Line (BIL), Band Sequential (BSQ), or Band Interleaved by Sample (BIS) formats. The data may have either 1 or 2 bytes per data value, and may have 4 to 16 bits per data value. In the case of 2 bytes per sample, the two bytes may be in either order.

the ability to display multispectral images in a variety of B/W or color formats using linear or equal area gray scales; display (internally generated) thematic images also in B/W or color, with an ability to control the color used for each theme. ArcView Shape Files may be overlaid on the images.

the ability to use histogram data in determining the gray-scale regime for a display or for listing and graphing.

the ability to reformat the data file in a number of ways, that is, by adding a standard header, changing from any one of the three interleave formats to either of the other two, editing out channels, combining files, adding or modifying channel descriptions, mosaicing data sets, changing the geometry of a data set, and a number of other changes.

the ability to create new channels of data from existing channels. The new channels may be the result of a principal components or feature extraction transformation of the existing ones, or they may result from the ratio of a linear

combination of existing bands divided by a different linear combination of bands.

the ability to cluster data using either a single pass or an iterative (isodata) clustering algorithm; to save the results for display as a thematic map. Cluster statistics can also be saved as class statistics. Use of clustering followed by ECHO spectral/spatial classification provides an effective multivariate scene segmentation scheme.

the ability to define classes via designating rectangular or polygonal training fields or mask image files, compute field and class statistics, and define test fields for use in evaluating classification results quantitatively. A feature called "Enhance Statistics" also allows one to improve the extent to which the defined class statistics fit the composite of all data in the data set. The result is a covariance estimation scheme.

the ability to determine the best spectral features to use for a given classification using (a) a search for the best subset of features using any of five statistical distance measures, (b) a method based directly upon decision boundaries defined by training samples, or (c) a second method based directly upon the discriminate functions. Also included are methods especially designed to search for narrow spectral features such as spectroscopic characteristics and for use of projection pursuit as a means of further improving the features extracted.

the ability to classify a designated area in the data file. Six different classification algorithms are available: use of minimum distance to means, correlation classifier (SAM), matched filter (CEM), Fisher linear discriminate, the Gaussian maximum likelihood pixel scheme, or the ECHO spectral/spatial classifier. Save the results for display as a thematic map, with or without training and test fields being shown. Apply a threshold to a classification and generate a probability/threshold map showing the degree of membership of each pixel to the class to which it was assigned.

the ability to list classification results of training or test areas in tabular form on a per field, per class, or groups of classes basis.

the ability to show a graph of the spectral values of a currently selected pixel or the mean \pm s for a selected area. Shows scatter diagrams of data from pairs of bands and ellipses of concentration for training sets and selected areas. Show a graph of the histograms of the class or field data values used for training. Show the coordinates of a currently selected area.

the ability to show a color presentation of the correlation matrix for a field or class as a visualization tool especially for hyperspectral data.

the ability to use several additional utility functions including listing out a subset of the data, for use externally, conducting principal component analysis, and so on.

the ability to transfer intermediate or final results, whether text, B/W image, or color image, to other application programs such as word processors, spreadsheet, or graphics programs by copying and pasting or by saving and then opening the saved file within another application.

the ability to provide for easy import of data in a variety of formats and easy export of results, both in thematic map and in tabular form.

Taken together, the capabilities noted can provide a state-of-the-art capability to the remote sensing analyst to analyze moderate and high dimensional data sets of practical size.

Image Processing

Image processing is the analyzing and manipulating of images. Currently, image processing generally involves three steps:

1. Import an image with an optical scanner or directly through digital photography.
2. Manipulate or analyze the image in some way. This stage can include image enhancement and data compression, or the image may be analyzed to find patterns that are not visible by the human eye. For example, meteorologists use image processing to analyze satellite photographs.
3. Output the result. The result might be the image altered in some way, or it might be a report based on analysis of the image.

Analog Image Processing

The first phase of any imagery application always includes a thorough visual inspection of the data. Even cursory visual analysis will reveal significant features within the imagery. The human image processing center is essentially an analog system well tuned to identifying patterns in the landscape. Photogrammetric techniques precisely measure landscape element(s) and provide ancillary information used in the feature identification. Table 2 notes the most commonly used elements of image interpretation.

Table 2

Elements of Image Interpretation

Element	Comments
Tone	lightness or darkness of a region
Texture	apparent roughness or smoothness
Shadow	shadows of buildings, trees, etc. help in feature identification and height
Pattern	the arrangement of individual objects into distinctive forms
Association	feature location among other objects without specific patterns
Shape	natural and man-made features have distinctive shapes
Size	allows absolute measurement and assists in feature identification

The choice of specific technique is often transparent to the analyst because all are used simultaneously in the interpretation process. Specific techniques can be applied as different objects of features are prioritized. An orderly approach to analog interpretation is especially

important when the researcher has only a limited knowledge of the study area. This scenario is quite common in preliminary studies. For example, if an analyst has little or no knowledge of the study area depicted in an image, he or she may use the pattern to distinguish between man-made and natural objects or arrangements, as in the case of an orchard versus a natural forest. The texture of an object is also useful in distinguishing objects. Different types of vegetation appear rough or smooth in a characteristic manner. Association combines general site knowledge with visual cues. For example, in many parts of the world, sport fields near large buildings generally identify schools, whereas winding roads in the middle of forests and fields signify cemeteries (Avery & Berlin, 1985). Data visualization techniques combine with the concept of examining remotely sensed data in multiple bands of the electromagnetic spectrum (multispectral), on multiple dates (multitemporal), at multiple scales (multiscale), and in conjunction with other scientists.

Digital Image Processing

Digital image processing is not only a step in the remote sensing process, but is itself a process consisting of several steps. It is important to remember that the ultimate goal of this process is to derive new data from an image and apply these new data towards resolving some question in the disease vector cycle. The steps taken in processing an image will vary between image types. Pope et al. (1992) used Advanced Very High Resolution Radiometer (AVHRR) optical imagery and airborne Synthetic Aperture Radar (SAR), an active and non-optical sensor, in their study of the Central Kenyan Rift Valley fever virus vector habitats. SAR and AVHRR are very different types of sensors with very different

processing requirements. However, as discrete data types they merged well and proved integral to the effective discrimination of habitat cover types. Regardless of the remotely sensed data type, three basic steps need to be addressed: *preprocessing*, *enhancement*, and *classification* (Jensen, 1996).

Preprocessing. Preprocessing is necessary because digital imagery as collected by the sensor invariably includes artifacts of the collection process. Bad lines skew due to orbital or flight path, atmospheric haze or pollution, and variability due to solar incidence; all need to be recognized if not always analyzed. Often data vendors apply corrections; it is simply important that the researcher be aware of their existence. Radiometric corrections may be applied to remove or mask bad lines and speckle as well as to improve the fidelity of the brightness value magnitudes. Gross geometric corrections are often applied to de-skew the data in order to correct for the earth's rotation under the platform while improving the fidelity of relative spatial or absolute locational aspects of image brightness values. The ancillary data collected during the time of acquisition (ephemeris data) are often used to model the hypothetically pure form of the image. Ideally, both radiometric and geometric preprocessing steps will reduce the influence of errors or inconsistencies in image brightness values that may limit one's ability to interpret or quantitatively process and analyze digital remotely sensed images. However, these variable environmental factors should be considered true data errors only when they obscure or create confusion among image brightness signals pertaining to surface cover types and conditions (Messina, Crews-Meyer, & Valdivia, 1998). Pope et al. (1992) did not correct for atmospheric conditions, while Wood et al. (1992) used a variety of preprocessing techniques to prepare the data for

analysis. The decision to correct or not may strongly affect the researcher's ability to use techniques such as image differencing for information extraction. However, any application of preprocessing techniques should be carefully considered and the results strenuously verified, as haphazard preprocessing could introduce false confidence. The sources of data error partially depend on the sensor and mode of imaging used to capture the digital image data. While certainly not exhaustive, sensors can be grouped into four general mechanical systems, each having their own characteristic sources of error: (a) scanned aerial photography (Beck et al., 1994); (b) optical scanners (Chwastek & Dworak 1990); (c) optical linear arrays (Wood, Washino, Beck, et al. 1991); and (d) side-looking radar (Pope et al.1992).

Radiometric noise generated by remote sensing instruments can take the form of random brightness deviations from electrical sources and coherent radiation interactions or more systematic variations having spatial structure or temporal persistence (Jensen, 1996).

There are five primary reasons or objectives for applying radiometric corrections to digital remotely sensed data, four of which pertain to achieving consistency in relative image brightness and one involving absolute quantification of brightness values. Relative correspondence of image brightness magnitudes may be desired for pixels: (a) within a single image (e.g., orbit segment or image frame); (b) between images (e.g., adjacent, overlapping frames); (c) between spectral band images, and (d) between image dates.

The key here is that brightness value inconsistencies caused by the sensor and environmental noise factors listed above are balanced or "normalized" across and between

image coverage and spectral bands. The other objective is the retrieval of surface energy properties, such as spectral reflectance, albedo or surface temperature, which require absolute radiometric processing. When using airborne thematic mapper (TM) data, Wood (Wood, Washino, Beck, et al. 1991) incorporated sun angle corrections in the preprocessing to improve the potential accuracy of the results by minimizing differences between scenes. This research also used the sun illumination corrections to improve the consistency of calculations. Geometric corrections are also common prior to any image analysis. If any types of area, direction, or distance measurements are to be made using an image, that image must have been rectified for those measurements to be accurate. Geometric rectification is a process by which points in an image are registered to corresponding points on a map or another image that has already been rectified. The goal of geometric rectification is to put image elements in their proper planimetric (*x and y*) positions. Digital images collected from airborne or space-borne sensors often contain systematic and unsystematic geometric errors. Some of these errors can be corrected by using ephemeris data and known internal sensor distortion characteristics. Other errors can be corrected only by matching image coordinates of physical features recorded by the image to the geographic coordinates of the same features collected from a map or global positioning system (GPS). Geometric errors that can be corrected using sensor characteristics and ephemeris data include scan skew, mirror-scan velocity variance, panoramic distortion, platform velocity, and perspective geometry. Geometric transformations modify the spatial relationships between pixels in an image. The transformation consists of two basic operations: (a) a spatial transformation, which defines

the rearrangement of pixels on the image plane, and (b) a gray-level interpolation, which deals with the assignment of gray levels to pixels in the spatially transformed image.

The imagery used for most medical remote sensing applications having pixel coordinates (x,y) is transformed with a geometric model into image g with coordinates (x', y') . The transformation may be expressed as $x' = r(x,y)$ and $y' = s(x,y)$. If r and s are known analytically, it is possible to reverse transform the image. However, in the majority of cases, tie points are used where the location of subset of pixels from the distorted input image is precisely known on the output image (Gonzalez & Woods, 1992). All scenes are generally corrected using quadrilateral regions with the vertices of the quadrilaterals being the corresponding tie points. The procedure runs either row major or column major with identical results.

$$x' = c_1x + c_2y + c_3xy = c_4 \text{ and } y' = c_5x + c_6y + c_7xy = c_8$$

The interpolation of the digital number (DN) values is separate from the spatial interpolation. Depending upon the coefficients c_i the procedure can yield non-integer values for \hat{x} and \hat{y} . Since the distorted image g is digital, its pixel values are defined only at integer coordinates. Consequently, using non-integer values for x' and y' causes a mapping into locations for which no digital number (DN) values are defined.

The technique most commonly used for the DN value interpolation is called the nearest neighbor approach or zero order interpolation. Nearest neighbor mapping is performed by mapping the integer coordinates (x, y) into fractional coordinates (x', y') by means of the

above equations. The selection of the closest integer coordinate neighbor to (x, y) is assigned the gray level of the nearest neighbor to the pixel located at (x, y) (Gonzalez & Woods, 1992). Errors are introduced in the DN value mapping process (ER-Mapper, 1995; Jensen, 1996). The root mean square error (RMSE) calculated should always be minimized and specified. The use of specific corner tie points found within the ephemeris data permits a high level of accuracy. However, many data types are not delivered with the appropriate ephemeris data or are over areas of few man-made features.

Elimination of the disease requires advance knowledge of the village's geography. The Lack of data led to the use of remotely sensed data. ...[T]he eradication effort has used two satellite-based technologies to assist in locating settlements. First remotely sensed data from Land Satellite (LANDSAT) have allowed the identification of remote and small settlements in dracunculiasis-endemic areas. ...Using Benin national maps, air photographs and other sources, this image was georeferenced to latitude and longitude using control points common to both the map and image. ...Selecting the control points was not simple, since the lack of paved roads, locational errors on the maps and the heterogeneity of the image made point identification difficult. ...The georegistered image was classified with supervised classification using a maximum likelihood classifier. ...The second use of satellite technology in the effort was the use of hand-held global positioning system (GPS) receivers in the field. (Clarke et al. 1991, pp. 230-32)

In this example Clarke et al. (1991) used the image to ground geocorrection method. This method is the correction of digital images to ground coordinates using ground control points collected from maps or ground GPS coordinates. It is often desirable to make sure that one image is referenced with another more so than with any particular planimetric system. In this scenario, image-to-image geocorrection is specified, which involves matching the coordinate systems of two digital images, with one image acting as a reference image.

Enhancement. Common enhancements include image reduction, image magnification; transect extraction, contrast adjustments (linear and nonlinear), band ratioing, spatial filtering, Fourier transformations, principal components analysis, and texture transformations (Jensen, 1996). Many mathematical operations can be used to enhance an image. Generally, they fall into two major categories: point operations and local operations.

Point operations change the value of each individual pixel independent of all other pixels, while local operations change the value of individual pixels in the context of the values of neighboring pixels. Three primary enhancement techniques for each operation are described here: *indices*, *principal components analysis*, and *spatial filtering*.

1. *Indices* measurement of vegetative characteristics and habitat is costly in terms of both time and money and often is not a practical approach. An alternative is the measurement of vegetative characteristics from remotely sensed data that allows for a more synoptic view as well as access to remote terrain. Assessment of vegetation has primarily involved satellite Landsat thematic mapper (TM) to compose vegetation indices, such as

the normalized difference vegetation index (NDVI), leaf area index (LAI), and the amount of photosynthetically active radiation (PAR). Vegetation indices fall under the category of image enhancements, which make an image more interpretable for a particular application. Vegetation indices generally can be thought of as algorithms using spectral bands as input and yielding scores reflecting vegetation characteristics while controlling the influence of spectral variance from other features. Most vegetation indices are built upon the knowledge that healthy, green vegetation normally reflects 40-50% of the incident energy in the near infrared part of the spectrum (0.7-1.1 μm) while absorbing 80-90% of the incident energy in the visible (0.4-0.7 μm) portion of the electromagnetic spectrum (Jensen, 1996). Dry or barren soil reflects more energy in the visible portion of the spectrum than healthy or green vegetation but less in the near infrared portion of the spectrum. Both dead and senescent vegetation have the highest reflectance in the visible spectrum but usually fall between dry soil and green vegetation in reflectance of near infrared energy. Theoretically, vegetation indices rely upon the spectral separability of dead or senescent vegetation, dry or barren soil, and healthy or green vegetation in the bands used to calculate such indices.

The most simple vegetation index is that of a single band; Jensen (1996) reported that historically Landsat Multispectral Scanner (MSS) Bands 4, 5, 6, and 7 have been used to represent vegetative biomass and/or ground cover. Lillesand and Kiefer (1987) profiled the use of AVHRR channels 1 and 2 in a simple vegetation index (VI), calculated as $VI = \text{Channel 2} - \text{Channel 1}$, found to be sensitive to green vegetation. With this index, green or healthy vegetation tends to yield a high value due to high reflectance in the near-infrared and low visible reflectance; snow, water, and clouds tend to have negative VI scores due to

having higher visible than infrared reflectance. Soil and rock, reflecting evenly in near infrared and visible spectra, tend to have scores near zero. One of the problems with the simple vegetation index (VI) is that it fails to account for systematic reflectance issues prevalent across bands in a given scene: surface slope, surface aspect, and changing illumination conditions, for example. Normalizing the simple vegetation index could compensate for some of those reflectance issues and is done with such indices as normalized difference vegetation index (NDVI), computed generally as follows: (near infrared - red)/(near infrared + red).

In general, there is a high amount of redundancy in the information contained by different vegetation indices, but there are nonetheless important differences to be gleaned when deciding which index is the most appropriate for the vegetation, study area, and available sensors.

Examples of vegetation indices applied in medical remote sensing are vast. By far the most commonly used vegetation index is NDVI. Because many disease vectors involve insects (often mosquitoes and flies), NDVI is used to approximate rainy conditions yielding high insect producing areas. Brady (1991) utilized NDVI to model the habitat of the trypanosome-carrying tsetse fly and found that interannual as well as intrannual variance in NDVI was a good predictor of trypanosomiasis across time and space. Ambrosia, Linthicum, Bailey, and Sebesta (1989) used NDVI to model *Aedes* mosquito habitat in conjunction with Rift Valley fever (RVF) outbreaks. A problem with NDVI was encountered in this research: in identifying dambos (flooded, low-lying areas), sisal

plantations, coffee plantations, and riparian areas were misclassified as dambos or mosquito habitat. To correct for this overclassification, a vegetation classification was combined with the NDVI scores to accurately represent dambo areas of mosquito habitat. Davies, Kilelu, Linthicum, and Pegram (1992), also researching RVF, found that AVHRR-based NDVI scores (especially 0.43 and higher) provided strong correlations between rainy seasons, dambo habitat, and RVF virus emergence in Zambia and Kenya. Linthicum, Bailey, Davies, and Tucker (1987) used NDVI to calculate their own potential viral activity factor (PVAF) and found both to be a reliable indicator of RVF in Kenya. Follow-up work (Linthicum et al., 1991) indicated the potential for real-time applications of AVHRR NDVI scores to select areas of possible outbreak that could then be more closely monitored with Landsat TM. Other applications of NDVI include habitat modeling of the following: (a) western-malaria mosquito in California (Wood, Beck, Washino, et al., 1991); (b) the parasite-carrying brown ear tick, responsible for East Coast Fever, Corridor Fever, and January disease in cattle (Perry, Kruska, Kundert, Lessard, & Norval, 1991); (c) Lyme tick disease in humans in Wisconsin (Kitron & Kazmierczak, 1997).

Despite the success of NDVI in each of these studies, there are drawbacks to using such an index. Huh (1991) pointed out that NDVI (and VI as well) is not an accurate vegetation classifier below 20% vegetative cover and, as such, may not be appropriate for delineating patchy vegetative areas. Jackson and Huete (1991) added the cautionary note that vegetation indices generally are not uniform but can instead be calculated from sensor voltage outputs, radiance values, reflectance values, and satellite digital numbers. Each is correct, but each also yields a different vegetation index value for the same surface

conditions. Thus it is important to maintain consistency in one's own applications and to read carefully the comparative results from other research to see how exactly each vegetation index is calculated. However, the general advantage of vegetation indices remains clear: They provide information important for detection, classification, and identification of many types of origins and destinations of disease vectors, whether dambo mosquito habitats or remote villages in endemic areas. NDVI specifically provides a surrogate for rainfall patterns where precipitation data of appropriate extent are lacking (Cross, Sheffield, Perrine, & Pazzaglia, 1984; Cross, Tucker, & Hyams, 1997).

Another index commonly used in medical remote sensing applications is the moisture index (MI). The index is calculated using satellite Landsat TM imagery as $(\text{band } 4 - \text{band } 7) / (\text{band } 4 + \text{band } 7)$. In order to correct for bare soil, a perpendicular vegetation index (PVI) is used as well, as in the Hugh-Jones et al. (1992) model of African bont tick habitat in Guadeloupe. In this study higher values of both MI and PVI were associated with more ticks in heterogeneous sites. Another method of representing moisture is through the modeling of temperature. Malone et al. (1994) created temperature difference (diurnal day-night surface temperature difference) images of Egypt that were found to significantly correlate with moisture areas associated with schistosomiasis risk in the Nile River area. The significance of this finding lies in the fact that moisture potential is often modeled using data from sensors such as Landsat TM that have low but fairly high spatial and spectral resolution (Hugh-Jones et al., 1992). This work utilized AVHRR data which has lower spectral and spatial resolution than TM data but which has much greater temporal

resolution as well as greater spatial extent. As such, diurnal (versus seasonal) variability in moisture as well as moisture variations across very large areas can be reliably assessed.

2. *Principal components analysis (PCA)* also falls under the category of image enhancements and is performed both to add to the interpretability of an image for a particular application and to streamline the redundancy present in multispectral data. PCA is commonly used as a means of compressing data and allows redundant data to be compressed into fewer bands that by definition are independent and noncorrelated. This compression is referred to as a reduction in dimensionality of the data, indicating a decrease in the number of bands of data required to yield applicable results (Jensen, 1996).

In performing PCA, the first step involves assessing the correlation of two or more bands of imagery for a particular scene. With Landsat thematic mapper (TM) data, the three visible bands (TM1, TM2, and TM3) are often highly correlated with each other, as in examples in Jensen (1996) and Ambrosia et al. (1989). Ambrosia et al. found in their work on modeling the *Aedes* Rift Valley fever (RVF) mosquito vector that, since TM data contains seven bands of information, which are intercorrelated to some degree, a reduction of total scene/band variance was needed to reduce processing time required for clustering. In order to reduce the number of possible redundant data bands while still retaining greatest scene variance, they performed PCA on TM1 (0.45-0.52 μm), TM2 (0.52-0.60 μm), and TM3 (0.63-0.69 μm) and then used the first PC in their clustering and class identification. Thus, three bands of visible spectrum information became one, and seven total bands became five: PC1, TM4, TM5, TM6, and TM7. Interestingly, Ambrosia et al. mentioned

that they would try this process with Satellite Pour l'Observation de la Terre (SPOT) data that may be more useful due to finer element resolution. Unfortunately, the increase in spatial resolution comes with a direct loss in spectral resolution, and it is unknown if the SPOT data will contain enough information to spectrally separate vegetation communities for *Aedes* habitat research because of waveband locations and lack of middle infrared bands (analogous to TM5 and TM7).

The second step after choosing correlated spectral bands involves the transformation of those bands in multispectral feature space, resulting in an uncorrelated multispectral dataset whose variance is ordered. Assume two highly correlated bands of pixel (brightness) values were used; due to the redundancy in those bands, the brightness values are too clustered (not easily separable) on a simple graph of Band 1 versus Band 2 brightness values. PCA transforms or rotates the axes of the original data such that variance is maximized and the axes remain perpendicular to each other. The first axis, also the first principal component (PC), is assumed to contain overall scene luminance (the global trend or first-order effect), whereas the other PCs represent intra-scene variance (the local trend or second-order effect). Where the goal of the transformation is data reduction, generally PC1 alone will be used, provided it accounts for some minimal threshold of total variance (usually above 90%), which is represented by a calculation from the eigenvalues and eigenvectors of the covariance matrix and computed as $\% \text{ variance explained by PC1} = (\text{eigenvalue PC1} * 100) / \text{sum (eigenvalues)}$.

The third step of PCA involves interpreting what these transformed axes mean for the new data (principal components). A matrix is created displaying the original bands used against the principal components and shows the factor loadings for each combination of original band and principal component. These loadings represent the correlation between original bands and principal components; for example, a principal component with loadings of 0.93, 0.84, 0.92, and 0.27 in Landsat TM Bands 1, 2, 3, and 4, respectively, would indicate a strong visible spectrum component but a weak near infrared component. At this point decisions are made regarding how many and which principal components to use in the next phase of analysis. If seven components are created, but the first three account for 95% of the variance in the data, processing time and space needs can be drastically reduced by using the three principal components rather than the seven original bands of information. In their work researching African bont tick habitat in Guadeloupe, Hugh-Jones et al. (1992) used PCA to separate habitat areas. They found that, unlike some insect vectors, the bont tick occupies several distinct habitat types, including dry meadows, pond, rocky grasslands, and dry scrub. Once imagery was masked to include areas identified with tick habitat, they successfully employed PCA to separate and later identify the various "distinct" habitat types.

3. *Spatial filtering* represents a series of mathematical techniques used to visually enhance the data. These spatial enhancements serve to improve the spatially dependent interpretability of the data. Spatial enhancement techniques focus on the concept of spatial frequency within an image and define homogeneous regions. These homogenous regions are identified by first locating edges. Similar edge effects are used to identify or classify

common landscape features. In terms of the letters on this page, for example, one identifies the outlines of the letters for recognition with respect to the spaces between the lines. It is the combination of both that permits character recognition. Spatial frequency broadly defines the relationship between edges and regions. For example, water bodies with slowly varying changes in their digital number (DN) values have low spatial frequency, while urban areas, which vary radically among adjacent pixels, exhibit high spatial frequency. The specific techniques used in spatial filtering are not often used in medical geography; this absence of use is due not to their ineffectiveness, but in part to their complexity and some real difficulty in quantitatively evaluating the output.

Classification. Unlike analog image processing, which uses all of the elements listed in Table 1, digital image processing primarily relies on the various recorded radiance characteristics within the elements of each individual image pixel. The radiance values are most easily ordered and classified with respect to each other in a nonspatial framework. As software programs develop, the ability to recognize spatial variation and account for statistical violations in sample distributions is improved. Consequently, information extraction is an important dynamic area of research. It is incumbent upon the analyst to stay abreast of the rapidly changing state of the art in systems.

Classification employs spatial pattern recognition to create thematic layers from image data; groups of pixels are sorted into classes based on their brightness values in one or more bands. Classes may be comprised of known feature sets (such as a land cover map with soil, water, and vegetation classes) or may consist of unknown feature sets that are

nonetheless distinguishable to the computer. Pattern recognition generally refers to enhancing data in order to discern patterns; while this process may be performed visually, currently spectral statistics are used to sort the image information and define and find patterns in that imagery. Pattern recognition consists of two phases: training and classification. Training involves defining the criteria or creating the set of decision rules used in the detection of pattern. Two types of training are used: *supervised* and *unsupervised*; sometimes a hybrid approach between the two methods is referred to as *training*.

1. *Supervised training* is characterized by greater analyst control. Supervised training works best in situations where the analyst is able to identify features in the imagery representative of each target class. These features may be identified on the image itself or with the use of data from other sources, such as aerial photography, maps, or various ground truth data. This approach requires a level of familiarity with both the landscape under study and the desired classes. By selecting pixels or an area of interest (AOI) representative of each class type, the analyst trains the computer to identify other pixels of similar characteristics. Thus the accuracy of the classification is highly dependent upon the quality and accuracy of training data. Hayes, Maxwell, Mitchell, and Woodzick (1985) used supervised, statistically based (parametric) classification to aid in identifying mosquito habitat in Nebraska and South Dakota in a low-cost identification and inventory of larger freshwater plant communities and wetlands.

2. *Unsupervised training*, on the other hand, requires much less analyst input and is more automatic. For this reason, it is sometimes referred to as *clustering*, as it is based on the natural groups of pixels as they fall within feature space. The computer relies upon spectral statistics to discern patterns in the spectral information, although, to some extent, the analyst does stipulate parameters for the statistical decision rules (e.g., specifying within-or-between class variance thresholds). These groups may still be merged, deleted, or manipulated at a later time as easily as with supervised classification. The patterns detected by the computer may or may not correspond to classes meaningful for the landscape under study but are instead groups of pixels that are more similar spectrally to each other than they are to all other pixels. While supervised training is heavily dependent upon the analyst, unsupervised training is dependent upon the data. Identification of thematic groups occurs after classification, and the usefulness of the classification is directly dependent upon the ability of the analyst to appropriately interpret and identify the classes. However, unsupervised training is very useful when less is known about the data prior to the analysis. For example, Hugh-Jones et al. (1992) used unsupervised classification (combined with Principal Component Analysis [PCA] and vegetation indices [VI]) to identify several distinct but previously unknown habitat types for the African bont tick in Guadeloupe. Clustering generally is a nonspatial statistical process whereby all or many of the pixels in the input data are used, regardless of their spatial proximity to one another. Two primary types of clustering are used: image density or iterative self-organizing data analysis (ISODATA) techniques (e.g., clustering algorithm) and Red-Green-Blue (RGB). The ISODATA clustering method uses spectral distance sequentially, iteratively classifying

pixels, redefining class rules, and classifying again so that eventually spectral distance patterns in the data are discernible. The RGB clustering method applies only to three bands, 8-bit data, and uses three-dimensional feature space by dividing that space into sections that then define the clusters. Note that both clustering methods are based on spectral distance and not spatial contiguity. ISODATA clustering can be quite slow due to the multiple iterations and does not take into account spatial homogeneity (spatial autocorrelation). However, with enough iteration this process will not be biased, based on the default starts point of clustering. With RGB clustering, the three input band limitation can be bothersome for some applications, but it is a fast processing method and is not biased based on the order of pixels analyzed. RGB clustering does not produce a signature set needed for comprehensive change-detection.

3. *Training*, whether supervised or unsupervised, results in a set of spectral signatures that define each class. Parametric signatures, based on statistical parameters (covariance matrix and mean), can be generated by either supervised or unsupervised training and can then be used to train a statistical classifier. Nonparametric signatures are created only by supervised training and are based upon discrete objects (polygons); that is, pixels are assigned to a class based on whether or not they belong to the feature space as defined by the analyst.

There are several advantages and disadvantages to each approach. Parametric signatures assume normality and are slower than nonparametric, but are able to classify each pixel into one and only one class and are often more easily interpretable than feature

space. Nonparametric signatures are useful for data that are distributed non-normally and also require less processing time; while they allow for overlap or unclassified pixels, the feature space extraction method (if used) can be difficult to interpret.

Classification is often performed using a specific classification scheme, where a set of target classes is known a priori. Classification schemes should contain classes that are either integral to the research or easily discernible in the data themselves. Classification schemes for various ecoregions have been popularized by researchers and provide a springboard for analysis when studies are conducted in areas similar to the one under which the classification scheme was built. Supervised classification is appropriate when the analyst needs only to identify relatively few classes, when the training sites have ground truth data available, or when clearly delineated and homogeneous regions can be identified to represent each class.

Unsupervised classification is best used when a large number of classes need to be identified, and it is particularly useful when classes do not fall in spatially contiguous areas that are easily discernible. It is not uncommon to generate a large set of classes using unsupervised classification and then to merge or reduce those classes via supervised classification (a hybrid approach). Additionally, the basic principles of supervised and unsupervised classification can be applied in sequence, through multiple iterations, or in some combination of order or area. For example, Beck, Wood, and Dister (1995) used a hybrid approach in Mexico to discover and then identify areas associated with malaria (primarily transitional swamp and unmanaged pasture). Lastly, just as there are parametric

and nonparametric signatures, there are also parametric and nonparametric decision rules. Decision rules are arguably the most important component of classification and involve the algorithm used in comparing pixel measurement vectors to signature sets. Pixels meeting a specified criterion are then assigned to that class. The predominant types of nonparametric decision rules are parallelepiped and feature spaced. Pixels unclassified after being put through the nonparametric decision rule are then either classified using a parametric rule or left unclassified. Pixels falling into more than one feature space object (overlap) can be decided by a parametric rule, by order, or left unclassified. The parametric rules commonly used include minimum distance and maximum likelihood.

In the parallelepiped decision rule, the pixel's values are compared to upper and lower limits of either each band in the dataset, a set number of standard deviations within the mean of each band, or customized user limits.

Parallelepiped is fast, does not assume normal distributions, and can also be used as an intermediary way of weeding out possible classes for a pixel before going to a more intensive processing procedure rule. However, parallelepipeds are rectangular in feature space, and therefore pixels in the corner of that space may be improperly classified due to its inclusion despite its distance from the center. Feature space, however, has elliptical-shaped classes and therefore does not tend to misclassify "corner" observations. This decision rule does allow overlap and unclassified pixels and may be difficult to interpret. Feature space can also be used on non-normally distributed data and is quite fast.

The minimum distance rule, a parametric decision rule, is also referred to as spectral distance and calculates the spectral distance between the measurement vector of a pixel and the mean measurement vector of a signature. There are no unclassified pixels, since every pixel is closer to one sample mean or another, and this rule is the next-to-fastest process (after parallelepiped). However, sometimes overclassification does occur when pixels that should not be classified into any group (spectral outliers) are classified anyway. Furthermore, this approach does not consider class variability, meaning that homogenous classes (like water) may be overclassified (errors of commission), whereas heterogeneous classes (like urban) will be underclassified (errors of omission).

Maximum likelihood (sometimes referred to as Bayesian) decision rule is based on the probability that a pixel belongs to a certain class. In its most simple form, this rule assumes that these probabilities are equal for all classes and that input data are normally distributed. This popular decision rule is one of the most accurate classifiers, but only if the data are normally distributed. Class variability is taken into account via the covariance matrix, but the required processing time is therefore quite long. Additionally, this method tends to overclassify when there is a large dispersion of the pixels in a cluster or training sample. Clarke et al. (1991) successfully used maximum likelihood in their work on the Guinea worm (*dracunculiasis*) in Benin and Nigeria. The use of supervised, maximum likelihood classification greatly aided in identifying and locating remote and small settlements in dracunculiasis-endemic areas for intervention with a settlement classification accuracy of 90%.

Another method of classifying the landscape involves using statistical analysis. For example, Wood, Beck, Washino, et al. (1991) used discriminate analysis to test their ability to classify the landscape into areas of high and low tick larval counts using spectral reflectance and distance to field measurements. Other researchers (Wood et al. 1992) have used regression analysis and multivariate statistical models to evaluate mosquito habitat. Regression analysis explains a dependent variable in terms of independent variables. Discriminate analysis, however, creates a linear combination of independent variables based on finding the greatest difference between classes of the dependent variables that were defined before data collection (as opposed to cluster analysis where classes are created based upon existing data trends) (Hugh-Jones et al., 1992). Whereas regression analysis assumes normality of the independent variable, discriminate analysis requires multivariate normality of the independent variables. While regression analysis can utilize categorical or interval data as independent variables, a categorical variable should not be used as an independent variable in discriminate analysis. Categorical data should not be used as independent variables in discriminate analysis because their non-normal distribution (often associated with noncontiguous data generally) causes model parameters estimates to be very positively biased, and a maximum likelihood estimator or logic should be used instead (Kleinbaum, Kipper, Muller & Nizam, 1998).

The other fairly serious data problem encountered in the use of statistics has to do with the spatially dependent nature of remotely sensed data. Landscape features tend to be related to those features closest to them; this phenomenon is referred to as spatial autocorrelation. For example, given a pixel with prairie land cover, the next closest pixel is

more likely to be prairie than some other type of land cover. Thus, spatial data are somewhat dependent upon each other, in a statistical sense. This lack of independence violates one of the most basic assumptions of regression analysis, which is that data are independent. A lack of independence causes an overestimation in the confidence of the results. There are statistical tests and corrective measures for spatial autocorrelation.

Remote Sensing Data Types

Aerial Videography

The potential of remote sensing for ecologically related applications was recognized as early as 1937, when Dalke (1937) reported the use of aerial photography for habitat mapping. Both photography and videography allow an investigator to see a site's geographic and cultural features in relation to other environmental measurements. Aerial photographs and videography provide information needed for updating maps of surface features and data bases, including the derivation of detailed digital terrain elevation data.

Aerial Video is a form of remote sensing where video is collected while flying transects over the area of interest. Aerial video has several benefits (Virginia, 2001): (a) covers a large area quickly and efficiently; (b) a large number of sampling locations can be collected; (c) unique or specially-managed areas (i.e., wildlife management areas, military bases) can be heavily sampled; (d) flights can be planned and completed quickly (weather permitting); (e) interpretation is done by a few highly-trained individuals, rather than many

field technicians; (f) flexibility of interpretation schedule, which occurs at a workstation; and (g) low cost of interpretation.

Digital Orthophoto Quadrangle (DOQ) Aerial Photography

Digital orthophoto quadrangle (DOQ) aerial photography combines the image characteristics of a photograph with the geometric qualities of a map and can be used in numerous geographic information system (GIS) applications, either alone or in combination with other digital data, such as digital line graphs (DLG) or digital raster graphics (DRG) (U.S. Geological Survey [USGS] 1996).

Geological Survey (USGS) is the lead federal agency for the collection and distribution of digital cartographic data. It began to produce DOQs in 1991 and currently has nearly 50,000 available for distribution. Complete DOQ coverage of the conterminous United States under this program is expected by the year 2004. Thereafter, the DOQs will be updated on a 10-year cycle for most areas and on a 5-year cycle in areas where land use change is more rapid.

The U.S. Department of Agriculture's Farm Service Agency (FSA), the Natural Resources Conservation Service (NRCS), the U.S. Forest Service (USFS), and the USGS are partners in the National Digital Orthophoto Program (NDOP). The USGS Earth Science Information Center (ESIC) distributes digital cartographic and geographic data produced through the USGS National Mapping Program (NMP).

Orthophotos combine the image characteristics of a photograph with the geometric qualities of a map. They serve a variety of purposes, from interim maps to field references for earth science investigations and analyses. The digital orthophoto is useful as a layer of a geographic information system (GIS) and as a tool for revision of digital line graphs and topographic maps. Unlike a standard aerial photograph, relief displacement in orthophotos has been removed so that ground features are displayed in their true ground position. This allows for the direct measurement of distance, areas, angles, and positions. Also, an orthophoto displays features that may be omitted or generalized on maps (USGS 1996).

Landsat

The processed product of Landsat satellite imagery is often contained in some type of thematic map. This in itself may be adequate, but such information may just be one of a series of data sources to be integrated, compared, and synergistically manipulated to arrive at a result or interpretation (Short, 1982). These analyses can be accomplished with Geographic Information Systems (GIS).

Geographic Information Systems are computer-based systems designed to facilitate the manipulation and analysis of spatial data. They were developed to assemble and analyze diverse data pertaining to specific geographic areas with Landsat satellite. Prior to their development, the most common medium for storing and analyzing such information was the basic analog map. The technology involved in the creation and display of such analog devices has reached a considerable level of sophistication but has never overcome some important obstacles. Even with the use of measurement tools such as scales and

planimeters, analyzing large amounts of data from large numbers of maps is a difficult and slow process (Marble et al., 1983). Garrison, Alexander, Bailey, and Marble (1965) first described the potential for integrating remotely sensed data with other nominal or attribute data (e.g., soil type, rainfall, species distribution). Since then, GIS Landsat satellite image technology has continued to develop, and there are now many such systems. Tomlinson (1984) reported that more than 1,000 GIS Landsat image and automatic cartography systems were in operation in North America. They generally contain the following four major capabilities: data encoding, data management, data manipulation, and data output. Data encoding or input is the component in a GIS Landsat image system that collects spatial data derived from existing maps or sensors. These data can be encoded with either manual (digitizing or keyboard) or automatic (scanning or digital data bases) techniques at a variety of scales. The second component of a GIS Landsat image system is data storage and retrieval. This subsystem organizes and stores data in a way to permit quick retrieval and rapid update. Each variable (termed *layer*) is stored digitally as a geographically referenced layer. When digital layers are geographically registered to one another, they form a data set of theoretically unlimited number of layers. These can then be queried according to user specifications.

The third component of a GIS Landsat image system, data manipulation, is the process used to extract information from databases and perform a variety of tasks. These include estimations of perimeters, areal calculations, search radius, distance calculations, and comparisons and evaluations of multiple data layers. The data output component of a GIS Landsat image system is the subsystem capable of displaying all or part of the original data,

as well as manipulated data and output in tabular or map forms. Output may be a hard-copy map or a listing of statistics scaled to any user-defined map dimensions (Jensen, 1986).

Definition of Terms

Aerial Photography: Photography produced in an airborne platform.

Aerial Videography: Videography produced in an airborne platform

Change Detection: Using a multivariate statistical procedure, a numerical value is assigned to each pixel of each image indicating the "severity" of "change" between the two points in time as evidenced in the data. These "severity" values range from 0 to 100; the higher the value, the more pronounced the change. For viewing, these are converted into scene images, one for each of several levels of a minimum severity. With this, insight can be attained in regard to severities and their geographical location over a particular area.

Digital Orthophoto Quadrangle (DOQ): a computer-generated image of an aerial photograph in which displacements caused by camera orientation and terrain has been removed.

Reflectance: is the fraction of the total radiant flux incident upon a surface that is reflected and that varies according to the wavelength distribution of the incident radiation.

Registration: the process of geometrically aligning two or more sets of image data such that resolution cells for a single ground area that can be digitally or visually superimposed on a map coordinate system may not be involved.

Remote Sensing: the science of obtaining information on the properties of an object or phenomenon through analysis of data acquired by a sensor not in physical contact with the object or phenomenon.

Resolution: the minimum distance between two adjacent features or the minimum size of a feature that can be detected by a remote sensing system.

Thematic Mapper (TM): a nonphotographic imaging system that utilizes an oscillating mirror and seven arrays of detectors that sense electromagnetic radiation in seven different bands. (The thematic mapper sensor is a derivative of the multispectral scanner [MSS] generation of scanners, achieving greater ground resolution, spectral separation, geometric fidelity, and radiometric accuracy.)

Research Question

A change-detection method that is accurate in determining change to the landscape over time can be critical in assisting an analyst in identifying the habitat of an infectious disease carrier. A reliable method will impact the health of a community by saving time and predict the possible spread of a land-based disease. Such a determination of the changes to surrounding landscape can be made for the purpose of epidemic disease prevention and/or the pre-inoculation of a human population. The research question emerging from these concerns and the availability of the more timely aerial videography technology is as follows: Will the method of analyzing change in remotely sensed data from multiple image technologies such as digital orthophoto quadrangle (DOQ) aerial photographs and aerial

videography and Landsat satellite detect changes in landscape between a time when the disease is not present and a time when the disease was present?

Hypothesis

The comparison of a remotely sensed frame from multiple image technology such as digital orthophoto quadrangle (DOQ) aerial photography and aerial videography and Landsat satellite prior to the disease outbreak with a multiple image technology frame after the outbreak of a disease will show a severity of changeⁱ in landscape.

Assumption and Limitation

For this study, the following assumption is made: The time expended between the two frames may show change unrelated to the procedure used. For this study the following limitation can occur: The Digital Orthophoto Quadrangle (DOQ) aerial photographs produced by governmental agencies could be up to three years between the time the image was made and the occurrence of the disease.

Chapter 2 presents a review of the relevant literature. Chapter 3 explains the proposed research methodology. Chapter 4 will discuss how the data was collected and analyzed. Chapter 5 will review both the findings and conclusions of the research.

ⁱ I.e. – If landscape index has changed, change will be measured in pixel values.

CHAPTER 2

REVIEW OF THE LITERATURE

For over 2 decades remote sensing technologies have proved to be valuable tools for describing the earth's landscape. Since the launch of the Earth Resources Technology Satellite (ERTS) in 1972, aggressive governmental programs have made attempts to integrate remote sensing technologies into the areas of forestry, agriculture, geology, and public health. Within a decade of that initial launch, operational programs had been developed in all areas—except public health. In large part, this was because priorities within the public health community were focused on vaccine development, therapies, and the continuation of traditional ground-based surveillance and control methods. Within 2 decades of the launch, however, health agencies began to reevaluate other approaches in the face of worsening health conditions around the world.

Simultaneous with this change have been significant advances in computer processing, improvements in data acquisition (i.e., additional remote sensing [RS] sensors with higher spatial, spectral, and temporal resolution), reduced hardware/software costs, and the development of computer-based remote sensing technology. All these factors have led parts of the public health community to question their earlier reservations regarding the utility of aerospace-based technologies.

Aerial Videography

Conventional aerial photographs have long been the mainstay for monitoring and inventorying the extent and change of the landscape. They are valuable because they can capture current land use and waste disposal practices. These images can later provide important clues to past cultural activities affecting the environment and supplement written records of construction and waste disposal. Aerial photographs provide a stable, scalable rendition of the surficial conditions of the earth. However, their cost and lack of digital format make them less attractive for integration with (the) GIS than digital forms of remote sensing (Petris, 2001). GPS-controlled aerial videography provides a geometrically stable, remotely sensed medium which is readily computer compatible and one that lends itself to georeferencing in a cost-effective and timely manner (Petris, 2001).

Aerial Videography Mapping–Wetland

Aerial Video imagery is proven effective in mapping vegetation with fine spatial detail in a riverine wetland (Sersland, 1994). A global positioning system (GPS) is used on the aircraft to compensate for lack of anthropogenic features within the wetland and satisfies the need for accurate georeferencing. Altitudes during this type of study can vary. However, if the individual frames are georeferenced, rectified, and mosaicked, it is successful. A classification routine is implemented in this type study. Using a single image classification can become inadequate since each requires a composite of two dates in order to improve spectral resolution. Accuracy assessment methods can compare classified point data with the reference data. This method is 60% accurate, with moderate success in

mapping plant communities. However, the most notable find in Sersland's (1994) study is that the use of GPS was essential in acquiring and analyzing the data. In addition, color video imagery loses its real-time advantage when attempting multitemporal classifications but gains substantial information content. Sesland concluded that multitemporal methods improve classification accuracy in the application of color composite airborne videography. Moreover, reliable classifications would not have been possible with single date data.

Aerial Videography Mapping - Soil Sampling

Airborne color infrared digital videography is a proven method of classifying fields without costly yield monitoring and grid soil sampling (Yang, Anderson, Everitt, 1995). Researchers feel that the new methods—also known as precision farming, variable-rate technology, and prescription farming—owe their rapid evolution to developments in the global positioning system (GPS), geographic information systems (GIS), and remotely sensed imagery (United States Department of Agriculture, 1995). The United States Department of Agriculture (USDA) state that the technology enables growers to improve crop production efficiency and reduce environmental pollution by adjusting the amounts of seed, fertilizer, and pesticide applied to meet the specific needs of discrete areas of a field. The USDA believe that a farmer can develop a site-specific management plan, although farmers need georeferenced information about the physical properties of their soil, as well as plan production and crop yield data. The USDA researchers concluded that, with this knowledge, growers could identify and distinguish relatively uniform regions (or zones) within a field and write a “prescription” for how each zone should be managed. The

USDA researchers' study site was located on Rio Farms in Monte Alto, Texas, using a multispectral digital video imaging system (green, red, and near infrared (NIR) of the spectrum) to acquire data about the field on different dates during the 1995-1996 growing seasons. They captured the imagery at altitudes ranging from 1,300 to 1,400 meters under sunny conditions. They registered, or georeferenced, the images with 14 ground-control points for the field, averaging the fixes to obtain coordinates for that location. After the second growing season, the researchers overlaid the sampling points on classification polygon maps along with the video images for spatial analysis. They extracted the video bands and calculated the normalized difference vegetation indices. Using GIS data, they then visualized and queried the data, generating statistics for each classification zone. A visual examination of the digital video images for the field indicated that there were three or four natural clusters. The study demonstrated that aerial digital videography in conjunction with ground sampling, global positioning system (GPS), geographic information system (GIS), and image-processing technologies can be a useful tool for monitoring within-field plant growth variation and establishing management zones for precision farming.

Aerial Videography Mapping - Data Collection

A recent study (Virginia, 2000) is of particular interest because it incorporated all the aspects of how video is collected in a research setting. It answered many question such as the various uses of aerial video, the reasons to use it, the video interpretation it is based on, the collection technique, the interpretation of the data, and the provision of many resources

for other studies currently underway and training that is being offered in the field. One of interest is an event titled Vegetation Survey Using Air Video Technology held in Lafayette, Louisiana. The study describes how, in 1994, the Fish and Wildlife Information (FWI) exchange began work on the federally funded Virginia Geographic Application Program (GAP) Analysis Project. The goal was to determine the biodiversity of the Commonwealth of Virginia and to help set conservation priorities. They decided to use aerial video for three reasons: (a) creation of models to predict land cover, (b) signature generation, (c) accuracy assessment of the Virginia Land Cover Map Layer.

The decision was made to use aerial video because the researchers rationalized that only two basic choices exist: field work (or ground truthing) and remote sensing. Remote sensing, they decided, is less expensive and faster because personnel are not required to sample land cover on the ground. Therefore, aerial video is a form of remote sensing in which the collection of video is made while flying transects over the area of interest. The GAP researchers rationalized that it had several benefits: (a) coverage of a large area quickly and efficiently; (b) collection of a large number of sampling locations; (c) a sampling of unique or specially managed areas (i.e., wildlife management areas, military bases); (d) flights that can be planned and completed quickly (weather-permitting); (e) interpretation by a few highly trained individuals, rather than many field technicians; (f) a schedule flexibility for interpretation; (g) low cost of interpretation.

The GAP researchers detailed how the video is collected and what the interpretations are based on. Many examples of other applications that use aerial video and references are included in this Internet Website.

Aerial Videography Mapping - Low Technology

A collaborated research effort (Garner & Grant, 1997) between the Center for Advanced Spatial Technologies (CAST) and the Arkansas Forestry Commission (AFC) was conducted to evaluate a low-tech, low-cost remote sensing platform known as Airborne Videography. The project described how it successfully uses airborne videography with the integration of global positioning systems (GPS) and video cameras to receive geographic information frame by frame. This unique project utilized a system that captured digital video frames overlaying the actual GPS coordinates directly upon each video frame. The video frames were analyzed to determine if a change in land mass had occurred. The State of Arkansas Forestry Commission uses the system to discriminate between active Southern pine beetle (SPB) sites and older inactive SPB sites. Another application was effective in mapping timber damage from an outbreak of tornadoes.

Aerial Videography Mapping–Inexpensive

A study conducted by Graham (1995) established that Airborne-Video is an inexpensive multistage imagery that provides geographically unbiased ground-truthing validation data for classification and accurate interpretation and validation of vegetation maps from satellite thematic mapper (TM) imagery. The goal for Graham's study included

four stages: (a) to refine GPS referenced airborne video imagery in Landsat Thematic Mapper Images for classification of land cover mapping and change assessment; (b) to develop and/or enhance the efficiency, cost effectiveness, utilize systematic and statistically valid ground-truthing, and validate methods using GPS referenced airborne video; (c) to provide enhancement of seasonal precipitation surface models to assist in land cover mapping, particularly in arid land systems; and (d) to develop a geostatistical vegetation/land cover mapping method that uses topographic, climatic, and GPS-referenced airborne video sample points, but with no multispectral satellite imagery or continuous coverage aerial photography.

Graham referenced an earlier state of Arizona analysis project and describes its successful use of GPS-reference airborne videography. The project rapidly acquired inexpensive multistage imagery from providing geographically unbiased ground-truthing and validation data for the classification, interpretation, and validation of vegetation maps from TM imagery.

Graham noted that development of a statistically sound strategy is possible for collecting and analyzing airborne video transect data for land cover mapping. She felt that the ability to survey and/or map large geographic regions without having to acquire expensive and often-unavailable satellite imagery or continuous aerial photographic coverage is of particular importance in the days of shrinking resource management dollars.

Digital Orthophoto Quadrangle (DOQ) Aerial Photography

Spatial Resolution of DOQ

The resolution is generally larger than the computed ground sample distance of the DOQ. The ground sample distance of the digital orthophoto is a result of the scanning aperture of the microdensitometer used to capture the digital image and the resampling algorithm. For example, if a scanning aperture of 25 micrometers is used on a 1:40,000 photo-scale images, the ground (pixel) sample distance is approximately 1 meter. A 7.5-micrometer scan yields a pixel size of 0.3 meters, while a 15-micrometer scan equates to a 0.6 meter. For the processed DOQ, the ground sampling distance (GSD) is 1 meter for quarter-quad digital orthophoto and 2 meters for quadrangle digital orthophotos (USGS, 1996). Digital orthophotos are produced at a finer sampling distance than 1 or 2 meters and processed by resampling to 1 or 2 meter horizontal ground sampling distance (GSD). Digital orthophotos produced at a coarser sample distances are not resampled to a finer horizontal ground sample distance (USGS, 1996).

The geographic extent of the digital orthophoto is equivalent to an orthophoto quarter-quadrangle (3.75-min of latitude and longitude), plus a minimum of 50 meters to a maximum of 300 meters of over edge is included (USGS, 1996), sufficient to offer coverage to encompass the four primary and secondary horizontal datum corner points. The over edge is useful for edge matching and mosaicking of quadrangles by offering areas outside the primary area of interest, which facilitate tonal matching between images. Every

orthophoto is a rectangle, but may not necessarily be the same size as its adjoining neighbor. The normal orientations of the data are by lines (rows) and samples (columns).

Spectral Range of DOQ

In order to assure that the image brightness values of the orthophoto closely portray the source imagery, very little image enhancement, other than a limited amount of analog dodging, is performed when preparing the photograph for scanning. Some deviation of brightness values may also occur during the scanning and rectification processes. Radiometric accuracy and quality are verified through visually inspecting and comparing the digital orthophoto to the original unrectified image.

Data Organization of DOQ

A gray-scale U.S. Geological Survey (USGS) digital orthophoto has several characteristics (USGS, 1996):

1. The data consist of an ASCII header, followed by 8-bit, binary image data.
2. The radiometric images store brightness data, which include 256 gray levels and represented as integers in the range of 0-255.
3. The ground sample distance of the 3.75-minute quarter-quadrangle is 1 meter.
4. The geographic extent of the digital orthophoto is equivalent to an orthophoto quarter-quadrangle (3.75-minutes of latitude and longitude). A minimum of 50 meters to a

maximum of 300 meters of over edge is included, sufficient to offer coverage to encompass the four primary and secondary horizontal datum corner points.

5. Standard digital orthophotos, 3.75 - minute coverage, are cast on the Universal Transverse Mercator (UTM) projection on the North American Datum of 1983 (NAD83) with coordinates in meters.

6. The ordering of the data is by lines (rows) and samples (columns), with each line containing a series of pixels ordered from west to east. The order of the lines is from north to south. When displayed on a computer, the image projection grid north is at the top.

7. The four primary datum (NAD83) corners are imprinted into the image as four solid white crosses with an image value of 255 and the four secondary datum corners as four dashed white crosses with similar intensity values.

Accuracy of DOQ

Digital orthophoto quadrangles and quarter-quadrangles must meet horizontal National Map Accuracy Standards (NMAS) at 1:24,000 scale and 1:12,000 scale, respectively (USGS, 1996). The NMAS specify that 90 % of the well-defined points tested must fall within 40 feet (1/50 in.) at 1:24,000 scale and 33.3 feet (1/30 in.) at 1:12,000 scale (USGS, 1996). The vertical accuracy of the source, the digital elevation model (DEM), must be equivalent to or better than a level-1 DEM, with a root-mean-square-error (RMSE) of no greater than 7 meters (USGS, 1989). The DOQ RMSE is the square root of the average of

the squared discrepancies. These discrepancies are the differences in coordinate (x and y) values derived by comparing the data being tested with values determined during aerotriangulation or by an independent survey of higher accuracy. All remaining inputs and processes (e.g., aerotriangulation control and methodology, scanner and sensor calibrations) used in digital orthophoto production must be sufficiently accurate to ensure that the final product meets NMAS specifications (USGS, 1989).

The National Aerial Photography Program (NAPP) imagery and NAPP-like photography are the primary sources of aerial photography used in the production of 1-meter digital orthophotos for the National Digital Orthophoto Program (NDOP). NAPP photography is quarter-quadrangle centered (3.75-min. of latitude by 3.75-min. of longitude in geographic extent) and taken at an aircraft altitude of approximately 20,000 feet above mean terrain using a 152-millimeter focal-length camera (USGS, 1996). The scale of the NAPP photography is approximately 1:40,000 (USGS, 1996). Orthophoto quadrangles may also be produced through the mosaicking of digital orthophoto quarter-quadrangles. Color infrared (CIR) photography may be used as a source. However, the resulting DOQ may either be a single black-and-white composite of all bands or a color DOQ with all three bands. Although NAPP is the primary image source, this does not prevent the use of additional aerial photographs or digital images in the future (USGS, 1996).

Landsat

Since the development of GIS Landsat image systems, potential for ecological research in a spatial context has improved tremendously. An GIS Landsat image system offers cost-effective techniques for addressing ecological planning, modeling, evaluation, and research efforts. For example, a GIS Landsat system approach has been used for the development of a program to model human intrusion into grizzly bear (*Ursus arctos*) habitat in Glacier National Park, Montana. This study successfully integrated locational data of human access areas, feeding preferences of the bears (which influence distribution of the bears), digital terrain information, and bear sightings (Martinka & Kendall, 1985). In another GIS Landsat system study, Ornsby and Lunetta (1987) identified food availability for whitetail deer (*Odocoileus virginianus*) in California, using thematic mapper (TM) scanner data and a GIS Landsat system. The GIS Landsat system was used to delineate areas of escape cover and food values on the classified TM image; habitat suitability was then assessed from these variables. A GIS Landsat system has also been used to link vegetative cover information with point coverage of radio telemetry locations to assess the preference for old growth vegetation by spotted owls (*Strix occidentalis*) in Washington (Young, Eby, Allen, Hewitt, & Dixon, 1988).

Wildlife and Resource Management Using A GIS Landsat Satellite System

More recently, satellite imagery has been utilized as an important component of many wildlife and resource management studies. This section focuses on the most relevant of these studies. Much of the pioneer work for such applications has focused on the use of

imagery for wetland applications (Cowardin & Myers, 1974; Work, Gilmer, & Klett, 1976; Work & Gilmer, 1976), especially with respect to the application and evaluation of habitat for waterfowl. With remote sensing, researchers have been able to map and assess the number of ponds in a given area. From this they can make evaluations about the quality of breeding habitat based on presence, abundance, and spatial relationship of various land cover classes.

Specifically, satellite imagery and meteorological satellite data were used by Kerbes and Moore (1975) to monitor snow clearance from nesting colonies of lesser snow geese (*Chen canagica*) in the Canadian Arctic. Based on rate of snowmelt, they found they could predict nesting success. In another study, Klaas, Anderson, and Fredereick (1978) used Landsat imagery to monitor food availability for snow geese (*Anser caerulescens*) on the DeSoto National Wildlife Refuge in Iowa and Nebraska. Food availability was estimated as a function of the acreage left unplowed following harvest. Digital classification has also been used to monitor and map habitat for reindeer (*Rangifer tarandus*) and moose (*Alcesalces*). LaPerriere Lent, Gassaway, and Nodler (1980) prepared vegetative maps for 13 million hectares in east-central Alaska which correlated vegetation type with moose habitat (George & Scorup, 1981; Laperriere et al., 1980). Their preliminary verification results found an overall accuracy of 77% in comparison to a reference data set collected independently (George & Scorup, 1981).

In Australia, kangaroo habitat has been identified with the use of Landsat data. Hill and Kelly (1987) found that standard cover habitat for gray kangaroos (*Macrovus giganteus*)

could be identified. Multispectral scanner (MSS) imagery was used to map habitat categories, which were then integrated with aerial census work and used to estimate population levels of the kangaroos. Also in Australia, Landsat MSS imagery has been used to map the distribution of the hairy-nosed wombat (*Lasiorhinus latifrons*). This was possible due to the wombat's propensity for building large and extensive mounds that can be detected on imagery. Hairy-nosed wombats are agricultural pests; thus, the ability to monitor their spread is a valuable management tool (Loffler & Margules, 1980).

In 1967, the National Aeronautics and Space Administration (NASA) initiated the Earth Resource Technology Satellite (ERTS) program. Through this program five satellites, carrying a variety of remote sensing systems, have been deployed. Their primary intention is to acquire earth resource information.

The Landsat 5, was launched in March 1984 and carries a four-band multispectral scanner (MSS) and a seven-band thematic mapper (TM) scanner. It orbits the earth at an altitude of 705 km in a sun synchronous near polar orbit. Its repeat coverage is 16 days (Slater, 1985).

MSS systems are so named because they simultaneously record energy from several portions of the electromagnetic spectrum. The MSS scans with a rapidly oscillating mirror, which directs reflected radiation through an optical system, thereby separating the radiation by wavelength bands. Each band is then focused on individual detectors with specific spectral sensitivities that convert the EMR to electrical energy (Estes, 1985). These four bands of EMR sensitivity are 0.5 to 0.6 μm (green), 0.6 to 0.7 μm (red), 0.7 to 0.8 μm

(reflective infrared), and 0.8 to 1.1 μm (reflective infrared). As with any optical system, the MSS is limited in its ability to distinguish surface features. The limit defines its spatial resolution. The area of spatial resolution for the MSS is square and covers an area of approximately 79 x 79 m. This area is termed a pixel (abbreviation for "picture element"). The reflective energy levels of individual pixels are recorded as digital values. Within one MSS image, data from all four spectral bands are a set of more than 30 million data values. The MSS has been carried aboard all five Landsat satellites and has provided consistent data, uninterrupted since 1972. Such data collected on multiple dates for the same area allow users not only to inventory but also to monitor. The TM scanner has been aboard only Landsats 4 and 5. TM discriminates reflected and emitted energy in seven bands: three in the visible, one in the near-infrared, two in the middle infrared, and one in thermal-infrared region of the electromagnetic spectrum. The TM scanner has higher spatial (30 m x 30m), spectral (7 bands), and radiometric resolution than the MSS. Because of these improvements, Solomonson (1984) suggested that the TM scanner is twice as effective in providing information as the MSS. Data from Landsat satellites are relayed to earth either directly or through a relay satellite to one of several acquisition centers around the world. Data are then sent to the NASA Image Processing Facility (IPF), part of the Landsat Ground Data Handling Systems at Goddard Space Flight Center. At the IPF, high-density digital tapes are produced and sent to Sioux Falls, South Dakota, where computer compatible tapes (CCTS) and film products are produced and distributed to users (Short, 1982).

Landsat 7

Not until 1992 was an attempt made to address the shortcomings in the program resulting from the 1984 law and to plan a future for the program beyond Landsat 6. In February, the Bush administration issued National Space Policy Directive No.5. The motivation for this policy initiative was, in part at least, a desire to address Landsat's ongoing problems and uncertainties. Perhaps of greater interest was the desire to ensure the continued availability of Landsat data after it had proved its utility in the Gulf War (U.S. National Aeronautics and Space Administration [USNASA], 2001).

The Directive established a management structure and goals for Landsat. It dealt with many issues surrounding the program as it outlined a strategy to (a) ensure the operation of Landsat 4 and 5 until the launch of Landsat 6; (b) acquire and launch Landsat 7 before the projected end of life of Landsat 6; (c) foster development of advanced land remote sensing systems and opportunities for commercialization; (d) minimize cost of Landsat data to USG users and in the U.S. global change research program; (e) limit the role of the USG in remote sensing to acquisition of data required for national security, foreign policy and public safety; (f) maintain an archive in the U.S. of existing and future Landsat type data; and (g) consider alternatives for maintaining data continuity after Landsat 7 (USNASA, 2001).

To implement this strategy, the Department of Commerce (DOC) was instructed to complete and launch Landsat 6 and to maintain, through Earth Observation Satellite Company (EOSAT), operation of Landsats 4 and 5 until Landsat 6 became operational.

Responsibility for Landsat 7 and beyond was assigned to the Department of Defense (DoD) and the National Aeronautics and Space Administration (NASA). The two organizations were instructed to develop and launch Landsat 7 with performance capabilities at least equal to Landsat 6 (USNASA, 2001).

In March 1992, DoD and NASA signed a management plan for Landsat 7. DoD accepted responsibility for the space segment and NASA assumed responsibility for the ground segment. A budget of \$880 million for the life of the program was established (USNASA, 2001).

In response to the management plan, DoD issued a Request for Proposals (RFP) for Landsat 7 in May 1992. The RFP specified performance requirements; it did not specify an instrument or instruments. The performance requirements included capabilities equivalent to Landsat 6. The RFP also included, for potential bidders to consider adding to their proposals, "prioritized enhancements" and "lower priority enhancements." The former included improved spatial resolution, improved absolute calibration, and stereo mapping capability. The latter included additional spectral bands, cross-track pointing, improved radiometric sensitivity, and improved line of sight (LOS) accuracy (USNASA, 2001).

The award for Landsat 7 was given to General Electric (subsequently sold to Martin Marietta, subsequently to merge into Lockheed Martin) The winning proposal called for two instruments—The Enhanced Thematic Mapper Plus (ETM+) as the "continuity" instrument (including the enhancement in ground resolution and absolute radiometric calibration noted above) and the High Resolution Multispectral Stereo Imager (HRMSI) as

an instrument that addressed many of the other performance enhancements described in the RFP. HRMSI performance characteristics included four, 10-meter VNIR bands, a 5-meter panchromatic band, stereo imaging capability, and off-track pointing (USNASA, 2001).

The contract for Landsat 7 accepted the two-instrument concept, but HRMSI was included as an option to be exercised, that is, selected or rejected, by February 1994. The option arrangement was necessary because neither DoD nor NASA had anticipated a second instrument on the platform or included funding for a second instrument in the budget for the baseline program defined in the management plan. Exercising the HRMSI option was dependent on both agencies securing additional funding to cover the anticipated additional cost (USNASA, 2001).

Land Remote Sensing Policy Act of 1992

The award for Landsat 7 brings us to October 1992 and enactment of the Land Remote Sensing Policy Act. That legislation contains the most recent directives from Congress on how the Landsat Program should proceed, and implementation of the provisions of the law has brought us to this meeting (USNASA, 2001).

The 1992 law conformed to the history of the program. It marked another shift in the philosophy of Landsat, returning control of the program to the government sector, at least through Landsat 7. It recognized officially the importance of data continuity by authorizing Landsat 7, but it left undefined the program after Landsat 7. The law was another attempt to

bring order and consistency to the program, but it did not resolve the fundamental policy issues (USNASA, 2001).

More specifically, the law ratified the actions and followed the recommendations of National Space Policy Directive No.5. It created an entity called Landsat Program Management, consisting of DoD, NASA, and any other agency the president selected to serve on it. LPM was directed to establish a management plan (it already existed) authorized to build, launch, and operate Landsat 7; maintain the Landsat system; provide data to the user community at the cost of fulfilling user requests; place copies of Landsat 7 data in the National Satellite Land Remote Sensing Data Archive at EDC; negotiate a new data policy for Landsat 4-6 with EOSAT; assume Landsat 6 program responsibilities from NOAA following launch of the system; initiate a Landsat Advisory Process; conduct a technology demonstration program; and assess options for a land remote sensing system after Landsat 7. All provisions of the law have been or are currently being implemented (USNASA, 2001).

Remote Sensing and Health

Even if a few general concepts are presented, remote sensing techniques discussed here are viewed from a health perspective. Thus, remote sensing is considered as a tool to assist in health research, in health education, and in the planning, monitoring, and evaluation of health programs.

Remote sensing techniques can be a useful tools for health researchers and planners because, as Scholten and De Lepper (1991) observed,

[h]ealth and ill-health are affected by a variety of life-style and environmental factors, including where people live. Characteristics of these locations (including socio-demographic and environmental exposure) offer a valuable source for epidemiological research studies on health and the environment. Health and ill health always have a spatial dimension therefore. More than a century ago, epidemiological remote sensing and other medical scientists began to explore the potential of maps for understanding the spatial dynamics of disease. (p. 160-170)

A variety of analytical tools are available with remote sensing, extending the capabilities of traditional data base management system (DBMS) to include the ability to analyze data based on their spatial characteristics. Eastman (1992) gives an example of the ability of remote sensing to analyze data based on their spatial characteristics:

Perhaps the simplest example of this is to consider what happens when we are concerned with the joint occurrence of features with different geographies. For example, find all areas of residential land on bedrock types associated with radon gas. This is a problem that a traditional DBMS simply cannot solve — for the reason that bedrock types and land use divisions simply do not share the same geography. Traditional data base query is fine so long as we are talking about attributes belonging to the same individuals. But when the entities are different it simply cannot cope. For this we need remote sensing. In fact, it is this ability to compare different entities based

on their common geographic occurrence that is the hallmark of remote sensing – a process called "overlay" since it is identical in character to overlaying transparent maps of the two entity groups on top of one another. (p. 43)

The example given in the preceding paragraph can be applied to the following situation: A map is needed that shows the districts where there is a health center and where fewer than 50% of 0-1-year-old children have received necessary vaccination. The map should show the hydrographic system (lakes, ponds, and rivers) of the area and the location of clean water sources and sanitation utilities. Also needed is remote sensing, because the immunization data, the water and sanitation data, and the hydrographic system data have different geographies. The analytical tools available within remote sensing are necessary to make possible the integration of data having different geographies.

It should be noted that the geographic analysis system can contribute to the extension of the database: By combining the areas where the immunization rate is low and the access to clean water is difficult, for example, the analyst defines zones and populations at greater risk. In this way, new knowledge of relationships between features is added to the database.

The overlay process is among the most fundamental aspects of remote sensing, but other processes are important and can be very useful in health research and planning. Of particular benefit to the investigation of illness at or near pollution and other hazardous sites is the ability to create buffer zones around the lines or points that represent those locations. The user can specify the size of the buffer and then intersect or merge this

information with disease incidence data to determine how many counts of the illness fall within the buffer (Twigg, 1990).

The association between proximity to nuclear power stations and the prevalence of childhood leukemia in northern England (Openshaw, Charlton, Wymer, & Crost, 1987) has been investigated in this way, and one can easily imagine similar applications with other diseases and other environmental causes or risk factors of disease. Buffer zones analysis can have useful applications in health services analysis and planning; it gives, for example, a quick and easy answer to the question "How many persons live within a 10 kilometer radius from this health care center? Within a 10 to 15 kilometer radius?" The generation of a distance/proximity surface (taking into account distance and "friction" of space — resulting in a cost, in money or time, of transportation) and allocation modeling (assignment of every point of an area to the nearest of a set of designated features, for example, health centers) are other geographic analysis tools that can be useful in health research and planning, where a nonspatial method could give a partial or even false answer. For example, remote sensing was used to study the difference in population per bed ratios between Blacks and Whites and the implications of open access to hospital services formerly reserved for Whites in Natal, South Africa. While the usual administrative boundary-based beds per capita ratios suggested that hospital bed resources in the province of Natal/Kwazulu were racially unequal, but nevertheless, as expressed by Zwarenstein, Krige, and Wolff (1991), adequate (264 people per general and referred bed for the whole population, 195 for Whites and 275 for Blacks), remote sensing analysis revealed widespread inadequacy, worse for Blacks. Of the estimated hospital catchments areas, half

have more than 275 Blacks per general and referral bed; half of these have more than 550 Blacks per bed. One third of the catchments areas estimated for Whites have ratios above 275 people per bed, and one half of these are also above 550 persons per bed. The remote sensing analysis shows that open access to beds previously reserved for Whites will make no difference to rural Blacks and almost none to urban Blacks, because there were relatively few such beds and these were concentrated in the cities. For the same reasons, the opening of private hospital beds would not alleviate the apparent bed shortages in priority areas. Because health is largely determined by environmental factors (including the sociocultural and physical environments, which vary greatly in space), it always has an important environmental and spatial dimension. The spatial modeling capacities offered by remote sensing can help in understanding the spatial variation in the incidence of disease and its covariation with environmental factors and the health care system. Remote sensing in health-related activities can play a role at three levels:

Remote Sensing and Health Research

By helping researchers to understand the distribution and diffusion of disease and its relationship to environmental factors (climate, water quality, sanitation, land use, agricultural, and other economic activities, rural-urban milieu, immunization rate, and so on), it is of value to etiology, epidemiology, and medical science in general.

Remote Sensing and Health Education

Because mapping is an excellent means of communication, remote sensing can be used, as Kabel (1990) suggested, in the preparation of educational material. In an article on participatory evaluation, Fuerstein (1987) described different methods for monitoring and evaluating community health projects, including mapping. Small or large maps may be drawn or painted by groups or individuals to represent the context in which they are living. These maps, showing location of houses by number and type, public and private buildings, water sources, sanitation, bridges, roads, social centers, neighborhood boundaries, and health centers give participants a wider view of where they are living. Maps can help discussion, analysis, decision making, management, and evaluation. Fuerstein suggested that these maps be posted in a public place and updated as changes occur to provide a permanent record. Remote sensing thus produces material that is both useful and conducive to public participation in community health projects. Remote sensing can contribute to community development in general by helping people understand their environment. Effects in the health domain are obvious. From this perspective, indicators developed with the people, such as the 32 indicators found in the basic minimum needs (BMN) database of Thailand (Nondasuta, 1988) or those measuring the 30 priority problems identified in the Recherche national essentielle en santé program in Bénin (Badou, 1994), deserve special attention (these indicators reflect the health level as well as social, economic, and environmental key determinants of health).

Remote Sensing and Health Planning

It is evident that many questions concerning the provision of health care are related to space. People are distributed in space, and they are not evenly distributed. Health problems vary in space; so do the needs of the people. Where should health care centers be situated, and what services should they offer to answer efficiently the needs of populations varying in numbers, densities, and health problems? These are problems that remote sensing can help resolve with their analysis tools.

Health officials as monitoring and evaluation tools, showing the spatial distribution and differential evolution of diseases, can also use maps produced by remote sensing. Monitoring and evaluation are essential parts of health programs, as well as other programs related to development. As the World Health Organization/United Nations Children Fund (WHO/UNICEF) Joint Monitoring Program (1993) pointed out, monitoring is defined as the periodic oversight of the implementation of an activity that seeks to establish the extent to which input deliveries, work schedules, other required actions, and targeted outputs are proceeding according to plan, so that timely action can be taken to correct the deficiencies detected. Closely linked to monitoring is evaluation, is a process by which program inputs, activities, and results are analyzed and judged explicitly against stated norms. These two terms are usually used in tandem as an integral part of every program. Monitoring is an essential element. By giving the managers, planners and policy makers' access to information on coverage, functioning, and utilization of the water and sanitation facilities, operation and maintenance, funding, and water quality, monitoring guides them in making

important decisions. Similarly, worthwhile evaluation of water and sanitation, as a result of effective monitoring, is necessary in ensuring rational utilization of investments allocated for the sector.

It is worth noting that the second version of the water and sanitation monitoring system (WASAMS) software contains a feature developed to facilitate links with remote sensing. The WHO/UNICEF Joint Monitoring Program's (1993) comments on the water and sanitation program could be said, with the same words, about health programs. Monitoring and evaluation are essential parts of health programs; remote sensing, by showing the spatial distribution of diseases in space and time, facilitates the monitoring and appraisal of the effectiveness of health programs.

Remote sensing is a relatively recent and complex technology, which explains why it has not been used to its full potential, especially in the health domain, where it is extremely promising. The possibilities are now more clearly seen. Hardware and software development has produced systems with functions and interfaces that make them much easier to use. This is good news, because remote sensing can be a tool of prime importance to health research and education and in the planning, monitoring, and evaluation of health programs.

Considerations in Landscape Epidemiology

Landscape epidemiology involves the identification of geographical areas where disease is transmitted. It is a holistic approach that involves the interactions and

associations between elements of the physical and cultural environments. First expressed by the Russian epidemiologist Pavlovsky (1966), the theory behind landscape epidemiology is simple: By knowing the vegetation and geologic conditions necessary for the maintenance of specific pathogens in nature, one can use the landscape to identify the spatial and temporal distribution of disease risk. Key environmental elements, including elevation, temperature, rainfall, and humidity, influence the presence, development, activity, and longevity of pathogens, vectors, zootomic reservoirs of infection, and their interactions with humans (Meade, 1988).

Vegetation type and distribution are also influenced by the environmental variables mentioned above and can be expressed as landscape elements that can be sensed remotely and whose relationships can be modeled spatially.

Remotely Sensed Landscape Epidemiology and Tick-Borne Diseases

Ticks vector an array of viral and rickettsial diseases, from Russian spring-summer encephalitis to boutonuse fever in India and South Africa. Each has its own characteristic landscape ecology. For example, landscape epidemiology of Central European encephalitis (CEE) was studied by Wellmer and Jusatz (1981). A toga virus vectored mainly by the common tick, *Ixodes ricinus*, causes the encephalitis. Wellmer and Jusatz delimited the temperature, humidity, precipitation cycles, vegetation, animal associations, and other environmental conditions needed for its maintenance. They found that lower mountain

elevations with dense and diversified vegetation are required by the disease system, limited by such factors as an annual isotherm of 46 °F (8 °C), homogenous vegetation that occurs in pine forest biomes, and light soils with humus for sufficient ground moisture. Wellmer and Jusatz determined that the foci are limited to a few square kilometers each and that a density of only 1 virus-infected tick per 1,000 ticks was required to maintain a natural nidus, which the German geographers called a standortraum (multifactor location space). They mapped and analyzed the geocological conditions for the repeated infection of specific regions by infected ticks and the recurrent infections of people who entered the natural nidus.

American studies have not been so specific, although there are two excellent candidate diseases to research: Rocky Mountain spotted fever (RMSF) and Lyme tick disease. Tick vectors pose a four-stage ecological puzzle. The “dog tick,” *Dermacentor variabilis*, and Lyme tick disease generally vector RMSF by the “deer tick,” *Ixodes scapularis* (now considered to include *I. Dammini*). There are important differences between the two disease systems, but they have the basics in common. Three separate blood meals are required for the life cycle. The adult female lays up to 10,000 eggs, which hatch into six-legged larvae vernacularly called “seed ticks.” These must feed for a week on a small animal, usually a rodent or rabbit. Then the larvae drop off and metamorphize into eight-legged nymphs. Nymphs climb grass and bushes and “quest” vigorously, waving their legs at vibration, heat, or carbon dioxide for an animal host and another blood meal. After feeding, they drop off and change into adult ticks. According to climate and species, nymphs instead of adults may crawl under leaves or other cover to spend the winter. In

spring they emerge hungry for another meal and are able to infect new rodents before new tick larvae hatch. Adult ticks need a third blood meal. They may also feed on rodents, rabbits, opossum, or even birds (useful for dissemination), but they tend to be more discriminating in what they latch onto. They prefer deer, raccoon, fox, or other large mammals, such as dogs, which can disperse them within a wide territory. Each of these blood meals provides an opportunity for the tick to become infected and in turn infect its next host. Here RMSF and Lyme tick diverge, however. The agent of Lyme tick is a spirochete, a bacterium, and *Borrelia burgdorferi*. Each larva must begin the cycle of infection again, and the enzootic disease is maintained by 25-50% infection rates of nymphs. The rickettsial agent of RMSF, *Rickettsia rickettsii*, can pass transovarially in ticks to the new generation. An even less than 10% infection rate among ticks seems adequate to maintain the disease system because the tick itself is the reservoir. The landscape epidemiology of both rests on the tick's habitat, abundance of small mammals, winter shelter, dehydrating sun, and bird predators. RMSF involves the intrusion of people into the natural nidus of the disease. Lyme tick seems to result from the expansion of the nidus to include people.

Rocky Mountain Spotted Fever (RMSF)

Rocky Mountain Spotted Fever (RMSF), sometimes called tick-borne typhus, was first identified in Montana and Idaho (hence its name), but it is most common today in the Piedmont region of the southeastern United States. It has been isolated from the chipmunk, meadow vole, pine vole, white-footed mouse, cotton rat, cottontail rabbit, opossum, and

snowshoe hare; however, serological (antibody) evidence for infection has been found in many other mammals (Burgdorfer, 1980). Better publicity has enabled RMSF to be diagnosed earlier and treated with antibiotics, sometimes even before diagnosis. Although this complicates statistical reporting, it has brought the case mortality down considerably. The onset of the disease involves fever, chills, aches, nausea, and a rash that spreads from the wrists and ankles. Many of these symptoms are common to other diseases and are easily misdiagnosed without antibody identification. Many cases are so mild they are asymptomatic. Of those who do develop disease symptoms and are not promptly treated, however, 6% to 15% may die.

Rocky Mountain Spotted Fever (RMSF) has been increasing in the United States, especially in the Southeast. It is hard to determine how much of the increase is due to greater incidence and how much to better reporting. The increase and focus of the disease are attributable to the rapid extension of suburbs into the wooded, open-field habitat of the dog tick. The Piedmont is one of the most rapidly growing regions within the southeastern United States. New suburban houses, parks, jogging paths, and family dog sojourns are in the natural nidus of RMSF. It is normally a zoonosis transmitted among small mammals, but when human residential land use and recreation are extended into the woods, infected ticks can be brought home to houses and yards.

Rocky Mountain Spotted Fever (RMSF) is caused by one of a group of organisms known as rickettsiae. These microorganisms are structurally related to bacteria, but in other characteristics they resemble viruses. The specific agent of RMSF is *Rickettsia* (named in

honor of Dr. Ricketts) and is transmitted by the bite of an infective tick. Unlike most arthropod-borne pathogens, the causative organism can be passed directly from one generation of ticks to the next (Riley, 1977). RMSF is a vectored "zoonosis" (i.e., a disease which primarily infects animals) and is communicable between vertebrate animals and humans and between various species of animals. The chain of disease transmission for RMSF is noted in Figure 1.

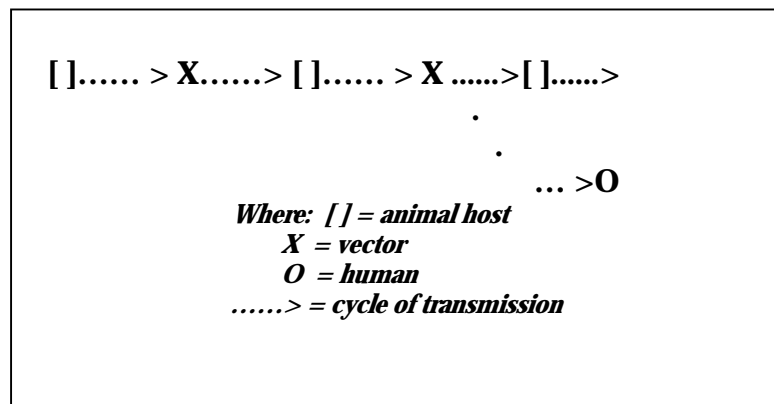


Figure 1. Human transmission chain for RMSF tick disease.

The animal host ranges from a variety of rodents to dogs and large hoofed mammals. Among the more common rodent hosts are the meadow and pine vole, chipmunk, white-footed mouse, cotton rat, cottontail rabbit, opossum, and snowshoe hare (Burgdorfer, 1980). The vector is the hard shell (*ixodid*) tick, which also serves as the main reservoir (Raoult & Walker, 1990). People serve as "dead end" hosts in the transmission chain, in that they cannot transmit the agent to a vector for transmission to another host. Figure 2 notes how the tick's lifespan proliferates itself as a carrier of RMSF to that of a human.

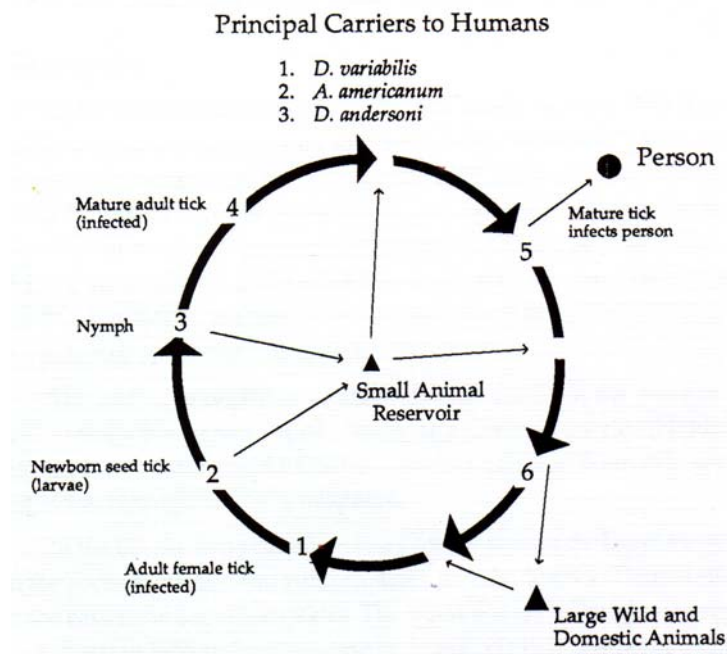


Figure 2. The 2-year cycle of RMSF is based on the lifespan of the tick.

A complete cycle is approximately 2 years, based on the lifespan of the tick. One can begin to analyze the cycle at the point where an infected, female adult tick has completed her third of three "blood meals" from a large wild or domestic animal, such as a horse, cow, deer, goat, or dog. The tick then falls to the ground and lays up to 10,000 eggs (many of which are transovarially infected) and subsequently dies. Within about 36 days the eggs hatch into larvae. These "seed ticks" subsequently cling to vegetation during their quest for small rodent hosts in order to consume the first of three blood meals during their lives. After feeding for approximately 1 week, each larva falls back to the ground and enters the nymph phase of its life cycle. After finding another small animal host, each nymph consumes the second of its blood meals, after which it falls to the ground and develops into an adult tick. At this point, the infected ticks are capable of transmitting the infectious agent

of RMSF to a human host. The infection cycle in humans commences when *Rickettsia* is inoculated into the skin from the saliva of a feeding tick, a process that can only occur after several hours of feeding (Weber & Walker, 1991). After being bitten by the infected tick, an incubation period ensues for a period of 3-10 days. During this time, the rickettsiae produce an inflammation of the inner linings of the blood vessels (USDHHS, 1985). The inflammation eventually becomes visible in the form of a rash, comprised of many red spots under the skin and often concentrated around the wrists and ankles (although it later spreads to the trunk and limbs). The rash often appears on the palms of the hands and the soles of the feet, a symptom that is unique to RMSF; consequently, the feeding tick provides a definitive diagnostic sign (Harden, 1990). Unfortunately, many people never become "clinical" (i.e., they never develop symptoms). The rash may be preceded by several days of chills, high fever, headache, and bone pain (USDHHS, 1985). Other symptoms may include spinal and muscle stiffness, nausea, loss of appetite, and vomiting. Fever may range from 104' to 107' F. and, if left untreated, may persist for several weeks, adversely affecting the central nervous system and resulting in delirium, convulsions, or coma by the end of the 1st week (Blank & Rake, 1955). Critical circulatory and pulmonary complications can occur by the end of the 2nd week (USDHHS, 1985). From the portal of entry in the skin, rickettsiae spread via lymphatics and the bloodstream to all body organs, including the heart, liver, kidneys, lungs, pancreas, gastrointestinal tract, and the brain (Weber & Walker, 1991).

In worst-case scenarios, death can result from toxemia, vasomotor weakness, shock, renal failure, or respiratory or cardiac arrest (McDonald, Anaker, and Garyan 1987). In the

U.S., case mortality has significantly declined, in part due to increased public awareness and early recognition of symptoms but also because of the introduction of antibiotic treatments in the early 1950s. Whereas mortality resulted from 73 % of the cases between 1895-1902 (Harden, 1990), by 1983 the case fatality rate of people who received antibiotic treatment had been reduced to 4 % (McDonald, 1987).

As one might expect, neither incidence nor case fatality rates are uniform across the U.S. population. Rates vary by age, sex, and race. Generally, the highest incidence of disease occurs in persons younger than 20 years of age (primarily 5-9-year-olds), with the largest occurrence among White males (USDHHS, 1981, 1991). The former trend may be related to increased suburbanization and exposure to ticks within wooded recreational areas and/or a close association with exposed household pets, specifically dogs (Newhouse et al. 1986; Pyle, 1979). The hypothesis is that the latter stages of the disease have more limited access to health care. It must be emphasized, however, that the data upon which the above generalizations are based are far from being complete. Since the majority of people infected never develop symptoms, the above statistics are based only on clinical reports.

Lyme Tick Disease

It is not certain whether Lyme tick disease was relocated from Europe, where the spirochete has been well established, or had an indigenous variant here. Lyme tick disease is caused by *Borrelia burgdorferi*, a corkscrew-shaped bacterium, which is transmitted primarily by ticks (*Ixodes scapularis*). The disease cycles among the white-footed mouse, white-tailed deer, other mammals (including humans), and birds. The first U.S. cases were

identified in 1962 in Lyme, Connecticut. Lyme tick already has spread to the Pacific and the southeastern coast of the United States. It has established three main foci: New England and the surrounding area, the upper Midwest, and the Pacific Northwest (Herrington, 1995).

In the East, the reforestation that has occurred since mid-20th century has brought the white-tailed deer and other forest animals back to levels not seen in 2 centuries. Deer, raccoon, and other animals favored by the deer tick (which is somewhat misnamed, as there are ticks that spend their entire life on deer, whereas *I. scapularis* needs three separate hosts) are those most quickly domiciliating to suburban residential areas. This is an intense nidus, with often more than half the adult deer ticks infected. Diffusion has occurred along deer paths in river corridors and electric pylon corridors into suburban areas. The West Coast nidus is not so dangerous for people. The main tick involved there, *I. neotamae*, is the most infected but feeds on the chief reservoir, woodland rats, not people. A second tick, *I. pacificus*, which feeds on many different hosts, including rats and people, is needed to connect the nidus to humans; but because of its diverse diet, as it were, the linking tick's prevalence of infection is low. The cultural interventions and buffers for the two systems are similar. Both diseases can be treated successfully with antibiotics if caught early. Theoretically, reservoir hosts can be destroyed; even if people wanted to exterminate the local deer, however, taking on the white-footed mouse would seem impossible. Because ticks become dehydrated on their questing roosts in the summer sun, by late afternoon they usually have climbed down to the cooler and humid ground. Therefore, walking into the nidus at the end of the day is protective. The different stages of the tick can be killed with

insecticide, a spray for the yard, or a treatment on the dog. Around the house, keeping the lawn mowed short and brushing out of the dog's fenced yard help keep the ticks away. Body buffers while walking into the nidus include long white pants (for easy viewing) tucked, even taped, into socks and boots, and long sleeves, as well as repellent, perhaps. Because it takes hours of feeding for the tick to pass the infection, the best single body buffer is careful examination of children and dogs twice a day. For the pinhead-size nymph of deer tick, this is, of course, easier said than done.

Lyme tick disease is currently the most commonly reported vector-borne disease in the United States. More than 80% of reported U.S. cases have occurred in the Northeast. In this region, residential development within recently reforested suburban areas has brought an increased number of people into closer contact with the tick vector, *Ixodes scapularis*, and consequently to the Lyme tick disease agent, *Borrelia burgdorferi*. This species of tick and its natural hosts (e.g., deer and small vertebrates) are associated with particular landscape features, the forested habitat, for example. High densities of white-tailed deer, the most important host of the adult-stage tick, are supported by the residential-forest landscape, which contains preferred forage in an abundance of edge habitat and ornamental plantings. Mice and other small vertebrates are common hosts of the juvenile stage of the tick, and many of these hosts also serve as reservoirs of the disease.

During 1992-1998, a total of 88,967 cases of Lyme tick disease were reported to the Centers for Disease Control (CDC) by 49 states and the District of Columbia, with the number of cases increasing from 9,896 in 1992 to 16,802 in 1998. A total of 92% of cases

was reported from 8 northeastern and Mid-Atlantic States and two North-Central states. Children aged 5--9 years and adults aged 45--54 years had the highest mean annual incidence (Campbell, Dennis, Hayes & Orlaski, 2000). Early symptoms of the disease usually include fatigue, chills and fevers, headache, muscle and joint pain, swollen lymph nodes, and a characteristic skin rash called erythema migran. Other late-stage symptoms might produce arthritis, nervous system abnormalities, and heart rhythm irregularities. The disease is treatable with antibiotics; however, on rare occasions death occurs.

Lyme tick disease is a vector-borne disease that is amenable to both geographical information systems and remote sensing techniques. Some studies have incorporated both techniques (Glass et al., 1995; Kitron & Kazmierczak, 1997), whereas other studies incorporate just remote sensing. Two approaches have been used to incorporate environmental variables into Lyme tick disease studies. The first approach included just variables assumed to have a high association with either tick distribution (Kitron, Bouseman, & Jones, 1991) or human cases (Glass et al., 1995; Kitron & Kazmierczak, 1997). The second approach includes several dozen environmental variables and then attempts to determine the importance of these through spatial statistics (Glass, Amerasinghe, Morgan, & Scott, 1994; Glass et al., 1995). Regardless of the approach, coverage of soil, vegetation, and water (hydrology, drainage basins) seems to be important. Thematic mapping has been used to describe patterns of selected environmental variables, tick distributions, or human cases. A wide range of spatial analytic techniques has been used within the referenced literature; these include: overlay analysis (Kitron et al., 1991), regression analysis (Glass et al., 1994, 1995), kriging (Nicholson & Mather, 1996), and

autocorrelation (Kitron & Kazmierczak, 1997). Several studies have produced risk maps based on adult *Ixodes scapularis* abundance per white-tailed deer (Glass, 1994), logistic regression of environmental variables (Glass et al., 1995), and the density of nymphal ticks

In 1992 a remote sensing study was initiated to identify landscape features related to exposure risk for Lyme tick disease in Westchester County, New York (U.S. National Aeronautics and Space Administration [USNASA], 1991). The United States National Aeronautical Space Administration (NASA) Ames Director's Discretionary Fund provided the funds necessary for this project. The Medical Entomology Lab (MEL) of the New York Medical College and the Westchester County Department of Health provided the investigative expertise for the epidemiological data. These data consist of canine seroprevalence rates (CSR) for antibodies to the Lyme tick disease agent. CSR were estimated from blood samples taken in 1991 by county veterinarians and represented by municipality, the percentage of domestic dogs exposed to Lyme tick disease. The canine data provided a useful measure of the distribution of Lyme tick disease risk within the county. The data showed that risk appeared to increase from south-to-north along an urban-to-rural gradient. Landsat Thematic Mapper (TM) data acquired on May 20, 1991, were used to characterize land cover along this gradient. The land cover map and epidemiological data were integrated in a geographic information system (GIS) to spatially relate landscape pattern and exposure risk (USNASA, 1991).

An image processing technique known as an *unsupervised classification* was used to develop the land cover map from the thematic mapper (TM) data (USNASA, 1991). Color-

infrared aerial photography and field visits provided the information necessary to accurately identify ("label") these classes by cover type. The land cover classification identified different types of residential areas, as well as different kinds of vegetative cover important for ticks and their hosts. GIS functions were used to define individual forest patches. The area of each forest patch was then determined. Based on proximity to a forest patch, vegetated residential areas were then categorized as adjacent or not adjacent to forest. The land cover composition of each municipality was quantified by overlaying a coverage containing the municipality boundaries, which also contained the CSR data, onto the new landscape map. As a result, a highly significant correlation was found between CSR and the proportion of vegetated residential area adjacent to woods (16K), as well as the proportion of forest within a municipality. Overall, the study showed that basic relationships between the type and placement of landscape elements; Lyme tick disease risk at the municipality level could be described using a remote sensing/geographic information system (RS/GIS) approach (USNASA, 1991).

The Westchester County, New York, study identified landscape features associated with exposure risk for Lyme tick disease at the municipality level (USNASA, 1991). The purpose of the second study was to determine if satellite could characterize landscape differences at a finer scale, which are related to variations in peridomestic exposure risk within communities. In this study, spectral indices derived from Landsat thematic mapper (TM) data were used to describe the landscape composition of residential properties in two Lyme tick disease-endemic communities of Westchester County, New York: Armonk and Chappaqua. The study included 337 properties, all of which contained some wooded area.

Woods are a more favorable environment for survival of tick larvae and nymphs than lawns or herbaceous vegetation. The wooded habitat on the properties had been previously sampled for *Ixodes scapularis* nymphs. Tick density determined from the sample data was used to assign risk levels to the properties. The locations of the sampled residences were digitized into the GIS from high-resolution black and white aerial photography (USNASA, 1991).

A transformation was performed on the TM data to generate "greenness" and "wetness" spectral indices for each pixel in the study area (USNASA, 1991). The means and standard deviations of these spectral indices, as well as topographic information, were determined for a 3x3-pixel area (approximately 90x90 meters) centered on each residence. Within a single community, a multivariate analysis of variance (ANOVA) showed that high-risk properties had significantly higher mean greenness and wetness than low-risk or no-risk risk properties. These properties appeared to contain a greater proportion of broadleaf trees, while lower risk properties were interpreted as having a greater proportion of nonvegetative cover and/or open lawn. The ability to distinguish these fine scale differences among communities and individual properties using satellite data illustrates the efficiency of a remote sensing approach for identifying eridomestic Lyme tick disease risk over a geographic areas (USNASA, 1991).

CHAPTER 3

METHODOLOGY

Research Question

Will the method of analyzing change in remotely sensed data from multiple image technologies such as digital orthophoto quadrangle (DOQ) aerial photographs and aerial videography and Landsat satellite detect changes in landscape between a time when the disease is not present and a time when the disease was present?

Research Design

The model noted in Figure 3 is used for the deductive process for the design of this research:

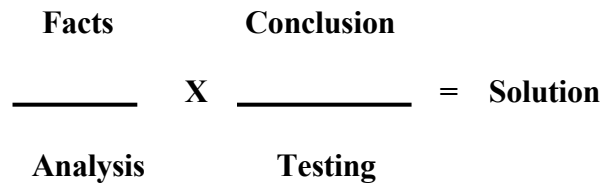


Figure 3. The deductive process model

The derivative of the model noted in Figure 3 is the design for this research. The model to design relationship is presented in Figure 4.

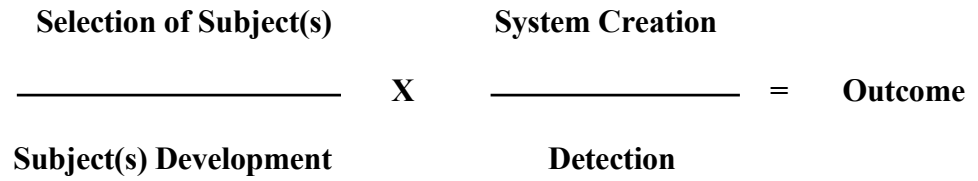


Figure 4. The relationship of model to research design.

The logical flow of the design for this research is noted in Figure 5.

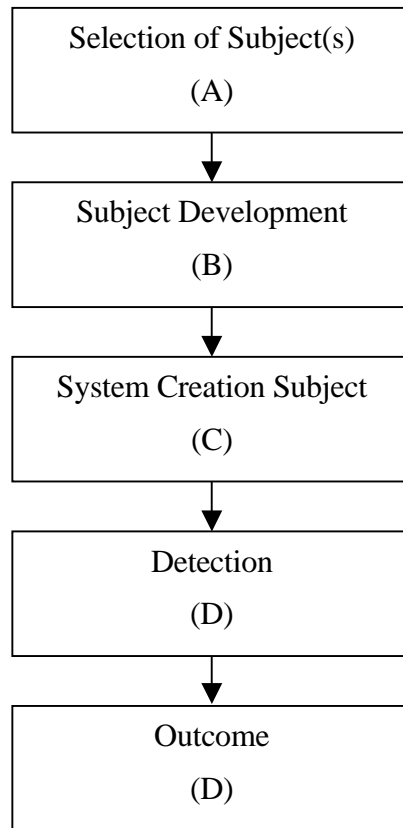


Figure 5. The research design flowchart.

The following section describes each step of the research design flowchart as noted in Figure 5:

Selection of Subject(s)

The selection of subject(s) flowchart, noted in Figure 6, was divided into two categories. The first category was the selection of an infectious disease candidate, and the second was the selection of aerial platform(s).

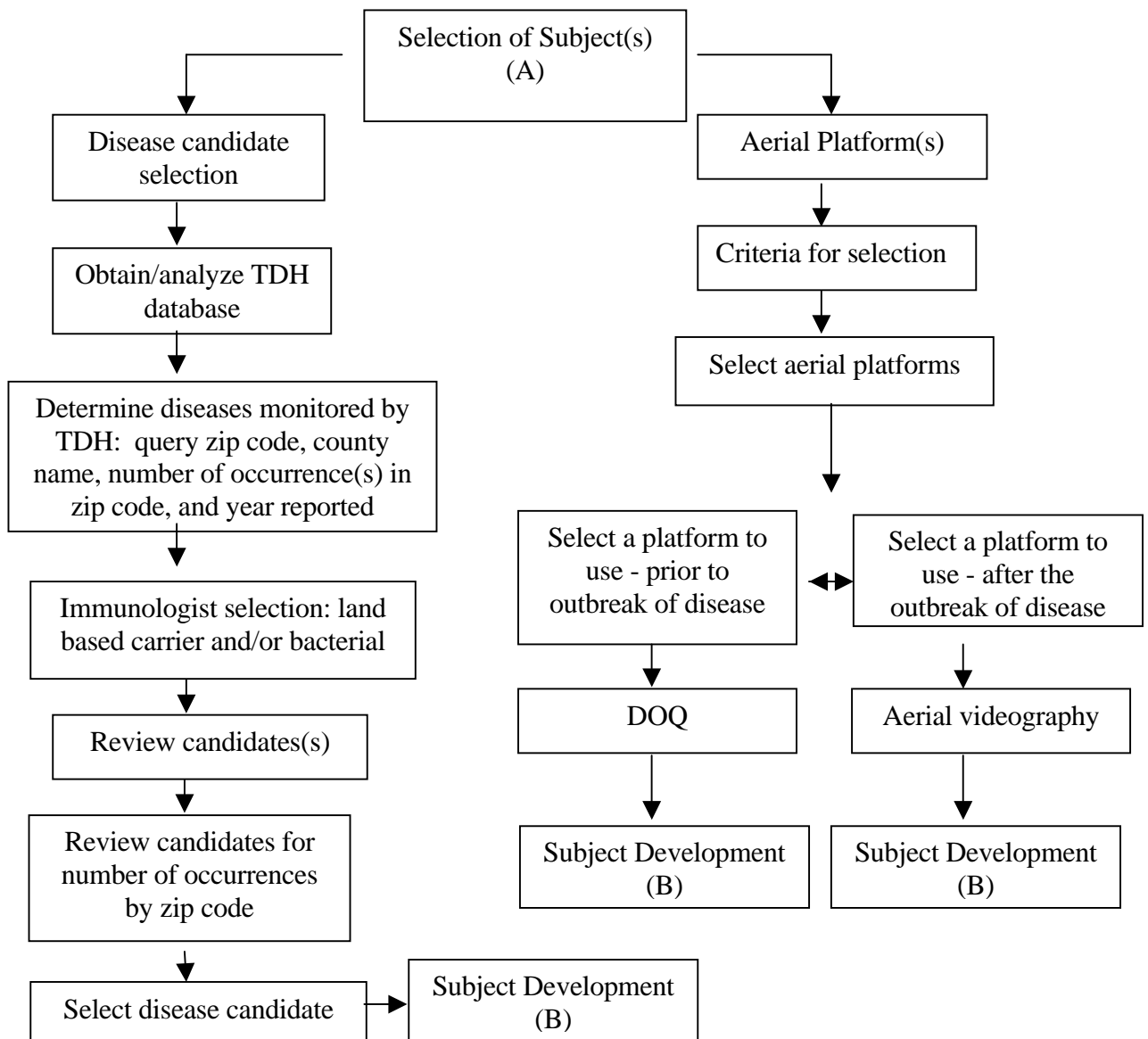


Figure 6. The selection of subject(s) flowchart.

Disease Candidate Selection

The researcher contacted the Texas Department of Health (TDH) to request their assistance in the selection of the most suitable infectious disease candidate for this research. They agreed to provide two immunologists to assist in the selection process. The immunologists selected are employees of the state of Texas and are charged with the responsibility and lawful monitoring for the control of communicable disease(s) in the state of Texas. A law within Chapter 97 of the Texas Administrative Code provides them with the authority to monitor and control all communicable diseases in the state. The provision is titled the Communicable Disease Prevention and Control Act, Health and Safety Code, Chapter 81. The law provides the Board of Health with the authority to adopt rules concerning the reporting of communicable diseases, section 12.001, and provides the Texas Board of Health with the authority to adopt rules for the performance of every duty imposed by law on the Texas Board of Health, the Texas Department of Health, and the Commissioner of Health.

The Texas Department of Health (TDH) provided the researcher with a comma-delimited zip file of all infectious diseases that occurred in the state of Texas. A conversion was made from the comma-delimited file to Microsoft Excel. It was determined that 55,308 line items were provided by TDH. The database contain(s) five fields (Figure 7): (a) zip code of the person contracting the infectious disease; (b) county of residence of the person contracting the infectious disease; (c) name of the infectious disease; (d) number of occurrences within the zip code, and (e) year the infectious disease was reported.

71386	HARRIS	SALMONELLOSIS	4	97
71854	BOWIE	SALMONELLOSIS	4	99
72708	JEFFERSON	SHIGELLOSIS	12	97
72795	VAL VERDE	SALMONELLOSIS	12	98
73059	HARRIS	E COLI 0157:H7	7	97
73149	DALLAS	SALMONELLOSIS	2	97
73951	JASPER	SHIGELLOSIS	10	98
75001	DALLAS	SALMONELLOSIS	4	99
75002	COLLIN	HEPATITIS A	11	97
75002	COLLIN	HEPATITIS A	11	97

Figure 7. Sample query of the TDH database.

Of the total items reported, 14,407 line items, or 26%, contained valid zip codes. A sort was made on the infectious disease field. It was found that the immunologists at the TDH monitored 51 communicable diseases.

The TDH immunologists selected known land-based habitats for carriers of the 51 reportable communicable diseases, resulting in a selection of 12 candidates for possible study (Figure 8).

Final Candidates for Land Base Infectious Diseases

CAMPYLOBACTERIOSIS	bacteria	meat products, water, milk
E COLI O157:H7	bacteria	water-food
LEGIONELLOSIS	bacteria	air/soil or water source
LISTERIOSIS	bacteria	soil or water
SALMONELLOSIS	bacteria	water-food
SHIGELLOSIS	bacteria	water-food
FLEA BORNE-typhus	flea borne	water-wetlands
MOSQUITO BORNE- dengue, arboviral encephalitis	mosquito borne	woods, wetlands
AMEBIASIS	parasite	water-food
CRYPTOSPORIDIOSIS	parasite	contaminated water
TICK BORNE- lyme, spotted fever, ehrlichios	tick borne	woods-grass
HEPATITIS A	virus	water-food

Figure 8. The candidates for land base infectious diseases.

Of the 12 candidates, the immunologists excluded 9 because the disease was transmitted indoors; the carriers also were viral in nature or bacterium or a parasite (Figure 8).

The three diseases that emerged were flea borne (typhus), mosquito borne (dengue, arboviral encephalitis), and tick borne (Lyme tick, spotted fever, ethritichois). Three queries were made to the TDH database to determine the number of occurrences and number of zip codes within each category (Table 3).





Table 3
TDH DATA 1995 TO 1999

<u>Disease Category</u>	<u>Occurrence(s)</u>	<u>No. Zip Codes</u>
Flea	134	51
Mosquito	124	47
Tick Borne	179	147

A search of the TDH database was conducted utilizing the software Geomedia Professional. The software was used first to overlay the United States Postal Department Zip Code boundaries on a map of the state of Texas. The researcher referenced the Geomedia Professional software to the TDH database and conducted queries of the data. The researcher queried each category and determined the frequency of occurrence within each zip code in the state of Texas. A standard rate of occurrence (ratio) was used for all three categories. A color scheme was added to display the ratio of occurrences by zip code. The results are depicted in Table 4. A pictorial view is provided in Figures 9a, 9b, and 9c.

Table 4

Occurrence(s) of a Disease in Zip Codes Affected 1995-1997

Occurrence(s) per Zip Code	Color Code	Number of Zip Codes affected for...		
		Flea	Mosquito	Tick
1 or more		21	32	119
2 to 5		16	10	2
6 to 10		6	1	0
11 to 14		1	1	0
Total zip code(s) affected		44	44	121

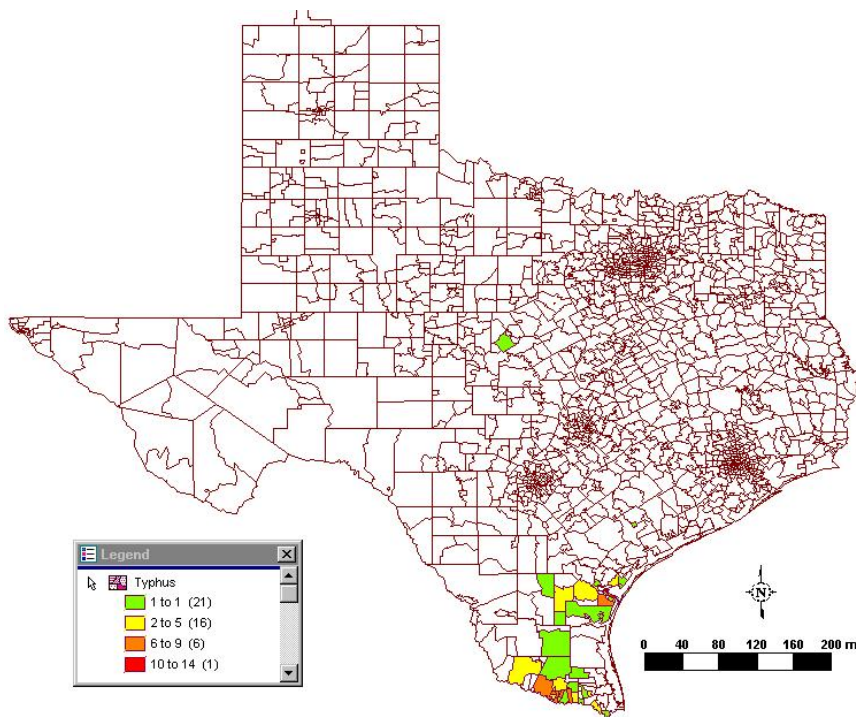


Figure 9a. Occurrences by zip code of flea borne diseases

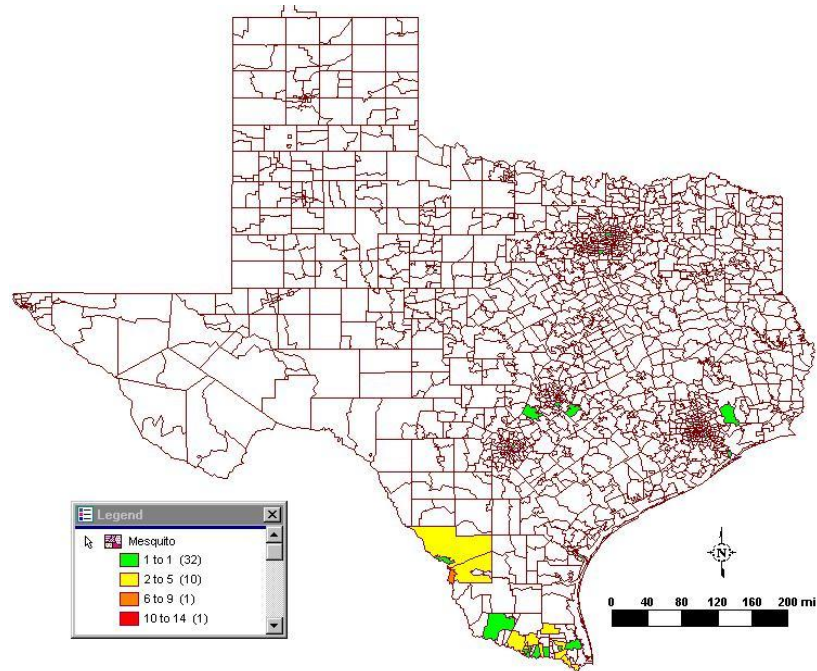


Figure 9b. Occurrences by zip code of mosquito borne diseases.

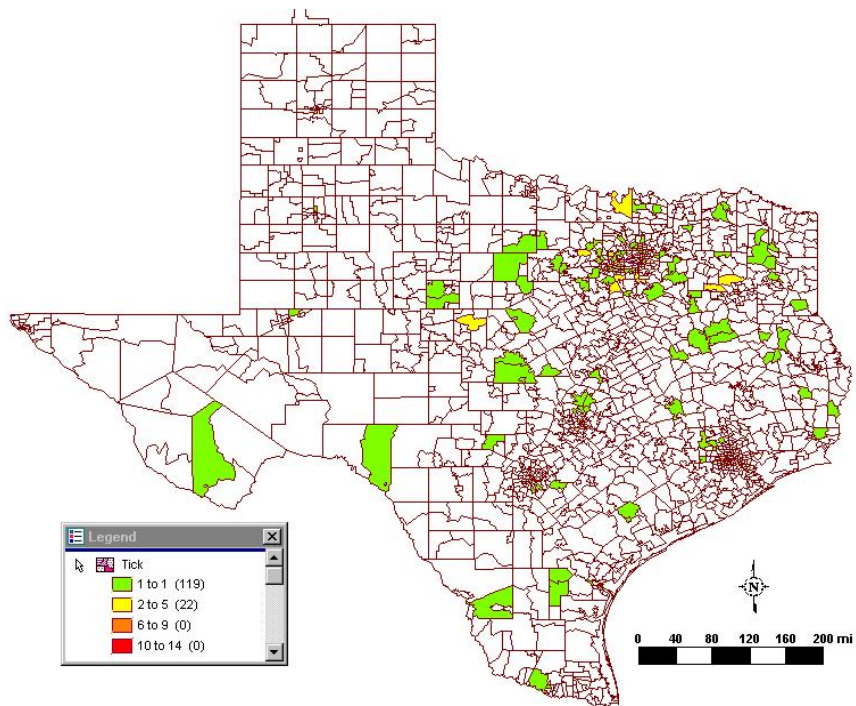


Figure 9c. Occurrences by zip code of tick-borne diseases.

Infectious tick-borne disease was selected for the study because it was prevalent within the state and tended to have a nonmobile population. The other two diseases (flea borne and mosquito borne), although they had a higher rate of incidence, were centered primarily near the border of Mexico, where the population is known to be more mobile.

Aerial videography.

Aerial videography is used to create models to predict land cover. Aspect, slope, elevation, and other physical features influence features of the landscape where definable land cover types are likely to appear. Aerial videography is an excellent medium that can define patterns that are useful in modeling and interpreting points that can predict land cover distribution. A digital video, by frame, can capture images that are reflected at the pixel level and thus permit measurable signatures by the researcher. Videography allows the researcher to identify land cover types and easy analysis and labeling of the spectral value of a scene. Finally, unlike satellite images, video provides accuracy without cloud cover and overlaid topography. The video image allows for manual identity of land cover and use of land cover map overlays. This remote sensing medium was chosen for study after the outbreak of a disease, because this medium of remote sensing has proven to be reliable, accurate, economical, consistent in digital output, and available.

Digital Orthophoto Quadrangle (DOQ) Aerial Photography

A DOQ is a computer-generated image of an aerial photograph in which displacements caused by camera orientation and terrain has been removed. These products combine the

image characteristics of a photograph with the geometric qualities of a map and can be used in numerous geographic information system (GIS) applications, either alone or in combination with other digital data. This remote sensing medium was chosen for study prior to the outbreak of a disease because the product was available and had aerial distortions removed in its final form.

Subject Development

The subject development, noted in Figure 10, is divided into three parts to be considered. The first is tick borne disease, the second is Digital Orthophoto Quadrangle (DOQ), and the last is aerial videography. Each of these parts is discussed below.

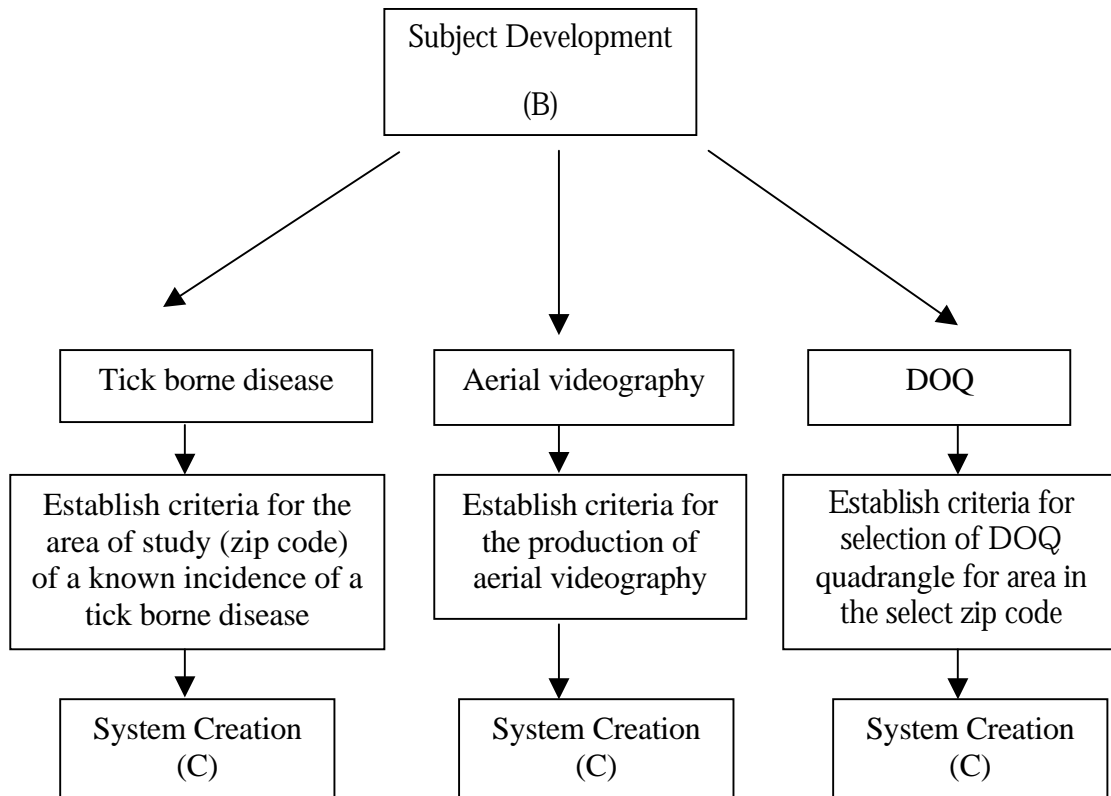


Figure 10. The development of subject(s) flowchart.

Tick-Borne Disease

The first subject to be developed is the specific disease under study or the tick-borne disease.

1. Select area of incidence: The occurrence of the disease in the area of study must occur after the area was documented with an aerial photograph (DOQ) and prior to that of the production of the aerial videography.
2. A select area is made up of a mixture of landscape composed of both rural and urban mix.
3. The area selected will have features that can be easily identifiable, like roads or other features, from an aerial platform.

Aerial Videography

The second subject to be developed is the second of two platforms or aerial videography. A system must be developed to capture aerial videography in an economical fashion. This includes a wing camera, digital video camera, GPS, and a method to transpose ground truthing data to each frame of the video.

A special converter box required development. It required that a constant flow of serial data in informational form be transposed on each video frame. (That is, the videography required a frame-by-frame overlay of latitude, longitude, altitude, heading of the aircraft,

speed of the aircraft, and real-time for documentation of the aircraft position relative to the exact registration of the two aerial platforms.)

The third subject to be developed is the first of two platforms or digital orthophoto quadrangle aerial photography.

Digital Orthophoto Quadrangle (DOQ) Aerial Photography

The Digital Orthophoto Quadrangle (DOQ) for the area of study must be available from the United States Geological Survey (USGS) or an authorized agent. The quadrangle(s) selected must be dated prior to the date the outbreak of the disease.

System Creation

The system creation (C), noted in Figure 11, represents the flow of the development of a system. It includes the selection of an area of an incidence of the outbreak of a tick-borne disease (C1), the selection of specific DOQ quadrangles required for the area of study (C2), the development of an aerial videography system (C3) that can detect the aircraft position (and other data) over landscape, and finally the precise registration (pixel- to-pixel) of the two aerial platforms that can show severity of changes in landscape.

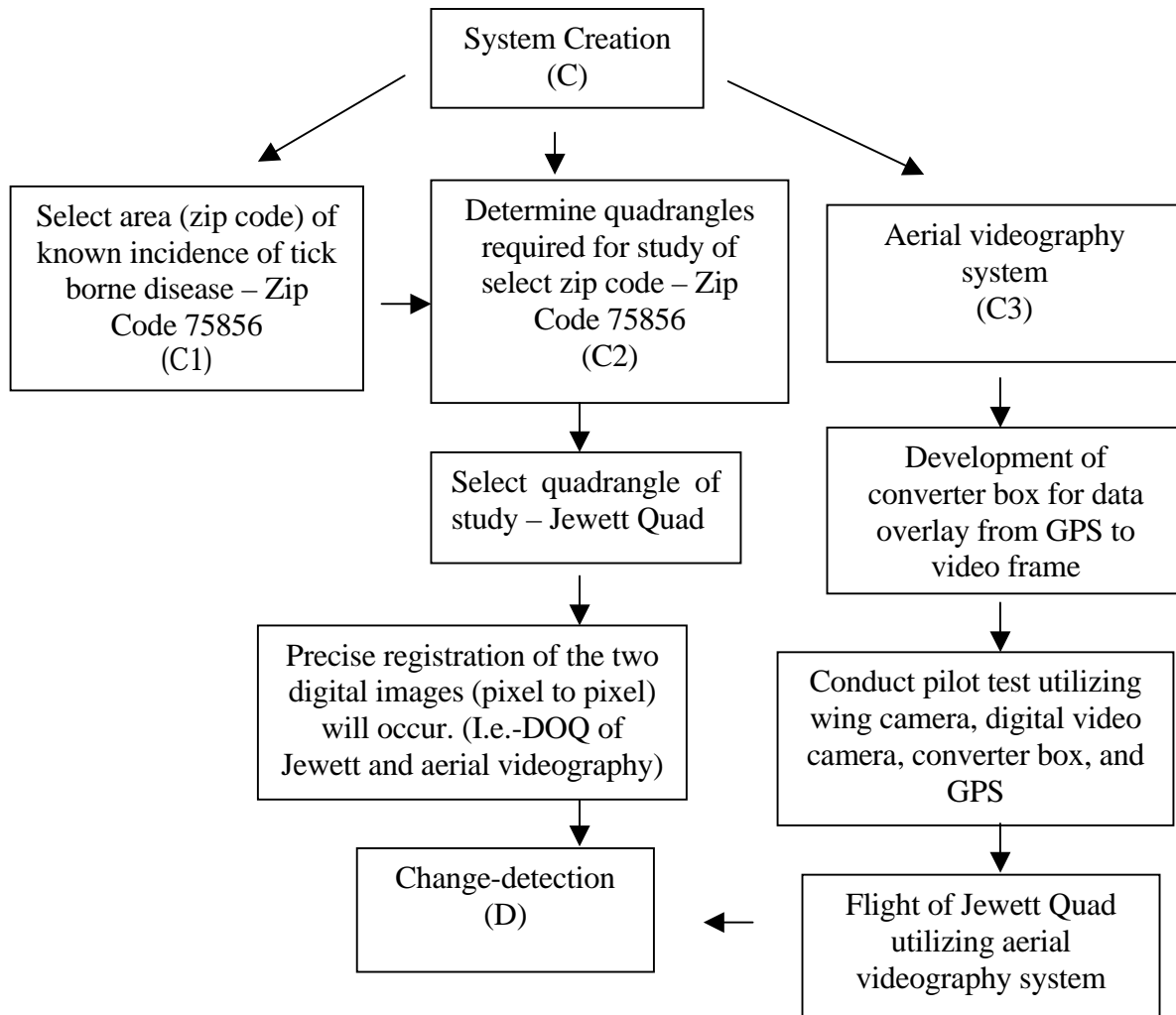


Figure 11. The system creation flowchart.

Zip code 75856 was selected because indicated incidences of the tick-borne infectious disease. The zip code is distributed over 496,656,702 square meters and includes Leon, Limestone, and Robertson Counties, Texas (Figure 12). It contains both rural and suburban landscape.

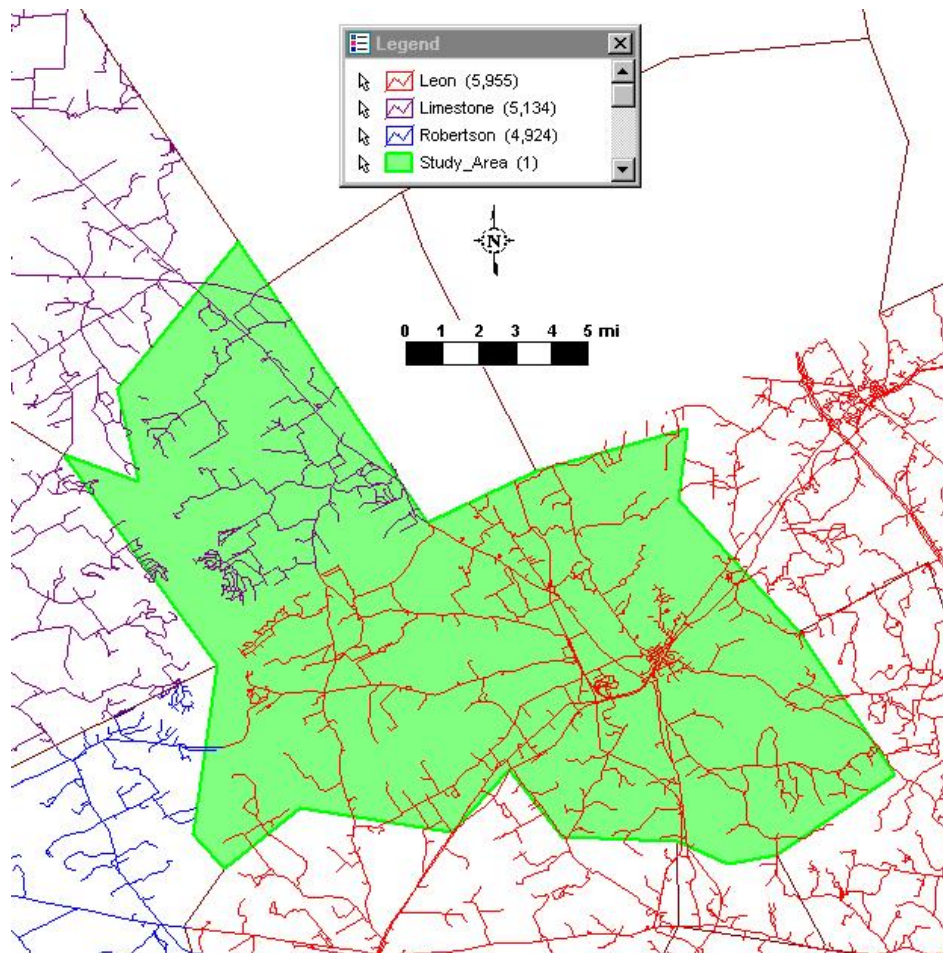


Figure 12. The selected zip code (75856).

The researcher was allowed to download from USGS Internet Web site 7 quarter-quadrangles of 2.5-meter resolution covering the select area of study (zip code 75856). The Quadrangles obtained, dated 1991, were USGS named Donie, Jewett, Farrar, Teague South, Round Prairie, Robbins, and Buffalo. Using Geomedia Professional software, the DOQ's were placed over the zip code for analysis (Figure 13).

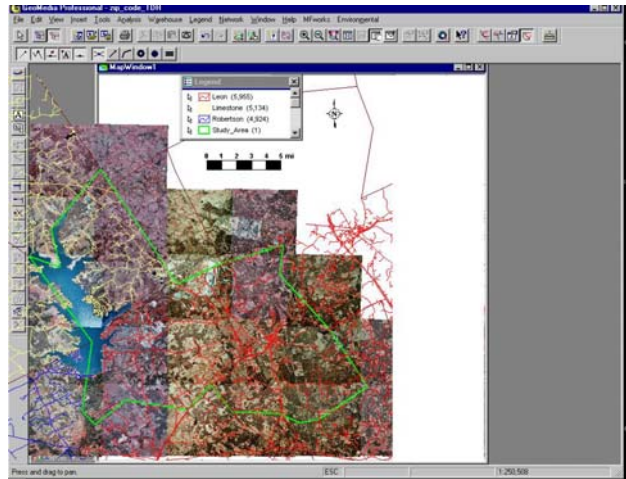


Figure 13. The seven quads DOQ overlay of zip code 75856.

It was determined that the Jewett quadrangle (circa summer 1991) was the best candidate for study since it was composed of both urban and rural landscape (Figure 14).

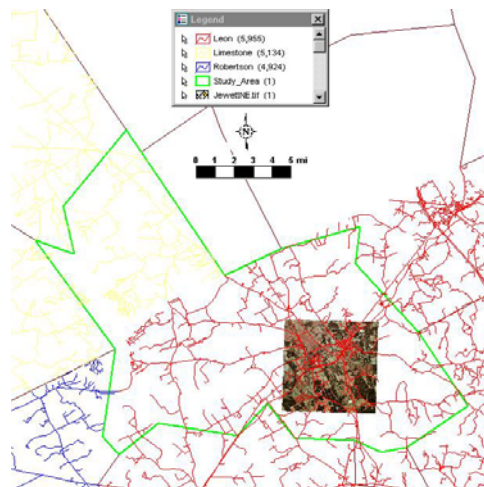
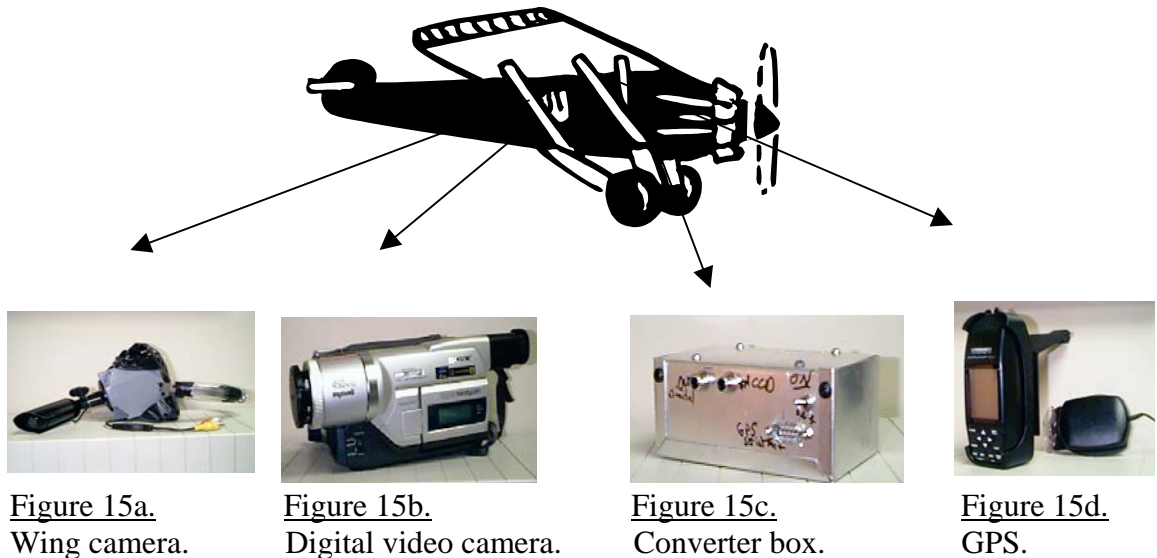


Figure 14. - DOQ Jewett Quad (2.5 meter depicted).

The researcher gathered the necessary parts and equipment, built the converter box, and mounted all equipment within and on the body of a small Cessna 172B fixed-wing aircraft.

The equipment included a digital video camera to record images onto digital High-8 videotapes (Figure 15b) and a strut-mounted remote camera (Figure 15a).



A global positions system (GPS) (Figure 15d) unit was designed to detect the latitude and longitude coordinates of the landscape directly below the aircraft, altitude of the aircraft, speed of the aircraft, what direction the aircraft is flying in relationship to true north, and denotes the real-time.

Converter Box

The flow of serial data from the GPS is unique because it conforms to a universal standard for all GPS units. It was necessary to capture a select portion of the data stream (Figure 16) and provide the researcher with the real-time data necessary for this study. A unique converter box was built to accept a flow of serial data and transpose the data logically onto each frame of the captured video.

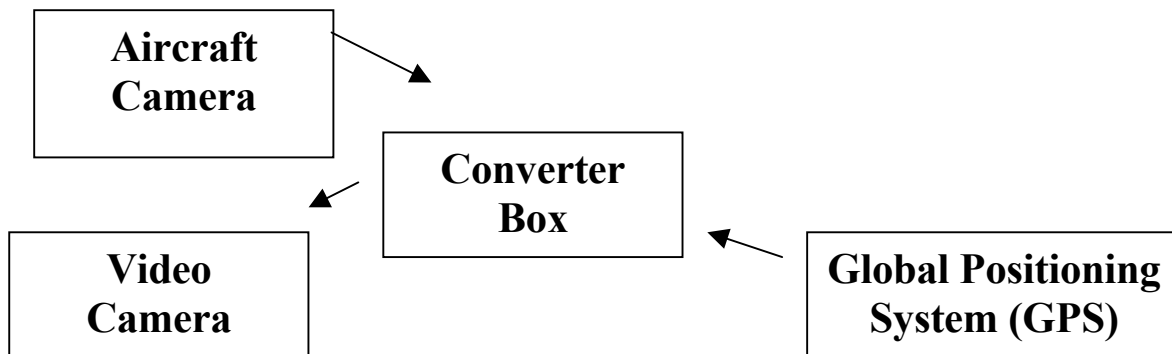


Figure 16. The data stream flow to and from the unique converter box.

An Internet e-mail query was made to the American Radio Relay League (ARRL), which made a referral to Steve Cerwin, a researcher with the Southwest Research Laboratory (SRL) of San Antonio, Texas. The researcher for SRL described a circuit board under development that would disseminate GPS serial data flow. He stated that the unit experiments were currently being used for toy radio controlled planes. SRL referred this researcher to an engineer in New Jersey. A contract was made to build a circuit board to order. The result of this specially built board is depicted in Figure 17.



Figure 17. The special built integrated circuit board.

It was necessary to design and build a special box to house the custom circuit board, house an internal power supply, allow the user to make adjustments to the data stream, provide special input and output ports for the data flow from the aircraft camera, the GPS,

and the digital camera (see Figure 16). In August 2000 the converter box was designed and built meet this criteria (see Figure 15c).

System Test

The Texas Research Institute for Environmental Studies (TRIES) of Sam Houston State University provided grant monies for the purchase of a wing camera and digital video camera. The GPS unit and converter box were purchased and built by the researcher. During the late summer and early fall of 2000 this researcher built and test flew the beta system. The beta flight test was a success. The result of the real-time serial data overlay to each video frame is depicted in Figure 18.



Figure18. The aerial video frame with serial data overlay.

Flight of Jewett Quad

A transecting flight over the Jewett DOQ quad was scheduled during April 2001. The images were collected while flying at a speed of approximately 90 MPH and at approximately 2000 ft above the ground. The video captured on the High-8 minivideotapes was fed to a Sony DV-2000 still image capture board to a personal computer. (Note: The Sony DV-2000 capture board was purchased under grant by TRIES of Sam Houston University for the purpose of this research.) Captured image(s) (Figure 19) were edited with the collaborated effort of a Sam Houston State University (TRIES) GIS analyst.

Registration of Images

The DOQ (Jewett Quad) (noted prior to the disease outbreak) was coregistered or sandwiched to a select aerial video frame (noted after the disease outbreak) utilizing the Geosystems, Inc., ERDAS Imagine software. The registration process was at a pixel-to-pixel level. The researcher applied the statistical change-detection procedure to the registered image.

Detection

The detection, noted in Figure 19, is divided into six categories that represent the procedure for multivariate change-detection.

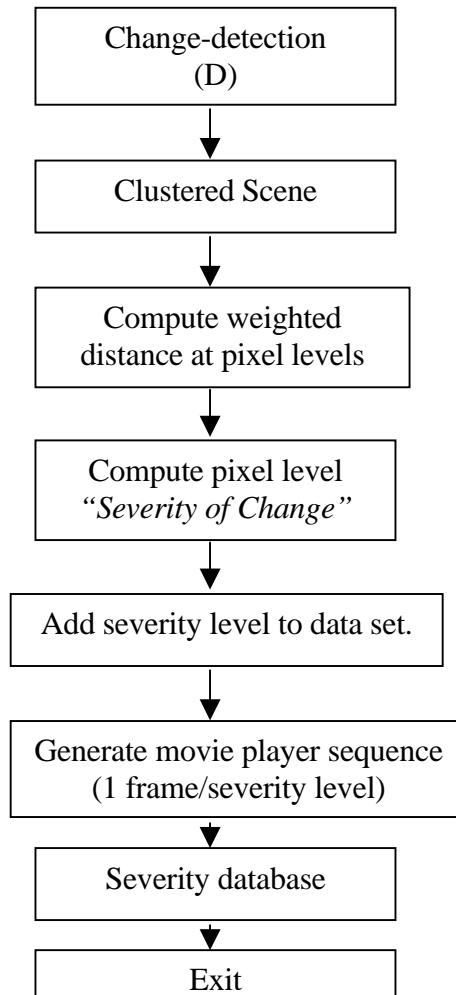


Figure 19. The change-detection flowchart - a multivariate statistical procedure.

Using a multivariate statistical procedure, a numerical value was assigned to each pixel of each image indicating the "severity" of "change" between the two points in time as evidenced in the data. These severity values range from 0 to 100; the higher the value, the more pronounced the change. For viewing, these were converted into scene images, one for

each of several levels of a minimum severity. With this, achieved insight can be attained in regard to severities and their geographical location over a particular area (Hallum, 1998).

Change-Detection-Differencing

Since the data is multidimensional, severities at the pixel level were determined through a multivariate statistical procedure that incorporates information from all the bands. The approach assumes data are available from two separate temporal passes of the remote sensing platform; the data of initial scene (i.e., at time 1) are clustered. This result, in a geographical assemblage of "similar" land cover types, serves as a "baseline" against which to compare the second scene of data. Since the two separate data sets have been registered to one another, the scenes are "differenced" at the pixel level, resulting in a "differenced" data set from which severities are determined. To generate the severity values, a statistical distance referred to as the Mahalanobis distance is computed for each pixel. To compute this distance at the pixel level required the following three entities: (a) the differences of each band (forming a vector); (b) the mean of these differenced vectors computed separately for each cluster in the scene; and (c) the variance-covariance matrix computed for each cluster.

The differences were computed at the pixel level, while the mean and variance-covariance matrix was computed at cluster-level. The distribution of the Mahalanobis distances computed at the pixel level tends to follow an approximate Chi-square distribution with "degrees of freedom" equal to the number of bands used in computing this

distance. By using this distribution, percentile ranks are determined for each pixel, yielding values ranging from 0 to 100.

Difference Characteristics–The Change Detector

The Change Detector is an additional tool for use by image analysts and/or other users. A user no longer needs to depend only upon qualitative indications and/or delineations to characterize changes; with this tool, the magnitude of the changes is quantified and, thereby, can be ranked and/or grouped by the user in whatever way is convenient to aid in an interpretation process for labeling the source of the changes, their geographical location, among others. Obviously, the change detector was limited as follows:

1. To cloud cover areas or areas where the data was of poor quality (e.g., due to stripping, extensive sun glint, and extensive haze).
2. In geographical areas where the Mahalanobis distances do not approximate a Chi-square distribution, the ordering of the distances was preserved by the percentile values (i.e., the severities) even though the severity values themselves may depart from theoretical exactness.

Severity Values

Under some rather simplifying assumptions, the distribution of the Mahalanobis distance function computed at the pixel level tends to follow an approximate Chi-square distribution with "degrees of freedom" equal to the number of bands used in computing this distance. By using this distribution, percentile ranks were determined for each pixel, yielding values ranging from 0 to 100. A severity value of 90 for a given pixel implies one

would expect approximately 10 %, by chance, of the pixels to experience a magnitude of change as extreme as that observed (or equivalently, approximately 90 % of the pixels, by chance, would be expected to have severity values less than that observed for a given pixel). Consequently, the higher the severity value, the more of an indication that the land cover has experienced a temporal change.

Ranking Severities

The pixel level severities represent percentile ranks of the magnitude of change in land cover as reflected in the data. For example, a severity of 90 for a given pixel implies that one would expect approximately 10 %, by chance, of the pixels to experience a magnitude of change as extreme as that observed (or equivalently, approximately 90 % of the pixels, by chance, would be expected to have severity values less than that observed for the given pixel). Consequently, the higher the severity value, the more of an indication that the land cover has experienced a temporal change.

A severity level was computed, for the purpose of this research, at the levels of eighty (80), ninety (90), ninety-five (95), ninety-nine (99), ninety-nine and nine tenth(s) power (99.9), and ninety-nine and ninety-nine thousand(s) power. A severity data set will then be created for each level of severity.

Severity Data Set

Once severities are computed for each pixel, they become a permanent add-on to the underlying data set. This add-on information is critical to the next step in the change-detection process, which is that of highlighting all pixels above a certain severity level.

This is accomplished by generating a graphical interchange format (GIF) file characterized as follows:

1. If the particular viewing frame is the 95th percentile frame, then all pixels are "blacked out" that had a severity rank less than 95.
2. Those greater than or equal to 95 severities were assigned the same color as the initial scene (i.e., the scene associated with time1). Once this information was put into a GIF file-imaging framework, it was ready for viewing in the frame viewer.

Viewing Severity Frames

The ability to view the various severity frames provides considerable insight into the geographic allocation and land cover change information over a particular area. The ability to do a "real-time" scan through the multitude of scenes, each at a different severity level (sequencing according to ascending order of severity magnitudes), with the actual initial full-site scene interleaved in, is an aid to the user in determining the location and severity of changes relative to the baseline scene (i.e., the one obtained in the first overpass at time 1). This was achieved by structuring the GIF imaging files into a frame viewer (1 GIF file per frame) that permits the user the freedom to scan forward and backward a frame (or multiple frames) at a time.

CHAPTER 4

DATA COLLECTION & ANALYSIS

In this chapter data collection and analysis are presented. The first section reports on four constructs: the DOQ and videography severity of change analysis; the two-band DOQ and videography severity of change analysis; videography to videography severity of change analysis; and finally, the Landsat to Landsat severity of change analysis. Each section describes the collection and analysis of data with aspects of importance to the process of developing and validating measurements of the construct. The second section presents an overview of the four related constructs, satisfaction and intention of the constructs, and the results of the analysis.

DOQ and Videography Severity of Change Analysis

Data Gathering

On May 24, 2001, an 8-inch strut-mounted camera was attached to a single engine Cessna-172 aircraft. Attached to the camera was a Sony DCR VRT120 digital camcorder. At 11:05 a.m. a transecting flight in the single engine aircraft was flown over a randomly selected area of the Jewett DOQ quad (circa summer 1991) near Jewett, Leon County, Texas. The digitally recorded Super 8 videotape was then edited using a Pentium II-350Hz personal computer. Installed in this computer was a Sony DV video image capture board.

A total of 14 images in .jpg format were captured at random from the video flown over the Jewett quad near Jewett, Leon County, Texas (Figure 20).

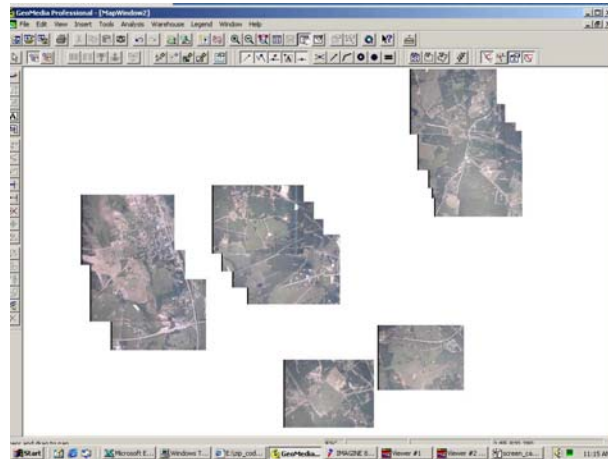


Figure 20. Random video frames captured over Jewett DOQ Quad area.

The 14 images were then placed in the hands of a GIS analyst on May 29, 2001, for registration of the Jewett DOQ (circa summer 1991) and aerial videography conducted on May 24, 2001.

Registration of Images

The first charge of analyst, Richard Rush of The Texas Research Institute for Environmental Studies of Sam Houston State University, Huntsville, Texas, was to orient both the DOQ image and video image to the ground. The ERDAS Imagine and Geomedia Professional software was used, as well as and a Dell OptiPlex GX1 Pentium III-400 MHz personal computer as the primary analytic tool to accomplish the orientation and later the final registration. Ground reference points such as buildings, roads, dams on ponds, or

objects that normally do not change over time were used to determine the direction or orientation for approximate image-to-image placement. Frame 14 of the video was selected as the frame of choice for the registration to the DOQ image (Figure 21).

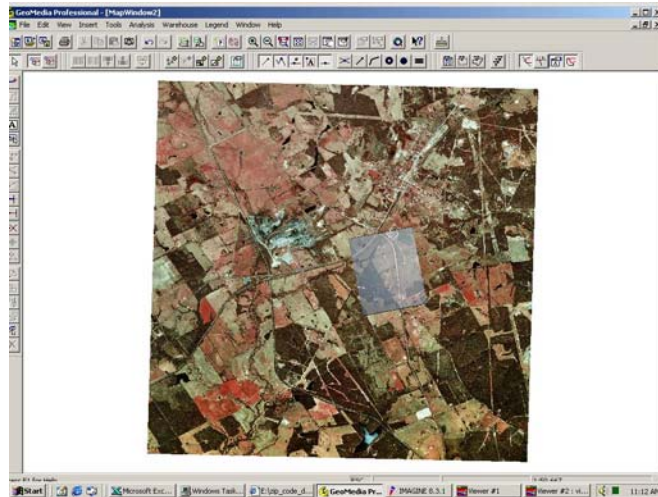


Figure 21. Video frame 14 selected to register with the Jewett DOQ.

Once it was determined that the video image to DOQ image was approximately matched, Rush began the fine tuning of the registration process of the video frame (Figure 22) to the DOQ frame (Figure 23). Next was to use the software to statistically register one image to the other. By request, the ERDAS Imagine software procedure spatial stretches the image and can resample it to a 1-meter pixel size. A polynomial correction procedure was requested of the software to achieve the perfect registration over the area of interest. The software completed a pixel-to-pixel match then digitized the area of interest (i.e., produces boundaries) that enabled the clipping of both images and provided a common boundary for the two. Next the image was converted to an ASCII file, that is, to ensure that the file be converted was not in a noncondensed 24-bit graphic format.



Figure 22. Registered video image. Figure 23. Registered DOQ image

File Conversion

A UNIX operating system.jpg utility converter was used by a computer software analysis to decode the .jpg graphic format to a Portable Picture Mat (PPM) file. The PPM format is binary and allows the programmer to create unique files by writing output through a C program. A special C program was written to convert the PPM files to ASCII format (Appendix A). The output results in four files being produced. All files are ASCII and display the row where the pixel is located, the column where the pixel is located, the intensity range (0-255) of the R or Red value of the pixel, the intensity range (0-255) of the G or Green value of the pixel, and the intensity range (0-255) of the B or Blue value of the pixel and is noted in a numeric format (Figure 24). The first file was named doqq.hal; the second, video.hal; the third, diff.hal; and the last, diff.ppm. The first two files contain the

ASCII row-column-R-G-B values of the DOQ and video image. The third file contains the ASCII row-column-R-G-B values of difference (i.e., change in intensity) between the DOQ and the video image. The last is a converted file from ASCII back to the PPM format.

```
847 1114
0 0 10 172 176
0 1 162 163 167
0 2 153 186 183
0 3 164 216 213
0 4 194 235 224
0 5 202 228 217
0 6 195 216 213
0 7 194 208 205
0 8 186 204 202
0 9 181 205 203
0 10 182 211 207
```

Figure 24. Example of ASCII output by row-column-R-G-B value.

Data Review DOQ to Video Registration

The researcher observed a notable difference in color saturation in both registered images (Figures 22 and 23). He made a comparison of the intensity values of the DOQ and video registered image. MS Paintshop software was used to measure the simultaneous intensity level of each image of a fixed reference point on the ground. The purpose was to determine if both were approximate in intensity range. The researcher noted that a difference of intensity was skewed (i.e., up to 50-60 levels of intensity difference). It was determined that no blue band existed in the DOQ image, but did in the video band. Also, it was observed that the DOQ image contained a near infrared band instead of the visible red band.

A view of the image was ultimately determined by the color gun assignment in the cathode ray tube of the monitor that accepts the input, and the researcher controlled the exact video sequence of bands. However, assignment of values within ERDAS Imagine only impacts how the image is viewed and not the actual data set. In other words, the video is reversible by ERDAS Imagine; the permanent data set is not.

Experimentation with ERDAS Imagine revealed that the DOQ image takes up band 2 instead of band 1 for normal video viewing. It was determined that the DOQ image created by the government “shifts” the bands of the spectrum to the right (Figure 25), allowing the Near Infrared Band (NIR) to become band 3, that is, normally the visible red band. This scheme then assigns the green band as the first in sequence instead of the normal visible red band, the red band becomes band 2, and the NIR is assigned to band 3. This is permanent and is created by the near infrared camera. It was concluded that the purpose of the government using such a scheme is to view vegetation and water for definitive labeling and frequency measurement (i.e., the light spectrum on foliage and water is at a higher range and therefore more visible).

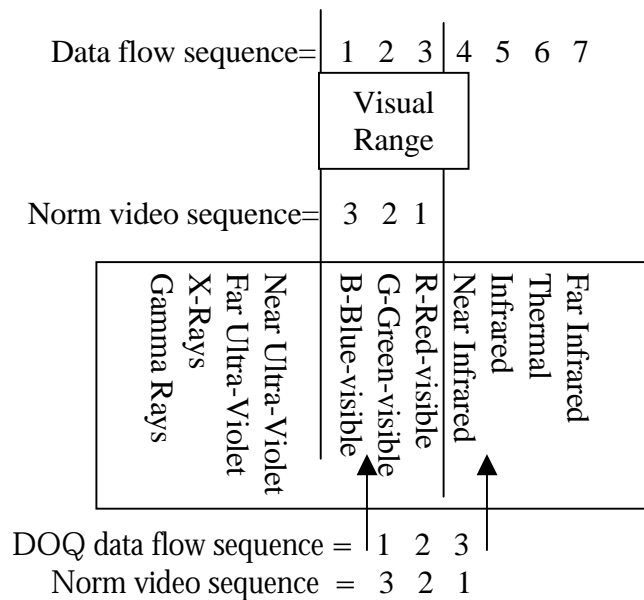


Figure 25. DOQ and video color sequencing.

Because each image is unique for the purpose intended, it was determined that the DOQ quad image could not be used except with another DOQ quad for total accuracy of registration. The only DOQ sequence for this area was completed in 1996. The next scheduled release is in 2003. It was determined that two bands of the spectrum can be registered, that is, the video image and the DOQ image with a severity of change analysis completed.

Two-Band DOQ and Videography Severity of Change Analysis

In light of the results of the video to DOQ registration results, two instead of three bands were suggested to be analyzed and to determine if there was any collation in intensity. The identical scene, as noted in Figures 22 and 23, was used for analysis, and a graphic (.jpg file) was created for each band for each image using ERDAS Imagine. The results are six images of all six bands (Figures 26-31), the Blue-Green-Red spectrum for

the true color video, and the Green-Red-NIR spectrum for the DOQ image.



Figure 26. DOQ Green Data Band 1.



Figure 27. DOQ Red Data Band 2.



Figure 28. DOQ NIR Data Band 3.

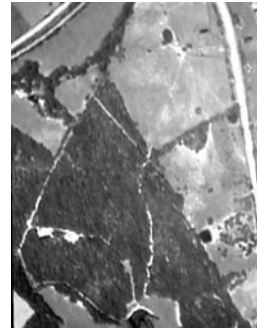


Figure 29. Video Blue Band 1.

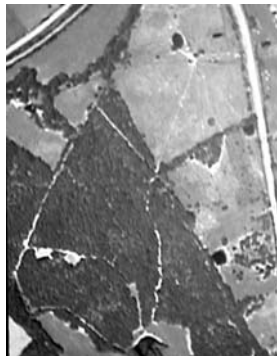


Figure 30. Video Green Data Band 2.

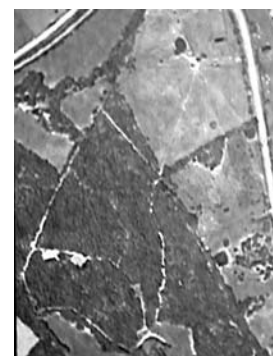


Figure 31. Video Red Data Band 3.

Registration of Images

The green data band 2 of video was registered (Figure 30) to the green data band 1 (Figure 26) of the DOQ. Second, the red data band 3 (Figure 31) of the video was registered to the red data band 2 (Figure 27) of the DOQ. The two images (Figure 28), the DOQ NIR band 3 and the Video Blue band 1 (Figure 29), would not be measured because no NIR spectrum existed on the video image, and there was no comparable video blue band on the DOQ image for comparison.

File Conversion

The same process used in the DOQ-to-video conversion was used in the file conversion for the two bands. A UNIX operating system.jpg utility converter was used by a computer software analysis to decode the .jpg graphic format to a Portable Picture Mat (PPM) file. The PPM format is binary and allowed the programmer to create unique files by writing output through a C program. A special C program was written to convert the PPM files to ASCII format. The output resulted in two files being produced. All files were ASCII and displayed the row where the pixel was located, the column where the pixel was located, the intensity range (0-255) of the B or Blue value of the pixel, and the intensity range (0-255) of the G or Green value of the pixel. The two files produced an ASCII row-column B-blue value and the G-Green values of the DOQ and video image from 0-255 intensity.

Data Review

The result of two-band DOQ and videography severity of change analysis was the same as the first results of the three-band analysis of the DOQ to videography. The researcher observed a notable difference in color saturation of the registered images (Figures 26-31). A comparison of the intensity values of the DOQ and video registered image was made. MS Paintshop software was used to measure the simultaneous intensity level of each image of a fixed reference point on the ground. The purpose was to determine if both were approximate in intensity range. The researcher noted that a difference of intensity was skewed (i.e., up to 50-60 levels of intensity difference). This two-band analysis was then abandoned.

Videography to Videography Severity of Change Analysis

The researcher determined that the intensity value of each pixel must not “shift” and result in a permanently skewed image like the DOQ to video aerial image. A decision was made to conduct a videography-to-videography severity of change analysis with the hope that the collection, registration, and measurement would result in an accurate measurement of severity on a pixel-to-pixel level.

Data Gathering

A second over flight was made on July 22, 2001, over the same area previously videotaped on May 24, 2001, in Leon County, Texas. The purpose of this flight was to duplicate and register the same area for change of severity analysis.

Contact was made with the Texas Department of Health and the regional office of St. Joseph Hospital District located in Bryan, Texas (this hospital district is the reporting agency for Leon County, Texas, to the Texas Department of Health) to determine whether an occurrence of a tick-borne disease had occurred over the zip code under study from May 24, 2001, until July 22, 2001. The hospital district infectious disease staff confirmed that one occurrence of a tick-borne disease was noted during the time period under study in Leon County, Texas.

Registration of Images

The image flown May 24, 2001 (Figure 32), and the image flown July 22, 2001 (Figure 33), were both registered. This registration allowed both file conversion and analysis of one videography frame to the other videography frame.



Figure 32. Registered Video Frame 5-24. Figure 33. Registered Video Frame 7-22.

File Conversion

The same process was used for the video-to-video frames as in the DOQ-to-video conversion and the file conversion for two bands. A UNIX operating system.jpg utility

converter was used by a computer software analysis to decode the .jpg graphic format to a Portable Picture Mat (PPM) file. The PPM format is binary and allowed the programmer to create unique files by writing output through a C program. Two special C programs were written to convert the PPM files to ASCII format and to convert ASCII format back to PPM format. The output resulted in an ASCII file to be produced (Figure 34). The program produced, by row and column, the intensity level (0-255) for the R-red value and the G-green value B-blue value of each pixel within the image. A total of 474,254 rows with three rows of intensity levels (Red-Blue-Green) were analyzed.

<u>Row /Column</u>	<u>Red</u>	<u>Green</u>	<u>Blue</u>
0 0	94	86	63
0 1	91	83	60
0 2	87	81	55
0 3	90	84	58

Figure 34. Example of data polled.

Data Review

A definitive collation of pixel intensity between the two registered images (Figures 32-33) was noted. A decision to proceed with the process of determining the severity of pixel level differences of the red, green, and blue values from two frames of videography was made. A definitive procedure is required to determine pixel intensity and produce a graphic that depicts the various levels of severity. This procedure is outlined below:

1. The two images must be registered to each other. The two images must contain the exact number of rows and columns, and each must represent the same geographical area.
2. The data set will contain five columns of data comprised of the variables, i , j , r , g , and b . The i and j are the row and column location, respectively, of the given pixel. The r , g , and b are the “red,” “green,” and “blue” value of the digitized image taken from the videography footprint covering the target area.
3. The row and column values in the two separate data sets must correspond to precisely the same location on the landscape.
4. The pixel level difference between the two data sets will be computed first (Appendix B). Consequently, the results in a single data set will have five columns (and will be the final data set to complete the determination of the severities): i j d_r d_g d_b (Please note: the subscripted d 's refer to the resulting pixel differences between the two data sets.)
5. The three differences for each pixel will be treated as a vector of length 3 (i.e., as a 3 by 1 column vector). Next, compute the overall average of the vectors for the entire scene (i.e., add all the d_r values and divide the total number of pixels—which will be the number of rows times the number of columns); the same method of computation will be repeated to obtain the result of difference for the green and blue pixels. This will result in an average vector of differences for the whole scene; the average vector should be denoted as follows:

$$\bar{\mathbf{d}} = \begin{bmatrix} \bar{d}_r \\ \bar{d}_g \\ \bar{d}_b \end{bmatrix}$$

The variance/covariance matrix is computed for the entire scene; the following is the substep for this procedure:

For each pixel, compute the following 3 x 3 matrix:

This is simply the multiplication of the 3 x 1 vector by the 1 x 3 vector

$$\begin{bmatrix} d_r \\ d_g \\ d_b \end{bmatrix} [d_r \ d_g \ d_b] = \begin{bmatrix} d_r d_r & d_r d_g & d_r d_b \\ d_g d_r & d_g d_g & d_g d_b \\ d_b d_r & d_b d_g & d_b d_b \end{bmatrix}$$

(Note: = 9 values for each pixel)

The 3 x 3 matrices above are added together and the sum divided by the total number of pixels (number of rows times the total number of columns).

Subtract the following 3 x 3 matrix from the one obtained in the step above:

$$\bar{\mathbf{d}} \times \bar{\mathbf{d}}^T = \begin{bmatrix} \bar{d}_r \\ \bar{d}_g \\ \bar{d}_b \end{bmatrix} \times [\bar{d}_r \ \bar{d}_g \ \bar{d}_b] = \begin{bmatrix} \bar{d}_r \bar{d}_r & \bar{d}_r \bar{d}_g & \bar{d}_r \bar{d}_b \\ \bar{d}_g \bar{d}_r & \bar{d}_g \bar{d}_g & \bar{d}_g \bar{d}_b \\ \bar{d}_b \bar{d}_r & \bar{d}_b \bar{d}_g & \bar{d}_b \bar{d}_b \end{bmatrix}$$

From the previous step, we now have the variance/covariance matrix (call it \mathbf{S}). \mathbf{S} must now be inverted (i.e., use any software package that obtains the inverse of a 3 x 3 matrix [denoted by \mathbf{S}^{-1}]).

Now, for each pixel (i.e.-for each row and column location), we compute the Mahalanobis distance as follows:

$$\begin{array}{ccc}
 (\mathbf{d}_r, \mathbf{d}_g, \mathbf{d}_b) & \times & \mathbf{S}^{-1} & \begin{bmatrix} \mathbf{d}_r \\ \mathbf{d}_g \\ \mathbf{d}_b \end{bmatrix} \\
 \uparrow & & \uparrow & \uparrow \\
 \mathbf{1} \times \mathbf{3} & & \mathbf{3} \times \mathbf{3} & \mathbf{3} \times \mathbf{1} \\
 \text{vector} & & \text{matrix} & \text{vector}
 \end{array}$$

So each Mahalanobis value at the pixel level is a single (i.e., 1 by 1) number.

The Mahalanobis distances will be sorted from smallest to largest and assigned the smallest value its rank value which is 1; assigned the second value the rank of 2, etc. for all the ranked values.

A division of each rank by the total number of pixels (i.e.-number of rows times the number of columns) will occur; the results are the value that indicates the proportion of Mahalanobis distances that is less than or equal to the given pixel value. Last, the value will be converted to a percentage by multiplying by 100 (Note: format, using two decimal places for accuracy.)

Create a final data set comprised of:

$$\mathbf{i} \quad \mathbf{j} \quad \mathbf{S}_{ev}$$

where \mathbf{i} and \mathbf{j} are the row and column numbers, respectively, of the pixel and " \mathbf{S}_{ev} " is the "severity."

A final step is to generate an image that depicts the geographical location of the pixels that experience the various levels of severity. The method for depicting this location is to work with the image (i.e., the r, g, b image) of the target area from the first over flight. This image will be the "baseline" scene. White out all those pixels with a severity of:

80 or less

90 or less

95 or less

99 or less

99.9 or less

99.99 or less

This mathematical detection procedure (Appendix B) allowed for the creation of a seamless interface (Appendix C) that would allow the researcher to control, validate, and observe the detection procedure as it was computed. The mathematical procedure was applied to the two registered video images (Figures 32, 33). The result was five graphic images (Figures 35, 36, 37, 38, 39) depicting the severity of change within the landscape. Each image produced pixels that were scattered and clustered. However, as the level of severity increased (i.e., 80:90:95:99:99.9:99.99) the background disallowed the intensity to

be produced. Hence, the pixels in the range requested or computed was representative of severity for that portion of the landscape.

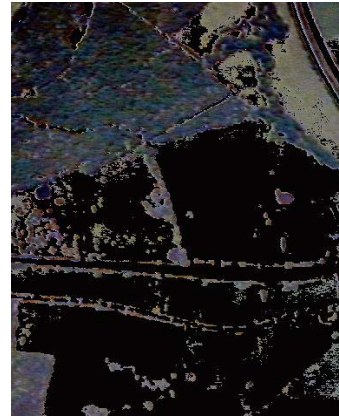
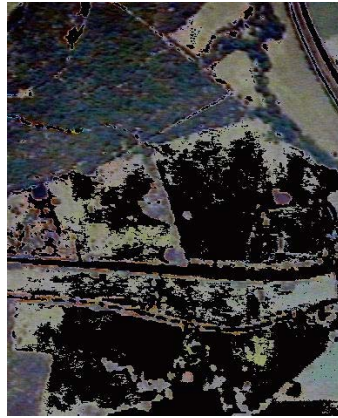


Figure 35. 80% Severity-Video.

Figure 36. 90% Severity-Video.

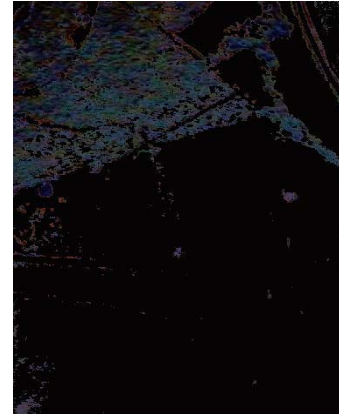
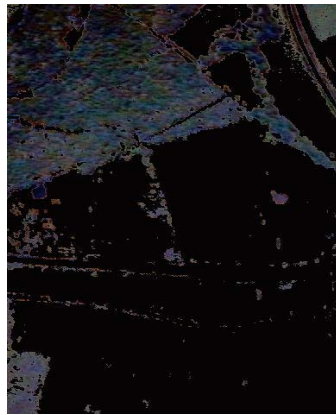


Figure 37. 95% Severity-Video.

Figure 38. 99% Severity-Video.



Figure 39. 99.9% Severity-Video.

Landsat to Landsat Severity of Change Analysis

The researcher decided to validate the severity of change procedure using another medium of remote sensing. A decision was made to conduct a Landsat to Landsat severity of change analysis with the hope that the collection, registration, and measurement would result in an accurate measurement of severity on a pixel-to-pixel level.

A measurement of Landsat to Landsat severity of change analysis was conducted on Nueces County, Texas from October 20, 1999, to April 6, 2000, to determine whether the collection, registration, and measurement would result in an accurate measurement of severity on a pixel-to-pixel level for a known occurrence of flea-borne typhus disease.

Data Gathering

A previous analysis of land-borne infectious diseases revealed that a high incidence of flea-borne typhus disease was noted in Nueces County, Texas (see Table 3). Contact was made with the Nueces County Public Health Office, Corpus Christi, Texas. The Nueces

County Public Health Office stated that for this time frame it found 10 cases of Typhus Murine disease in Nueces County.

The researcher contacted the Center for Space Research at the University of Texas-Austin for Landsat 7 Enhanced Thematic Mapper Plus imagery for Nueces County, Texas. The center had Landsat images over two time frames, October 20, 1999, and April 6, 2000 (Figures 40, 41). The researcher was allowed to use each image for research and analysis.

Registration of Images

A portion of each scene (Figures 43, 44) was selected and cropped from each of the two Landsat images (Figures 41, 42) furnished by the University of Texas-Austin. The first cropped image (Figure 42) was then registered to the second cropped image (Figure 43) for further analysis.



Figure 40. Nueces Landsat (10-20-1999).

Figure 41. Nueces Landsat (4-6-2000).



Figure 42. Nueces Registered Landsat.



Figure 43. Nueces Registered Landsat.

File Conversion

The same process was used for the video-to-video frame conversion. A UNIX operating system.jpg utility converter was used by a computer software analysis to decode the .jpg graphic format to a Portable Picture Mat (PPM) file. The PPM format is binary and allows the programmer to create unique files by writing output through a C program. Two special C programs were written to convert the PPM files to ASCII format and to convert ASCII format back to PPM format. The output results in an ASCII file to be produced (Figure 35). The program produces, by row and column, the intensity level (0-255) for the R-red value and the G-green value B-blue value of each pixel within the image

Data Review

Like the videography-to-videography mathematical procedure, the change-detection procedure (Appendix B) was carefully followed for the Landsat to Landsat images and incorporated in the seamless user interface (Appendix C).

A comparison of the two registered Landsat images of Nueces County (Figures 42, 43) revealed a definitive collation of pixel intensity within each image. The mathematical procedure was applied to the two registered video images (Figures 42, 43). The result was five graphic images (Figures 44, 45, 46, 47, 48) depicting the severity of change within the landscape. It was noted that the pixels of high intensity were both scattered and clustered and as the level of severity increased (i.e., 80:90:95:99:99.9:99.99). The background disallowed the intensity to be produced. Hence, the pixels in the range of requested or computed was representative of severity for that portion of the landscape.



Figure 44. 80% Severity–Landsat.

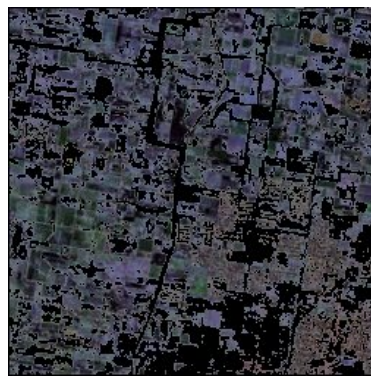


Figure 45. 90% Severity–Landsat.

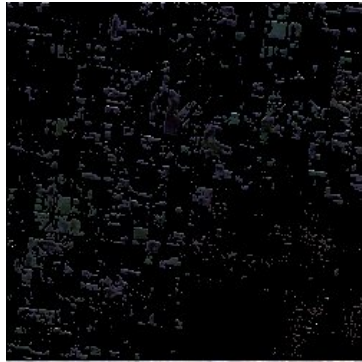


Figure 46. 95% Severity – Landsat.



Figure 47. 99.9% Severity–Landsat.



Figure 48. 99.99% Severity – Landsat.

CHAPTER 5

FINDINGS AND CONCLUSIONS

Findings

DOQ and Aerial Videography

A validation technique was applied to the registered DOQ frame prior to the disease outbreak and an aerial videography frame after the disease outbreak, and a broad range of intensity values was noted. A comparison of the intensity values of the registered DOQ and video images revealed that the level of intensity was skewed and not reliable. Further research revealed that, during the production of the DOQ image by the government, the DOQ image was created with a near infrared camera. This proprietary method adopted by the government is permanent in the production of the film and cannot be changed. Specifically, the method utilized causes a shift in the spectrum band to the right. This shift is then embossed permanently on the film. Hence, the detection of severity for the registered videography and DOQ graphic image(s) was not possible because a shift in the spectrum of the DOQ image caused two levels of color to be improperly matched over one another, causing the intensity to be skewed and inaccurate.

Two-Band DOQ and Videography

It was found that a two-band analysis might be possible. That is, a blue band did not exist in the DOQ image but did in the video band. The DOQ image contained a near infrared band instead of the visible red band. But since the DOQ and video image is ultimately determined by the color gun assignment in the cathode ray tube of the monitor that accepts the input, the analyst can exercise control over the exact video sequence of bands. The Blue and Red bands of the spectrum (i.e., Blue-Green-Red spectrum for the true color video and the Green-Red-NIR spectrum for the DOQ image) could be measured. The same process used in the DOQ-to-video conversion was used in the file conversion for the two bands. The result of two-band DOQ and videography severity of change analysis was the same as the results of the three-band analysis of the DOQ to videography. Again, the researcher observed a notable difference in color saturation of the registered images. A comparison of the intensity values of the DOQ and video registered image was made, and the results showing a notable skewed intensity difference, as in the previous study, were noted. This two-band analysis was then abandoned.

Videography to Videography

It was found that the procedural process and the outcome of videography-to-videography image collection, registration, and measurement would result in an accurate measurement of severity on a pixel-to-pixel level. After a second flight was made on July 22, 2001, over the area previously videotaped on May 24, 2001, contact was made with two sources to verify whether a case of tick-borne disease had occurred in Leon County

over this time frame. The source reported that an occurrence had been reported during the time frame under study. The researcher found that the videography-to-videography remote sensing images would become a valid vehicle for proving or disproving the method of severity analysis detection procedure. The result was the creation of a seamless interface that allows the researcher to validate each detection procedure as it is computed and provide a graphical image of the severity of change within the landscape. The detection procedure revealed changes in the landscape utilizing the severity of change-detection method.

Landsat to Landsat

The measurement of Landsat to Landsat severity of change analysis in Nueces County, Texas, from October 20, 1999, to April 6, 2000, was successful in determining the collection, registration, and measurement severity on a pixel-to-pixel level for a known occurrence of flea-borne typhus disease.

A comparison of two registered Landsat images of Nueces County revealed a definitive collation of pixel intensity with each image. Like the videography-to-videography mathematical procedure, the change-detection was carefully followed for the Landsat to Landsat images and incorporated in a seamless visual basic program written in MSAccess. The result was a validation of the detection procedure of the severity of change within the landscape.

Conclusions

The method of analyzing remotely sensed data by the comparison of digital orthophoto quadrangle (DOQ) aerial photographs and aerial videography to detect changes in landscape between a time when the disease is not present and a time when the disease was present is reported as a negative. It was determined that the two remote sensing mediums (i.e., DOQ and aerial videography) were incompatible with each other since pixel-to-pixel registration could not occur due to a permanent shift in the color bands. Further, the shift could not be accomplished without a skew of the dataset.

However, the method of detection of change in landscape between a time when the disease is not present and a time when the disease was present using like remote sensing mediums (i.e., videography to videography and Landsat to Landsat) was comparable with each other. The result in this research was positive.

The comparison of a digital orthophoto quadrangle (DOQ) frame prior to the disease outbreak with an aerial videography frame after the outbreak of a disease cannot be shown as a severity of change in landscape. However, the method of detection of change in landscape between a time when the disease is not present and a time when the disease was present using like remote sensing mediums was positive.

Summary

Although the mediums used varied, the instrument or procedure developed has produced data that are valid and reliable. Therefore, the procedure appears to be useful in judging the severity range of landscape in the setting of biomedical medical professionals and environmental specialist. Second, the outcome is not variable and can be used as an economical and fast tool of measurement to determine whether landscape change is directly attributable to an infectious land-borne disease. The ability to control any interfering variables is the key for using land-based detection analysis as a predictive measure. Even so, the severity of change methodology noted here should prove to be a useful adjunct to any health professional dedicated to the analysis of the environment with remote sensing if the data are recent and are used judiciously and repeatedly.

Health professionals or immunologists or environmental professionals are no different from many other industries in that they find themselves in the position of having progressively increasing costs of doing business coupled with decreasing revenues. The detection and prevention of disease must stay competitive, and the comparison of videography frames is a better solution than costly satellite analysis.

Future Research

An extension of this research would be to conduct a similar study with larger, more diverse areas of landscape. A large library of video frames taken prior to the outbreak of a known land-borne disease over known areas of outbreak would yield opportunities for

greater interpretations of correlations and would be ideal for instant 1-day analysis. Questions for future studies are: Can the measurement (severity of change) of landscape apply to other groups of interest? (These could include the clear cutting of trees by the U.S. Forest Service, drug (plant) growth by the U.S. Drug Enforcement Agency, or still-water distribution or erosion by the U.S. Corps of Engineers).

Other researchers may find this severity of change methodology for the measurement of landscape utilizing aerial videography useful for judging the service quality of their own remote sensing techniques. If this methodology is validated among users with specialized knowledge and skills, its use could be extended to information professionals with a more generalized scope of work. While research on remote sensing is very much in evidence in the noncommercial sector, it is much less developed in commercial sectors of industry. Using this methodology in a more general population of information specialists may spur investigation of the techniques used here by other for-profit organizations. The predictive value of this methodology should also be tested more extensively with populations having greater autonomy, stability in and control over their working environment. It may be inferred that the decrease in uncertainty inherent in such a situation leads to concomitant decrease in variability. In situations where many variables may be controlled, a path analysis would prove to be a useful tool for predicting future use in a given population.

Probably the most intriguing adjunctive shock to arise from this research is that occurring as part of the nature of the portal. The design of the seamless interface (a) allows one to observe and determine whether the mathematical conclusions are valid, and (b)

allows one to view the image as the level of severity occurs. This idea is utilized by using multiple frames imbedded in popular movie software (i.e., RealPlayer, Apple QuickTime, etc.) by displaying the severity levels in “real” time. This could allow immediate conclusions to be drawn over an area analysis.

The answers uncovered in this research lead logically and inevitably to more questions about severity of change analysis from an aerial platform. And how do we, as information and remote sensing professionals, view the remote sensing products of our colleagues who are in the business of serving us? Do we hold them to the same standard of judgment for remote sensing services that we use to those we offer to the remote sensing consumer? Is it a high or low standard? Remote sensing providers should have as their objective a continual monitoring and adjustment of the product so as to maintain a high level of quality. Using this methodology for the detection of the severity of change in landscape appears to be an ideal way to do this.

APPENDIX A

PROGRAM UTILITY TO CONVERT AND CREATE INTENSITY DATA

A UNIX operating system.jpg utility converter was used by a computer software analysis to decode the .jpg graphic format to a Portable Picture Mat (PPM) file. The PPM format is binary and allows the programmer to create unique files by writing output through a C program. Two special C programs were written to convert the PPM files to ASCII format and from ASCII back to PPM format. The output results in four files being produced. The first program displays ASCII formatting by the row where the pixel is located, the column where the pixel is located, the intensity range (0-255) of the R or Red value of the pixel, the intensity range (0-255) of the G or Green value of the pixel, and the intensity range (0-255) of the B or Blue value of the pixel and is noted in a numeric format. The first two files contain the ASCII row-column-R-G-B values of the video images. The third file contains the ASCII row-column-R-G-B values of difference (I.e.-change in intensity) between the video scenes. And the last is a converted file from ASCII back to the PPM format. The program is as follows:

```
#include<stdio.h>
#include<string.h>
#include<stdlib.h>

void doit (char *fn1, char *fn2);
int main()
{
    char fname1[120];
    char fname2[120];

    printf ("Enter name of first ppm file: ");
    scanf ("%s",fname1);

    /*
    // both args must be ppm or pgm files...
    */
    if ((strstr(fname1, ".ppm") == NULL) && (strstr(fname1, ".pgm") == NULL)) {
```

```

    fprintf(stderr,"file %s didn't have .ppm or .pgm extension!\n", fname1);
    exit(1);
}

printf ("Enter name of second ppm file: ");
scanf ("%s",fname2);
if ((strstr(fname2,".ppm") == NULL) && (strstr(fname2,".pgm") == NULL)) {
    fprintf(stderr,"file %s didn't have .ppm or .pgm extension!\n", fname2);
    exit(1);
}

/*
// once we have the prefix names... doit!
*/
doit(fname1, fname2);
printf("done!\n");
}

void doit (char *fn1, char *fn2) {
    char p6[3];
    int range;
    int rows1,cols1;
    int rows2,cols2;
    int row,col,band;
    int ival1, ival2;
    FILE *halF1, *ppmF1;
    FILE *halF2, *ppmF2;
    FILE *halF3, *ppmF3;
    char hal_fn1[120], ppm_fn1[120];
    char hal_fn2[120], ppm_fn2[120];
    char hal_fn3[120], ppm_fn3[120];
    char line[120];
    int bands, bands2;

    /*
// open first input (ppm) and output (hal) files...
*/
    strcpy(ppm_fn1,fn1);
    strcpy(hal_fn1,fn1);
    strcpy(hal_fn1+strlen(fn1)-4,".hal");

    if ((ppmF1=fopen(ppm_fn1,"r"))==NULL) {
        fprintf(stderr,"Couldn't open ppm file: %s\n",ppm_fn1);
    }

```

```

    exit(1);
}
if ((halF1=fopen(hal_fn1,"w"))==NULL) {
    fprintf(stderr,"Couldn't open hal file: %s\n",hal_fn1);
    exit(1);
}

/*////////////////////
// open second input (ppm) and output (hal) files...
*/
strcpy(ppm_fn2,fn2);
strcpy(hal_fn2,fn2);
strcpy(hal_fn2+strlen(fn2)-4, ".hal");

if ((ppmF2=fopen(ppm_fn2,"r"))==NULL) {
    fprintf(stderr,"Couldn't open ppm file: %s\n",ppm_fn2);
    exit(1);
}
if ((halF2=fopen(hal_fn2,"w"))==NULL) {
    fprintf(stderr,"Couldn't open hal file: %s\n",hal_fn2);
    exit(1);
}

/*
// make sure that the two input ppm files start with 'P6'
// (3-band) or 'P5' (1-band)
*/
p6[0] = fgetc (ppmF1);
p6[1] = fgetc (ppmF1);
p6[2] = '\0';
printf("First P value: %s\n",p6);
if (strcmp(p6,"P6") == 0)
    bands = 3;
else if (strcmp(p6,"P5") == 0)
    bands = 1;
else {
    fprintf(stderr,"Hey, incorrect P value!!!\n");
    exit(1);
}

p6[0] = fgetc (ppmF2);
p6[1] = fgetc (ppmF2);
p6[2] = '\0';
printf("Second P value: %s\n",p6);

```



```

if (strcmp(p6,"P6") == 0) {
    bands2 = 3;
} else if (strcmp(p6,"P5") == 0) {
    bands2 = 1;
} else {
    fprintf(stderr,"Hey, incorrect P value!!!\n");
    exit(1);
}

if (bands != bands2) {
    fprintf(stderr,"Hey, bands don't match!!!\n");
    exit(1);
}

printf ("bands, bands2: %d %d\n",bands, bands2);

/*
// read rows, columns and range values from the first input ppm files
// skip the line if it starts with a '#'
*/
fscanf(ppmF1,"%s",line);
if (line[0] == '#') {
    while (fgetc (ppmF1) != '\n');
}
/*sscanf (line, "%d",&rows1);*/
fscanf (ppmF1, "%d",&rows1);
fscanf (ppmF1, "%d",&cols1);
fscanf (ppmF1, "%d",&range);

/*
// read rows, columns and range values from the first input ppm files
// skip the line if it starts with a '#'
*/
fscanf(ppmF2,"%s",line);
if (line[0] == '#') {
    while (fgetc (ppmF2) != '\n');
}
fscanf (ppmF2, "%d",&rows2);
fscanf (ppmF2, "%d",&cols2);
fscanf (ppmF2, "%d",&range);

/*
// check to make sure rows match...
*/

```

```

if ( (rows1 != rows2) || (cols1 != cols2) ) {
    fprintf(stderr,"Ack! not the same number of rows or columns!\n");
    fprintf(stderr,"R1,C1,R2,C2 %d %d %d %d\n",rows1,cols1,rows2,cols2);
    exit(1);
}
else {
    fprintf(stderr,"Rows is:%d\n",rows1);
    fprintf(stderr,"Cols is:%d\n",cols1);
    fprintf(stderr,"Rnge is:%d\n",range);
}

/*
// open the difference files (ppm & hal)...
*/
sprintf(hal_fn3, "diff.hal");
if (bands == 3)
    sprintf(ppm_fn3, "diff.ppm");
else
    sprintf(ppm_fn3, "diff.pgm");
if ((ppmF3=fopen(ppm_fn3,"w"))==NULL) {
    fprintf(stderr,"Couldn't open ppm file: %s\n",ppm_fn3);
    exit(1);
}
if ((halF3=fopen(hal_fn3,"w"))==NULL) {
    fprintf(stderr,"Couldn't open hal file: %s\n",hal_fn3);
    exit(1);
}

/*
// print row colum at the top of the hal files...
*/
fprintf(halF1,"%d %d\n",rows1,cols1);
fprintf(halF2,"%d %d\n",rows1,cols1);
fprintf(halF3,"%d %d\n",rows1,cols1);

/*
// write P6 at the top of the difference ppm file...
*/
if (bands == 3)
    fprintf(ppmF3,"P6\n%d %d\n%d\n",rows1,cols1,range);
else
    fprintf(ppmF3,"P5\n%d %d\n%d\n",rows1,cols1,range);

/*

```

```

// 1. read binary from the two input ppm files
// 2. find difference (is abs. value what is needed?)
// 3. write ascii to the hal files and binary to the difference ppm
*/
for (row = 0; row < rows1; row++) {
  if (row%100 == 0) {
    printf (" Row: %d\n",row);
  }
  for (col = 0; col < cols1; col++) {

fprintf(halF1,"%d %d ",row,col);
  fprintf(halF2,"%d %d ",row,col);
  fprintf(halF3,"%d %d ",row,col);

  for (band = 0; band < bands; band++) {
    ival1 = fgetc (ppmF1);
    ival2 = fgetc (ppmF2);

    fprintf (halF1,"%d ",ival1);
    fprintf (halF2,"%d ",ival2);
    fprintf (ppmF3,"%c",abs(ival2-ival1));
    fprintf (halF3,"%d ",abs(ival2-ival1));

  }
  fprintf(halF1,"\n");
  fprintf(halF2,"\n");
  fprintf(halF3,"\n");
}
}
fclose(ppmF1);
fclose(halF1);
fclose(ppmF2);
fclose(halF2);
fclose(ppmF3);
fclose(halF3);

printf ("output files: \n %s\n %s\n %s\n %s\n",hal_fn1, hal_fn2, ppm_fn3, hal_fn3);
}

```

APPENDIX B

NOTES FOR MATHEMATICAL PROCEDURE OF DIFFERENCE CONVERSION

Equals:	3.10938	1.64063	2.4375								
	1.64063	2.23438	1.3125								
	2.4375	1.3125	3.75								
We call the Variance / Covariance Matrix "S".											
In the next step, S must be inverted.											
Finding an inverse Matrix											
To obtain A^{-1} for any $n \times n$ matrix A for which A^{-1} exists, follow these steps.											
Step 1	From the augmented matrix $[A I_n]$, where I_n is the $n \times n$ identity matrix										
Step 2	Perform row transformations on $[A I_n]$ to get a matrix of the form $[I_n B]$.										
Step 3	Matrix B is A^{-1} .										
To confirm that two $n \times n$ matrices A and B are inverses of each other, it is sufficient to show that $AB = I_n$. It is not necessary to show also that $BA = I_n$.											
Finding the inverse of a 3×3 Matrix											
Find A^{-1} if $A =$											
				1	0	1					
				2	-2	-1					
				3	0	0					
Step 1 Write the augmented matrix $[A I_3]$											
				1	0	1	1	0	0		
				2	-2	-1	0	1	0		
				3	0	0	0	0	1		
Step 2 Since 1 is already in the upper left-hand corner as desired, begin by using the row transformation that will result in 0 for the first element in the second row. Multiply the element of the first row by -2, and add the result to the second row.											
				1	0	1	1	0	0		
				2	-2	-3	-2	1	0		
				3	0	0	0	0	1		
To get 0 for the first element in the third row, multiply the elements of the first row by -3 and add to the third row.											
				1	0	1	1	0	0		
				0	-2	-3	-2	1	0		
				0	0	-3	-3	0	1		
To Get 1 for the second element in the second row, multiply the elements of the second row by -1/2											
				1	0	1	1	0	0		
				0	1	3/2	1	-1/2	0		
				0	0	-3	-3	0	1		
To get 1 for the third element in the third row, multiply the elements of the third row by -1/3.											
				1	0	1	1	0	0		
				0	1	3/2	1	-1/2	0		
				0	0	1	1	0	-1/3		
To get 0 for the third element in the first row, multiply the elements of the third row by -1 and add to the first row											
				1	0	0	1	0	1/3		
				0	1	3/2	1	-1/2	0		
				0	0	1	1	0	-1/3		
To get 0 for the third element in the second row, multiply the elements of the third row											

I	J	dr	dg	db						
1	1	5	3	3						
2	1	3	3	4						
3	1	1	2	2						
4	1	2	4	0						
1	2	1	4	2						
2	2	0	0	1						
3	2	3	3	5						
4	2	2	4	3						
	Sum:	17	23	20						
	Ave:	2.125	2.875	2.5						
Difference Set = Dif										
Step 6: Treat each of the three differences for each pixel as a vector of length 3.										
This makes each pixel's difference values into a 3 row by 1-column vector.										
For the "dr" average, add all the "dr" values and divide by the total number of pixels.										
I used 4 rows with 2 columns in the example above, so the denominator is 8.										
The sums and averages are shown at the bottom of the (Dif) set.										
I also haven't figured out how to draw matrix brackets around the d bar vector below.										
		2.125								
d bar =		2.875								
		2.5								
The "d bar" vector above is an average vector of differences for the whole scene.										
Step 7 is on the next worksheet. That step calculates the variance / covariance matrix.										

	2					4	8	6				
d bar =	4	times	2	4	3	equals	8	16	12			
	3						6	12	9			
Step 7 a.	For each pixel, construct a 3 x 3 matrix											
	The example above used the STEP 6 Differences for the pixel for Row 4, Col 2 data											
	The formulas in column H, which use \$ signs, can be copied and pasted into columns I and J.											
	Described verbally below is the pattern of multiplication to use on every difference pixel in step 6.											
	Row 1											
	Row 2	X	Col 1	Col 2	Col 3	equals	R1 X C1	R1 X C2	R1 X C3			
	Row 3						R2 X C1	R2 X C2	R2 X C3			
							R3 X C1	R3 X C2	R3 X C3			
The steps described above are to be performed on the step 6 difference matrix below:												
I	J	dr	dg	db		5				25	15	15
1	1	5	3	3		3	5	3	3	15	9	9
2	1	3	3	4		3				15	9	9
3	1	1	2	2								
4	1	2	4	0								
1	2	1	4	2								
2	2	0	0	1								
3	2	3	3	5								
4	2	2	4	3								
To illustrate how to automate this, I selected and copied the 3 values (5,3,3) for Row 1 Col 1, in C14:E14 and												
	First I pasted them the same order into H14:J14, and then											
	Selected cells G13:G15, used Edit, Paste Special, and Transpose to paste them vertically.											
	These kinds of things can be automated with macros or used to show programmers the patterns.											
	In cell K13 I typed the formula =G13*H\$14 and stretched it downward two cells with its fill handle.											
	I stretched the K13 formula cell's fill handle to the right two cells, then down two cells again.											
	The K13 formula was checked by hand in all nine cells, K13:M15 and was shown to work.											
	The formula works with the sideways "T" pattern used above.											
	The 3x3 matrix in K13:M15 is repeated for all the pixel differences, of step 6.											
This ends step 7a.												
	Step 7 b is on the next worksheet or tab. The following notes precede that sheet:											
	The way the pattern is set up, the general formula in the upper left corner											
	of each 3x3 matrix becomes =G(n)*H\$(n+1) meaning that											
	pasting it in a "T" shape makes the line number after the asterisk one greater than											
	that before the asterisk, which contains the top row of that pixel.											

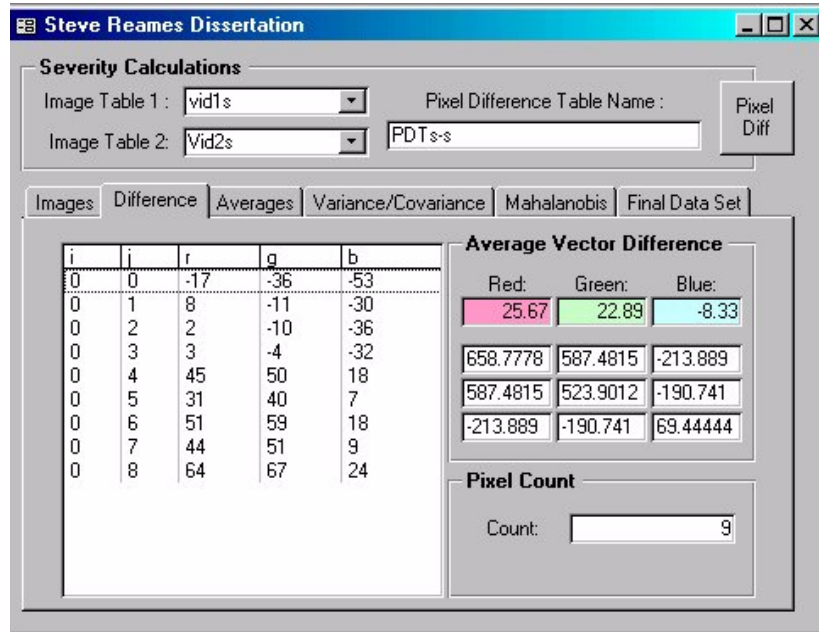
5				25	15	15
3	5	3	3	15	9	9
3				15	9	9
3				9	9	12
3	3	3	4	9	9	12
4				12	12	16
3				9	9	12
3	3	3	4	9	9	12
4				12	12	16
2				4	8	0
4	2	4	0	8	16	0
0				0	0	0
1				1	4	2
4	1	4	2	4	16	8
2				2	8	4
0				0	0	0
0	0	0	1	0	0	0
1				0	0	1
3				9	9	15
3	3	3	5	9	9	15
5				15	15	25
2				4	8	6
4	2	4	3	8	16	12
3				6	12	9
			Sum:	61	62	62
				62	84	68
				62	68	80
			Sum / 8:			
				7.625	7.75	7.75
				7.75	10.5	8.5
				7.75	8.5	10

7 b Sum/8:	7.625	7.75	7.75		The 3x3 sum/8 pixels matrix from the previous step, recopied.				
	7.75	10.5	8.5		If there are 10,000 rows by 20 columns, the dividend is 200,000 not 8.				
	7.75	8.5	10						
then from step 6 we multiply d bar with d bar transposed:									
	2.125				4.515625	6.109375	5.3125		
d bar =	2.875	2.125	2.875	2.5	6.109375	8.265625	7.1875		
	2.5				5.3125	7.1875	6.25		
Subtracting the pink 3 x 3 matrix in F5:H7 from the black matrix in B1:D3									
gives us the Variance / Covariance Matrix .									
7 b Sum / 8:	7.625	7.75	7.75						
	7.75	10.5	8.5						
	7.75	8.5	10						
minus									
d bar transpose:	4.51563	6.10938	5.3125						
	6.10938	8.26563	7.1875						
	5.3125	7.1875	6.25						
Equals:	3.10938	1.64063	2.4375						
	1.64063	2.23438	1.3125						
	2.4375	1.3125	3.75						
We call the Variance / Covariance Matrix "S".									
In the next step, S must be inverted.									

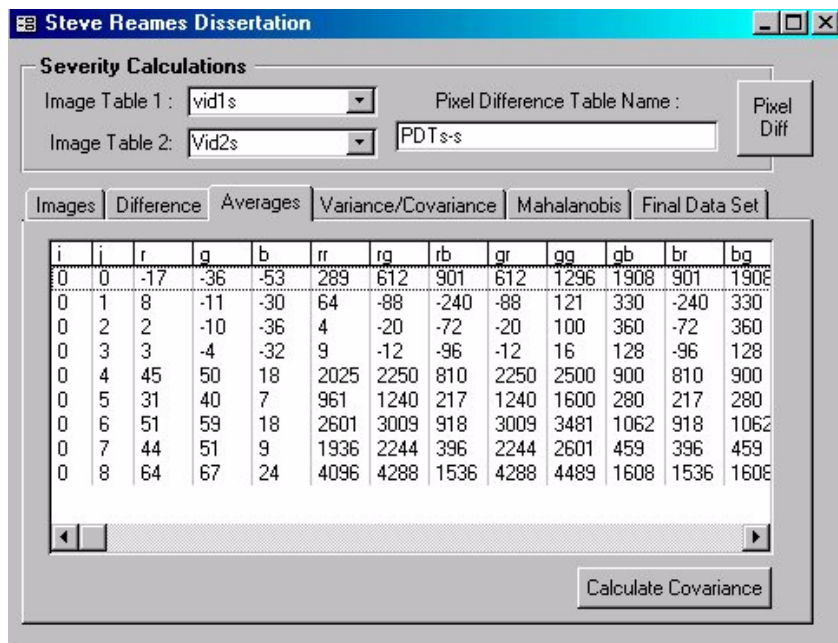
Equals:	3.10938	1.64063	2.4375																	
	1.64063	2.23438	1.3125																	
	2.4375	1.3125	3.75																	
We call the Variance / Covariance Matrix "S".																				
In the next step, S must be inverted.																				
Finding an inverse Matrix																				
To obtain A^{-1} for any $n \times n$ matrix A for which A^{-1} exists, follow these steps.																				
Step 1	From the augmented matrix $[A I_n]$, where I_n is the $n \times n$ identity matrix																			
Step 2	Perform row transformations on $[A I_n]$ to get a matrix of the form $[I_n B]$.																			
Step 3	Matrix B is A^{-1} .																			
To confirm that two $n \times n$ matrices A and B are inverses of each other, it is sufficient to show that $AB = I_n$. It is not necessary to show also that $BA = I_n$.																				
Finding the inverse of a 3×3 Matrix																				
Find A^{-1} if $A =$																				
				1	0	1														
				2	-2	-1														
				3	0	0														
Step 1 Write the augmented matrix $[A I_3]$																				
				1	0	1	1	0	0											
				2	-2	-1	0	1	0											
				3	0	0	0	0	1											
Step 2 Since 1 is already in the upper left-hand corner as desired, begin by using the row transformation that will result in 0 for the first element in the second row. Multiply the element of the first row by -2, and add the result to the second row.																				
				1	0	1	1	0	0											
				2	-2	-3	-2	1	0											
				3	0	0	0	0	1											-2R1+ R2
To get 0 for the first element in the third row, multiply the elements of the first row by -3 and add to the third row.																				
				1	0	1	1	0	0											
				0	-2	-3	-2	1	0											
				0	0	-3	-3	0	1											
To Get 1 for the second element in the second row, multiply the elements of the second row by -1/2																				
				1	0	1	1	0	0											
				0	1	3/2	1	-1/2	0											
				0	0	-3	-3	0	1											-1/2R2
To get 1 for the third element in the third row, multiply the elements of the third row by -1/3.																				
				1	0	1	1	0	0											
				0	1	3/2	1	-1/2	0											
				0	0	1	1	0	-1/3											
To get 0 for the third element in the first row, multiply the elements of the third row by -1 and add to the first row																				
				1	0	0	1	0	1/3											
				0	1	3/2	1	-1/2	0											
				0	0	1	1	0	-1/3											
To get 0 for the third element in the second row, multiply the elements of the third row																				

APPENDIX C

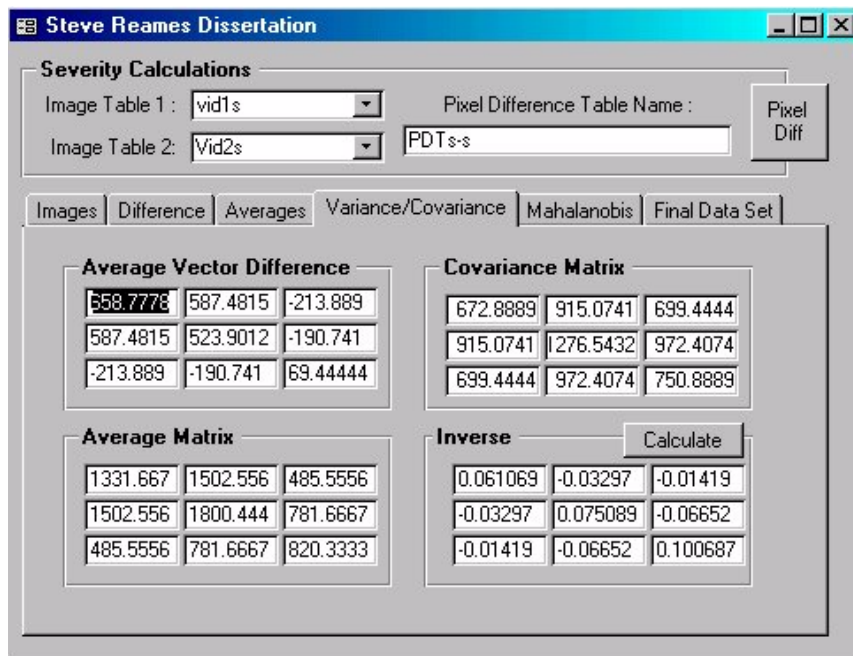
VISUAL BASIC WINDOWS AND PROGRAMMING CODE FOR DIFFERENCE CONVERSION



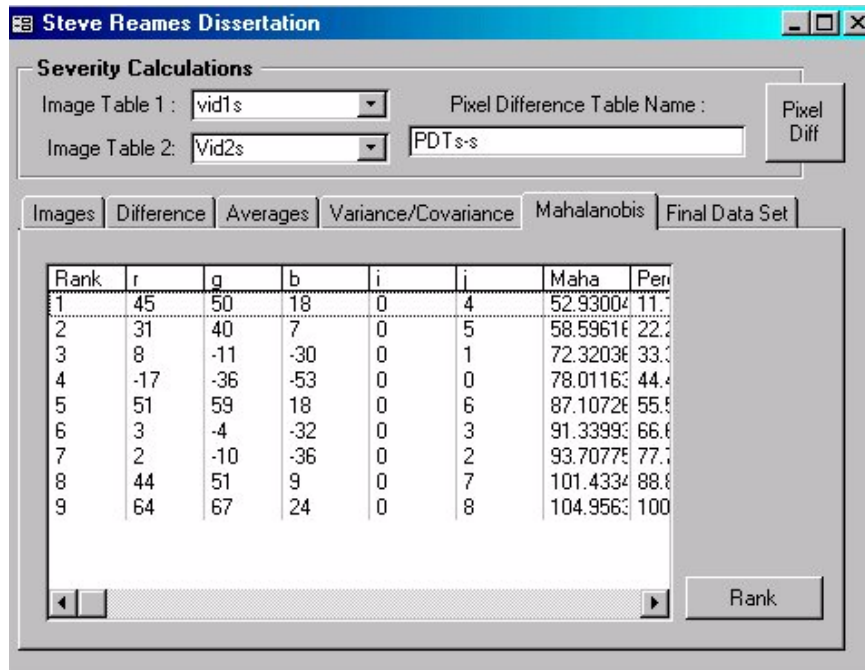
Pixel Difference Calculation Window



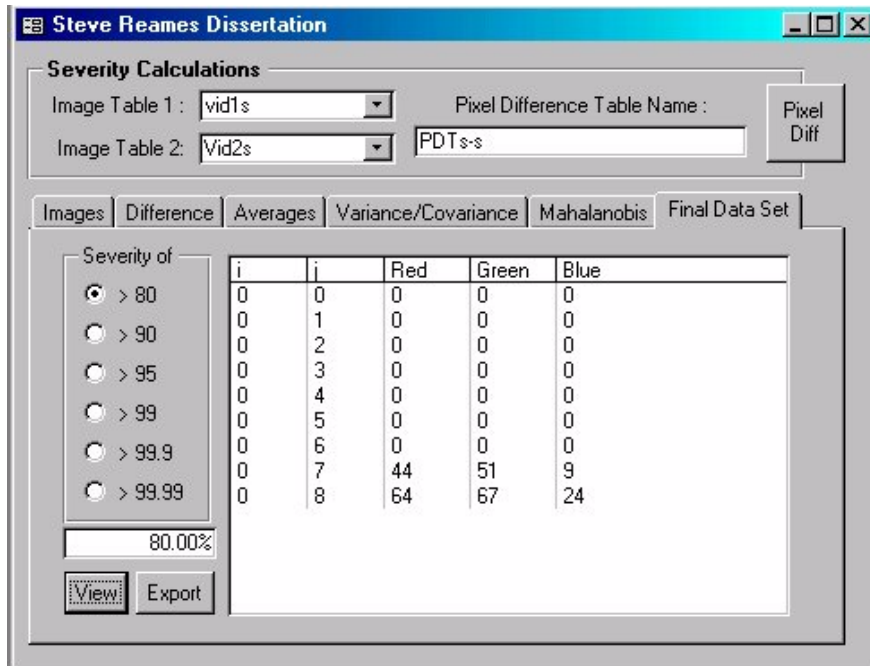
Averages Calculation Window



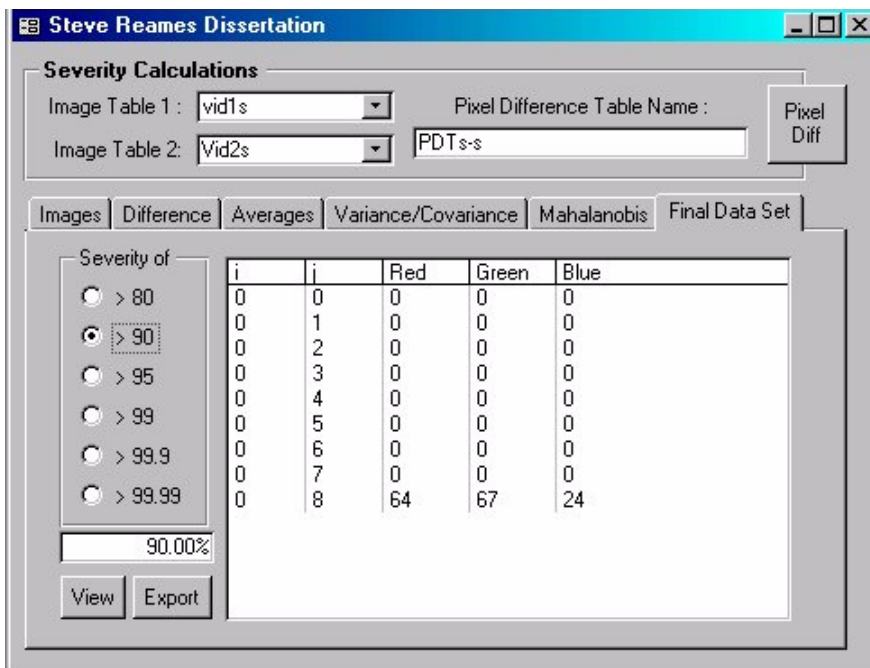
Variance/Covariance Window



Mahalanobis Window with Rank



Final Data Set (80 percentile) Window



Final Data Set (90 percentile) Window

Visual Basic Code written to MSAccess

Option Compare Database
Option Explicit

Private Sub CalcCova_Click()

Call MultDifTable ':Uncoment this line to calculate

Dim multsql As String

multsql = "SELECT TOP 10 [tmpmultdif].* FROM [tmpmultdif] ORDER BY
[tmpmultdif].i,[tmpmultdif].j;"

 Debug.Print multsql

 ' check to skip preview values

 Me.VarianceList.RowSource = multsql

 Me.VarianceList.Requery

 Me.srr = DSum("rr", "tmpmultdif") / Me.PixelCount

 Me.srg = DSum("rg", "tmpmultdif") / Me.PixelCount

 Me.srb = DSum("rb", "tmpmultdif") / Me.PixelCount

 Me.sgr = DSum("gr", "tmpmultdif") / Me.PixelCount

 Me.sgg = DSum("gg", "tmpmultdif") / Me.PixelCount

 Me.sgb = DSum("gb", "tmpmultdif") / Me.PixelCount

 Me.sbr = DSum("br", "tmpmultdif") / Me.PixelCount

 Me.sbg = DSum("bg", "tmpmultdif") / Me.PixelCount

 Me.sbb = DSum("bb", "tmpmultdif") / Me.PixelCount

 DoCmd.GoToControl ("Page4")

End Sub

Private Sub Command126_Click()

Dim mahasql As String

Call MakeMahaTable

Me.MahaList.RowSource = "Select tmpmaha.* FROM tmpmaha;"

Me.MahaList.Requery

End Sub

Private Sub Command127_Click()

Dim dblocal As Database

Dim rst1 As Recordset

Dim SS As String

Dim PixelCount As Integer

Dim RankNo As Integer

Dim SS2X As String

'Call MakeMahaTable :NO NEED TO RECALCULATE

Set dblocal = DBEngine.Workspaces(0).Databases(0)

Set rst1 = dblocal.OpenRecordset("SELECT tmpmaha.r, tmpmaha.g ,tmpmaha.b, tmpmaha.i, tmpmaha.j, tmpmaha.Mahalanobis FROM tmpmaha ORDER BY tmpmaha.Mahalanobis;")

rst1.MoveLast

'PixelCount = rst1.RecordCount :GRAB FROM FORM INSTEAD

rst1.MoveFirst

DoCmd.RunSQL "Delete tmpRankedMaha.* FROM tmpRankedMaha;"

SS = "INSERT INTO [tmpRankedMaha] (rank,r,g,b,i,j,maha,percentage) VALUES ("

RankNo = 1

Do While Not rst1.EOF ' Start the loop.

SS2X = SS & RankNo & "," & rst1.Fields(0) & "," & rst1.Fields(1) & "," & rst1.Fields(2) & "," & rst1.Fields(3) & "," & rst1.Fields(4) & "," & rst1.Fields(5) & "," & RankNo / Me.PixelCount & ");"

'Debug.Print SS2X :Display entry only if it was added

dblocal.Execute (SS2X)

rst1.MoveNext

RankNo = RankNo + 1

Loop

rst1.Close

Me.MahaList.ColumnCount = 8

Me.MahaList.RowSource = "SELECT TOP 10 tmpRankedMaha.* FROM tmpRankedMaha;"

Me.MahaList.Requery

rst1.Close

'dblocal.Close

End Sub

Function MakeMahaTable()

Dim makemaha As String 'Original Ranking Method using increments on a query see global module function "getNextcounter"

```
makemaha = "SELECT [" & Me.it2create & ".i, [" & Me.it2create & ".j, [" & _
    Me.it2create & ".r, [" & _
    Me.it2create & ".g, [" & _
    Me.it2create & ".b, " & _
    "[r]*([r]* " & Me.inv11 & " +[g]* " & Me.inv21 & " +[b]* " & Me.inv31 & ") AS
rval," & _
    "[g]*([r]* " & Me.inv21 & " +[g]* " & Me.inv22 & " +[b]* " & Me.inv23 & ") AS
gval," & _
    "[b]*([r]* " & Me.inv13 & " +[g]* " & Me.inv23 & " +[b]* " & Me.inv33 & ") AS
bval," & _
    "[rval]+[gval]+[bval] AS Mahalanobis INTO tmpmaha " & _
    "FROM [" & Me.it2create & "];"
```

DoCmd.RunSQL makemaha

End Function

Function MultDifTable()

Dim makemult As String

```
makemult = "SELECT [" & Me.it2create & ".i, [" & Me.it2create & ".j, [" & Me.it2create
& ".r, [" & Me.it2create & ".g, [" & Me.it2create & ".b, " & _
    "[r]*[r] AS rr, [r]*[g] AS rg, [r]*[b] AS rb, " & _
    "[g]*[r] AS gr, [g]*[g] AS gg, [g]*[b] AS gb, " & _
    "[r]*[b] AS br, [b]*[g] AS bg, [b]*[b] AS bb, " & _
    "([rr]+[rg]*2+[gg]+[gb]*2+[bb]+[br]*2)/" & Me.PixelCount & _
    " AS Avgrgb Into tmpmultdif FROM [" & Me.it2create & "];"
Debug.Print makemult
DoCmd.RunSQL makemult
```

End Function

Private Sub Command148_Click()

```
Me.SeverityList.RowSource = "SELECT tmpRankedMaha.i, tmpRankedMaha.j,
IIf([percentage]<=" & Me.Severity & ",0,[r]) AS Red, IIf([percentage]<=" & Me.Severity
& ",0,[g]) AS Green, IIf([percentage]<=" & Me.Severity & ",0,[b]) AS Blue FROM
tmpRankedMaha ORDER BY tmpRankedMaha.i, tmpRankedMaha.j;"
```

Me.SeverityList.Requery

End Sub

Private Sub CreatePixeldif_Click()

On Error Resume Next

Dim dblocal As Database

```
Dim rstImg1 As Recordset
Dim rstImg2 As Recordset
```

```
Dim strsql2 As String
Dim strsql1 As String
Dim SS As String
Dim SS2X As String
Dim r As Integer
Dim b As Integer
Dim g As Integer
'Dim A(1, 1) As Integer
```

```
Set dblocal = DBEngine.Workspaces(0).Databases(0)
Set rstImg1 = dblocal.OpenRecordset("SELECT " & Me.it1 & ".* FROM " & Me.it1 & "
ORDER BY " & Me.it1 & ".i, " & Me.it1 & ".j")
Set rstImg2 = dblocal.OpenRecordset("SELECT " & Me.it2 & ".* FROM " & Me.it2 & "
ORDER BY " & Me.it2 & ".i, " & Me.it2 & ".j")
```

```
rstImg1.MoveLast
rstImg2.MoveLast
rstImg1.MoveFirst
rstImg2.MoveFirst
```

```
SS = "CREATE TABLE [" & Me.it2create & "] (i INTEGER, j INTEGER, r
INTEGER, g INTEGER, b INTEGER);"
dblocal.Execute ("Drop Table [" & Me.it2create & ");")
dblocal.Execute (SS)
```

```
SS = "INSERT INTO [" & Me.it2create & "] (i,j,r,g,b) VALUES("
```

```
Do While Not rstImg1.EOF ' Start the loop.
```

```
If rstImg1.Fields(0) = rstImg2.Fields(0) And rstImg1.Fields(1) = rstImg2.Fields(1)
Then
```

```
r = rstImg1.Fields(2) - rstImg2.Fields(2)
g = rstImg1.Fields(3) - rstImg2.Fields(3)
b = rstImg1.Fields(4) - rstImg2.Fields(4)
```

```
SS2X = SS & rstImg1.Fields(0) & "," & rstImg1.Fields(1) & "," & r & "," & g & ","
& b & ");"
dblocal.Execute (SS2X)
```

```
Else
MsgBox "Check Datasets", vbCritical, "Out Of Sink"
```

```

Exit Sub

End If
rstImg1.MoveNext
rstImg2.MoveNext
Loop
Me.AVD.RowSource = "Select [" & Me.it2create & "].* From [" & Me.it2create &
"];
Me.AVD.Requery
rstImg1.Close
rstImg2.Close

DoCmd.GoToControl ("Page2")
End Sub

Private Sub CreateTextFile_Click()

On Error Resume Next
Dim Filename As String
Dim dblocal As Database
Dim rstsev As Recordset
Dim rstImg1 As Recordset
Dim RetVal
Dim strsqlsev As String
Dim strsqlimg As String

Dim SS As String
Dim SS2X As String

Set dblocal = DBEngine.Workspaces(0).Databases(0)
Set rstsev = dblocal.OpenRecordset("SELECT tmpRankedMaha.i, tmpRankedMaha.j,
Iif([percentage]<=" & Me.Severity & ",0,[r]) AS Red, Iif([percentage]<=" & Me.Severity
& ",0,[g]) AS Green, Iif([percentage]<=" & Me.Severity & ",0,[b]) AS Blue FROM
tmpRankedMaha ORDER BY tmpRankedMaha.i, tmpRankedMaha.j;")
Set rstImg1 = dblocal.OpenRecordset("SELECT " & Me.it1 & ".* FROM " & Me.it1 & "
ORDER BY " & Me.it1 & ".i, " & Me.it1 & ".j")

Filename = InputBox("Enter File Name:", "Create File")
Filename = "C:\REAMES\" & Filename & ".txt"

Open Filename For Output As #1 ' Open file for output.

rstImg1.MoveLast
rstsev.MoveLast

```

```
rstImg1.MoveFirst
rstsev.MoveFirst
```

```
Do While Not rstsev.EOF ' Start the loop.
```

```
  If rstsev.Fields(2) + rstsev.Fields(3) + rstsev.Fields(4) = 0 Then
    Print #1, rstsev.Fields(0); Spc(1); rstsev.Fields(1); Spc(1); rstsev.Fields(2); Spc(1);
rstsev.Fields(3); Spc(1); rstsev.Fields(4)
  Else
    Print #1, rstsev.Fields(0); Spc(1); rstsev.Fields(1); Spc(1); rstImg1.Fields(2); Spc(1);
rstImg1.Fields(3); Spc(1); rstImg1.Fields(4)
  End If
  rstsev.MoveNext
  rstImg1.MoveNext
Loop
```

```
  Close #1 ' Close file.
```

```
  rstsev.Close
  rstImg1.Close
 RetVal = Shell("C:\WINDOWS\notepad.exe " & Filename, 1) ' Run Calculator.
```

```
End Sub
```

```
Private Sub it2_AfterUpdate()
End Sub
```

```
Function SumLP(X As Integer) As Currency
  'Dim LedTrans As Integer
  'Dim ni As Integer
  '
  ' LedTrans = Me.AVD.ListCount - 1
  ' If LedTrans > 0 Then
  ' For ni = 1 To LedTrans
  '   SumLP = SumLP + Me.VarianceList.Column(X, ni)
  ' Next
  ' SumLP = SumLP / LedTrans
  ' End If
End Function
```

```
Private Sub Command1_Click()
' Get a normal random 3x3 matrix
Dim A As Variant
Dim b As Variant
```

```

A = randn(3)
Call asignval(A, Me.cm11, Me.cm12, Me.cm13, Me.cm21, Me.cm22, Me.cm23,
Me.cm31, Me.cm32, Me.cm33)
' Find its inverse matrix
b = inv(A) 'Using MatrixVB on mmatrix.dll
Call asigninv(b)
DoCmd.GoToControl ("Page5")
End Sub
Private Sub asignval(mat, a1, a2, a3, b1, b2, b3, c1, c2, c3 As Long)

    mat.r2(1, 1) = a1
    mat.r2(1, 2) = a2
    mat.r2(1, 3) = a3
    mat.r2(2, 1) = b1
    mat.r2(2, 2) = b2
    mat.r2(2, 3) = b3
    mat.r2(3, 1) = c1
    mat.r2(3, 2) = c2
    mat.r2(3, 3) = c3
End Sub

Private Sub asigninv(bmat As Variant)

    Me.inv11 = bmat.r2(1, 1)
    Me.inv12 = bmat.r2(1, 2)
    Me.inv13 = bmat.r2(1, 3)
    Me.inv21 = bmat.r2(2, 1)
    Me.inv22 = bmat.r2(2, 2)
    Me.inv23 = bmat.r2(2, 3)
    Me.inv31 = bmat.r2(3, 1)
    Me.inv32 = bmat.r2(3, 2)
    Me.inv33 = bmat.r2(3, 3)

End Sub

Private Sub Command130_Click()
On Error GoTo Err_Command130_Click

    DoCmd.Close

Exit_Command130_Click:
Exit Sub

```

```
Err_Command130_Click:
    MsgBox Err.Description
    Resume Exit_Command130_Click
```

```
End Sub
```

```
Private Sub Rank_Click()
    Dim dblocal As Database
    Dim rst1 As Recordset
```

```
    Dim SS As String
    Dim RankNo As Long
    Dim SS2X As String
```

```
    Call MakeMahaTable 'NO NEED TO RECALCULATE
    Set dblocal = DBEngine.Workspaces(0).Databases(0)
    Set rst1 = dblocal.OpenRecordset("SELECT tmpmaha.r, tmpmaha.g ,tmpmaha.b,
    tmpmaha.i, tmpmaha.j, tmpmaha.Mahalanobis FROM tmpmaha ORDER BY
    tmpmaha.Mahalanobis;")
```

```
    rst1.MoveLast
    'PixelCount = rst1.RecordCount :GRAB FROM FORM INSTEAD
    rst1.MoveFirst
```

```
    DoCmd.RunSQL "Delete tmpRankedMaha.* FROM tmpRankedMaha;"
    SS = "INSERT INTO [tmpRankedMaha] (rank,r,g,b,i,j,maha,percentage) VALUES ("
    RankNo = 1
```

```
    Do While Not rst1.EOF ' Start the loop.
```

```
        SS2X = SS & RankNo & "," & rst1.Fields(0) & "," & rst1.Fields(1) & "," &
        rst1.Fields(2) & "," & rst1.Fields(3) & "," & rst1.Fields(4) & "," & rst1.Fields(5) & "," &
        RankNo / Me.PixelCount & ");"
```

```
        'Debug.Print SS2X :Display entry only if it was added
        dblocal.Execute (SS2X)
        rst1.MoveNext
        RankNo = RankNo + 1
```

```
    Loop
```

```
    rst1.Close
    'Uncheck following to preview in list
```

```
Me.MahaList.ColumnCount = 8
Me.MahaList.RowSource = "SELECT TOP 10 tmpRankedMaha.* FROM
tmpRankedMaha;"
Me.MahaList.Requery
```

```
DoCmd.GoToControl ("Page6")
'dblocal.Close
End Sub
```

```
Private Sub SeverityChoices_AfterUpdate()
Select Case Me.SeverityChoices
Case 1
Me.Severity = 0.8
Case 2
Me.Severity = 0.9
Case 3
Me.Severity = 0.95
Case 4
Me.Severity = 0.99
Case 5
Me.Severity = 0.999
Case 6
Me.Severity = 0.9999
End Select
End Sub
```


REFERENCES

- Agmon, N., & Ahituv, N. (1987). Assessing data reliability in an information system. Journal of Management Information Systems, 4, 34-44.
- Ambrosia, V., Linthicum, K., Bailey, C., & Sebesta, T. (1989). Modeling Rift Valley fever (RVF) disease vector habitats using active and passive remote sensing systems. IGARSS '89 Remote Sensing. An Economic Tool for the Nineties, 2758-2760.
- Avery, T., & Berlin, G. (1985). Interpretation of aerial photographs (4th ed). New York: Macmillan.
- Badou, J. (1994). Les populations béninoises mobilisées. Le CRDI Explore, 22, 1.
- Beck, L., Rodriguez, M., Dister, S., Rodriguez, A., Rejrnkova, E., Ulloa, A., et al. (1994). Remote sensing as a landscape epidemiological tool to identify villages at high risk for malaria transmission. American Journal of Tropical Medicine and Hygiene, 51, 271-280.
- Beck, L., Wood, B., & Dister, S. (1995). Remote sensing and GIS: New tools for mapping human health. Geo Info Systems, 5(9), 32-37.
- Biel, L., & Landgrebe, D. (2001). Multispec© - A multispectral image data analysis system. School of Electrical and Computer Engineering and Lab for Applications of Remote Sensing (LARS). Retrieved February 1, 2002, from Perdue University School of Engineering Website: <http://www.ece.purdue.edu/~biehl/MultiSpec/description.html>
- Blank, H., & Rake, G. (1955). Viral and rickettsial diseases of the skin, eye and mucous membranes of man. Boston: Little, Brown..
- Brady, J. (1991). Seeing flies from Space. Nature, 351, 695.
- Burgdorfer, W. (1980). Spotted fever-group diseases. In J.H. Steele (Sec. A Ed.), CRC Handbook Series in Zoonoses: Vol. 4. Bacterial Rickettsial, and mycotic disease (pp. 279-304). Boca Raton, FL: CRC Press.

- Campbell, G., Dennis, D., Hayes, E., & Orloski, K. (2000). Surveillance for Lyme disease--United States, 1992-1998. Center for Disease Control, MMWR Surveillance Summaries, 49, (SS03), 1-11.). Retrieved February 1, 2002, from the U.S. Center of Disease control Website: <http://www.cdc.gov/mmwr/preview/mmwrhtml/ss4903a1.htm>
- Chwastek, J., & Dworak, T. (1990). Satellite remote sensing of industrial air pollution in the Cracow Special Protected Area. IEPTO, 10(6), 288-289.
- Clarke, K. C., Osleeb, J. R., Sherry, J. P., Meert, J. P., & Larsson, R. W. (1991). The use of remote sensing and geographic information systems in UNICEF's dracunculiasis (Guinea worm) eradication effort. Preventive Veterinary Medicine, 11, 229-235.
- Cowardin, L., & Myers V., (1974). Remote sensing for identification and classification of wetland vegetation. Manage, 38, 308-314.
- Cross, E., Sheffield, S., Perrine, R., & Pazzaglia, G. (1984). Predicting areas endemic for Schistosomiasis using weather variables and a Landsat database. Military Medicine, 149, 542-44.
- Cross, E., Tucker, C., Hyams, K. (1997). The use of AVHRR and weather data to detect the seasonal and geographic occurrence of *Phebotomus papatasi* in Southwest Asia. In Proceedings of the International Symposium on Computer Mapping in Epidemiology and Environmental Health Alexandria, VA: (pp.24-26).
- Dalke, R. (1937). The cover map in wildlife management. Wildlife Management, 1, 100-106.
- Davies, F., Kilelu, E., Linthicum, K., & Pegram, R. (1992). Patterns of Rift Valley fever activity in Zambia. Epidemiology and Infection, 108(1), 185-191.
- Eastman, J. (1992). IDRISI user's guide, Version 4.0 rev.1. Worcester, MA: Clark University Graduate School of Geography.
- ER-Mapper. (1995). ER-Mapper 5.0 reference. Earth Resource Mapping Pty.
- Estes, R. (1985). Remote sensing fundamentals. In R Holz (Ed.), The surveillant science (pp 12-27). New York: Wiley.
- Fuerstein, M. (1987). Partners in evaluation leadership. ASEAN Training Centre for Primary Health Care Development, 1(1), 34-35.

- Garner, M., & Grant, J. (1997). Using airborne videography to image the land. Center for Advanced Spatial Technologies (CAST) and the Arkansas Forestry Commission (AFC). Retrieved February 1, 2002, from CAST Arkansas Forestry Commission Website: <http://www.cast.uark.edu/cast/projects/video/ppframe.htm>
- Garrison, W., Alexander, R., Bailey, R., Dacey, M., & Marble, D. (1965). Data systems requirements for geographic research in scientific experiments for manned orbital flight. In Proceedings of the American Astronautical Society's Third Goddard Memorial Symposium (pp 326-327) Laurel, MD.
- George, T., Scorup, P. (1981, June), Reindeer range inventory: Use of winter Landsat imagery for stratification of digital classification. In Proceedings of Machine Processing of Remotely Sensed Data with Special Emphasis of Range, Forest and Wetland Assessment (pp. 416-418).
- Glass, G., Amerasinghe, F., Morgan, J., & Scott, T. (1994). Predicting *Ixo scapularis* abundance on white-tailed deer using geographic information systems. American Journal of Tropical Medicine and Hygiene, 51(Suppl.), 538-544.
- Glass, G., Schwartz, B., Morgan, J., Johnson, D., Noy, P., & Israel, E. (1995). Environmental risk factors for Lyme disease identified with geographic information systems. American Journal of Public Health, 85 (7), 944-948.
- Gonzalez, R., & Woods, R. (1992). Digital image processing. Reading, MA: Addison-Wesley.
- Graham, L. (1995, October). Developing geostatistical vegetation mapping methods. Texas Regional Institute for Environmental Studies (TRIES) (Technical Rep. No. 95-06. Huntsville, TX: Sam Houston State University.
- Grendon, M., & D'Onofrio, M. (2001). Data quality in the healthcare industry. Data Quality, 7(1). Retrieved February 1, 2002, from Data Quality Magazine Website: <http://www.dataquality.com/901GD.htm>
- Harden, V. (1990). Rocky Mountain spotted fever. Baltimore: The John S. Hopkins University Press.
- Hallum, C. (1998, January). A sampling and estimation strategy for monitoring the Sam Houston National Forest. Report on reassigned time. Huntsville, TX: Sam Houston State University.
- Hayes, R., Maxwell, E., Mitchell, J., & Woodzick, T. (1985). Detection, identification, and classification of mosquito larval habitats using remote sensing scanners in earth-orbiting satellites. Bulletin of the World Health Organization, 63(2), 361-374.

- Herrington, J. (1995). An update on Lyme disease. Health and Environmental Digest, 9, 29-32.
- Hill, G., & Kelly, T. (1987). Habitat mapping by Landsat for aerial census of kangaroos. Remote Sensing of the Environment 21: 53-60.
- Hugh-Jones, M. (1989). Applications of remote sensing to the identification of the habitats of parasites and disease vectors. Parasitology Today 5(8), 244-251.
- Hugh-Jones, M., Barre, N., Nelson, G., Wehnes, K., Warner, J., Gavin, J., et al. (1992). Landsat-TM identification of *Amblyomma variegatum* (Acari: Ixodidae) habitats in Guadeloupe. Remote Sensing Environment, 40, 43-55.
- Huh, O. (1991). Limitations and capabilities of the NOAA satellite advanced very high-resolution radiometer (AVHRR) for remote sensing of the earth's surface. Preventive Veterinary Medicine, 11, 167-184.
- Jackson, R., & Huete, A. (1991). Interpreting vegetation indices. Preventive Veterinary Medicine, 11, 185-200.
- Jensen, J. (1986). Digital image processing. A remote sensing perspective. Upper Saddle River, NJ: Prentice Hall.
- Jensen, J. (1996). Introductory digital image processing. A remote sensing perspective. Upper Saddle River, NJ: Prentice Hall.
- Kabel, R. (1990, July). Predicting the next map with spatial adaptive filtering. In Proceedings of the Fourth International Symposium in Medical Geography, University of East Anglia, Norwich, UK.
- Kerbes, R., & Moore, H. (1975, November 18-20). Use of current satellite imagery to predict the nesting success of lesser snow geese. Paper presented at Workshop on Remote sensing of Wildlife, Quebec, Canada.
- King, W., & Epstein, B. (1983). Assessing information value: An experimental study. Decision Sciences, 14, 34-35.
- Kitron, U., Bouseman, J., & Jones, C. (1991). Use of the ARC/INFO GIS to study the distribution of Lyme disease ticks in an Illinois county. Preventive Veterinary Medicine, 11, 243-248.
- Kitron, U., & Kazmierczak, J. (1997). Spatial analysis of the distribution of Lyme disease in Wisconsin. American Journal of Epidemiology, 145(6), 558-566.

- Klaas, E., Anderson, W., Frederick, R. (1978, October). Use of Landsat imagery for estimating food available to refuging lesser snow geese. In Proceedings Pecora IV Symposium: Application of Remote Sensing Data to Wildlife Management National Wildlife Federation Technical, Series 3.
- Kleinbaum, D., Kupper, L, Muller, K., & Nizam, A. (1998). Applied regression analysis and multivariable methods. Pacific Grove, CA: Duxbury Press.
- LaPerriere, A. Lent, P., Gassaway, W., & Nodler, F. (1980). Use of Landsat data for moose habitat analyses in Alaska. Wildlife Management, 44, 881-887.
- Lillesand, T., & Kiefer, R. (1987). Remote sensing and image interpretation. New York: Wiley.
- Linthicum, K., Bailey, C., Angleberger, D., Cannon, T., Logan, T., Gibbs, P, et al. (1991). Towards real-time prediction of Rift Valley fever epidemics of Africa. Preventive Veterinary Medicine, 11, 325-334.
- Linthicum, K., Bailey, C., Davies, F., & Tucker, C. (1987). Detection of Rift Valley fever viral activity in Kenya by satellite remote sensing imagery. Science, 235(4796), 1656-1659.
- Loffler, E, & Margules, C. (1980). Wombats detected from space. Remote Sensing of the Environment, 9, 47-56.
- Malone, J., Huh, O., Fehler, D., Wilson, P., Wilensky, D., Holmes, R., et al. (1994). Temperature data from satellite imagery and the distribution of schistosomiasis in Egypt. American Journal of Tropical Medicine and Hygiene, 50, (6): 714-722.
- Marble, D., Peuquet, D., Boule, A., Bryant, N., Calkins, H., Johnson, T., et al. (1983). Geographic information systems and remote sensing. In R. Collwell (Ed.), The manual of remote sensing (p. 923-958). Falls Church, VA: American Society of Photogrammetry and Remote Sensing.
- McDonald, G., Anacker, R., & Garjian, K. (1987). Cloned gene of Rickettsia rickettsii surface antigen: Candidate vaccine for Rocky Mountain spotted fever. Science, 235, 83-84.
- Martinka, C., Kendall, K. (1985, April 30-May 2). Grizzly bear habitat research in Glacier National Park, Montana. In Proceedings--Grizzly Bear Habitat Symposium, Missoula, MT (Gen. Tech. Rep. INT-207, p. 150-157, Ogden, UT: U.S. Department of Agriculture, Forest Service, Intermountain Research Station.

- Messina, J., Crews-Meyer, K., & Valdivia, G. (1998). The evaluation of preprocessing levels and classification techniques towards the better discrimination of suspended particulates. In Proceedings of the 1998 ASPRS Conference, Tampa, FL.
- Newhouse, V., Choi, K., Holman, R., Thacker, S., D'Angelo, L., & Smith, D. (1986). Rocky Mountain spotted fever in Georgia, 1961-75: Analysis of social and environmental factors affecting occurrence. Public Health Reports, 101(4), 419-428.
- Nicholson, M., & Mather, T. (1996). Methods for evaluating Lyme disease risks using geographic information systems and geospatial analysis. Journal of Medical Entomology, 33 (Suppl.), 711-720.
- Nondasuta, A. (1988). Nutrition studies in Thailand: The basic minimum needs guiding principles the foundation for quality of life. American Journal Clinical Nutrition, 48(5), 1214-8
- Openshaw, S., Charlton, M., Wymer, C., & Craft, A. (1987). Building a Mark I geographical analysis machine for the automated analysis of point patterns cancer and other spatial data. Research Rep., 12. Economic and Social Research Council Northern Regional Research Laboratory, University of Newcastle upon Tyne, Newcastle, UK.
- Ornsby, J., & Lunetta, R. (1987). Whitetail deer food availability maps from thematic mapped data. Photogrammetric Engineering and Remote Sensing, 53, 1585-1589.
- Pavlovsky, E. (1966). The natural modality of transmissible disease. Urbana: University of Illinois Press.
- Perry, B., Kruska, R., Kundert, K., Lessard, P., & Norval, R. (1991). Estimating the distribution and abundance of *Rhipicephalus appendiculatus* in Africa. Preventive Veterinary Medicine, 11, 261-268.
- Petris Technology, Inc. (2001). GeoReferenced Videography for Monitoring and Mapping Corridor Rights-of-Way. Retrieved February 1, 2002, from Petris Technology, Inc. Website: <http://www.petris.com/news/documents/wpUtility.shtml>
- Pope, K., Sheffnert, E., Linthicum, K., Bailey, C., Logan, T., Kasischke, E., et al. (1992). Identification of central Kenyan Rift Valley fever virus habitats with Landsat TM and evaluation of their flooding status with airborne imaging radar. Remote Sensing of Environment, 40, 185-196.
- Pyle, G. F. (1979). Applied medical geography. Washington, DC: V.H. Winston.

- Raoult, D., & Walker, D. (1990). *Rickettsia Rickettsii* and other spotted fever group rickettsiae (Rocky Mountain spotted fever and other spotted fevers). In Mandell, D. Gordon, & J. Bennett (Eds), Principles and practice of infectious diseases (3d ed). New York: Churchill Livingstone.
- Riley, H. (1977). Rocky Mountain spotted fever. Hospital Practice, April 1977, 51-57
- Roberts, D., Savage, H., Legters, L., Rodriguez, M., Rodriguez-Ramirez, A., Rejmankova, E., et al. (1991). Overview of field studies for tile application of remote sensing to tile study of malaria transmission in Tapachulah, Mexico. Preventive Veterinary Medicine, 11, 269-76.
- Scholten, H., & De Lepper, M. (1991). The benefits of the application of geographical information systems in public and environmental health. WHO Statistical Quarterly, 44(3), 53-62.
- Seddon, J. (1997). In pursuit of quality. Dublin: Oak Tree Press.
- Sersland, C. (1994). Assessing wetland vegetation with GPS-linked color video image mosaics. In Proceedings, 15th biennial workshop on color photography & videography in resource assessment, American Society for Photogrammetry & Remote Sensing (pp. 53-62). Bethesda, MD.
- Short, N. (1982). The Landsat tutorial workbook: Basics of satellite remote sensing; (NASA Reference Publication 1078), Washington, DC: Applied Information Sciences Branch, NASA/Goddard Space Flight Center, Greenbelt, Maryland.
- Slater, R. (1985). Survey of multispectral imaging systems for earth observations. Remote Sensing of the Environment, 17, 85-102.
- Solomonson, V. (1984). Landsat 4 and 5 status and results from thematic mapper data analyses. In Proceedings symposium on machine processing of remotely sensed data. Purdue University, West Lafayette, IN, 4B-3- 4B-12.
- Tomlinson, R. (1984). Geographic information systems - a new frontier. The Operational Geographer 5, 31-35.
- Twigg, L. (1990). Health base geographical information systems: Their potential examined in the light of existing data sources. Social Science and Medicine, 30, 1.
- U.S. Department of Health and Human Services (1985). Rocky Mountain spotted fever. (NIH Publication No. 85-400). Bethesda, MD: National Institute of Health.

- U.S. Department of Health and Human Services, Public Health Service, Centers for Disease Control. (1981). Annual summary 1980: Reported morbidity & mortality in the United States. Morbidity and Mortality Weekly Report, 29, No. 54.
- U.S. Department of Health and Human Services, Public Health Service, Centers for Disease Control. (1991). Summary of notifiable diseases, United States 1990. Morbidity and Mortality Weekly Report 39, No. 53.
- U.S. Geological Survey (1989). North American datum of 1983--map data conversion tables: U.S. Geological Survey Bulletin 1875, Vol. 3. Retrieved February 1, 2002, from U.S. Geological Survey Website: http://edcwww.cr.usgs.gov/glis/hyper/guide/usgs_doq
- U.S. Geological Survey. (1996). Standards for digital orthophotos, specifications: Reston, Virginia. Retrieved February 1, 2002, from U.S. Geological Survey Website: http://edc.usgs.gov/glis/hyper/guide/usgs_doq
- U.S. National Aeronautics and Space Administration, Ames Research Center, The Center for Health Applications of Aerospace Related Technologies (CHAART) (1991). Lyme Disease in Westchester, New York. Retrieved February 1, 2002, from the NASA CHAART Website: <http://geo.arc.nasa.gov/sge/health/projects/lymeny/lymeny.html>
- U.S. National Aeronautics and Space Administration, Ames Research Center, The Earth Sciences Division, Ecosystem Science and Technology Branch (2001). History of Landsat. Retrieved February 2, 2002, from the NASA Landsat Website: <http://geo.arc.nasa.gov/sge/landsat/landsat.html>
- United States Department of Agriculture, Agricultural Research Service. (1995). A view from above: Characterizing plant growth with aerial videography precision farming. Retrieved February 1, 2002, from the Precision Farming Website: http://www.precisionfarming.com/features/0199pf_yang.html
- Virginia Gap Analysis Project at the Fish and Wildlife Information Exchange 2001. Aerial videography. Retrieved November 29, 200, from the Virginia Tech University Website: <http://fwie.fw.vt.edu/www/video/>
- Wang, R., & Strong, D. (1996). Beyond accuracy: What data quality means to data consumers. Journal of Management Systems, 12, 5-34.
- Weber, D., & Walker, D. (1991). Rocky Mountain spotted fever. Infectious Disease Clinics of North America, 5, 19-35.
- Wellmer, H., & Jusatz, H. (1981). Geocological analysis of the spread of tick-borne encephalitis in Central Europe. Social Science and Medicine, 5D, 159-162.

- WHO/UNICEF Joint Monitoring Program. (1993). Water and sanitation monitoring system, guide for managers, Geneva, Switzerland: World Health Organization.
- Wood, B., Beck, L., Washino, R., Hibbard, K., & Salute, J. (1992). Estimating high mosquito-producing rice fields using spectral and spatial data. International Journal of Remote Sensing, 13(15), 2813-2826.
- Wood, B., Beck, L., Washino, R. Paichick, S., & Sebesta, P. (1991). Spectral and spatial characterization of rice field mosquito habitat. International Journal of Remote Sensing, 12(3), 621-626.
- Wood, B., Washino, R., Beck, L., Hibbard, K., Pitcairn, M., Roberts, D., et al. (1991). Distinguishing high and low anopheline-producing rice fields using remote sensing and GIS technologies. Preventive Veterinary Medicine, 11, 277-288.
- Work, E., & Gilmer, D. (1976). Utilization of satellite data for inventorying prairie ponds and lakes. Photogrammetric Engineering and Remote Sensing, 42, 685-694.
- Work, E., Gilmer, D., & Klett, A. (1974). Utility of ERTS for monitoring the breeding habitat of migratory waterfowl: NASA Goddard Space Flight Center. In Proceedings Symposium on the Earth Resources Technology Satellite-1, 3rd, in Washington, DC, December 1973, Proc., v.1, 1671-1685 and v.2, 102-113.
- Yang, C., Anderson, G. & Everitt, J. (1995). A view from above: Characterizing plant growth with aerial videography precision farming. Retrieved February 1, 2002, from the Precision Farming Website:
http://www.precisionfarming.com/features/0199pf_yang.html
- Young, T., Eby, J., Allen, H., Hewitt, M., & Dixon, K. (1988). Wildlife habitat analysis using Landsat and radio telemetry in a GIS with application to spotted owl preference for old growth. In Proceedings Geographic Information System, San Francisco, 12-14.
- Zwarenstein, M., Krige, D., & Wolff, B. (1991). The use of a geographical information system for hospital catchment area research in Natal/KwaZulu. South African Medical Journal, 80, 10.