# Use Of Small Format Aerial Photography In NPS Pollution Control Applications

**by**

**Youtong Fu**

Dissertation submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirement for the degree of
Doctor of Philosophy
in
Biological Systems Engineering

Approved

| | |
|---|---|
| _____ | _____ |
| Saied Mostaghimi, Chair | V. O. Shanholtz, Co-Chair |
| _____ | _____ |
| G. V. Loganathan | D. L. Gallagher |
| _____ | |
| Bill Carstensen | |

February 24, 2003
Blacksburg, Virginia

**Keywords**: 35-mm aerial slides, poultry facility, image processing, PCA, shape modeling, GIS, Remote Sensing, TMDL

# Use Of Small Format Aerial Photography In NPS Pollution Control Applications

by

Youtong Fu

S. Mostaghimi, Chair, V.O. Shanholtz. Co-Chair

Biological Systems Engineering

(ABSTRACT)

An automated procedure was developed to identify and extract confined poultry facilities from color 35-mm slide imagery collected by the United States Department of Agriculture/Farm Service Agency (USDA/FSA). The imagery is used by the USDA/FSA to monitor compliance with various farm support programs and to determine crop production acreage within a given county. The imagery is generally available for all counties within the state on an annual basis. The imagery, however, is not flown to rigid specifications as flight height, direction, and overlap can vary significantly. The USDA/FSA attempts to collect imagery with reasonably clear skies, as visual interpretations could be drastically impacted by cloudiness.

The goal of this study was to develop procedures to effectively utilize this imagery base to identify and extract poultry facilities using automated techniques based on image processing and GIS. The procedure involved pre-screening the slides to determine coverage, geopositioning to USGS quadrangle base, color scanning to convert slide image to a digital format and archiving each data file with a naming convention that would allow rapid retrieval in later analysis. Image processing techniques were developed for identifying poultry facilities based on spectral characteristics. GIS tools were used to select poultry facilities from an array of features with similar spectral characteristics. A training data set was selected from which the spectral characteristics of poultry facilities were analyzed and compared with background conditions. Poultry facilities were found to have distinguishable characteristics. Descriptive statistics were used to define the range of spectral characteristics encompassing poultry facilities. Thresholding analyses were

then utilized to eliminate all image features with spectral characteristics outside of this range. Additional analyses were made to remove noise in the spectral image due to the sun angle, line of sight of camera, variation in roof reflectance due to rust and/or aging, shading by trees, etc. A primary objective in these analyses was to enhance the spectral characteristics for the poultry facility while, at the same time, retaining physical characteristics, i.e. the spectral characteristic is represented by a single blue color with a high brightness value. The techniques developed to achieve a single blue color involved the use of Principal Component Analysis (PCA) on the red color band followed by RGB to Hue and RGB to Saturation analyses on the red and green color bands, respectively, from the resulting image. The features remaining from this series of analyses were converted into polygons (shape file) using ArcView GIS, which was then used to calculate the area and perimeter of each polygon.

The parameters utilized to describe the shape of a poultry house included width, length, compactness, length-width ratio, and polygon centroid analysis. Poultry facilities were found to have an average width of approximately 12.6m with a low standard deviation indicating that the widths of all houses were very similar. The length of poultry facilities ranged from 63m to 261m with and average length of 149m. The compactness parameter, which also is related to length and width, ranged from 30 to 130 with a mean value of approximately 57.

The shape parameters were used by ArcView GIS to identify polygons that represent poultry facilities. The order of selection was found to be compactness followed by length-width ratio and polygon centroid analysis. A data set that included thirty 35-mm slide images randomly selected from the Rockingham County data set, which contained over 2000 slides, was used to evaluate the automated procedure. The slides contained 182 poultry houses previously identified through manual procedures. Seven facilities were missed and 175 were correctly identified. Ninety-seven percent (97%) of existing poultry facilities were correctly identified which compares favorably with the 97 % accuracy resulted by manual procedures. .

The manual procedure described by Mostaghimi, et. al.(1999) only gave the center coordinates for each poultry facility. The automated procedure not only gives the center coordinate for each poultry building but also gives estimates for geometric parameters area, length and width along with an estimate of the capacity of building (i.e. number of birds), and waste load generated by birds including nutrient and bacteria content. The nutrient and bacteria load generated by each poultry facility is important information for conducting TMDL studies currently being developed for impaired Virginia streams. The information is expected to be very helpful to consultants and state agencies conducting the studies. Agricultural support agencies such as USDA/NRCS and USDA/FSA, Extension Service, consultants, etc. will find the information very helpful in the development of implementation plans designed to meet TMDL target water quality goals. The data also should be useful to Water Authorities for selection of appropriate treatment of water supplies and to county and local government jurisdictions for developing policies to minimize the degradation of water supplies.

## Acknowledgements

I would like to thank my committee chairman and co-chairman Dr. Saied Mostaghimi and Dr. V. Shanholtz for their many years of guidance, support, encouragement and paticence. I would like to thank Dr. Mostaghimi's continuing support and advice to get me through my studies. I want to thank my committee members for their support throughout my studies. I would like to thank Dr. Gene Yagow for sitting as a committee member for my final examinational and his suggestions and guidance.

None of my research would have been possible without the help of the crew at the GIS and Remote Sensing Project at BSE Department of Virginia Tech. Particularly, Jennifer Miller, later became Jennifer Miller-McClellan, who helped me from the beginning to end, and managed the project with me to deliver quality products. These products were very useful for my research. All of the staff in the BSE Department gave me a hand at one time or another, but special thanks go to Jan Car, Jeff Wynn, Kevin Brannan, Linda Altizer and Jack Davis.

I would like to thank Dr. Jim Kern and Mr. Byron Petrauskas for helping me to prepare the final defense. I appreciated your guys' some tough and useful questions so that I felt much more confident for final examination. Special thanks goes to Chris Harrel for helping me to format the draft of my dissertation.

I must not forget to thank Mary Shanholtz, for her kind support, encouragement and her hosptitality. I would like to thank Mrs. Harshbarger and the late Dr. Harshbarger for their encouragement to continue my study.

I would like to thank Mr. Phil McClellan, for his many years of support, encouragement and friendship. I am grateful for his thoughtful advice and kind helps for many years.

Again, I would thank Dr. Shanholtz, for his over ten years' help, advice and nonstop support of me even after he retired from Virginia Tech. It rarely happens in today's

world when a professor retires from University and still continues to support his former students. I am grateful for his endless care and encouragement. Thank you very much again, Dr. Shanholtz.

And, I want to thank my wonderful wife, Qun , who stood by me through it all. Thank you for your love and support. Thank you for giving me a beautiful daughter, Jenny. The love from you and Jenny is always driving me to move on and keep going for new levels of professional goals. I also want to thank my family in China, who had offered encouragement and support throughout.

# TABLE OF CONTENTS

# LIST OF ILLUSTRATIONS

# LIST OF TABLES

# 1. INTRODUCTION

## 1.1 Background

Agricultural non-point sources are blamed for making significant contributions to the degradation of both surface and ground waters in the United States (U.S.) (USEPA, 1996). Nutrients such as nitrogen and phosphorus from agricultural activities are carried by runoff into rivers and their downstream water bodies and can cause serious damage to water quality. Fertilizers and animal wastes are the typical sources of nutrients applied to agricultural lands. These nutrients can pollute the drinking water and potentially threaten human health (USEPA, 1996). Studies have shown that nitrate is the primary constituent in animal wastes, which can impair the quality of water and may be toxic to infants. Phosphorus in animal wastes can be attached to sediment and carried to streams by runoff and sediments (VADCR, 1998).

Nitrogen from confined animal sites could be a serious source of non-point source (NPS) pollution. Animal waste is a major source of bacteria and parasites. The applicable Virginia state standard specifies that the number of FC bacteria shall not exceed a maximum allowable level of 1,000 colony forming units (cfu) per 100 milliliters (ml) (Virginia State Law 9VAC25-260-170). However, fecal coliform counts up to 424,000 colonies/100ml were found during a survey of a local Virginia stream (Lipton, 1997). Dairy and beef cattle operations also are important contributors to NPS pollution. Runoff from the fields, where the animal waste is spread, can carry the eroded soil particles, along with the attached P and $NH_4$ and soluble nutrients, to downstream rivers and lakes. The waste in cattle grazing cool-season grasses can cause the nitrate toxicity and other N-related diseases. The $NO_3$ and microorganisms from the waste can be leached into groundwater or move to streams either overland or to sub-surface through soil, fractured rock or sinkholes.

The year 1997 has been called the year of the chicken (Lipton, 1997), because farmers were blamed for pollution in the upper Potomac River. The mass production of poultry is

not only threatening the local water quality, but also the water quality of downstream rivers which may be located hundred of miles from confined animal facilities. For instance, the poultry industry in West Virginia has affected the Potomac River and could eventually threaten the Chesapeake Bay (Lipton, 1997). West Virginia's chicken industry, along with other animal operations such as cattle farming, has already harmed the Potomac's headwaters (Lipton, 1997). In response to the increased demand for poultry, beef and pork around the world, the scale and number of confined animal facilities have expanded. For example, in West Virginia, the poultry industry has increased 200 percent over the last decade (Lipton, 1997). In Virginia, the poultry industry, which is mainly concentrated in the Shenandoah Valley, produces over 260 million birds a year (Lipton, 1997). In Moorefield, West Virginia, one million pounds of chicken are processed daily, which is sufficient for 3 million meals. The processed chicken is shipped to places as far away as Russia and Japan (Lipton, 1997). The increase in poultry production has resulted in a corresponding increase in waste to be disposed.

In response to public concern about the contamination of waste from livestock operations, the EPA and United States Department of Agriculture (USDA) developed a Unified National Strategy for Animal Feeding Operations (AFO) in March 1999, as part of the Clean Water Action Plan (Cleanwater, 1999). This strategy includes a national goal that "*All AFOs should develop and implement technically sound, economically feasible, and site-specific comprehensive nutrient management plans (CNMPs) to minimize impact on water quality and public health."* The release of the USDA-EPA draft strategy to control animal waste runoff was a key action item under then President Clinton's plan. In Virginia, with the provisions of the Clean Water Act (CWA), approximately 140 watersheds are identified as having potential problems or are impaired by animal waste (Mostaghimi, et. al., 1999).

For several decades, scientists and engineers have tried to understand the complexity of NPS pollution. Modeling is a very effective and powerful way to understand the cause and pathways of NPS pollution. Computer models have been developed to estimate the magnitude of NPS pollution. The lack of availability of data sources, however, is almost

always a “bottleneck” for obtaining successful model simulations. The data required by these models include land use/land cover, soil, topographic characteristics and accurate pollutant source and load information. For example, to simulate the pollutants from confined animal waste, the location of the confined animal facility, the surrounding land use/land cover, soil, management, type of animal waste and waste production rates are important inputs to NPS models.

Large volumes of information are typically required to estimate and/or to track NPS pollution from watersheds. Federal and state agencies such as EPA and the Virginia Soil and Water Conservation Division are beneficiaries of these types of information. Images obtained from remote sensing techniques can provide current information for updating land use/land cover map data layers (Steven, 1993; Jaggard and Clark, 1990). The information can be used to identify the relationship between potential NPS pollution and the pathways of pollutants. Satellite images and aerial photography are major data sources for remote sensing. Aerial photography is the oldest and most widely used method of remote sensing (Campbell, 1987). The two types of films typically used in aerial photography include large format (e.g. 23cm x 23cm) and small format, (e.g. 35mm and 75 mm) film. The large format film has been widely used for civil and military purposes. Small format film, which can be used by less expensive aircraft (Eastman, 1997), is becoming an alternative to conventional large format film for aerial surveys. Slides created by 35-mm films are generally developed as a final product for applications. For example, the USDA Farm Service Agency (USDA/FSA) has developed 35-mm slides for monitoring agricultural activities (Pane, 1997). In Virginia, 35-mm slides provide the most updated information for use by USDA Natural Resource Conservation Service (USDA/NRCS) (Pane, 1997). It also is the primary data source for landuse/land cover used by local USDA/NRCS offices in Virginia.

Currently, 35-mm slides can be economically scanned into a digital image at very high resolutions. Graphic and image processing software is available to analyze the slide images. This capability significantly reduces the time required to capture and compile land use/land cover information. The satellite images generally are too expensive for

public use. The 35-mm aerial photo slides developed by USDA/FSA have been used to extract confined animal sites for modeling NPS (Mostaghimi et. al., 1998). The 35-mm slides are generally available to public research institutions at no cost.

For the past ten years, Geographic Information Systems (GIS), Remote Sensing (RS), and Global Position Systems (GPS) have been used to collect, store and analyze data. Current GIS have very robust functions for processing spatial information important to identifying NPS pollution. GPS can provide accurate navigation (Stafford, et. al. 1994) by using satellite technology and computers to compute positions anywhere on earth. GPS can be used to locate watershed and field boundaries as well as data sampling positions.

The trend in livestock production has been toward fewer operators but more concentrated production systems (Lipton, 1997). As the demand for meat production increases, so does the number of confined animal facilities. Large companies contract with farmers to grow most poultry (e.g. broilers or turkeys). A greater concentration of poultry is located in relatively small lots or buildings. A high degree of industry integration, standardized rations, and complete confinement has made poultry operations much more manageable. The building shape and size can be used to estimate the number of birds being raised because of industry standards for bird density. Information on poultry population is necessary to estimate waste load, which can be used to give an estimate of potential nutrient and bacteria loads applied to the land. The location of the facility with physical and geographic information can be used to estimate potential loads delivered to nearby streams.

An existing research project being conducted in the Department of Biological Systems Engineering, Virginia Tech, Blacksburg, Virginia, is focusing on the identification of confined animal facilities, including poultry, by manual interpretation of 35-mm aerial slides (Mostaghimi, et. al., 1998). The manual procedure identifies the type and geographic location of each facility. The process is time consuming and could significantly benefit from reliable automated procedures.

## 1.2 Justification

Significant animal waste is being generated from individual confined animal facilities. The control of the pollution due to animal waste was one of the key goals of President Clinton's administration. These wastes contribute to NPS pollution problems when they are spread on adjacent land areas. Cost effective procedures for identifying and tracking confined animal operations are needed to develop effective control alternatives. Development of an automatic procedure for identifying confined animal facilities from 35-mm slide imagery could overcome current issues with manual image interpretation methods. An automated procedure is expected to result in the location of the confined animal sites, building footprints (area), an estimate of number of animal units, and animal waste production.

The rapid development of technologies such as GIS and RS facilitates the development of an automated procedure to identify confined animal facilities. Developments in GIS allow for rapid manipulation of complex data in different formats. Robust spatial analysis tools such as overlay, reclassification and neighborhood functions are available for performing a wide array of analysis. The development of image processing technology for remote sensing has enabled signature development and classification schemes for land use/land cover that support NPS modeling. A comprehensive literature review conducted for this study did not yield results that involve the adaptation of remote sensing technologies to 35mm imagery for NPS pollution studies. It is difficult to recognize confined animal facilities on the image, because they tend to be mixed with complex topographic, land use/land cover and other ground features. The fingerprint of a confined animal facility is very small when compared with land use/land cover features. Often it may best be described as a point distribution. The facilities usually do not have relatively large uniform spectral or texture distributions and are much more easily overshadowed by surrounding features. Developing the algorithms, combined with the technologies of GIS and image processing for shape modeling, could be useful for identifying confined animal facilities.

## 1.3 Goal and Objectives

The overall goal of this study was to utilize pattern recognition and GIS technologies to identify and extract confined poultry facilities from 35-mm slide imagery. The extracted information can be utilized to estimate potential stream pollutant loads from confined animal facilities. This study is expected to provide an efficient approach for obtaining important data for Total Maximum Daily Load (TMDL) analysis. The results of this research are expected to reduce the cost for data preparation and improve decision-makers' abilities to better understand the potential impacts of pollutants on receiving water systems and to develop effective alternative management strategies.

The specific objectives of the study were to:

- Develop spectral and geometric signatures to identify poultry houses;
- Develop pattern recognition algorithms to extract confined poultry facilities; and
- Evaluate the accuracy of the procedure using an independent data set.

# 2. LITERATURE REVIEW

## 2.1 Non-point Source Pollution

Non-point source pollution originates from diffuse land areas that intermittently carry pollutants to surface and ground waters. Soil erosion and sedimentation from land use activities, nutrients and organic materials from agricultural areas, and storm water from urban areas are the major sources of such pollution. The effect of NPS pollution includes beach closures, unsafe drinking water, fish kills and other severe environmental health problems that result in the expenditure of millions of dollars to restore and protect damaged areas (USEPA, 1997). NPS is the leading remaining cause of water quality problems in the U.S. (USEPA, 1997). Across the U.S., NPS pollution is the most pervasive cause of water quality problems (USEPA, 1997). It is the primary reason that approximately 40% of surveyed rivers, lakes and estuaries in the U.S. are not clean enough to meet basic uses such as fishing or swimming (USEPA, 1996).

NPS pollution differs from point source (PS) pollution, which is defined as those sources that enter a watershed from a specific point. For example, effluent from industrial and sewage treatment plants, pollutants from farm building or solid waste disposal sites are typical point sources of pollution. PS pollution is generally defined as discharge from a pipe such as from a municipal or industrial source.

Novotny and Chesters (1981) state that NPS pollution *"originates from non-point sources which are commonly regarded as those sources discharged to a watershed in a way that they depend upon the vagaries of the hydrologic cycle to transport them to the stream system."* It is the result of land use interaction with the hydrological transport system. Sedimentation occurs when wind or water runoff carries soil particles from an area, such as a farm field, and transports the eroded material to a water body. The random hydrologic process results in highly random patterns of NPS pollution. Significant research has shown that reducing the hydrological activities of land is often an effective strategy for controlling excessive pollution.

The U.S. is one of the leading agricultural nations in the world. With over 330 million acres of agricultural land under cultivation, American agriculture is noted worldwide for its tremendous productivity, quality, and efficiency in delivering goods to consumers. This productivity, however, is not without environmental cost. As readily acknowledge by farmers and environmentalists, when improperly managed, agricultural activities can directly affect water quality. Although NPS pollution can occur anytime, any land disturbing activity (such as forestry, grazing, septic systems, recreational boating, urban runoff, etc.), can potentially impact NPS.

The most recent National Water Quality Inventory concludes that agricultural NPS pollution is the leading source of water quality impacts to surveyed rivers and water (Table 2-1) (USEPA, 1996). NPS pollution from agriculture originates from several different sources: non-irrigated and irrigated croplands, orchard/fruit crop production, and animal production on pastureland and feedlots, and livestock facilities. Agriculture has been recognized as a significant source of nutrient pollution. Nutrient and bacteria are the two leading factors to degrade the water quality. The primary pollutants from non-irrigated and irrigated croplands are sediment, nutrients and pesticides. Orchard/fruit crop production areas are particularly significant sources of pesticides and nutrients with little or no sediment delivery. Runoff from barnyards, dairies and feedlots primarily contribute nutrients, organic matter, ammonia, fecal bacteria and other microorganisms to receiving waters. Livestock grazing freely along stream banks leads to direct contamination of streams with animal waste. Overgrazing of pastureland contributes sediment and nutrient pollution through runoff. The surface disruption and reduction in natural cover associated with overgrazing also increases the erosion of these lands.

Confined animal facilities allow farmers and ranchers to efficiently feed and maintain livestock. However, these areas can become major sources of animal waste. Runoff from poorly managed facilities can carry pathogens (bacteria and viruses), and oxygen-demanding substances that contaminate shellfish areas and cause other water quality problems. Some bacteria present in animal waste are extremely harmful to the human immune system.

**Table 2-1. Leading sources of water quality impairment related to human activities.**

| Rank | Rivers | Lakes | Estuaries |
|---|---|---|---|
| 1 | Agriculture | Agriculture | Industrial discharge |
| 2 | Municipal point source | Unspecified non-point sources | Urban runoff/storm sewers |
| 3 | Hydrologic modification | Atmospheric deposition | Municipal point sources |
| 4 | Habitat modification | Urban runoff/storm sewers | Upstream sources |
| 5 | Resource extraction | Municipal point sources | Agriculture |

Source: The 1996 National Water Quality Inventory Report to Congress, U.S. EPA (1996).

## 2.2 Confined Animal Facilities

The issue of pollutants from animal feeding operations is one of great concerns to the U.S. Government and the public. The importance of waste management from animal feeding operations has been addressed in the introduction of Senate Bill 1323, the *"Animal Agricultural Reform Act"*. Because animal waste is a major source of agricultural NPS pollution, many state legislatures introduced or passed legislation in the 1996-98 period to control how livestock and poultry operators manage waste.

The impairment of air and water quality by confined animal operations is increasingly becoming more of a public concern. The waste from confined animal sites is very significant. It is estimated that manure from a 200-cow diary operation produces as much pollution as the sewage from a community of 5000-10,000 people. The annual litter from a typical broiler house, which houses 22,000 birds, contains as much phosphorus as the sewage from a community of 6,000 people (USDA/NRCS, 1995).

Waste production by animals is based on a unit weight of 1000 pounds. For example, a diary cow weighing 1400 pounds equals 1.4 animal units. A broiler weighing 5 pounds equals 0.005 animal units (i.e. it takes 200 broilers to equal one animal unit). Table 2-2 summaries the amount of manure produced by different types of livestock.

Livestock feed is brought to confined animal sites and animal products such as meat, eggs or milk are transported from the sites while the manure remains at these sites. The quality of rations fed to the animals in confined operations has resulted in an increase in the amount of manure. For example, the typical daily nitrogen produced in the manure from a dairy cow has increased in the past 30 years from 0.37 pounds per day per animal unit to 0.45 pounds per day, an increase of about 20 percent (USDA/NRCS, 1995).

Confined animal facility operations can improve the management of manure. However, a USDA survey shows that 75 percent of farmers in West Virginia have inadequate storage facilities for manure (Lipton, 1997). Many farmers merely store the manure outside in

piles unprotected from weather, which results in a high potential for pollutants to be carried to nearby streams by storm runoff.

The Potomac River watershed, with its headwaters in West Virginia, has approximately 900 poultry farmers (Lipton, 1997). The U. S. officials estimate that these farmers dispose of about 4.6 million pounds of dead chicken carcasses a year. Many carcasses are buried on the surrounding farms and become cesspools of bacteria (Lipton, 1997). Animal waste also is a source of bacteria. Fecal coliform is often used as an indicator bacteria for assessing potential water quality impacts and/or problems. An estimate of production rates and fecal coliform content for various animals is given in Table 2-2.

As shown in Table 2-2, the volume of waste and associated pollutants is proportional to the number of animals in confinement. Several methods are typically utilized to dispose of waste. When dairy or beef animals are in pasture, waste is disposed naturally, i.e. through direct deposition to the land surface and/or where access to streams exists, through direct deposition to streams. Waste from diary facilities is often applied as liquid manure to agricultural land, usually to the cropland and/or pastureland within the watershed where the diary facility is located. The land application of waste increases the potential for significant pollutants (e.g. nutrients and bacteria) to be delivered to receiving water-bodies by storm runoff. In general, confined diary cows spend 25% of their time in loafing lots between March and November. The Virginia Department of Conservation and Recreation (VADCR) (1999) assume that 90 percent of the cows grazing on pastureland with a 0 and 5 percent slope have access to streams compared with 65 percent of cows grazing on pastureland with a slope greater than 5 percent (VADCR, 1998).

Confined chicken operations also generate significant chicken litter. As given in Table 2-2, the volume of waste is directly proportional to the number of poultry. Poultry litter is utilized as a feed source for cattle or applied to agricultural land as an organic fertilizer. It is estimated that the land application rate of chicken litter is equivalent to 1.86 ton/acre/year of manure to 100 percent of cropland in the US (The Muddy Creek TMDL

Establishment Workgroup, 1999). The pollutants from poultry waste (e.g. nutrients and bacteria) also are available for transport to receiving water bodies by storm runoff.

**Table 2-2. Amount of manure produced by different types of livestock in the USA.**

| Livestock type | Total manure | Nitrogen | Phosphorus |
|---|---|---|---|
| ------ lbs/day/ animal unit------ | | | |
| Beef | 59.1 | 0.31 | 0.11 |
| Diary | 80.0 | 0.45 | 0.07 |
| Chickens (layers) | 60.5 | 0.83 | 0.31 |
| Chickens (broilers) | 80.0 | 1.10 | 0.34 |
| Turkeys | 43.6 | 0.74 | 0.28 |
| Source: Natural Resources Conservation Service, Agricultural Waste Management Handbook (1992) | | | |

**Table 2-3. Estimates for fecal coliform production rates and fecal content for selected animals.**

| Animal | Fecal Coliform Production Rate | Fecal Coliform Content of Feces | Reference |
|---|---|---|---|
| Cow | $5.4 \times 10^9$ cfu/day | $0.23 \times 10^6$ cfu/g | Metcalf and Eddy, 1991 |
| Chicken | $0.24 \times 10^9$ cfu/day | $1.7 \times 10^4$ cfu/g | Metcalf and Eddy, 1991; NCSU, 1990 |
| Turkey | $0.13 \times 10^9$ cfu/day | $0.29 \times 10^6$ cfu/g | Metcalf and Eddy, 1991 |

A USGS study (Goolsby and Battaglin, 2000) reported that animal manure was the source of up to one-third of the nitrogen load leading to the hypoxia condition in the Gulf of Mexico. Major fish kills have occurred along the eastern coast of US by *pfiesteria piscicida*. It has been found that *pfiesteria picicida* (USGS, 1998) can be caused by excess nutrient.

Pathogens from animal manure can be transferred to drinking water supplies or food. With the increase in "*organic*" production of fruits and vegetables, there is the potential for more manure application on these crops. A new Food and Drug Administration (FDA) initiative (FDA/CFSAN, 1998) has suggested that pathogens from animal manure can be transmitted to fresh fruits and vegetables. The people who participate in water contact sports can be affected by pathogens. Some of the organisms, which are from warm-blooded animals, that are of concern include: *Escherichia coli*, *Salmonella*, *Giardia*, *Campylobacter*, and *Cryptosporidium parvum*.

Odors from confined animal sites are often more of a concern than water quality problems. The nearest human residence and water bodies can be impacted by the location of the facilities (Rozum and Huminek, 1998). For example, nitrogen in the form of ammonia gas can be transported downwind and be re-deposited onto water bodies adding to the nutrient load.

The management of animal manure has become a critical issue for animal agriculture in the U.S. The USDA and EPA released a Draft of the National Strategy for Animal Feeding Operations (AFO's), September 21, 1998 Federal Register, in response to an action item in President Clinton's Clean Water Action Plan of February, 1998.

In general, climate, farm size, topography and proximity to water bodies, manure management strategy, and soil type where nutrients are applied are the important factors determining the fate of pollutants from livestock.

## 2.3 Federal Non-point Source Pollution Control Programs

Congress enacted section 319 of the Clean Water Act (CWA) (33 U.S.C. 1329 (b) (2) (F) and (k)) in 1987, establishing a national program to control NPS pollution. According to the Section 319 of the CWA, states must address NPS pollution by developing NPS assessment reports that identify NPS pollution problems and the pollutants responsible for the water quality impairments. States then must develop management programs to control NPS pollution. All states now have EPA-approved NPS assessment reports and

management programs, and are implementing their management plans. EPA has a key role in helping states to control NPS pollution. Section 319 provides *"for State review of Federal assistance applications and development projects to determine their consistency with the requirements, goals, policies and other provisions of the State's non-point source management program"* (EPA, 1998). In May 1996, the U.S. EPA published a national NPS Program and Grants Guidance document (EPA, 1997). The mutual commitment between EPA and the States, as described in the report, is to share the vision to improve water quality by implementing effective NPS programs. EPA switched its role from grants oversight and administration towards a partnership with States to implement well-designed programs. EPA provides technical assistance to State agencies. EPA also initiated cooperative programs with other Federal agencies to promote consistency with and support for state programs at the Federal level (EPA, 1997), e.g. USDA and the National Oceanic and Atmospheric Administration (NOAA). The interagency cooperative program has promoted enhanced used of Fedplan (Federal Agency Environmental Program Planning Guidance) to help ensure that Federal agencies plan and budget for NPS pollution control activities (EPA, 1997).

The Water Quality Planning and Management Regulation (40 CFR 130), which links a number of CWA sections, including section 303(d), forms the water quality-based approach to protecting and cleaning up the nation's water (EPA, 1997). Specific procedures are developed in the approach to control the amount of pollution entering a water-body based on the intrinsic conditions of the water body. Water quality standards are also established to protect the water body for its intended use. The first step in applying the water quality-based approach is to determine if the water quality of a water-body meets or is expected to meet water quality standards after the implementation of technology-based controls. An overall plan is then set up to calculate Total Maximum Daily Loads (TMDLs) (Section 2.5), which estimates the pollutants loads that enter the water bodies.

EPA works closely with states and other partners to quickly develop TMDLs for waters affected by NPS. EPA is providing improved technical tools and training to State and local agencies and is working to develop and distribute tools on NPS screening modeling,

monitoring and the selection management measures. EPA is providing details on all aspect for developing TMDLs for nutrients, sediment, pathogens, and variable flow issues. To support the TMDL effort, EPA is continuing to improve BASINS (EPA, 1997).

The Clean Water Action Plan was developed by federal agencies in February 1998. The purpose of the Clean Water Action Plan is to develop and implement a comprehensive plan that would help revitalize the nation's commitment to water resources and to build a new framework for watershed protection in the next century (Cleanwater, 1999). Under The Clean Water Action Plan, EPA is fully expanding its administrative authority by cooperating with other federal agencies to better implement NPS pollution program. EPA and USDA are jointly developing a *"broad, unified national strategy to minimize the environmental and public health impacts of Animal Feeding Operations"* (Cleanwater, 1999).

Agriculture is a major contributor to NPS pollution in rivers and lakes. USDA is responsible for environmentally sound agriculture and is actively working with federal, state, tribal and private partners to *"establish by 2002 two million miles of conservation buffers to reduce pollutant runoff and protect watershed"* (Cleanwater, 1999). The U.S. Geological Survey (USGS) is working with state and tribes to improve monitoring and assessment of water quality, focusing on nutrient and related pollution (Cleanwater, 1999). The U.S. Department of the Interior (DOI) working with USDA is developing a Unified Policy to *"enhance watershed management for the protection of water quality and the health of aquatic systems on federal lands and for federal resource management"* (Cleanwater, 1999).

The watershed approach was proposed as a collaborative effort by federal, state, tribal, and local government. The approach was the *"key to setting priorities and taking action to clean up rivers, lakes, and coastal waters"* under The Clean Water Action Plan (Cleanwater, 1999). This approach balances the efforts to control point source pollution and NPS pollution at the watershed scale. For the past 25 years, most water pollution control efforts have been broadly applied to national programs. Recently it has been

recognized that efforts should be focused on special watershed conditions. Therefore, the approach is helpful to identify the most-effective pollution control strategies to meet clean water goals (Cleanwater, 1999).

EPA has developed the Index of Watershed Indicators and has made it available on the Internet to better serve the public who are interested in water quality conditions in their community. The USDA/FSA is using GPS to support precision farming with goals to increase agricultural production while reducing NPS pollution. USDA/FSA has also used digital geographic imagery with GIS to manage land and crop-based information for water and soil conservation (FSA, 1999).

## 2.4 Non-point Pollution Control Program in Virginia

As required under the CWA of 1987, the Virginia NPS Pollution Management Program was developed in August 1988 and revised in May 1989. The Department of Conservation and Recreation (DCR) in Virginia has developed the NPS Management Program in cooperation with the Non-point Source Advisory Committee (NPSAC). NPSAC is an interagency committee, chaired by DCR staff, but made up of representatives of all federal and state agencies that share responsibilities for NPS pollution program implementation in the commonwealth. It received approval from EPA in July 1989. The NPS management program includes all statewide programs, which quantify, control and limit effects of NPS pollution on the state's surface and ground waters (VADCR, 1997). The Virginia Department of Conservation and Recreation (DCR) coordinates and directs programs and services to prevent degradation of the Commonwealth's water quality and quantity. In Virginia, DCR is the lead agency for developing and implementing statewide NPS pollution control programs and services. These programs support both individual natural resources stewardship and assist local governments with resource management. State agencies fund these statewide programs. The federal NPS Pollution Control Program provides grants such as the 319 Program that provide technical and financial assistance, educational programs, and funds for research.

Virginia has been divided into 494 unique hydrologic units or watersheds (Virginia DCR, 1999) for planning purposes. Each watershed has a committee comprised of citizens, government and private sector representatives from various interest groups. Identification of the causes of NPS pollution is the major mission of the committee (VADCR, 1999). Virginia DCR completed Virginia's first NPS pollution assessment in 1988, with subsequent updates and refinements in 1993 and 1996. The 494 watersheds in Virginia have been ranked by their NPS potential (VADCR, 1999). The NPS pollution assessment is based on land use, livestock population, forest harvesting, disturbed acreage, best management practices (BMPs) implementations and potential erosion rates.

## 2.5 Total Maximum Daily Loads

Section 303(d) of the CWA (40 CFR 130.7) describes TMDL as *"a tool for implementing State quality standards and is based on the relationship between pollution sources and in-stream water quality conditions"* (USEPA, 1991). The objective of a TMDL is to allocate allowable loads among different pollutant sources so that the appropriate control actions can be taken and water quality standards could be achieved. The TMDL provides an estimate of pollutant loads from all sources and predicts the resulting downstream pollutant concentrations. The effect of all activities or processes that cause or contribute to water quality-limited conditions of a water body should be considered in the calculation of TMDLs. The pollution source, pollutant characteristics and geographic scale of the pollution are important factors considered in calculating TMDLs.

TMDLs are plans designed to direct management actions so that polluted water bodies are restored to a level that achieves state water quality standards. The following equation is used to express each component of a TMDL.

**TMDL=WLA + LA + MOS** (2-1)

*Where:*
*WLA is waste load allocation from point sources.*
*LA is load allocation from NPS.*

*MOS is margin of safety, which is used to account for lack of knowledge concerning the relationship between pollution control mechanisms and water quality.*

The TMDL process is a rational method for weighing competing pollution concerns and developing an integrated pollution reduction strategy for point sources and NPS. The development and implementation of control actions to the attainment of water quality standards is linked with TMDL. The TMDL provides the basis for States to establish water quality-based controls. Estimation of NPS pollutant contributions is the key component of the TMDL process. The MOS usually increases when NPS pollution is involved because of the complexity and uncertainty in estimating NPS pollution. The general steps involved with TMDL development (USEPA, 1997) include identification of the pollutant to consider, estimation of the water body assimilative capacity, and estimation of the pollution from all sources contributing to the water body.

The development of the TMDL results in the allocation (with a margin of safety) of pollutants among the different pollution sources in a manner that water quality standards are achieved. The 40 CFR 131.2 states that the water quality standards *"serve the dual purpose of establishing the water quality goals for a specific water body and serve as the regulatory basis for the establishment of water-quality-based treatment controls and strategies beyond the technology-based levels of treatment required by sections 301(b) and 306 of the Act".*

A phased approach is appropriate for the calculation of TMDLs when LA is considered (USEPA, 1997). This is because there must be assurance that NPS control measures will achieve expected load reduction. When the phase-approach is used, the TMDL has LAs and WLAs calculated with margin of safety to meet water quality standards. The phase-approach provides for NPS control tradeoffs (40 CFR 130.2(I)). For example, if the implementation of BMPs for reducing NPS pollution makes more stringent load allocations practical, then WLAs could be made less stringent.

A monitoring program is needed to support the TMDL process to better describe pollutant contributions from point and NPS pollution. Modeling procedures are usually

used to conduct TMDLs. The data from monitoring is very important for calibration of the TMDL model. Because NPS pollution is both spatially and randomly distributed over a watershed area, large volumes of data are typically required to conduct the modeling procedures. GIS is often used to integrate with models. The new developments in GIS technology often make it much easier to apply TMDL models.

## 2.6 Geographic Information Systems

GIS is a computer-based tool for the storage, retrieval, management, manipulation, analysis and display of spatial and non-spatial data. A GIS has robust tools for effective and efficient storage and manipulation of different sources of data, including the remote sensing data. Diverse data sources can be processed into a unified database.

Map registration is one of the basic functions of GIS. The process uses ground control points (GCP) (Campbell, 1987) to transform a cartographic map into known geographic reference coordinate systems. The different data layers from various sources can be registered into a common coordinate system so that they will overlay correctly at all locations. When multiple data sources are used, registration is always the first step before performing other GIS analyses.

Overlay is a unique GIS function that provides a solution by combining all information from different sources. It is defined as a subset of a spatial data set dealing with one thematic topic. The overlay function is used to combine registered data layers such that the results contain selected information from all the data layers. The overlay operator is the important feature that distinguishes GIS from other mapping software such as AutoCAD. Mathematical arithmetic and Boolean operators are often used to examine the spatial patterns caused by the interaction of one map with another. For example, Fu (1994) used overlay functions to derive the Soil Conservation Service (SCS) curve number by combining map coverages for land use/land cover with soil mapping units. Another example is the overlay of a soils map onto slope to gain a better understanding of the interaction between the two maps.

Spatial and nonspatial queries are important GIS functions. Spatial query is used to find the specific locations of the spatial data objects. The spatial query function can also be used to locate spatial objects which meet some specified criteria. The neighbourhood of spatial objects can be evaluated with spatial query functions. The attributes associated with spatial data objects can be displayed after the spatial query function is used. In general, spatial query searches the locations of spatial objects in a map first and then locates the attributes in a table and displays the attributes. However, nonspatial query starts with locating nonspatial attributes from one or more tables, then displays the results on a map. For example, nonspatial query can be used to search all polygons that satisfy selected criteria and display the selected polygons on the map.

Vector and raster data models are used in GIS analyses. The vector data model is based on the *"spaghetti model"* (Bonham-Carter, 1994). In the spaghetti model, points are represented as pairs of spatial coordinates, lines as strings of coordinate pairs, and areas as lines that form closed polygons. In the more sosphicated topologic vector data model, the boundaries of polygons are separated into a series of arcs and nodes, and the spatial relationships between arcs, nodes and polygons are explicitly defined in attribute tables. The vector data model contains the topological relationship between arcs, nodes, and polygons. For example, a polygon to the left and right of each arc is explicitly defined. Each topographic attribute is assigned to an arc, node or polygon through a process of building topology. The area and perimeter of a polygon are calculated after the topological relationship is built. The vector data model is well suited for the representation of graphic objects. The arcs can be very smooth curves, which are better approximated by vertices that are not constrained by the coarse Cartesian coordinates of a lattice. The arcs and nodes can be represented in numbers with sufficient precision. Therefore, the area and perimeters of polygons are estimated much more accurately. However, for overlay procedures, the vector data model can take significant computing resources because complex and often tedious geometric calculations are required.

The raster data model represents spatial objects as a regular matrix of numbers. The grids of the matrix are called the pixels. A regular array is used to store the values (attributes) of pixels. The row and column are used to define a particular location within the array.

Spatial analysis such as algebraic map calculations and overlay are relatively easily performed within the raster data model. Satellite and aerial photographic images are represented in raster format. Mathematical algorithms developed for matrix calculations can be quickly adapted to process these type of images.

## 2.7 Remote Sensing

Remote sensing is the science and art of *"obtaining information about an object, area, or phenomenon through the analysis of data acquired by a device that is not in contact with the object, area, or phenomenon under investigation"* (Lillesand and Kiefer, 1994). It is the *"science of deriving information about the earth's land and water areas from images acquired at a distance. It usually relies upon measurement of electromagnetic energy reflected or emitted from the feature of interest"* (Campbell, 1987). The electromagnetic energy emanating from distant objects is detected and measured from remote sensing and can be categorized by class or type, substance, and spatial distribution.

Remote sensing is used to study the spectral response pattern of objects, such as vegetation, soil and water, etc. It is now a rapidly growing source of global geo-spatial information and has become a very important tool in many operational programs involving resource management, engineering, environmental monitoring, and geological exploration. Campbell (1987) and Lillesand and Kiefer (1994) give details about the historic, current development and research in remote sensing. An important advantage of the remote sensing is the ability to provide current information, particularly with land use/land cover, building footprints, etc.

The two general types of remote sensing images are those obtained from satellites and those obtained from aerial photography. Satellite images such as Spot or MSS (Campbell, 1987) provide large regional images of the earth. These images are often used to provide information useful for studying global carbon cycles, hydrologic balance and bio-geochemistry of critical greenhouse gases. Satellite images are obtained in earth orbits.

Imagery from aerial photography when compared with satellite imagery is taken at relatively low flight heights at large scale and high resolution. Aerial photos can provide images, which cover much smaller areas and are generally taken to support small area applications.

### 2.7.1 Image Scanner

Equipment to rasterize documents has been available for sometime. The technology has been used to support the archival of all types of documents. With advances in software such as pattern recognition, the technology has become a powerful support tool for a wide array of applications. Specific to this research, image scanners have been developed to support the utilization of 35-mm slide film. For example, the Nikon LS-1000 (Super CoolScan) (Nikon Corporation, 1996) was utilized in this study.

It takes 40 seconds for Nikon LS-1000 to scan a slide for 2592 x 3888 pixels. It has a precision of 12-bit A/D (Analogue-to-Digital) conversion for wide dynamic range, up to 4096 depths of color. The Nikon LS-1000 has special optics, which preserve definition and overcome problems of geometric distortion. This scanner incorporates Nikon's new high-power light emitting diode (LED) illumination technology. This technology can speed the process of scanning and also can contribute substantially to a unique color reproduction from negatives and slides. The LED-based design affords very lower power consumption, virtually no heat dissipation or noise, as well as long life and durability. It is a single-pass Red-Green-Blue (RGB) line sequential scanner that uses LED strobe operation to produce RGB triplets with a monochromatic charge coupled device (CCD). In general, scanners that use monochromatic CCDs are inherently less noisy and have fewer color and spatial artifacts than single-pass designs that use a color CCD with a white-light source. Super CoolScan is reliable and low cost. The scanner runs extremely quite (no fan) and compact. Today's retail price for Super CoolScan is about US $1,500. The scanner settings for resolution and physical dimensions can vary up to 2700 dot per inch (dpi) or pixel per inch at a resolution of 0.0001 inch. The scanner has an optional automatic slide feeder for high-speed archiving. This equipment is compatible with personal computers, which is a very convenient function for scanning images. It is considered the best scanner in the market for 35-mm slides.

### 2.7.2 Image Color and Mode

Hue is the wavelength of light reflected from or transmitted through an object. More commonly, hue is identified by the name of the color such as red, orange, or green. Saturation, sometimes called chroma, is the strength or purity of the color. Saturation represents the amount of gray in proportion to the hue and is measured as a percentage from 0% (gray) to 100% (fully saturated). Brightness is the relative lightness or darkness of the color and is usually measured as a percentage from 0% (black) to 100%(white). Intensity refers to the total brightness of a color.

The red-green-blue (RGB) color model is widely used by commercial software packages to represent an image. In the RGB model, an intensity value ranging from 0 (black) to 255 (white) for each of the RGB component is assigned to each pixel to create an index value for brightness. Each image can be represented as Pic(i,j). Each point (i, j) of an image is called a pixel. Intensity-Hue-saturation (IHS) is another color model, which is often used to represent an image. The IHS model is better to represent images from different sources (Lillesand and Kiefer, 1994). The advantage of using the RGB model is that each data layer (color) can be processed independently. The ranking of the data layers based on importance can be determined through Principal component analysis (PCA). The disadvantage of the RGB model is that interpretation becomes more difficult when there are high values in two or more layers of data at the same time (Banham-Carter, 1994). The advantage of IHS model is that the three data color layers can be combined together to represent true color values of an object. The disadvantage of IHS model is that the relative importance of each data layer is very difficult to determine.

## 2.8 Image Processing

Image processing refers to the procedures that are used to enhance, contract, transform, resample and reclassify the image. It is a very broad subject. It is used to manipulate and interpret a remote sensing digital image. Four common basic operations that are involved in the image processing include image enhancement, image rectification, image restoration, and image classification.

An image obtained from remote sensing equipment often contains noises. Noise is created due to limitation in the sensing, signal digitization, or data recording process. Periodic drift or malfunction of a detector, electronic interference between sensor components and intermittent *"hiccups"* in the data transmission and recording sequence are the potential sources of noises. Several methodologies and/or procedures that have been developed to remove noises include image enhancement, restoration and rectification.

Image rectification and restoration processes aim to correct distorted or degraded image data to create a more accurate representation of the original scene. Image restoration is concerned with the correction and calibration of images in order to obtain as accurate representation of the earth surface as possible. It also can remove the noise resulting from atmospheric conditions (Campbell, 1987). Image restoration attempts to remove known value distortion. Geometric and radiometric corrections can be made by image restoration. These procedures are often referred to as preprocessing analysis before any manipulation and analysis of the image data. Geometric distortion exists in raw image data due to variation in latitude, attitude, and velocity of the sensor platform. The purpose of geometric correction is to compensate for the distortions due to such factors as earth curvature, atmospheric reflection, relief displacement, etc. The corrected image has the geometric integrity of a map. Image restoration is concerned with the correction and calibration of images in order to obtain a true representation of the earth surface. Image restoration is used to remove the noise resulting from atmospheric conditions (Campbell, 1987). Image restoration is an attempt to remove known value distortion. Geometric and radiometric corrections can be made by image restoration.

The theory and procedures related to geometric correction of raw image, or raw data, are given by Campbell (1987) and Lillesand and Kiefer (1994). Periodic drift or malfunction of a detector, electronic interface between sensor components, and intermittent *"hiccups"* in the data transmission and recording sequence often create image noises in raw data. The image noise masks the true characteristics of a digital image. A digital image could be restored to as close to an approximation of the original scene as possible by image

restoration process. Destriping (Lillesand and Kiefer, 1994) is often used for noise removal.

Image enhancement is to *"improve the visual interpretation of an image by increasing apparent distinction between the feature in the scene"* (Lillesand and Kiefer, 1994). It is often conducted after image rectification and restoration. The purpose of image enhancement is to make a selected segment of an image more observable. Contrast, spatial feature, and multiple-image manipulations are the procedures used in image enhancement.

Gray-level thresholding is often used in contrast manipulation. It is used to segment an input image into two classes, one for those pixels having values below a selected gray level and one for those above this value. The segmented image can be further processed independently. Additional effort can be focused on a specific segment. Spatial filtering is a method to manipulate spatial feature. Spatial frequencies are emphasized or de-emphasized by spatial filtering. Spatial frequency is the measure of the tonal variation in an image. High spatial frequency refers to abrupt changes in gray level within a small number of pixels, for example, across a river. Low spatial frequency refers to when gray levels vary only gradually for a large number of pixels, such as cropland. Low pass filters are designed to emphasize low frequency features, which means that low spatial frequency is kept. High pass filters function in the opposite direction. Edge enhancement is used to contrast high spatial frequency while maintaining low spatial frequency (Lillesand and Kiefer, 1994).

Multi-image manipulation is the procedure that is used to combine the spectral information contained in several bands to enhance an image. The advantage of the procedure is that it conveys the spectral or color characteristics of an image feature, regardless of variations in illustrated conditions such as topography. For example, the ratio of the correlation between digital numbers (DNs) in one band with the corresponding values in another band can be used to enhance an image (Lillesand and Kiefer, 1994).

Redundant information exists in multiple-band images because an image generated from digital data from various wavelength bands often appear and convey essentially the same information. Although ratioing can be used to enhance an image, it cannot be used to reduce the same information contained in several bands. Principle component analysis (PCA) is used to remove or reduce the redundancy in multi-spectral data. PCA is a transformation function that transfers an original image, which has an n-channel data set, into fewer components. The components are thus statistical abstractions of the variability inherent in the original band set. The components produced by a PCA transformation are not correlated with each other. Each component carries different information, which usually represents the rank of respective influences. The first few components include the most of the real information. The later components probably contain minor information and sometimes the noise. The PCA analysis illustrates the relative importance of variables by calculating the variance-covariance matrix (Davis, 1986). This analysis is very helpful when little prior information concerning a scene is available.

Intensity-Hue-saturation transformation is a procedure to transform an image represented by RGB model into IHS model to achieve better color enhancement (Lillesand and Kiefer, 1994). After an image is represented by IHS model, each IHS component can be operated independently without affecting the others. The Hexcone model (Lillesand and Kiefer, 1994) is often used to transform the RGB model into IHS model.

Image classification is used to categorize all pixels in an image into defined classes. Each pixel in an image is treated as an individual unit composed of values in several bands. Image classification is one of the most important components of remote sensing, pattern recognition and image analysis. The two general methods that are used in image classification include supervised classification and unsupervised classification (Lillesand and Kiefer, 1994) and Campbell (1987).

Supervised classification is defined as *"the process of using samples of known identity to classify pixels of unknown identity"* (Campbell, 1987). The three steps that are usually involved with supervised classification include the creation of training data, the creation of signature files from the training data, and the application of a classification procedure

to the image bands using the signatures created from the training data. The result from supervised classification is an image feature category, for example, a land use type. The commonly used supervised classification methods are parallelepiped, minimum distance, maximum likelihood, Bayer's classification, echo classification (Ketting and Landgrebe, 1975) and isodata (Duda and Hart, 1973).

In unsupervised classification, the unknown pixels in an image are aggregated into a number of classes based on natural groupings or clusters present in the image. The dominant spectral response patterns that occur within an image are identified. The spectral classes are groups of pixels that are uniform with respect to the brightness in each spectral channel. The classified image features are verified by ground truthing. Each procedure has unique advantages and disadvantages. The unsupervised classification procedure does not require training data. An image analyst will determine the image feature classes based on each relationship with spectral classes. Supervised classification is a very efficient procedure to classify an image when verified training data is available. Its main disadvantage is the excessive time and resources required to develop training data. Additionally, many spectral signatures vary over seasons.

The integration of GIS and remote sensing has been demonstrated as an effective way to analyze, monitor and manage natural resources (Campbell, 1987). A fully functional GIS has many robust functions, which have been extensively used in image processing. For example, on-line digitizing functions have been used to create training data for supervised classification. Raster to vector conversion utilities have been used to convert classified land use images into polygons for further analysis. Overlay functions can be used to extract topographic and land use information from remote sensing images. Spectral signature of an object is the unique and absolute response pattern. Understanding the spectral signature of an object is a key step in understanding an image.

## 2.9 Small Format Aerial Images

Small format aerial photography (SFAP) can be used for new topographic surveys, provided substantial control points are available (Warner, et. al., 1996). SFAP has been

integrated into a statewide field-testing of forestry inventory in the U.S. (Baston, 1974). In Canada, several organizations have integrated SFAP into their basic operation, such as Ontario's Ministry of Natural Resources (MNR). SFAP has provided low cost and reliable data for forestry management. It can meet the foresters' accuracy requirements at a much cheaper cost compared with the traditional aerial survey. The SFAP can be used to identify sampling locations for pest management, and fire warning and control (Kirby, 1980; Hall and Aldred, 1992; Willett and Ward, 1978; Stevens, 1972; and Clements et. al., 1983). The SFAP has been efficiently used for wildlife management, such as, monitoring population, evaluating habitat of wildebeest (Norton-Griffiths, 1973, 1988; Hiby et. al., 1988) and to measure food consumption of swans (Lee and McKelvey, 1984). Meyer et al. (1982) used SFAP to define a specific niche for monitoring and mapping rangeland. Tueller et al. (1988) employed SFAP to monitor vegetation changes. Geologists have developed comprehensive procedures for interpreting land forms, surface features, and mining operations, and have determined iceberg volumes (Hall and Walsh, 1977; Wracher, 1973; Farmer and Robe, 1977). The most successful application of SFAP (Warner et. al., 1996) in urban areas was the use of 35-mm color photography to record kampong developments in Indonesia. A 35-mm Nikon camera mounted on light aircraft provided 1:2500 scale mosaics of overcrowded kampongs. Useful information could be obtained fast and economically from printed 35-mm photos.

## 2.10 Shape Modeling and Analysis

Shape is defined in The New Shorter Oxford English Dictionary (1993) as an *"external form"* or *"contour"* or *"that quality of a material object or geometrical figure which depends on constant relations of position and proportionate position among all the points composing its outline or external surface"*. The shape of an object has external boundary to distinguish it from another object. Studies to develop procedures for identifying objects have been generally classified into two categories: shape analysis and shape modeling.

Shape analysis is the technique used to describe shape dimensions of objects. Shape modeling, however, consists of shape analysis and the construction or restoration of

shapes. A specific object is identified through shape analysis first, then a procedure is developed to model the object. An important class of shape modeling techniques is based on the representation of the outer boundaries of objects. An identified object can be quantitatively described by defined shape modeling. It provides important features critical to the recognition and order of objects in the world around us.

The methodologies developed for shape analysis have been applied to medical diagnosis, earth science, automatic banking, etc. For example, geologists use the shape analysis methodology to locate oilfields (Davis, 1986). Shape analysis is often used to define the parameters for hydrologic and NPS models. For example, geometric parameters such as watershed area, channel and overland flow length and burification ratio are normally calculated by shape analysis. GIS functions are often used to conduct shape analysis of watersheds (Fu, 1994).

Shape modeling has been used to match the image of an ATM card user with digital copies stored in a database (Dejesus, 1995), to identify cancer (Georgia Tech, 1997) and to identify building shapes such as the Pentagon (Weng, 1993).

The general purpose of shape modeling is to describe the shape of an object under investigation and then match it with the shape of a known object. A mathematical equation is often used to describe the shape of an object. The equation can be simple or a very complicated mathematical description. The measure of the area of a shape is an example of simple mathematical equation. The Fourier series and moment analysis are much more complicated mathematical representations.

It is often very difficult to measure or define the shape of an object in a precise mannor (Davis, 1986). Although many methods have been developed to model shape, none have proven satisfactory for completely modeling a shape. The following sections describe three methods often used for shape modeling.

## 2.11 Measurement of Moments

One of the most efficient methods for shape modeling is the measurement of moments. Ullmann (1973) describes the utility of moments for shape measurements in the following statements. *"For recognition purpose it is of course impractical to work with an infinite sequence of moments, but at the cost of losing uniqueness, we can use the values of a chosen set of moments. Given an input pattern, we can evaluate, for instance, 22 moments. The values of these moments can be regarded as the components of a 22-component vector, and ordinary pattern recognition techniques can be applied to this vector. One possible advantage of applying pattern recognition techniques to this vector rather than directly to the input pattern is that the total number of components of the vector may be less than that total number of elements of the input pattern, and yet the vector may contain enough information for recognition purpose. A further advantage of the vector cab be easily normalized so that, if the overall blackness, position, height, width, orientation, or even the perspective of the input pattern is altered, the vector obtained from it remains unchanged."* The moment analysis uses a few key parameters to represent the characteristic of a shape in a mathematical manner. The mathematical equation is constructed based on normalization of shapes. The mathematical equation then can be used to model shapes.

The general form for a non-central moment of a two dimensional continuous function is

$$M_{pq} = \iint i^p j^q g(i, j) dxdy \qquad (2.2)$$

Where

*$M_{pq}$ is non-central moment. The g(i, j) is any function of two independent variables, i.e., the continuous picture function, and p, q are nonnegative integers from 0,1, 2,…n. The integration is over the entire plane. I,j represent the number of row and column of an image.*

Hu (1962) has shown that the equation under certain very general conditions is an infinite sequence of the coefficients, where p and q are nonnegative integers, uniquely

determined by g(i,j) and, conversely, g(i, j) is uniquely determined by the infinite sequence. The remark from Hu (1962) states the discrimination power of moments. For a given region, the non-central p q moments for this region is:

$$M_{pq} = \sum_{(i,j)\in R} i^p j^q \tag{2.3}$$

The moments determined from equation 2.2 are not translation invariant (Otterloo, 1991). To obtain translation invariant moments, the central moments, i.e., moments of the form (Otterloo, 1991) are:

$$\mu_{pq} = \sum_{(i,j)\in R} (i-\bar{i})^p (j-\bar{j})^q \tag{2.4}$$

Where:

$$\bar{i} = M_{10} / M_{00} \tag{2.5}$$

$$\bar{j} = M_{01} / M_{00} \tag{2.6}$$

The following equations are the central moments of order of 3:

$$\mu_{02} = \sum_i \sum_j (i-\bar{i})^0 (j-\bar{j})^2 g(x,y) = M_{02} - \frac{M_{01}^2}{M_{00}} \tag{2.7}$$

$$\mu_{03} = \sum_i \sum_j (i-\bar{i})^0 (j-\bar{j})^3 g(x,y) = M_{03} - 3\bar{j} M_{02} + 2\bar{j} M_{01} \tag{2.8}$$

$$\mu_{10} = \sum_i \sum_j (i-\bar{i})^1 (j-\bar{j})^0 g(x,y) = M_{10} - \frac{M_{10}}{M_{00}}(M_{00}) = 0 \tag{2.9}$$

$$\mu_{11} = \sum_i \sum_j (i-\bar{i})^1 (j-\bar{j})^1 g(x,y) = M_{11} - \frac{M_{10} M_{01}}{M_{00}}(M_{00}) \tag{2.10}$$

$$\mu_{12} = \sum_i \sum_j (i-\bar{i})^1 (j-\bar{j})^2 g(x,y) = M_{12} - 2\bar{j} M_{11} - \bar{i} M_{02} + 2\bar{j}^2 M_{10} \tag{2.11}$$

$$\mu_{20} = \sum_i \sum_j (i-\bar{i})^2 (j-\bar{j})^0 g(x,y) = M_{20} - \frac{2M_{10}^2}{M_{00}} + \frac{M_{10}^2}{M_{00}} + M_{20} - \frac{M_{10}^2}{M_{00}} \tag{2.12}$$

$$\mu_{21} = \sum_i \sum_j (i-\bar{i})^2 (j-\bar{j})^1 g(x,y) = M_{21} - 2\bar{i} M_{11} - \bar{j} M_{20} + 2\bar{i}^2 M_{01} \tag{2.13}$$

$$\mu_{30} = \sum_i \sum_j (i - \bar{i})^3 (j - \bar{j})^0 g(x,y) = M_{30} - 3\bar{i} M_{20} + 2M_{10} \bar{i}^2 \quad (2.14)$$

Hall (1979) describes the details of normalization of the central moments. For pattern recognition, a small set of low-order moments is used to distinguish between different patterns. In general, the central moment of order 3 is sufficient to identify shapes (Hu, 1962). The general procedures to use moment analysis can be described as follows:

First, central moments are calculated for an object of interest. Then first, second, third and even higher order moments (as necessary) are calculated based on equations 2.7 through 2.14. The calculated moments are then compared with those of a known object to determine if the shapes match. An excellent source for the application of moment analysis for modeling shapes is given in Tomita and Tsuji (1990).

## 2.12 Fourier Measurement of Shape

The term of Fourier series is well known to engineers. In 1807, the French Mathematician Joseph Fourier (1780-1830) submitted a paper to the Academy of Sciences in Paris. The ideas contained in the paper have been developed into an important area of mathematics named in his honor-- Fourier analysis. The Fourier transformation based on Fourier analysis has been widely used to model shapes. For example, the Fourier transformation also forms the basis for noninvasive medical diagnostic protocols such as computer-assisted tomography and for analytical techniques such as nuclear magnetic resonance. Fourier series have been proposed to serve various functions in shape analysis procedures. (Otterloo, 1991). Weng (1993) used Fourier transformation to detect the Pentagon building from an image. The general form of Fourier series is to represent a function f(x) by a series of combinations of the cosine and sine if f(x) is continuous on [-L,L].

$$f(x) = a_0 + \sum_{n=1}^{\infty} [a_n \cos(\frac{n\pi x}{L}) + b_n \sin(\frac{n\pi x}{L})] \quad (2.15)$$

The coefficients are estimated as

$$a_0 = \frac{1}{2L} \int_{-L}^{L} f(x) dx \quad (2.16)$$

$$a_n = \frac{1}{L}\int_{-L}^{L} f(x)\cos(\frac{n\pi x}{L})dx \quad (2.17)$$

$$b_n = \frac{1}{L}\int_{-L}^{L} f(x)\sin(\frac{n\pi x}{L})dx \quad (2.18)$$

The above equations can also be represented in the polar coordinate system as:

$$a_k = \frac{2}{k}\sum_{j=1}^{n} r_j \cos k\theta_j \quad (2.19)$$

$$b_k = \frac{2}{k}\sum_{j=1}^{n} r_j \sin k\theta_j \quad (2.20)$$

Where:

$a_k, b_k$ are defined in equations 2.19 and 2.20. respectively; $r_j$ is the radii.

Davis (1986) gives details on the use of Fourier series as a Fourier descriptor (FD) for modeling shape. Some conditions for using Fourier transformations include: 1) the boundary of a region or object should be known; (2) the centroid of each region and object can be determined; and that 3) the shape of a region or object can be represented by Fourier series in a polar system (Davis, 1986). A polar system is the angular orientation of a radius extending from a point within the outline or the distance along the radius from the central point to the outline.

Otterlo (1991) gives detail steps to normalize an arbitrary FD so that it can be matched to a test set of FDs without considering its original size, position, and orientation. Therefore, FD can be used to conduct similarity test as illustrated by Otterlo (1991). However, the FD based on polar coordinate system also has limitations. The FD can be used to analyze single-valued outlines. The single-valued outline is defined (Davis, 1986), as a case in which *"a radius drawn from centroid must intercept the perimeter only once"*. Convoluted shapes cannot be represented and analyzed by FD (Davis, 1986).

## 2.13 Measures of Shape Used by Earth Scientist

Earth scientists, such as geologists and geographers, have used shape analysis techniques to understand research objects. Geologists have studied a broad spectrum of shapes of rocks. Geographers often study the shape of a city boundary. The shape of a river channel sometimes is studied for flood analysis.

Often the basic mathematical measures of area, perimeter, length and width are used to identify specific objects. The objects studied by geologist and geographers are much more irregular than for example object shapes typically modeled in the medical and banking industries. Moment and Fourier techniques are difficult to apply to earth sciences. For example, the many shapes of rocks have more than single-value outlines. Table 2-4 compares the different methods used to model object shapes. The challenge is to find procedures to calculate shape-modeling parameters (area, perimeter, length and width, etc.) quickly and accurately for a large numbers (thousands) of objects.

Shape modeling methods used by earth scientists require less strict mathematical conditions and less computational time compared with moment analysis and Fourier analysis (Davis, 1986). Moment analysis and Fourier analysis have been successfully applied to medical studies, and criminal investigation (Liang and Clarson, 1989; Mok and Boyer, 1986; Tello, et. al. 1986; Cipra, 1993). However, it is very difficult to use moment analysis and Fourier analysis to model the shape of rock types and other earth study objects from remote sensing images because of computational time and strict mathematical requirement. For example, in order to use the method of moment analysis for shape-modeling rocks, the central moments must be determined.

An extensive literature search did not find any references relating to the automatic detection and extraction of specific types of object from remote sensing images for use in hydrologic and NPS models.

## Table 2-4. Measures of Shapes Used by Earth Scientists

| | |
|---|---|
| **Measures based on axial ratios** | |
| Form (Length and width ratio) | $F = \frac{l}{w}$ |
| Elongation | $E = \frac{w}{l}$ |
| Circularity | $C_1 = \sqrt{\frac{lw}{l^2}}$ |
| Measures based on perimeters | |
| Gain shape index | $GSI = \frac{p}{l}$ |
| Shape Factor | $SF_1 = \frac{p_c}{p}$ |
| **Measures including both perimeters and areas** | |
| Circularity | $C_2 = \frac{4A}{p^2}$ |
| Compactness | $K = \frac{p^2}{A}$ |
| Size | $SI = \frac{2A}{p}$ |
| Thinness Ratio | $TR = 4\pi(\frac{A}{p^2})$ |
| **Measures based on areas** | |
| Circularity | $C4 = \sqrt{\frac{A}{A_c}}$ |
| Shape Factor | $SF_2 = \frac{A_l}{A}$ <br> $SF_2 = \frac{A}{A_C} x100$ |
| **Measures based on areas and aerial length** | |
| Form Ratio | $FR = \frac{A}{l^2}$ |
| Other measures | |
| Mean side | $\bar{S} = \sum S_j / n$ |
| Variance of Sides | $s_s^2 = \frac{\sum(S_j - \bar{S})^2}{n}$ |

The sources for these measures are given in Davis (1986), and Tomite and Tsuji (1990). Keys:

$A$ – Area of object

$A_c$ – Area of smallest enclosing circle

$A_i$ – Area of largest inscribed circle

$L$ – Length of long axis

$w$ – Width of object perpendicular to long axis.

$p$ – Perimeter of object

$S_j$ – Length of $j$th side of object, considered as a polygon

$n$ – Number of sides, considered as a polygon

## 2.14 Pattern Recognition for 35mm Slide Imagery Applications

The 35-mm slide imageries developed by USDA/FSA could be used to locate confined animal facilities. Automatic identification of these facilities from slide images could save time and reduce the cost. Furthermore, the results obtained from automatic identification could be much more consistent compared with the manual interpretation method. There is no standard procedure to process the 35-mm slides. New procedures are needed for handling, scanning and processing the 35-mm slides. The techniques of GIS, remote sensing and pattern recognition could be combined to develop automatic procedures of identification and extraction.

Much research has been conducted on the application of pattern recognition techniques to identify land use/land cover classification and crop yield based on the spectral characteristics. The most common application of pattern recognition to NPS pollution control has been limited to identifying land use, geological features, soil types, and topographic features for local areas, for example using unsupervised classification procedures to class a satellite image (Campbell, 1987). These features typically have distinct and relatively uniform spectral characteristics and represent large areas in the image, which makes classification of these objects much easier (Campbell, 1987).

A literature review did not reveal any research relating to the application of pattern recognition to locating confined animal facilities. It is much more difficult to recognize facilities such as confined animal lots on an image because they typically are intermingled with complex topographic, land use/land cover and other ground features. Compared with the area-distribution features such as land use and soils, these facilities are very small and can be considered as point distributions. They do not have relatively

large uniform spectral or texture distributions and consequently are much more easily overshadowed by surrounding features.

Shape identification for specific objects on the images is generally not considered in the traditional remote sensing application in environmental studies (Campbell, 1987) because the distinct features of land use/land cover, crop, etc. are usually satisfactory for making a classification. However, the automatic identification of a shape is expected to play a very important role in locating confined animal facilities. The shape of confined animal facilities may be the only unique feature available to recognize a facility. For example, the rectangular shape of a building is expected to be a footprint for a poultry facility. A diary loafing lot generally has a more circular shape.

Development of automatic recognition algorithms for shape description is expected to be important for identifying confined animal facilities. Although the shapes of the buildings are very important indicators of the type of the confined animal facilities, it is often difficult to distinguish the building from similar non-confined animal facilities. For example, a poultry house has a shape that is similar to a farm machine shed. Both buildings could have very similar spectral characteristics, which would make them difficult to be separated. For these situations, other knowledge about the buildings and surrounding area may be required to achieve successful identification.

GIS is expected to be a useful tool for handling the large volume of spectral data and for performing the spatial data analysis. Over the past decade, significant research has been conducted on GIS system design, stability tests, and new module development (Goodchild, 1999). The emphasis in GIS software development has gradually shifted more toward applications. The results from GIS analyses could provide critical indicators, for example, recognition of road networks (Wang and Howarth, 1987). The challenge is to incorporate intelligence into the computer and merge technologies to arrive at an acceptable solution.

## 2.15 Applications of 35-mm Aerial Slides for NPS Pollution Control

On April 1, 1996, the Virginia Department of Environmental Quality (DEQ) submitted the State's 303 (d) TMDL priority list to the USEPA. Approximately 50 water bodies were identified as impaired, mainly as a result of bacteria from agricultural sources. Combined with the provisions of the CWA, one hundred and forty hydrologic units (HUs) were identified as either being the source of the impaired stream or were at risk of future bacteria standard violations. In an attempt to address legal implications of developing TMDLs for the listed impairments, the Virginia DCR initiated a project with the Department of Biological Systems Engineering (BSE) of Virginia Tech (VA Tech) in 1997. The goal of this project was to identify confined animal facilities, a significant potential agricultural source for bacteria and nutrients (Mostaghimi and McClellan, 1997). The project started during the summer of 1997 and continued through 2001. The main task of the project was to manually identify confined sites for poultry, dairy, beef cattle and swine from 35-mm slides and derive a map layer for confined animal sites (Figure 2-1). The initial phase of the project focused on the Shenandoah Valley region of Virginia.

## 2.16 Confined Poultry Operations

Poultry waste could potentially cause substantial water quality problems in Virginia (VDEQ, 1999). There are 1300 unregulated poultry operations in Virginia. These poultry facilities produce more than 220 million birds and over 4.4 billion pounds of manure each year. Poultry is the largest component of the state’s farm economy, producing one–third of all agriculture revenues (Chesapeake Bay Foundation, 1998).

A typical poultry house can hold 20,000-30,000 broilers. Each poultry house annually generates over 225 tons of manure. The manure or litter, is usually collected, mixed with other materials, and stockpiled before being spread onto fields. One poultry house can generate the waste equivalent of a town. For example, the Virginia poultry industry

generates the equivalent of human waste generated by 1,800 towns the size of Ashland or Warrenton, Virginia (Chesapeake Bay Foundation, 1998).

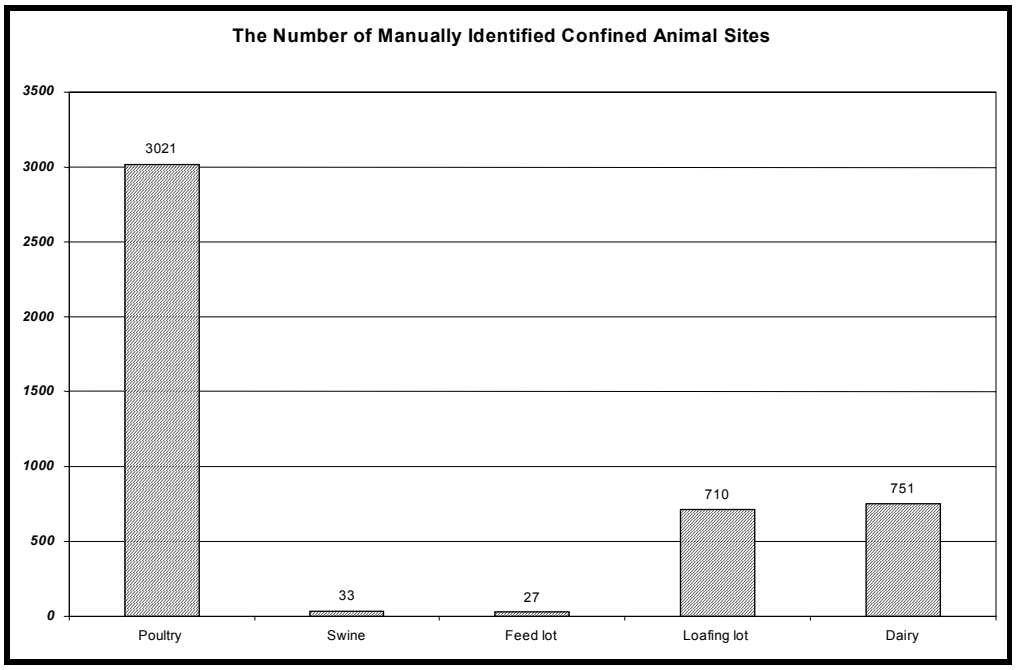


**Figure 2-1. The number of manually identified confined animal sites by BSE Dept., Virginia Tech (Mostaghimi, et. al., 1999).**

In the US, *"Ten companies produce 92 percent of the national poultry" (*Kratz, 1998) Contract poultry farming is when a company and a farmer sign a contract to produce broilers, breeders or replacement pullets and/or commercial laying hens. This has been the most dominant practice in poultry industry for more than 50 years. A poultry company furnishes feed and chicks that have been prepared genetically and otherwise for livability, growth and feed efficiency. Cooperation between the company and poultry producers is the key to success for both. An agreement to grow chickens must be reached with a poultry company before proceeding with any plans to build or produce a poultry operation. Buildings constructed by a poultry producer must conform to industry standards as to type, amount, installation, etc. Even when a producer purchases an existing poultry farm, a perspective company and/or company representative will inspect the premises to confirm that facilities meet industry standards. Other issues such as site selection, farmstead, building materials and shapes are normally discussed between a

producer and company. This type of cooperation has lead to similarities among buildings in terms of shape, construction materials and surrendering environment As a result, modern poultry houses tend to have similar image footprints.

In summary, numerous projects are using small-format aerial photography for visual interpretations to solve a multitude of problems. In one research project, Mostaghimi et. al., (1998), used 35-mm imagery to manually located poultry facilities. However, no research was found that attempted to automatically identify a poultry facility using 35-mm slides. The current state and national efforts to calculate TMDLs for all water bodies that do not meet state water quality standards highlight the need for a relatively rapid automated procedure that could be used to locate poultry facilities and calculate geometric characteristics needed to estimate nutrient and bacteria loads.

# 3. METHODOLOGY

As previously stated, the primary goal of this study was to investigate the feasibility of developing an automatic procedure for identifying poultry facilities using 35-mm color slides available from the USDA/FSA (Figure 3-1). Some experience with 35-mm color slides was gained from a project that involved manual identification of confined animal facilities using 35-mm color slides, which provided insight into poultry facility signatures and possible approaches that might be used to develop an automated procedure.

A typical signature of a poultry facility is shown in Figure 3-2. The two poultry buildings have clear and visible rectangular shapes. The buildings are often located parallel to each other. The facility always has road access for shipping the birds, which often completely circumvents the poultry house. A turn-around lot is usually located at one end of the building. The facility is normally located near a main road. Often one small rectangular or square shape building is located close to the facility for litter storage. The physical dimensions of poultry houses vary from approximately 9 to 12 m wide and 92 to 308 m or more in length. An industry standard size is about 12 m wide and 120 m long. Roofs of poultry houses are normally coated with white finish (Figure 3-3), which reflects heat. Heat stress can result in an adverse effort on rate of production and mortality of birds. Agricultural lands will often, but not always, surround the building (Figure 3-3).

It has been verified that a confined poultry facility can be manually identified from a 35-mm color slide with a high level of confidence (Mostaghimi, et. al., 1998). However, it is a big challenge for artificial intelligence to extract confined poultry facility houses from a 35-mm aerial image. As can be seen in Figure 3-1, many objects exist with white color signatures and some have shapes similar to poultry houses.


**Figure 3-1. An example of a 35-mm aerial slide.**


**Figure 3-2. Aerial view of confined poultry houses.**

**Figure 3-3. A modern confined poultry house surrounded by agricultural lands.**

This section describes the methodology developed to automatically extract poultry housing facilities from 35-mm slide imagery. The procedure involved developing techniques for creating digital images from the 35-mm slides; developing signatures for poultry housing facilities; refining image processing techniques; adapting GIS models; and developing appropriate techniques for shape analysis. The methods explored and procedures utilized are discussed in the following sections.

## 3.1 Development of Poultry House Signature of Poultry Signature

The development of a poultry house signature involves creating an image database, sampling and conducting analysis of sample data. The image database involves the selection of a study area, and the collection and processing of 35-mm slides. Sampling includes the development of sampling strategies and the creation of training data. Sampling analysis includes measuring the slide image's resolution, brightness, pattern, association, size, shape, contrast, and checking the date when slides were taken.

### 3.1.1 Study Area

The study area is located in Rockingham County (Figure 3-4), which has been subdivided into 27 Hydrologic Units (HUs) by VADCR for NPS control planning. The county, which covers 546,308 acres, has the largest concentration of poultry operations in Virginia. Rockingham County was also selected for this study because the 35-mm color slides were being used in a companion project (Mostaghimi, et. al., 1998) to identify confined animal facilities through manual techniques. Procedures developed to handle and process slides were shared by the two projects. Land use/land cover, soils, transportation and other topographic information were obtained from the DCR/DSWC. The local USDA/FSA has extensive records on poultry operations within the county including 35-mm slide coverage.

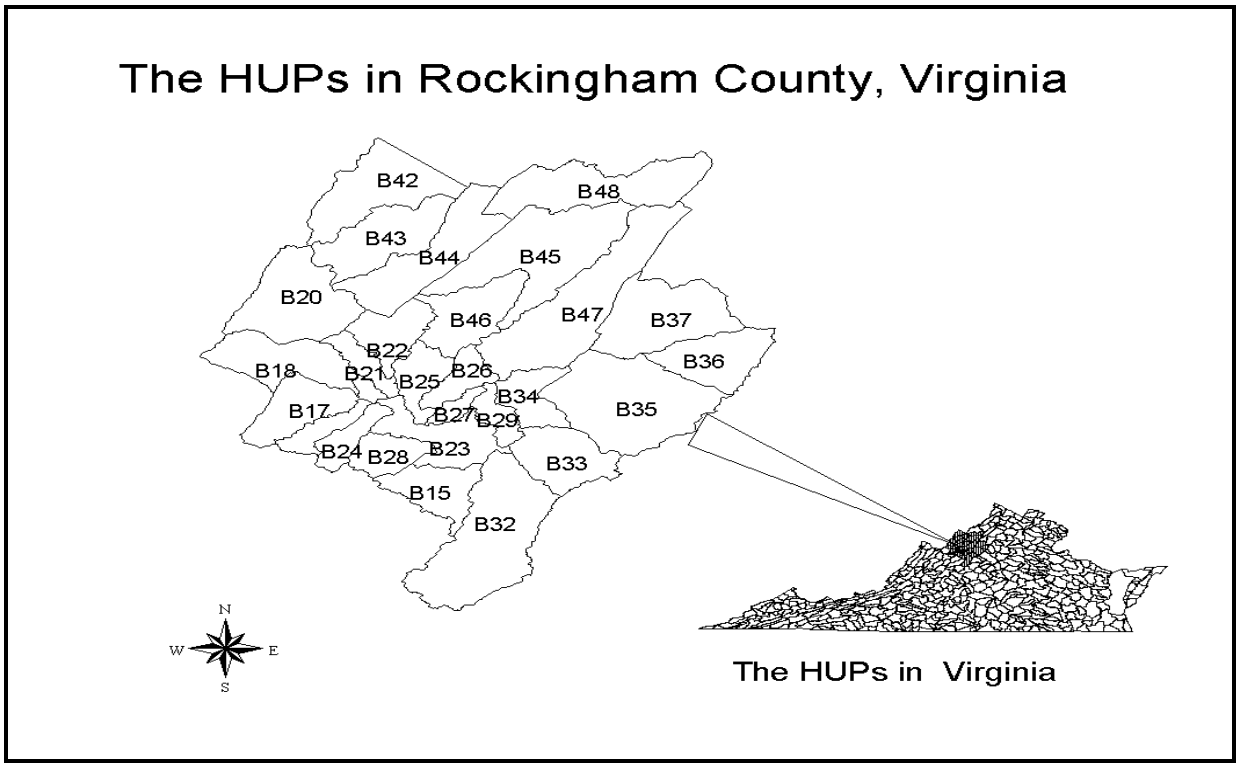


**Figure 3-4. Map of Rockingham County used as the study area.**

### 3.1.2 Data Management Protocols

The slides and relevant support data, such as flight height and camera focal length, were obtained from the local USDA/FSA office in Rockingham County. The date a slide was taken was printed on each slide image (Figure 3-1, bottom right). In order to efficiently process the slides, a data management protocol was established. The USGS 1:24000 scale quadrangle maps were used as the base map for registering the 35-mm slides converted to digital format. An index scheme was designed to easily reference the slide location or position on quadrangle maps. The USGS topographic maps representing coverage for Rockingham County were each sub-divided into five segments (columns) and were numbered one to five from left to right. Within each column, slides were numbered in ascending order from top to bottom. Each slide was assigned a unique filename, which provided specific information relative to the location of the slide on the USGS quadrangle. The filename contained 11 characters based on the following convention. The first three characters represented the HUC within which the slide was predominately located. The fourth, fifth and sixth characters, were the geographic reference number assigned to each USGS 7-1/2-minute quadrangle map. The seventh and eighth positions were used to represent the column location of the slide on the quadrangle map (west to east). The last three positions (9-11) were used to assign the North-South location of the slide within a column. For example, the filename B2539r02005 represents a 35-mm slide located in hydrologic unit B25, on 7-1/2 minute USGS quadrangle 39r column 2 and position 5 within column 2.

The approximate position of each slide image was identified manually on the topographic map. A slide projector was used to display an enlarged image of the slide on a screen and the location of the slide on the quadrangle map was determined from a manual inspection of the projected image and the topographic map. The outline for the slide

image was then manually sketched onto the topographic map. After the slides were manually referenced to the topographic maps, filenames were assigned based on the convention outlined above.

### 3.1.3 Converting Slides to Digital Images

The process of converting color 35-mm slides into digital format was accomplished with a slide scanner. The procedure, however, involved the evaluation of numerous factors to insure that the digital image would meet the needs of this study. These included the determination of the slide image scale, slide resolution, scan resolution, scan calibration, selection of scanner to meet project requirements, and the procedures for scanning each slide image.

#### 3.1.3.1 Resolution and Scale

The scan resolution (i.e. the digital image resolution) is related to the scale and resolution of 35-mm color slides. The processing procedure includes the selection of scanner, determination of the 35-mm slide image scale and its resolution, and scanning resolution.

Resolution is defined as *"the dimensions of the smallest objects recorded on an image"* (Campbell, 1987). The resolution of the 35-mm slide film is the dimension of the smallest object recorded on the film. Likewise, the resolution of the digital image from scanning of the slide film is the dimension of the smallest object recorded on the image scanned.

An image resolution is defined by the number of pixels displayed per unit of printed length in an image, usually measured in pixels per inch (ppi). The resolution of slide film determines the resolution of the scanning of slide film. The Rockingham County slides were taken with 35-mm Kodachrome film (Kodak, 1998). The highest resolution for the film is 50 line pairs per mm where one line pair is roughly equal to two pixels. Therefore, the theoretical resolution for the film was 2240 pixels per inch (ppi). However, due to variations in camera exposure and file development, for Kodak Kodachrome film, the maximum resolution is 2150 ppi (Kodak, 1998). Thus, the resolution for the 35-mm

slides was assumed to be 2150 ppi. This also implies that the maximum resolution available to convert the 35-mm slides into digital format is 2150 ppi

The scale of the digital slide was also an important parameter for this research. The representation fraction (RE) (Campbell, 1987) is widely used for determining image scale. It is expressed as:

$$Rrepresentation\ \ fraction = \frac{Focal\ length}{Altitude} \tag{3.1}$$

Where:

*Altitude is* flight height above the terrain, and

*Focal length* is camera focal length.

The altitude and focal length must be expressed in the same units. For example, a camera with a focal length of six inches used at an altitude of 10,000 feet would give a photo scale of 1:20,000 (i.e. 0.5/10000).

The focal length of the camera used for obtaining the 1995 Rockingham County 35-mm imagery was 24 mm and the flight height was 9500 feet (Pane, 1997), which sets the scale at about 1:120065. With this scale and a resolution of 2150 ppi, one pixel on the slide film is roughly equal to 1.42 m ground distance.

When printed, an image with a high resolution contains more, and therefore smaller, pixels than an image with a low resolution. For example, a 1-by-1-inch image with a resolution of 72 ppi contains a total of 5184 pixels (72 pixels wide x 72 pixels high). The same 1-by-1-inch image with a resolution of 300 ppi contains a total of 90,000 pixels. Higher resolution images usually produce more detail and subtle color transitions than lower-resolution images. The resolution required in the scanned image is directly proportional to computer storage requirements. Computer storage requirements increase as the resolution becomes finer. A 35-mm slide scanned at 500 ppi generates a 10-megabyte digital file when formatted in the window bitmap format (bmp). A 43-megabyte file resulted when the slide image was scanned at the maximum resolution of

2150 ppi. The goal was to scan at a resolution sufficient to meet the needs of the research study. A slide can be scanned at a resolution less than 2150 ppi, for example, 500, 1000 or 1500, etc. The smaller the resolution results in the larger the ground distance represented by a pixel. The ground distance is important because of the need to preserve (i.e. recognize) the poultry building footprints. The following equation was developed to calculate ground distances for different scanning resolutions.

$$GD = \frac{FA}{Fl * SR} * 25.41 \qquad (3.2)$$

*Where:*
*GD is ground distance (m);*
*FA is flight altitude above the terrain (m);*
*Fl is the camera focal length (mm);*
*25.41 is the conversion factor; and*
*SR is the scanning resolution, 500, 1000 or 1500, ppi etc.*

The length of a poultry building can vary from 30 m to over 100 m. The width of a poultry house is typically 10 m. A slide scanned at 500ppi with a resolution at 6.1 m should be sufficient to represent confined poultry houses. This was confirmed by the BSE project (Mostaghimi, et. al., 1998). The 35-mm slide shown in Figure 3-1 was scanned at 500 ppi.

### 3.1.3.2 Selection of Scanning Equipment

The Nikon LS-1000 (Super CoolScan) (Nikon Corporation, 1996) was selected for the study. It has a precision of 12-bit A/D (Analogue-to-Digital) conversion for wide dynamic range and for up to 4096 depths of color. The Nikon LS-1000 has special optics, which preserve definition and overcome problems of geometric distortion. This scanner incorporates Nikon's new high-power light emitting diode (LED) illumination technology. This technology can speed the process of scanning and also can contribute substantially to a unique color reproduction from negative and slides. Its LED-based

design results in lower power consumption, virtually no heat dissipation or noise, as well as long life and durability. It is a single-pass Red-Green-Blue (RGB) line sequential scanner that uses LED strobe operation to produce RGB triplets with a monochromatic charge coupled device (CCD). In general, scanners that use monochromatic CCDs are inherently less noisy and have fewer color and spatial artifacts than single-pass designs that use a color CCD with a white-light source. Super CoolScan is reliable and is low cost. The scanner runs extremely quiet (no fan) and is compact. The retail price for Super CoolScan is approximately $1,500. The scanner settings for resolution and physical dimensions can vary up to 2700 dot per inch (dpi) or pixel per inch at a resolution of 0.0001 inch. The scanner has an optional automatic slide feeder for high-speed archiving.

#### 3.1.3.3 Calibration of Scanner

There is no standard guidance for scanning slides (e.g. scanner settings for range of colors). The settings for color range were first made based on suggestions from the Nikon LS-100 user's manual. A procedure was then developed to calibrate the scanner color settings to insure that the digital image represented the colors of the original slide. The brightness is an important parameter in developing a signature for poultry facilities.

Five or more slides were randomly selected from the Rockingham County slide database. The scanned image was displayed on the computer screen. The digital image was compared with the original slide visually and the RGB colors (i.e. red, green and blue) adjusted until a satisfactory match was achieved. The same operator performed all calibration adjustments to minimize operator's error. The same steps were duplicated for the selected slides. The final color setting was determined based on the average of settings for the individual slides. All slides were scanned with the same color settings to insure that spectral characteristics were based on the same ranges. After calibrating the scanner, the same color settings were used for all slides. The resolution of the scanner was set to 500 ppi for this study.

### 3.1.3.4 Pre-processing of 35-mm Slides

The main role of pre-processing was to correct problems caused by such factors as atmospheric interference and system noise (Campbell, 1987). The two common issues encountered with 35-mm slides during a manual interpretation study (Mostaghimi, et. al., 1998) were related to atmospheric (e.g. cloud cover) and geometric (e.g. variable flight altitude and attitude) factors. Atmospheric conditions can impact aerial photography by reducing the energy emitted from a ground object and/or adding a scattered radiance to the signal detected by the sensor. In general, the atmospheric conditions can change the spectral distribution of energy of ground objects to the sensor. However, studies conducted during manual interpretation found that the spectral characteristics of poultry houses did not vary from one season to another. The major atmospheric problem that was encountered was related to cloudiness. This issue was resolved by careful selection of slides. Most 35-mm slides are normally taken on a sunny day to minimize the effects of cloud cover. However, even on a sunny day, some cloud cover or haze often occurs. Clouds impact the resolution and contrast of a 35-mm slide and those slides that are significantly impacted by cloud were excluded from this study (Figure 3.5).

Systematic and random distortions (Campbell, 1987) represent two types of geometric problems. Systematic error is predictable and can be corrected by applying formulas derived by mathematically modeling the sources of the distortions, for example the deskewing procedure (Lillesand and Kiefer, 1994). Flight data was not available to make any attempts to resolve image problems related to different flight heights and/or attitude. This was one of the conditions of 35-mm slides that were readily available for this study. Therefore, the geometric correction was only made during geo-referencing of the digital image. As noted by Campbell (1987), ground control points (GCPs) are often used to correct random distortions through a rectification or resampling procedure. Rectification is the process of geo-referencing a non-referenced image. GCPs are known ground locations that can be located on both the photography and the digital imagery. GCPs were selected from USGS 24000 topographic maps and were carefully checked by site visits

and verified with a Global Positioning System (GPS). The GCPs were used by ArcView 3.1 (ESRI, 1997) to perform rectification.

### 3.1.3.5 Development of the Training Database

A database was selected and used to develop spectral characteristics, shape characteristics and background conditions. This database is referred to herein after as training database because it was used to develop signature characteristics that were used to identify poultry houses for the remainder of the county. The following sections discuss the procedures used to select a training database and the types of analysis that were performed. Campbell (1987) gives a detailed description of the significance of training data. The selection of training data is very important for recognition and classification (Scholz, et. al. 1979; Hixson, et al. 1980). The training data set should contain the range of characteristics expected in the signatures for poultry facilities. The characteristics were selected to include the range in location of the facilities, landuse conditions surrounding the facilities, the geometric configuration of buildings, and the spectral and texture properties of the facilities.

The creation of a representative training data set involved a thorough understanding of the characteristics of poultry facilities and signatures that they leave on 35-mm slide imagery. This information was obtained from a review of literature, site investigations, and consultation with poultry growers and personnel from service agencies such as USDA/FSA and USDA/NRCS. Special attention was made on geographic location, geometric representation, surrounding land use and cropland relative to poultry facilities

After a thorough review of poultry facility characteristics, a sampling plan that consisted of slide data set and a facility data set was designed to create a training data set. The slide data set was defined as a complete slide image. Whereas the facility data set included a small sub-set of the entire slide image and contained no more then one poultry facility. The slide samples were randomly selected from the Rockingham County slide data set.

The slides had been manually interpreted for poultry facilities and the results verified by field checks. The facility samples were selected from the slide data set.
The criteria used for the selection of samples included the following:

- The samples were selected so as to be representative of the region (county), i.e. the slide samples were not clustered in one section of the county used as the study area nor were the facility samples collected from one region of the slide image;
- Facility samples were not selected from the edge or boundary of a slide image to minimize the effect of photo distortion;
- Facility samples were distributed uniformly throughout the slide image to incorporate any diversity within the scene;
- The spectral property of the facility sample should be homogeneous. Data from each facility sample should have unimodal frequency distribution, i.e., only one peak in a histogram; and
- The building footprints for facilities in facility samples should be clear and measurable.

This selection strategy provided for three techniques to describe poultry signatures and to automatically identify them using 35-mm imagery. The spectral characteristics and shape identification procedures were developed for a facility sample, evaluated on the slide sample, and refined and/or modified as required, and then evaluated on the larger Rockingham County data set. The procedures utilized to evaluate spectral characteristics are discussed in the following sections.

Adobe Photoshop (1996), which is a popular commercial software package, was used to process digital images. The system tool "*magic wand*" was quite useful for this study. The scanned images, which were in Microsoft Windows BMP format, were converted into RGB data format because the RGB format represents the three primary colors (red, green and blue) in separate data layers. The spectral distributions of the three colors were analyzed independently.


**Figure 3-5. An example of a 35-mm aerial slide with significant cloud cover**

The "*magic wand*" tool was used to extract the poultry facility houses from their background. This tool "*lets you select a consistently colored area (for example, a red flower) without having to trace its outline*" (Adobe, 1996). Important options with this tool include "*Tolerance*" (setting up color range) and "*Anti-aliased*"(defining a smooth edge) (Adobe, 1996). The "*magic wand*" was very useful for extracting a regular or irregular polygon from an image without digitizing its boundary. It was used to extract images that included poultry buildings from slide samples. The extracted poultry facility buildings were saved as individual image files to form facility samples.

### 3.1.3.6 Spectral Analysis of Training Data

The spectral characteristic analyses on slide and facility data sets were conducted using the "*Histogram*" function of Adobe Photoshop (Adobe, 1996). As illustrated in Section 2.7.2, each pixel contains brightness intensity for one or more color bands. The "*Histogram*" function was used to understand the frequency distribution of the pixels in an image by graphing the number of pixels at each color intensity level. Important options within the "*Histogram*" function included mean, standard deviation, pixel, level, count and percentile. The mean option gave the average brightness intensity for each color band. The deviation option showed the variation of brightness intensity for each color band. The pixels option gave the total number of pixels used to calculate the histogram. The level option displayed the intensity level of the area underneath the pointer. The count option showed the total number of pixels corresponding to the intensity level underneath the pointer. The percentile option displayed the cumulative number of pixels at or below the level underneath the pointer and was expressed as a percentage of all the pixels in the image, from 0% at the far left to 100% at the far right.

The "*Histogram*" analysis described the spectral characteristics for individual slide and facility samples. The distribution of spectral characteristics provided an indication of the grouping of poultry house spectral signatures, which can be used to separate them from other signatures. The separation values could be selected manually from a visual inspection of the histogram graph. The separation values varied from slide to slide.

Statistical procedures were investigated to provide an overall estimate of the separation value for the entire training dataset. The Anderson-Darling normality test (Minitab, Inc. 1995) was used to test both slide and facility samples for normality. This provided an indication of how to select the separation value. For example, if the normality test shows that spectral characteristics for facility samples are normally distributed then the mean spectral characteristics can be used as the separation value. The normality tests for slide samples showed that the spectral characteristics were normally distributed while the spectral characteristics for facility samples were not normally distributed. Therefore the minimal value was used for the separation value. Minitab (Minitab, Inc. 1995) was used to calculate the descriptive statistics for both slide and facility samples.

Each slide and facility sample has three primary color bands (red-green-blue). Each color band records different spectral information. To determine the correlation between color bands, the statistical package ER Mapper (1995) was used to calculate the variances-covariance matrix and report eigenvectors with eigenvalues. This statistical procedure was based on factor analysis theory, a well-known statistical methodology (Davis, 1986). For the three-color bands, the procedure involves calculating a three x three matrix of variances and covariances [$s^2$]. From [$s^2$], three eigenvalues and three eigenvectors are extracted. Because the variance-covariance matrix was symmetrical, the three eigenvectors were mutually orthogonal. The values in each eigenvector are the correlations among the three-color bands.

### 3.1.3.7 Geometric Analysis of Training Data

The purpose of the geometric analysis of training data was to quantify the shape of poultry houses. To define shape characteristics of the facility sample, Photoshop 5.0 software (Adope, 1996) was used to assign the same brightness intensity to each color band to form a uniform image. ArcView 3.1 (ESRI, 1997) was used to import the uniform image of the poultry facility sample and convert it into a shape file. ArcView functions were then used to calculate area, perimeter, length and width of facility samples. To determine the length and width, the "*Measurement*" tool of ArcView was

used to manually measure length and width to provide the best approximation to the rectangular shape. The length and width were measured three times, then averaged.. ArcView was then used to calculate compactness and the length-width ratio. The statistical package Minitab (Minitab, Inc. 1995) was used to calculate descriptive statistics (i.e. mean, maximum, minimum, standard deviation, and normality test) for compactness, width, length and length-width ratio.

## 3.2 Imaging Processing of Slide Samples

Image processing was used to classify the image from the slide samples into similar categories and to group the categories based on spectral characteristics. Almost without exception, the image was significantly improved when functions to enhance, to show contrast or to transform the image were applied. These functions tended to overcome problems within each color band. The brightness intensities were recalculated after image processing. The two general image-processing procedures used included image classification and segmentation. Image classification has been widely applied to land use/cover classification (Campbell, 1987) while image segmentation has been extensively used in medical applications.

Significant differences in brightness intensity existed between the slide and facility samples (i.e. dominated by poultry house). Image segmentation procedures were used to separate objects with spectral characteristics similar to poultry facilities from background information. To achieve these separations, thresholding through the use of histogram analysis and descriptive statistics was used.

There also existed differences in the brightness intensity in the three primary color bands. To attain additional separation, image processing functions, such as PCA, RGB to Hue and RGB to Saturation, were used to represent the most dominant information in the three primary bands. The following sections describe these functions.

### 3.2.1 Thresholding Procedures

Thresholding was used to divide an image into homogeneous regions. The procedure uses image gray level or color intensity to cluster picture points that have similar brightness values independent of their positions. Point properties (i.e. individual pixel values) are used to obtain an image that only contains objects with brightness intensity within predetermined levels.

As illustrated in the literature review section, an image can be represented as a function represented by Pic(i,j). The thresholding function g(i,j) is defined as:

$$g(i,j) = \begin{cases} Pic(i,j), Pic(i,j) \geq T \\ 0(null), Pic(i,j) < T \end{cases} \quad (3.3)$$

Where:

*T: is the thresholded value.*

The values of Pic(x, y) greater that T were kept. The values of Pic(x, y), which were less than T, were assigned to numerical zero. The new image after the thresholding procedure only contains the object, of which the brightness intensity is greater than or equal to the threshold value.

Histogram analysis was used to determine the distribution of the brightness intensity of each image. The result was used to study the brightness levels of different objects. The intensity on the edges was used as the thresholding value. For single distribution, the valley corresponding to the intermediate brightness levels was used as the thresholding value. When multiple distributions exist in the histogram (i.e. several peaks and valleys will exist), the peaks and valleys were mathematically defined as local maximum and local minimum, respectively. The valleys in a histogram are often the locations to define the thresholding values (Tomita and Tsuji , 1990).

Procedures to automatically determine thresholding values were also explored. The equation for calculating entropy is given by Tomita and Tsuji (1990) as:

$$Entropy = -n\sum_{i=1}^{n} h(i)\log h(i) \quad (3.4)$$

*Where:*

*n is the number of pixels,*

*h(i) is the value of a pixel.*

When the distribution of a histogram is uniform, entropy is at maximum. The value of entropy approaches zero when there is a significant cluster in the histogram.

Directivity is used as the indicator for edge detection for a histogram. The entropy value gives an approximate measure of directivity. The directions of directivity are given by detecting clusters in the histogram (Tomita and Tsuji, 1990). For example, 2N (N>0) clusters detected in a histogram indicate the directivity in N directions. The uniform distribution of a histogram means there is no directivity in the image.

The entropy and directivity are effective tools for calculating the thresholding values for an image when the spectral variability is small (Tomita and Tsuji, 1990). However, significant spectral variation was found in the 35-mm slide images. For example, the spectral distribution of a slide sample is shown in Figure 3-6, which has many valleys. The spectral characteristics associated with confined poultry facility houses where not greatly different from background spectral conditions. It was very difficult to identify the location of valleys that identify distinct clusters, which would be associated with poultry facilities in a histogram. Therefore, it became very difficult to develop a totally automated procedure to determine the thresholding value for poultry houses.

The results of spectral analysis on poultry facility samples represent spectral ranges of poultry facilities for the three primary colors. The lowest brightness values for the three colors were used for thresholding values for image segmentation.

The following procedure was used to find the thresholding values for objects, which have spectral characteristics similar to poultry houses. The ER Mapper, 1995 software package was used for all image-processing activities. ER Mapper has been widely used by different scientific disciplines for image processing applications and has reliable functionality necessary for this study, e.g. image conversion to polygon, PCA analyses, image resampling, mosaic, etc. The thresholding function in ER Mapper was developed based on equation 3.3.

The histogram function of the ER Mapper was utilized to display the brightness intensity of a slide sample, which is in a RGB model. Figure 3-6 gives an example histogram from a slide sample. The lowest value for each primary color was selected as the thresholding value for each color band. For example, the brightness value in Figure 3-6 varies from 19 (bottom left) to 255 (bottom right) with the lowest value around 240. The difficulty is that each color band generally will have a different value and these values will vary from one slide image to another.

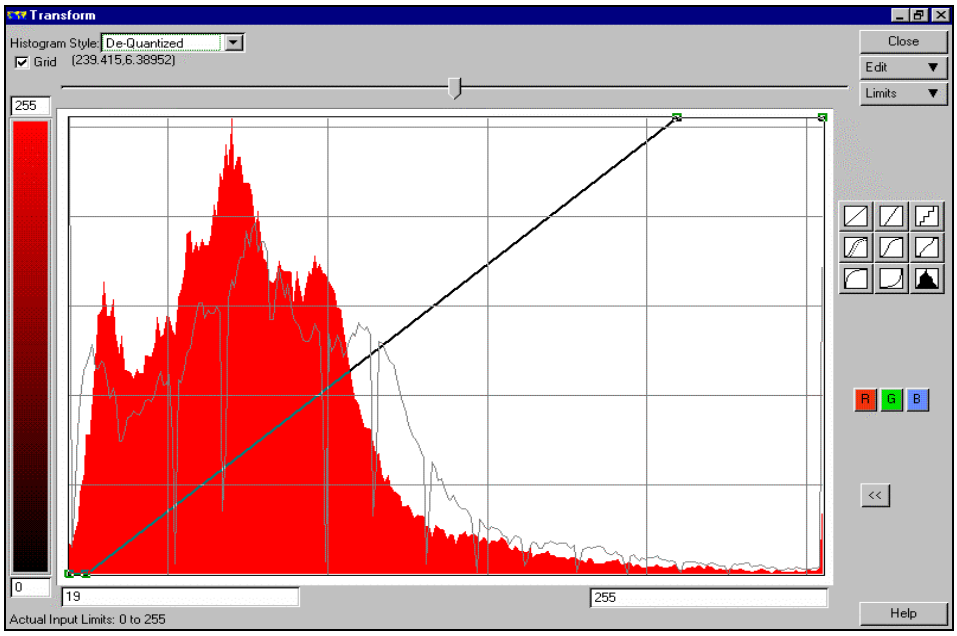


**Figure 3-6. Example of histogram with multiple peaks and valleys**

Because the values vary across colors and from slide to slide, statistical procedures were investigated to provide an overall estimate of the separation value for the entire slide training dataset. These procedures were identical to those used for describing spectral characteristics of the facility training dataset. For continuity purpose, they are repeated in this section. The Anderson-Darling normality test (Minitab, Inc. 1995) provided an indication of how to select the separation value. The normality tests for slide samples showed that the spectral characteristics were normally distributed while the spectral characteristics for facility samples were not normally distributed. Therefore, the minimal value was used for the separation value.

The selected values were entered into the thresholding function of ER Mapper to process each color band of a slide sample. The new image resulting from the threshold analysis was saved for additional image processing, which will be described in the next sections. Figure 3-7 gives the menu layout for ER Mapper thresholding analysis. Figures 3-8 and 3-9 give an example of the thresholding procedure applied to a slide and to a poultry facility sample, respectively.

### 3.2.2 Image Enhancement and Transformation

The image resulting from threshold analysis had narrow spectral ranges similar to confined poultry facilities. However, the image still contained other information not related to poultry houses (e.g. Figure 3-10), which required additional image processing methods to remove the non-related information. Image enhancement and classification methods illustrated in Chapter 2, Section 2.8, were evaluated for potential to remove redundant information. Image classification methods were developed to categorize information over large areas. Supervised classification requires training data while unsupervised classification does not. With the narrow spectral ranges, neither the supervised nor unsupervised image classification scheme was appropriate for extracting confined poultry facilities from slide samples.

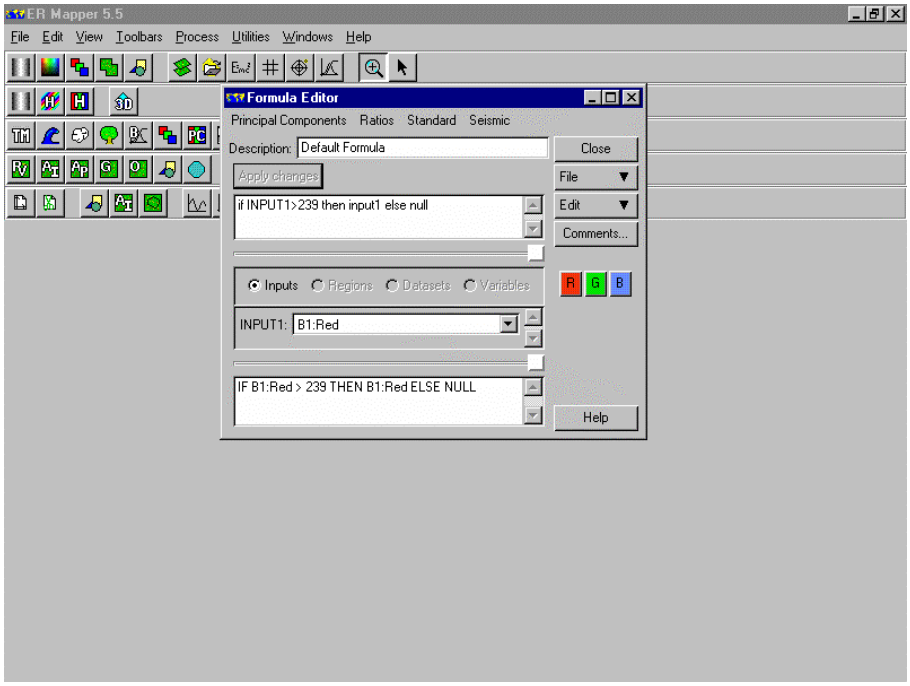


**Figure 3-7. Example ER Mapper menu for thresholding analysis**


**Figure 3-8. Slide image B3542R01006 before thresholding**.

**Figure 3-9. Slide sample B3542r01006 after thresholding.**

In order to properly extract a poultry house, its spectral information must be preserved (i.e. the variability in the brightness intensity must be combined such that the physical geometry for the building is not changed). The variability appears as color changes with the physical footprint for the building (Figure 3-10). For the same color band, differences also existed among the poultry facility samples that were discussed in an earlier section. Principle component analysis (PCA) analysis, Red Geen Blue (RGB) to Hue and RGB to Saturation were used to analyze the new images created from threshold analysis. These three procedures are discussed in the following sections.

### 3.2.3 Principal Component Analysis

The intensities for each primary color band of slide and facility samples had different spectral characteristics. Red and green color bands had the widest spectral distributions while the blue color band had the narrowest distribution. The blue color band had a cluster of highest brightness intensity. Because the spectral distributions for each color band and the gray band (average of the three color bands) still contain significant variability, it was not possible to separate the poultry facilities from background noise without further processing. Thus, the Principal component analysis was used to determine the most influential information from the multiple color bands.

Principal component analysis (PCA) is a linear transformation technique related to Factor Analysis (Davis, 1986). When the channels (bands) of an image are given, PCA is used to produce a new set of images, known as components, which are not correlated with one another and are ordered in terms of the amount of variance calculated from the original band (Eastman, 1997). In remote sensing, the components are thus a statistical abstraction of the variability inherent in the original band set. Each component carries new information. The results of PCA usually are represented in the ranks of their influences. The first few components carry most of the real information. The later components probably contain minor information including some noise (Eastman,

**Figure 3-10. Thresholding procedure applied to a poultry facility sample.**

1997). Satellite images have multiple color bands. It is very common to find that the first two or three components explain virtually all of the original variability in reflectance values (ER Mapper, 1995). Other components are dominated by noise. By rejecting these components, the volume of data is reduced with no appreciable loss of information. With each component of PCA, the sequence of important variables can be determined (ER Mapper, 1995). The 35-mm slides contain three bands: red, green and blue. Each band contains relevant information with noise. PCA was used to extract the most important information from 35-mm slides based on combination of the three bands.

The ER Mapper software package (ER Mapper, 1995) was used to conduct PCA on each slide image. The variance-covariance matrix was used for PCA calculation (Davis, 1986). The slide sample has 3 x 3 variance-covariance matrix. When the covariance matrix is symmetrical, the sum of the eigenvalues of a matrix is equal to the trace of the matrix. Therefore, the eigenvalues can be used to decide the rank of influences of color bands.

The variance-covariance matrix was calculated for each slide sample and the results (e.g. Table 3-1) were used to determine the most influential components. Typically, the three primary color bands are highly correlated (correlation matrix summary, Table 3-1). Also, the correlation eigenvectors for Principal Component 1 contain most of the spectral information for all bands. Principal Component 2 is dominated by the blue color band and Principal Component 3 is dominated by the green color band. This analysis demonstrated that no single band could be used to identify the poultry house.

Another problem with the imagery that had to be resolved was the practice of stamping the flight date on the film negative, which consequently became imbedded into the digital image when the slide was scanned (e.g. Figure 3-11 for flight date 95 9 23). This created background noise that had to be removed because characters such as 1, l or I could be misclassified as a poultry house.

**Table 3-1. Summary of statistical analysis of a slide sample by ER Mapper.**

| | Band1 | Band2 | Band3 |
|---|---|---|---|
| **Non-Null Cells** | 3486800 | 3486800 | 3486800 |
| Minimum | 3.000 | 6.000 | 0.000 |
| Maximum | 255.000 | 255.000 | 255.000 |
| Mean | 80.860 | 80.906 | 90.602 |
| Median | 78.000 | 78.000 | 88.000 |
| Std. Dev. | 37.232 | 27.813 | 26.950 |
| Std. Dev. (n-1) | 37.232 | 27.813 | 26.950 |
| Corr. Eigenval. | 2.796 | 0.174 | 0.030 |
| Cov. Eigenval. | 2691.200 | 170.203 | 24.651 |
| **Correlation Matrix** | **Band1** | **Band2** | **Band3** |
| Band1 | 1.000 | 0.935 | 0.826 |
| Band2 | 0.935 | 1.000 | 0.932 |
| Band3 | 0.826 | 0.932 | 1.000 |
| Determinant | 0.014 | | |
| **Corr. Eigenvectors** | **PC1** | **PC2** | **PC3** |
| Band1 | 0.570 | -0.702 | -0.427 |
| Band2 | 0.592 | -0.009 | 0.806 |
| Band3 | 0.569 | 0.712 | -0.411 |
| **Inv. of Corr. Ev.** | **PC1** | **PC2** | **PC3** |
| Band1 | 0.570 | 0.592 | 0.569 |
| Band2 | -0.702 | -0.009 | 0.712 |
| Band3 | -0.427 | 0.806 | -0.411 |
| **Covariance Matrix** | **Band1** | **Band2** | **Band3** |
| Band1 | 1386.209 | 968.084 | 828.903 |
| Band2 | 968.084 | 773.568 | 698.620 |
| Band3 | 828.903 | 698.620 | 726.277 |
| Determinant | 11291207.246 | | |


**Figure 3-11. An example of a slide sample containing the flight date before its removal.**

**Figure 3-12. Image slide from Figure 3-10 after PCA analysis.**

The spectral characteristics of the flight date were determined by extracting the flight date signature using Adobe software to create a flight date image. Minitab statistical software was used to calculate the descriptive statistics for each color band. From this

analysis, the red color band dominated flight date. Based on the PCA analysis on slide and flight date samples, PCA was performed on the red color band to remove the flight date and to enhance the slide sample image. The results from PCA analysis were saved as a new image and used for the RGB to Hue and RGB to Saturation image processing steps. Figure 3-12 illustrates enhancement achieved through the PCA analysis. Note that the variation in color has been removed, however some background signature still remains within the poultry house footprint.

### 3.2. 4 RGB to Hue and RGB to Saturation

The image format used in ER Mapper is the RGB model (i.e. red, green and blue colors). Each color takes 8-bits to represent the depth of a color and each color is represented by values ranging from 0 to 255. The RGB color model uses separate red, green, blue and intensity color receptors to display up to four (gray color band) separate layers of data at the same time (ER Mapper, 1995). The gray intensity is determined by averaging the values of the three colors. For example, if the same data set is loaded into each of red, green and blue layers, the resulting image appears as a shade of gray. Areas in an RGB image that look predominantly red would have high values in the red layer of data and low values in the other layers. This also applies to blue and green colors (ER Mapper, 1995). The purpose of RGB to Saturation transformation was to enhance the image. After the image was processed by the RGB to Hue transformation, the RGB to Saturation transformations were performed, the image only contains a single color data layer. The objects remaining in the image only contain uniform information.

As illustrated in the literature review, the Hue, Intensity and Saturation (HIS) model is a different RGB model. A layer of each type must be present to obtain an image because the hue, intensity and saturation layers are interdependent. The data in hue layer determine the color of the pixel. The values in HIS model are mapped around a color wheel so that 0 is red, increasing values move through yellow, green, blue, violet and back to red again at a value of 255 (Bonham-Carter, 1994). The data in the Saturation layer determine how much of the color there is, that is, how saturated it is. The data in the

intensity layer determine the brightness or darkness and ranges between 0 for black and 255 for white. Details on HIS can be found in ER Mapper (1995) and Bonham-Carter (1994).

As stated before, the purpose of RGB to Hue transformation was to enhance the image. This transformation was performed on the images created by the PCA procedure. Each pixel value was modified to better represent the true value of objects after the transformation. This was required because the new data layer was the result of a combination of three RGB layers in HIS setting. This transformation was conducted on the red color band and the new image will retain the dominant wavelength in the red color band.

RGB to Saturation transformation was used to purify dominant color band. The spectral signature of a poultry facility sample is represented by the combination of the primary colors. Each color band can represent a majority but not 100% of the spectral information for a poultry facility. The intent of the RGB to Saturation transformation on red color was to remove the influences by red color. This procedure removed the flight date from the image. The remaining information included the spectral characteristics of the green and blue color bands.

The green color band was removed by the RGB to Saturation transformation on green color band. The green color band had a larger standard deviation in brightness intensity compared with the blue color band. After this transformation, the image only contained a single color band (i.e. the blue color band). The new image was saved for shape analysis. The process is illustrated in Figure 3-13.

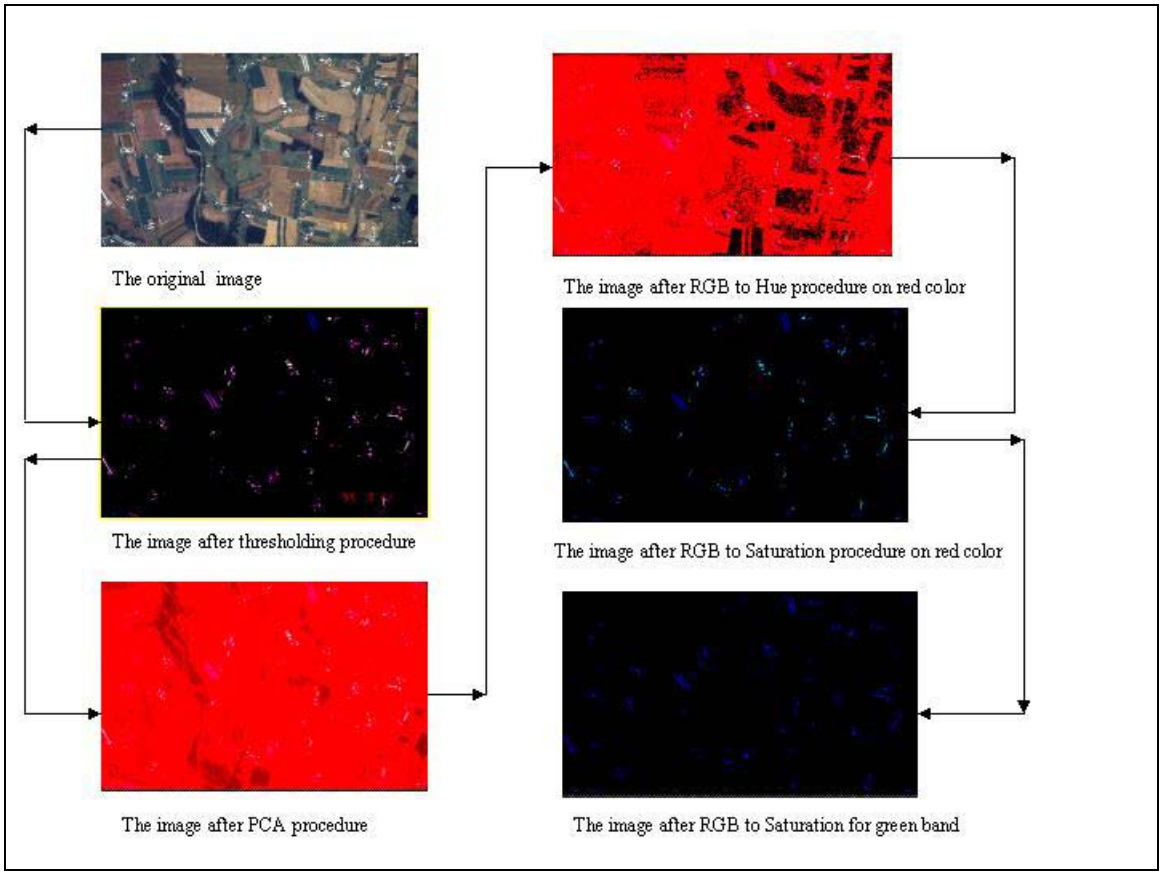


**Figure 3-13. Illustration of image processing procedures to identify poultry houses.**

## 3.3 Shape Modeling

The objects (polygons) in the images processed with the procedures discussed in the previous section have the same spectral characteristics and include poultry facilities. The image, however, still contains hundreds of other objects. Additional analyses were required to identify which objects were poultry facilities. The shape of confined poultry facility building was used as an indicator during manual identification. However, as previously described in the literature review, no automated procedure existed to identify poultry facilities from 35-mm slide imagery. The following section discusses the shape modeling methodology developed to identify poultry facilities from other objects with similar spectral characteristics.

The confined poultry facility buildings were captured in an image with distortions because of photographic distortions due to variations in conditions when the slides were taken. These distortions include shape rotation, scale and change in shape. In order to minimize the impacts caused by these distortions, a dimensionless shape modeling method was needed. Unlike other characteristics of an image such as texture, spectral, color and brightness, the shape modeling procedure for a confined animal facility should be invariant for translation, rotation and scaling. The shape model should preserve the property of information to be used to reconstruct the shape of a confined poultry facility.

Shape similarity tests (Ullmann, 1973; Otterloo, 1991) often have been used to identify an unknown object by comparing it with a fully described object. The similarity test measures objects in an image. For pattern recognition, a small set of low-order moments is used to distinguish between different patterns. Tomita and Tsuji (1990) used the first and second moments to match the shape. In general, scalar transform measurements are often used in similarity test (Pavlidis, 1978; Tomite and Tsuji, 1990). Scale transformation measurements include area, perimeter, circularity, size, moment, and principle axis (Tomite and Tsuji, 1990). Scale transformation measurements should be

invariant to orientation and scale of objects (Tomite and Tsuji, 1990). In order to use similarity test, a shape model, which can represent both a known sample with an object to be identified, was needed. A poultry facility building has a distinguishably rectangular shape with generally standard dimensions. Modern poultry facilities, particularly, are similar in terms of shape and dimensions. This similarity is important to develop a shape model for poultry facility buildings. The results of the shape analyses on training data were used to define the similarity among the confined poultry buildings. The shape similarity test then is used to compare these objects in the processed image with the training data to determine if these objects belong to poultry facility buildings. In order to use similarity, however, a proper shape model is needed to describe an object in an image. Also, the shape model could be used to conduct scalar transformation measurement (i.e. it should be invariant for transformation, rotation and scaling).

As discussed in the literature review, most shape modeling analyses are conducted on objects, which are presented in a vector data mode, for example, moment analysis and Fourier transformation. Vector data usually can be used to describe the object more precisely. The objects remaining after image processing, however, are represented in a raster data mode. To conduct shape analysis, the raster images resulting from spectral analysis were converted to vector format by using raster to vector conversion functions in ER Mapper (ER Mapper, 1995). Figure 3-14 shows an image that was converted into a vector format. The vector polygons resulting from the conversion were complex and irregular. The polygons were varied due to their orientations. Use of the Fourier transformation requires that samples be representative of all orientations of confined poultry facility buildings. The same size and shape of confined poultry facility buildings in different images would be expressed in different representations by Fourier transformation due to the different orientations in the image. The application of Fourier transformation has very strict mathematical requirements, i.e. the tolerance on samples is very low. As previously discussed in the literature review section, the Fourier transformation is not only rotation variant but also requires significant computer resources. Thus, the Fourier transformation was determined to be inappropriate for modeling the shape of confined poultry facilities in this study.

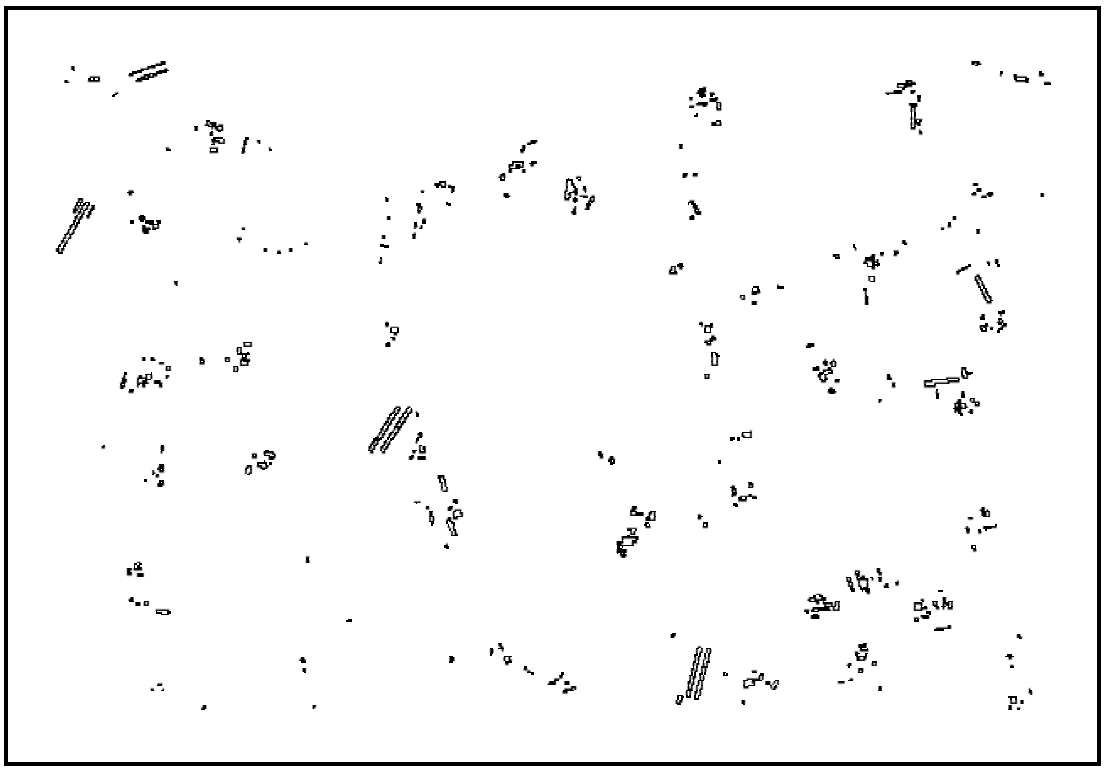

**Figure 3-14. A slide sample image after conversion to vector format.**

Moment analysis was also evaluated for its potential to be used for the shape model. Unlike Fourier transformation, moment analysis is rotation invariant. Poultry facilities have a distinguishably rectangular shape. Three orders of moments are often significant to represent a shape. Three orders of moments are often calculated for both samples and the object. Then similarity test is used to identify the object by comparing it with the samples. Although all the confined poultry facility buildings have similar shapes, it is impossible to build the sample database, which can be used to conduct the moment analysis. Also like the Fourier transformation, there are usually hundreds of polygons in a 35 mm slide and significant computing resources are required for conducting moment analysis. The moment analysis approach was also determined to be inappropriate for this study.

The shape modeling methods used by earth scientists were also evaluated as an alternative procedure for identifying the poultry facility buildings. These methods are represented in relatively simple mathematical expressions, compared with moment analysis and Fourier transformation. Objects are represented in vector data model as points, lines or polygons. The point is the minimum unit. The line and polygon are made of points. The vector data can be used to describe the object more precisely. Each polygon can be identified by a unique identification code. Each unique identification code can be associated with a database with different types of attributes. Through the database management function, different attributes can be manipulated to perform different GIS analyses. A polygon has basic geometric attributes: length, width, area and perimeter. These attributes can be manipulated by GIS functions. A defined polygon can be used to represent a specific object. The polygon can be described by a specific mathematical function. For example, a confined poultry facility can be represented by a rectangular polygon. The methods represented in Table 2-4, Chapter 2 can be integrated with GIS functions.

Some GIS software packages such as ArcInfo and ArcView already have functions, which can be applied to some methods. For example, GIS is based on topographic relationships. The area and perimeter of polygons are automatically calculated. A shape

in a GIS environment can be recognized as a polygon when the shape is presented in vector format. A polygon has perimeter, area, length and width. The GIS functions for polygons can be used for shape identification and analysis. In general, length, width, area and perimeter are often regarded as the basic elements for shape analysis by using these methods in Table 2-4 of Chapter 2.

To apply the shape modeling method, shape parameters were defined as follows:

- Area (A): The area was defined as the total number of pixels in the region or object when data are represented in raster mode. In the vector mode, the area was the area of the polygon or object.
- Perimeter (P): The perimeter of the object was defined as the number of pixels on the boundary of the region or objects when data are represented in a raster model. In the vector mode, the perimeter was the sum of all the arcs on the boundary of a polygon.

Compactness is a dimensionless variable. It is an invariant variable for shape rotation and scaling (Tomita and Tsuji, 1990). Compactness can offset the impact of distortion of shape, which was found as a necessary criterion for identifying confined poultry facilities from 35-mm slide imagery. The compactness was defined after Davis (1986) as:

$$Compactness = \frac{P^2}{A} \qquad (3.5)$$

*Where:*
*P is the perimeter of an object; and*
*A is the area of an object.*

The compactness has been shown to be an effective indicator for identifying many shapes (Davis, 1986). For example, a circular shape has the lowest compactness ($4\pi$). A square shape has a fixed compactness value of 16. The calculation of compactness for a polygon

can be quickly determined with available GIS functions (e.g. ArcView GIS (ESRI, 1997)).

The compactness for rectangular shaped poultry buildings was expressed as:

*Compactness of a rectangular shape* = 8+4*(L/W)+4*(1/(L/W)) (3.6)

*Where:*
*L is the length of an object; and*
*W is the width of an object.*

From equation 3.6, it can be shown that the length-width ratio for a rectangular shape is non-linearly related to compactness. Length-width ratio is also a dimensionless variable and invariant to shape rotation and scaling. However, it is very difficult to program a function to determine length and width of an arbitrary polygon. After extensive literature review of commercial GIS packages, none were found with the capability to calculate the length and width for an arbitrary polygon. For example, ArcView GIS has functions, which can be used to calculate the perimeter of a polygon or the length of a polyline.

Tomita and Tsuji (1990) described the following procedures to define the length and width of a polygon. Their procedure uses the eigenvalues of the variance and covariance matrix of moment of a polygon (Tomita and Tsuji, 1990). The matrix is expressed as:

$$v = \begin{vmatrix} M_{20} & M_{11} \\ M_{11} & M_{02} \end{vmatrix} \quad (3.7)$$

*Where:*
*$M_{02}$, $M_{11}$, $M_{20}$ are defined in equation 3.2.*

*The two eigenvalues of the matrix, λ1 and λ2, are length and width of the polgyon, respectively (assuming $\lambda 1 > \lambda 2$). The smaller value can be used for the width. The above matrix, however, may not have a solution.*

From a review of literature and field surveys, it has been found that modern poultry facilities have similar widths but vary in length. The field surveys were verified from analyses of training samples, which show that the widths of confined poultry facilities follow normal distribution based on the Anderson-Darling Normality Test. Based on the results of samples and engineering knowledge about the confined poultry operation, a mean value of 8.92 pixels for width, which was found from an analysis of poultry facility samples, was assigned to the polygons for shape analysis. The length of each polygon was calculated by dividing the polygon area by its width using ArcView GIS functions.

The compactness values for the training facility data were used to set the range for query. ArcView GIS query functions were used to select all polygons with compactness values within the specified range. These polygons were converted into a new shape file for further processing using the length-width ratio. The ranges for length-width ratio from sample analysis were used to conduct a query to select polygons. ArcView GIS functions were used to search for polygons that fell within the specified length-width ratio. These polygons were converted into a new shape file and represent confined poultry facility building polygons. The polygons were first selected based on compactness. Then selected polygons were selected based on length-width ratio.

## 3.4 Validation

This step involved the evaluation of the procedures on a dataset independent of the training data used to develop signatures for the automated procedure. The dataset included 30 samples randomly selected from the Rockingham County database. The slides included 182 poultry facilities that had been verified by site visits (Mostaghimi, et. al., 1998). The 35-mm imagery was processed following the procedure developed from the training database. The extracted polygons (poultry buildings) were overlaid on the

original slide images and compared with poultry building footprints that had been verified from field inspection.

The accuracy of the procedure was evaluated as follows:

- "Interpreted" represented those confined poultry facilities that correctly matched those identified manually.
- "Missed" represented the confined poultry facilities which were manually identified, but were not found by the automated procedure
- "Misinterpreted" represented the objects extracted by the automatic procedure, which were not confined poultry facilities.

The statistical methods such as confusion matrix analysis and Kappa coefficient are most commonly used to evaluate results of image interpretation (Campbell, 1987). The confusion matrix analysis identifies not only the overall errors for each category, but also misclassifications by category. The Kappa coefficient is "*a measure of the differences between the observed agreement between the two maps and the agreement that might be contributed solely by chance matching of the two maps*" (Campbell, 1987). The Kappa coefficient provides a measurement of agreement that is adjusted for chance agreement. The equation for calculating the Kappa coefficient:

$$\hat{\kappa} = \frac{Observed - Expected}{1 - Expected} \tag{3.8}$$

The Kappa coefficient in effect adjusts the "percentage correct" measure by subtracting the estimated contribution of chance agreement. This coefficient ranges between +1.0 and -1.0. A positive 1.0 indicates perfect effectiveness of pattern recognition. The negative value represents poor pattern recognition.

The Kappa coefficient is very effective for evaluating several categories of image classifications, for example, land use classification. For this study, values were either mis-classified or correctly selected; therefore the chance of positive selection was not

deemed appropriate. The footprint for poultry buildings was well defined leaving little opportunity for a chance selection. Therefore, the following equation for number of poultry facilities correctly identified was utilized.

$$Accuracy(\%) = \frac{Number\ of\ poultry\ facilities\ identified}{Total\ poultry\ facilities} * 100 \qquad (3.9)$$

## 3.5 Data for NPS pollution applications

Identification of NPS pollutants is critical for calculating TMDLs. The identification normally includes geographic location and estimation of pollution loadings. From this study, the location and an estimate of the number of birds can be made. The location and number of birds are very important data sources for calculation of pollutant loadings or estimation of input parameters for models, which are used for poultry related water quality simulations. They provide the basis for ranking each facility based on NPS pollution potential (Map Tech, Inc., 1999).

### 3.5.1 Estimating Number of Birds and Location of Buildings

A procedure was developed to estimate the waste generated by each poultry facility. The number of birds in each facility was based on density data obtained from *Structure and Environment Handbook* (Midwest Plan Service, 1987). The floor space required for birds at 4 weeks and 10 weeks of age was 00.05 $m^2$and 0.19$m^2$, respectively. Area and density data were used to calculate the total number of birds in each identified facility. The following relationships were developed to calculate the number of birds:

$$TNB = \frac{A_1 * fs^2 * 10.758}{srq} \qquad (3.10)$$

Where:

*TNB is the total number of birds in each facility during one growth cycle;*

*$A_1$ is the area of each facility in pixel$^2$;*

*fs is the scale factor representing one pixel size to ground distance, e.g., one pixel of an image taken at 500 ppi is equal to 6.1 meters; and*
*srq is the area in ($m^2$), which is occupied by each bird, e.g. 0.19 $m^2$/bird when birds reach 10 weeks of age.*

The manure produced from each confined poultry facility was estimated from the total number of birds based on data from Midwest Plan Service (Midwest Plan Service, 1987). Nutrients, e.g. total nitrogen and total phosphorus, and bacteria production rates, e.g. fecal coliform were also estimated. The following relationship was developed to calculate waste:

$$PWST = \frac{TNB * PW}{AUNIT} * AMWST \tag{3.11}$$

*Where:*
*PWST is the total waste generated from each facility (kg/day);*
*TNB was previously defined in equation 3.10;*
*PW is the weight of each bird in pounds, e.g., a 4-week-old bird weighs about 0.49kg;*
*AMWST is the amount of manure generated in kg/day/per animal unit;*
*AUNIT is animal weight normalized to 450kg (i.e. a 450kg animal is equal to one animal unit).*

The geo-referenced location of poultry facility is very useful for determining the non-point source pollution potential of the facility (Map Tech, Inc. 1999). The geo-referenced information can be used to link the relationships between confined poultry facilities and surround land use/cover. The geo-referencing procedure was used to register a 35-slide image into Universal Transversal Mercator (UTM) coordinates.

### 3.5.2 Determining Model Parameters

Computer models have been widely used in NPS pollution calculations. The Hydrological Simulation Program-Fortran (HSPF) (Bicknell et al, 1997) has been applied to a variety of scenarios in the United States. The US EPA Chesapeake Bay Program

continues to use HPSF to simulate total watershed contributions of flow, sediment and nutrients to the tidal regions of Chesapeake Bay (Donigan et. al., 1990). HSPF is a key component of BASINS 2.0 (USEPA, 1999), which has been extensively used for calculating TMDLs.

The rates of accumulation of pollutants (fecal coliform, nutrients) (ACQCP) and Maximum storage of pollutants (SQOLIUM) are two important model parameters of HSPF. The two parameters are key components of water quality in HSPF model (Donigan et. al., 1990). The two parameters are used to represent the accumulated rate and maximum storage rate before wash-off of pollutants, respectively. The two parameters determine the amount of pollutants reduction for TMDLs. Poultry waste is a significant contributor to fecal coliform loads to Virginia water bodies. The number of birds and the location of the facility are important to making reliable estimates of values for the two parameters of HSPF.

The following equation is used to calculate the SQLIUM for die-off fecal coliform.

$$N_t = N_0(10^{-kt}) \qquad (3.12)$$

*Where*

*$N_t$ : number of fecal coliforms at time t;*

*$N_o$ : number of fecal coliforms at time 0;*

*k = first order die-off rate constant; Typical values for warm months = 0.51/day and for cold months = 0.36/day (Horsley and WhittenInc., 1996).*

From equation 3-12, the following assumption was made to illustrate the relationship between ACQCP and SQOLIUM: The maximum buildup (SQOLIUM) during warm month is 1.5 times greater than ACQCP. During cold month, SQOLIUM is 1.8 times greater than ACQCP (Horsley and Whitten Inc., , 1996).

The number of birds in a facility building was calculated from equation 3.10. The production rate of manure and fecal coliform per chicken was obtained from ASAE

(1998). The example was conducted on 4-week chicken. Two types of land use/cover, cropland and pasture, were used to calculate the values of ACQCP and SQOLIUM.

## 3.6 Summary

The automated system required the development of procedures to pre-screen 35-mm slides, determination of scan resolution required to identify poultry building footprints, data management protocol digital images, scanning, and rectification. These procedures, with the possible exception of scan resolution, were also necessary for the manual procedures utilized by Mostaghimi et al (1998). The system proposed for automation of the identification of poultry facilities includes the development of training databases and defining their spectral and geometric properties. Image processing tools were developed to cluster features through thresholding and enhance images to include all features with spectral characteristics similar to poultry facilities as bounded by the training data set. Shape modeling procedures were developed to identify poultry facilities and determine selected geometric characteristics. Procedures were also presented to illustrate the use of geometric detail for estimating nutrient and bacteria loads from each poultry facility.

# 4. RESULTS AND DISCUSSION

Procedures were developed to locate and extract confined poultry facilities from 35-mm slide photography and to generate data important in conducting watershed pollutant source assessments, the development of TMDLs, and determining the pollution potential of individual poultry facilities. The different steps involved in the process along with their interactions are illustrated in Figure 4-1. Additional procedures were developed to create digital imagery from 35-mm slides and to manage the information as it was generated. The procedures were used to process 23 slide samples that contained 87 poultry facility samples. The following sections discuss and summarize the results.

## 4.1 Data Management System

A data management system was developed to organize, track, archive and retrieve 35-mm slides. The USGS 7-1/2 minute 1:24000 scale topographic maps served as the base map for the study. The index scheme previously discussed in the Methods and Procedures Chapter was based on the slides location on a topographic map. The index codes gave the location of the slide in the county (i.e. Hydrologic Unit), topographic map, and relative location on the topographic map. For example, a slide with filename B2138r02007 is located in hydrologic unit B21, on topographic map 38r, at local coordinate column 2 and row 7 on the topographic map, which locates the slide left of center near the bottom of the topographic map. The index provides a rapid means for organizing slides, e.g. by Hydrologic Units within the County or by topographic map. It also provides a rapid method for random selection of slides for further analyses. The system was easily implemented and it is very helpful when searching for location of a slide in a topographic map.

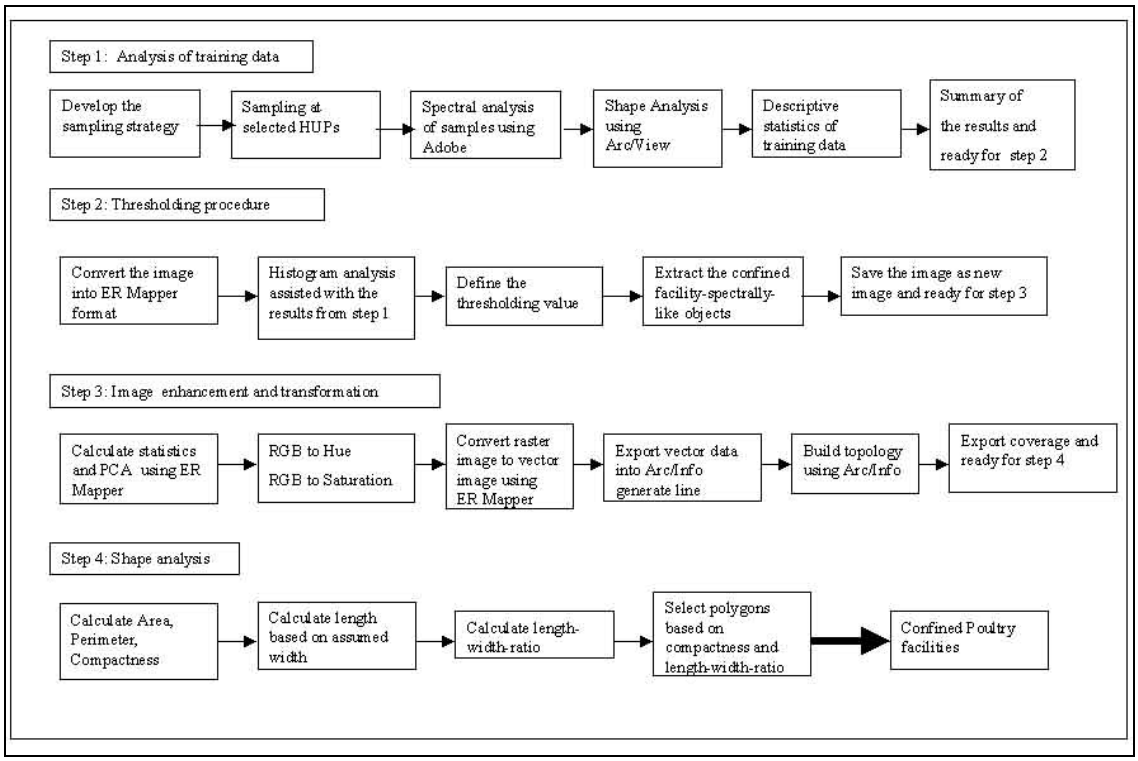


**Figure 4-1. Summary of the image procedures developed for the study.**

## 4.2 Scale and Resolution

The scale and resolution were very important parameters for effective use of 35-mm photography. The ability to develop a signature that would allow automated identification of poultry facilities required an understanding of the scale and resolution of the 35-mm slide photography and the minimum resolution required for the digital 35-mm slide imagery. The relationship among the scale, flight height, focal length of camera and scanning resolution for 35-mm slides is given in Table 4-1.

As listed in Table 4-1, the flight height for Rockingham County 35-mm imagery was about 9500 feet. The scale for the 35-mm slide was calculated as 1:120680. The scale and resolution of the photography fix the maximum resolution of the digital image created from scanning the slide. A slide scanned at 300 ppi gives a ground distance of 10.2 meters per pixel while a slide scanned at the maximum resolution of the film (i.e. 2150 ppi) is approximately equal to 1.42 m ground distance. Most modern confined poultry facilities are approximately 146 meters in length and 12.6 meters in width. The 2150 ppi resolution met the requirements of this study. From Table 4-1, we can find that the flight height for Rockingham County was about 9500 feet.

The scale was determined by the ratio of ground distance and flight height. The scale for the 35-mm slide was calculated as 1:120680. The scale and resolution of the photography fix the maximum resolution of the digital image created from scanning the slide. A slide scanned at 300 ppi gives a ground distance of 10.2 meters per pixel while a slide scanned at the maximum resolution of the film (i.e. 2150 ppi) is approximately equal to 1.42 m ground distance. Most modern confined poultry facilities are approximately 146 m in length and 12.6 m in width. Thus, the 2150 ppi met the requirements of this study.

**Table 4-1. The scale and scanning resolution of 35-mm imagery for Rockingham County, Virginia.**

| Country: Rockingham, Virginia | Date: 1995 |
|---|---|
| Focal length of camera: 24mm | Flight height: 9500 feet |
| Scale: 1:120680 | |
| Scanning Resolution (ppi) | Ground Truth for a pixel size (meter) |
| 300 | 10.2 |
| 500 | 6.10 |
| 1000 | 3.00 |
| 1500 | 2.00 |
| 2150 | 1.42 |

## 4.3 Spectral Analysis of Training Data

The purpose of analysis of the slide and poultry facility samples was to understand the spectral characteristics of poultry facility buildings. Twenty-three slide samples with 87 poultry facility samples were randomly selected from the Rockingham County database. The spectral characteristics of each primary color band (i.e. RGB) of slide and facility samples were obtained using Adobe tools. The descriptive statistical tools of Minitab (Minitab Inc., 1995) were used to analyze the spectral characteristic of the samples. The results on mean brightness intensities are shown in Figures 4-2 for the slide samples.

In general, the brightness for each color channel of an image ranged from 0 to 255. The level of brightness increased from 0 to 255. The number closest to 0 represents the darker color. The highest brightness value, 255, represents the color white. The mean brightness (intensity) was about 87 (Figure 4-2). The minimum and maximum intensities of slide samples were 76 and 96, respectively. The standard deviation was about 6. The standard deviation indicates that the spectral differences among the slide samples were very small. Based on the Anderson-Darling Normality Test (Figure 4-2), the intensity of the slide samples followed a normal distribution. The normal distribution and small standard deviation suggest that the areas covered by the slide samples had similar spectral characteristics. It should be noted that the areas covered in the slide samples were principally agriculture and forest.

The results of spectral analysis of poultry facility samples are shown in Figure 4-3. The mean intensity of the poultry facility samples was about 222, which represents a very bright color. The maximum intensity was about 251. The maximum intensity was very close to the white color. Most poultry facility samples have brightness greater than 202. The high mean intensity suggest that poultry facility buildings have a very high brightness. A field survey indicated that the roofs of poultry facility buildings are made of aluminum materials, which have very bright reflection. The standard deviation of poultry facility samples was 27. This illustrates that the spectral differences between each poultry facility sample were significantly larger than the spectral differences of the slide

samples. The larger standard deviation of the poultry facility sample could be caused by differences in building materials or by the age of the building. A comparison of spectral characteristics between slide and poultry samples is summarized in Table 4-2. From Table 4-2, it can be found that the mean intensities of the three primary colors for slide samples were 78, 88 and 104, respectively. However, the intensities of the three primary colors of poultry facility samples were 219, 220 and 246, respectively. The intensity of poultry facility samples was approximately 3 times greater than the intensity of slide samples. The mean intensity of poultry facilities was significantly different than background features in the slide samples, which gives a strong spectral signature pattern for identification of poultry facilities and other buildings with similar spectral signatures. The results indicate that poultry facilities are very distinguishable from neighboring features. The spectral analysis for each primary color of both slide and poultry facilities are shown in Figures 4-4 through 4-9, respectively. The mean intensity for color red of the slides was 78. The color green had a mean intensity of 87. The highest mean intensity was 103 for the blue color. These values form the spectral signature of confined poultry facilities. The color blue band had highest brightness values. The second highest brightness value was color green band. The color red had the lowest brightness values among the three primary colors. The standard deviation of brightness for color red was the highest among the three primary colors. The standard deviation of brightness for color green was the second highest while brightness for color blue had the smallest standard deviation. There also existed similarity in the three primary colors. Most of the color samples were clustered in the third quartile of the histogram. Rusted roofs on the poultry facility building caused the large standard deviations for the red and green color bands. As noted previously, the roof of a poultry facility building is constructed with metal material, usually aluminum. Over time the roof will lose its brightness due to the impact of weather conditions. The older buildings have more spectral information in the red and green color bands, which means these color bands contain more noises.

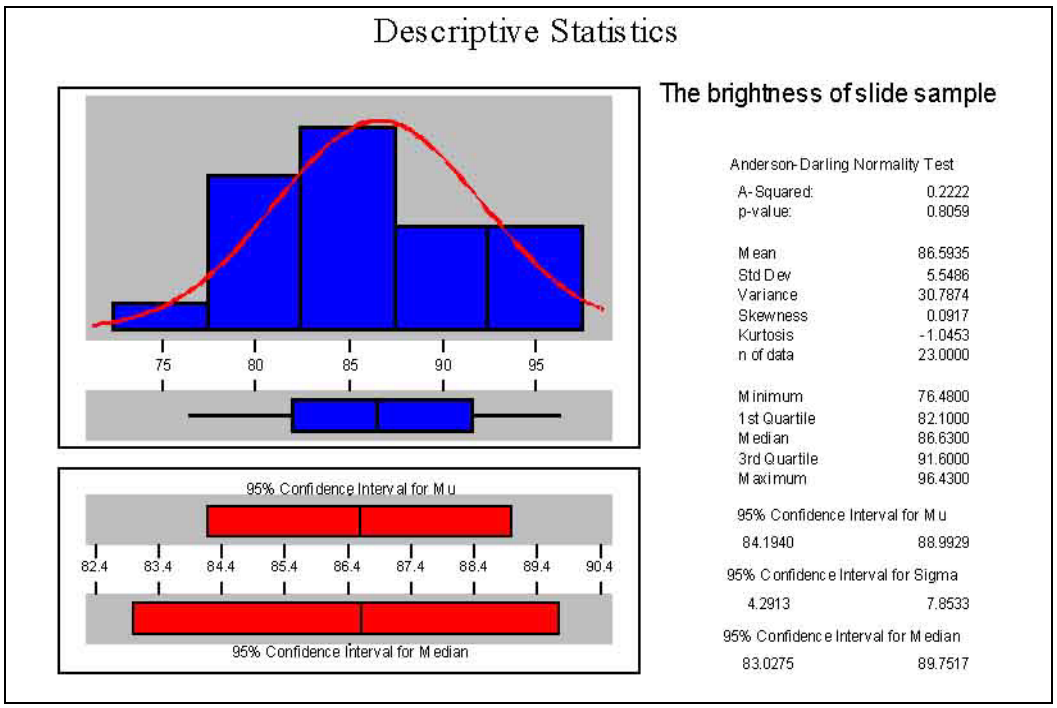


**Figure 4-2. The descriptive statistical analysis of brightness intensity of slide samples.**

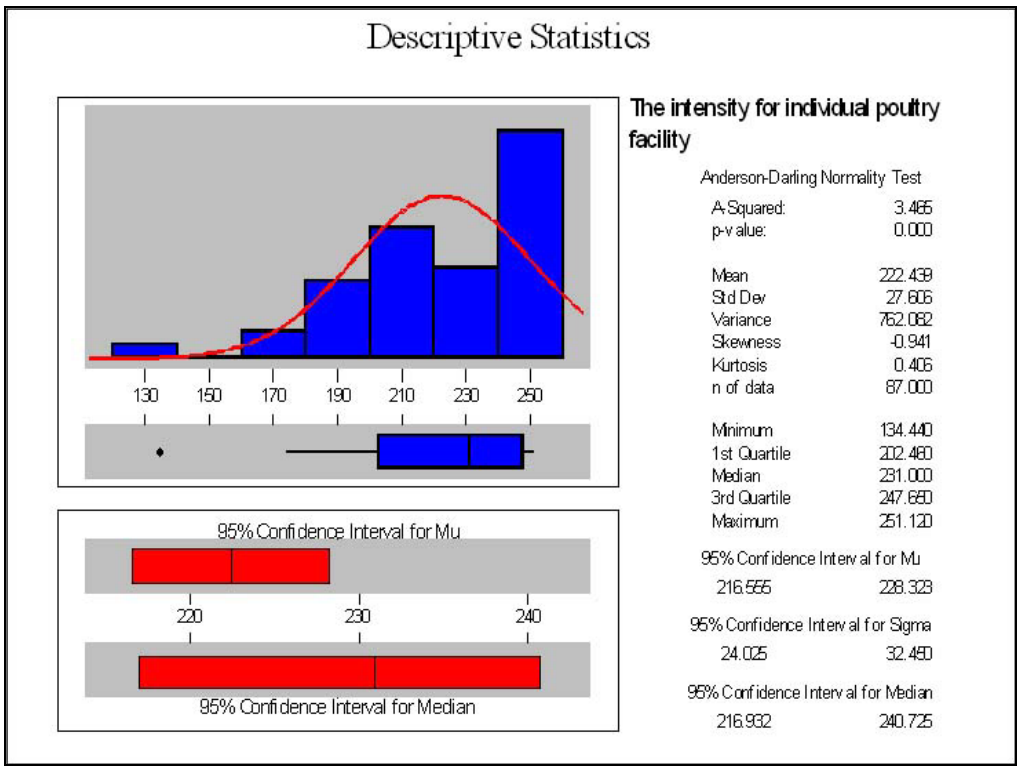


**Figure 4-3. The descriptive statistical analysis of brightness intensity of poultry facility samples.**

**Table 4-2. Statistical analyses of mean brightness for slide and facility samples.**

| Color Bands | Mean Brightness For: | | Ratio of Facility Sample/Slide Sample brightness |
|---|---|---|---|
| | Slide Sample | Facility Sample | |
| All | 87 | 222 | 2.6 |
| Red | 78 | 219 | 2.8 |
| Green | 88 | 220 | 2.5 |
| Blue | 104 | 246 | 2.4 |

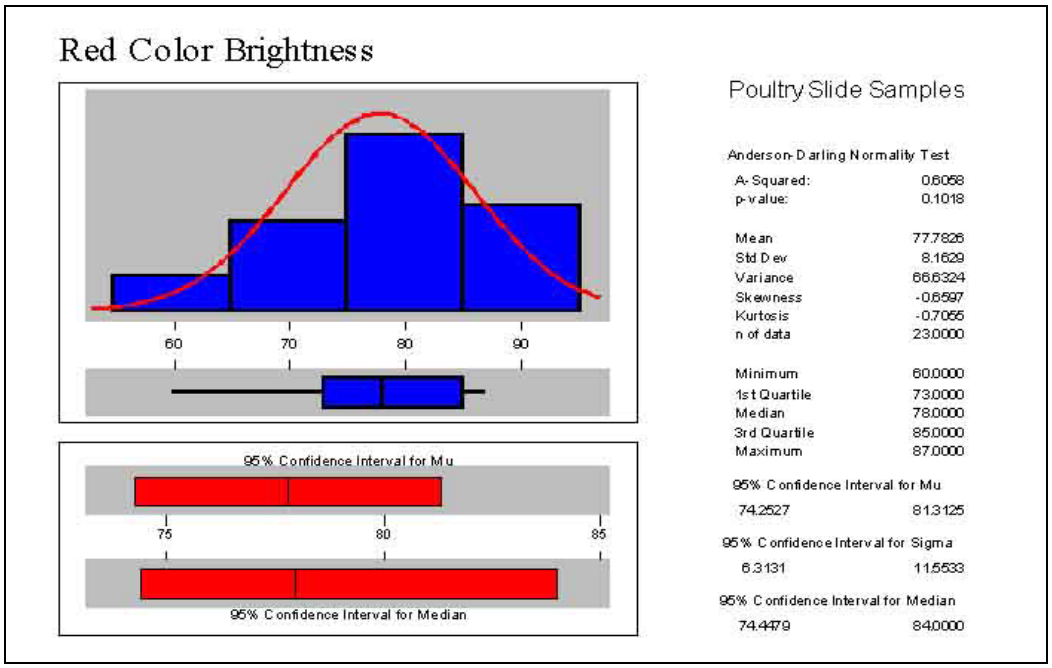


**Figure 4-4. The statistical analysis of mean intensity of red color band for slide samples.**

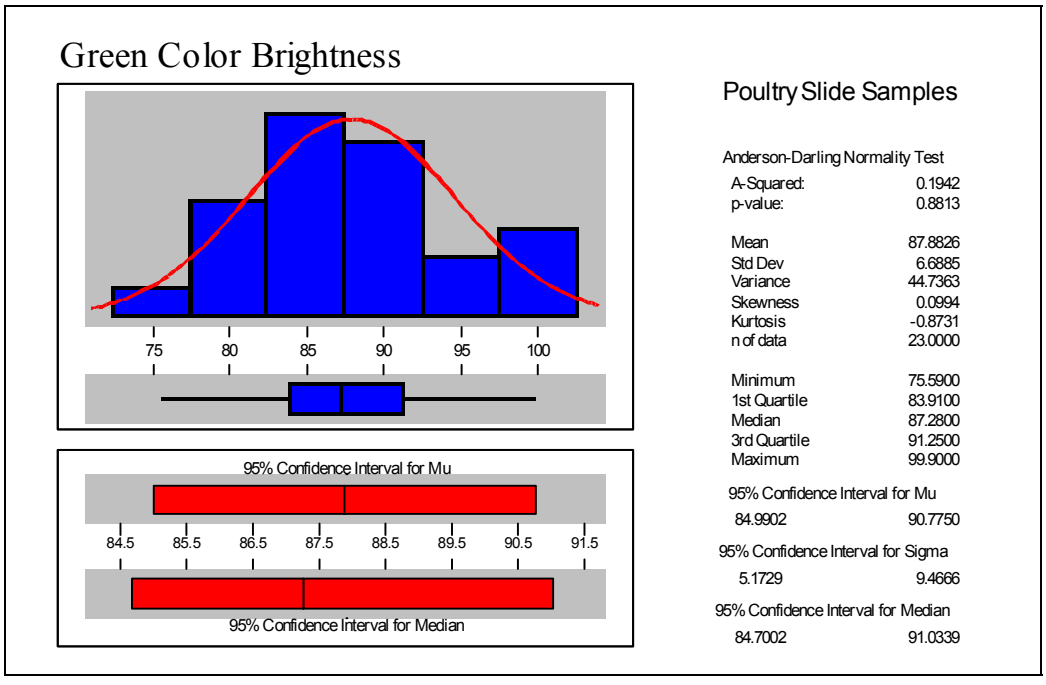


**Figure 4-5. The statistical analysis of mean intensity of green color band for slide samples.**

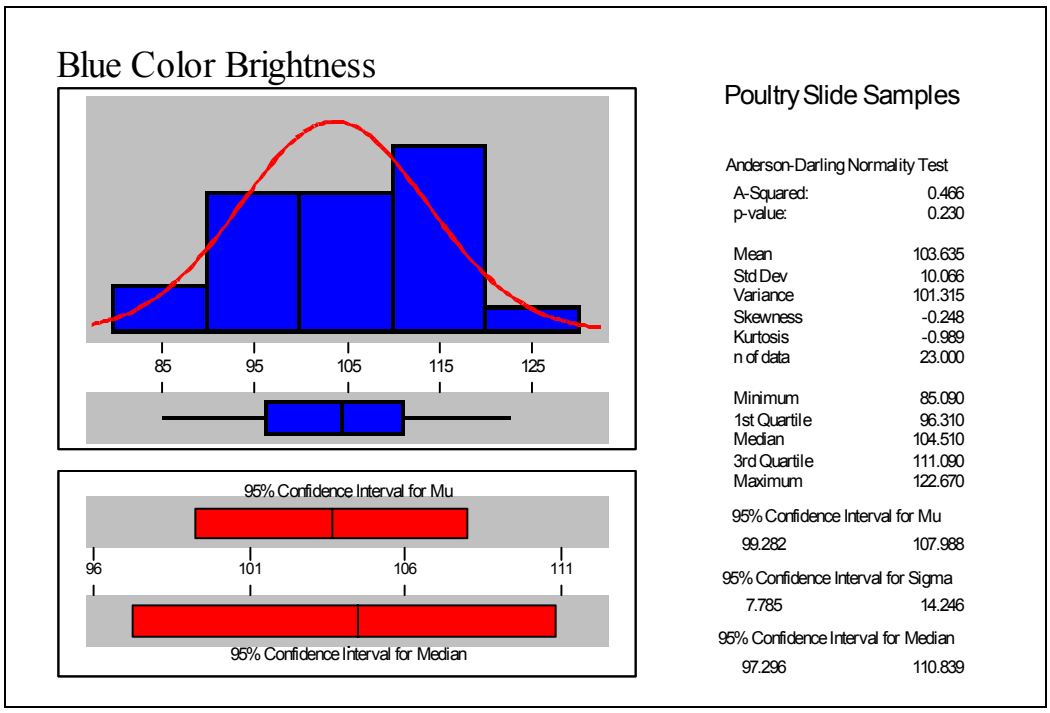


**Figure 4-6. The statistical analysis of mean intensity of blue color band for slide. samples.**

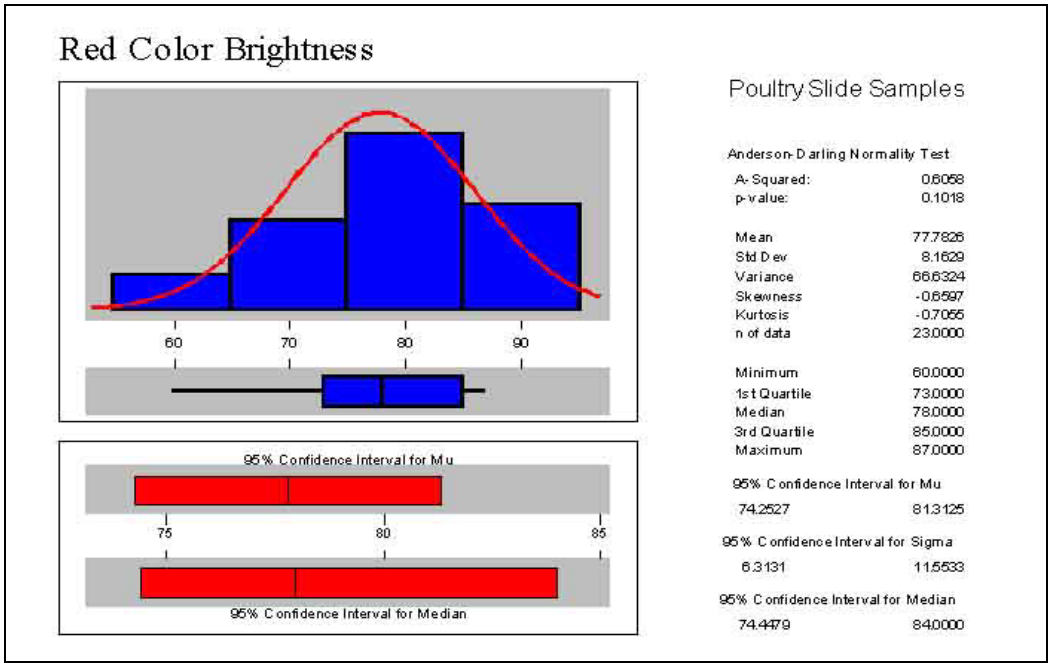


**Figure 4-7. The statistical analysis of mean intensity of red color band for poultry facility samples.**

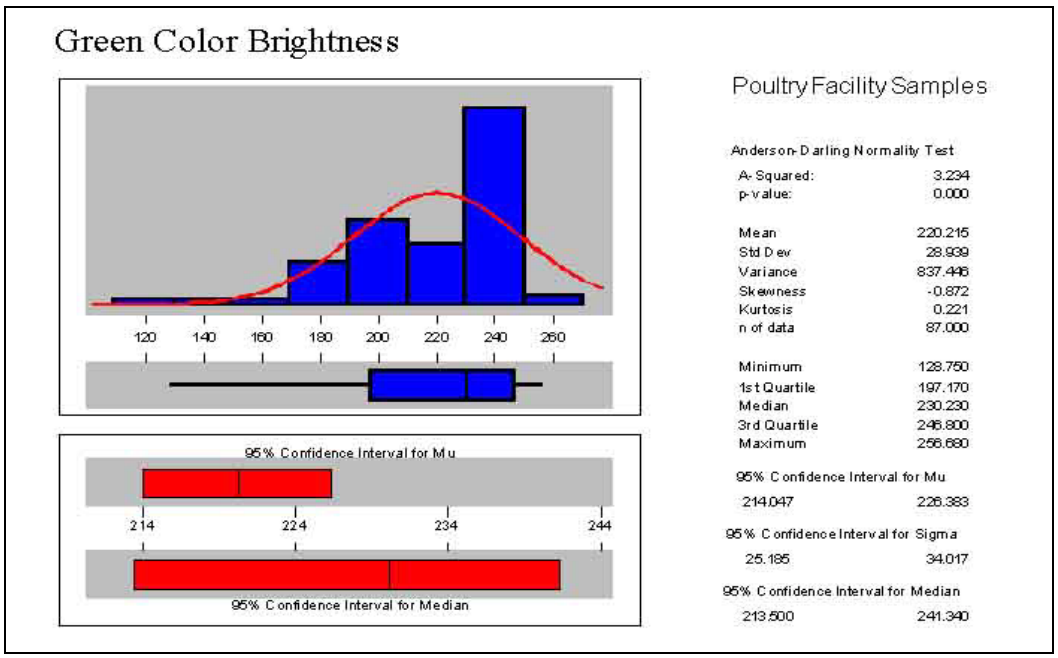


**Figure 4-8. The statistical analysis of mean intensity of green color band for poultry facility samples.**

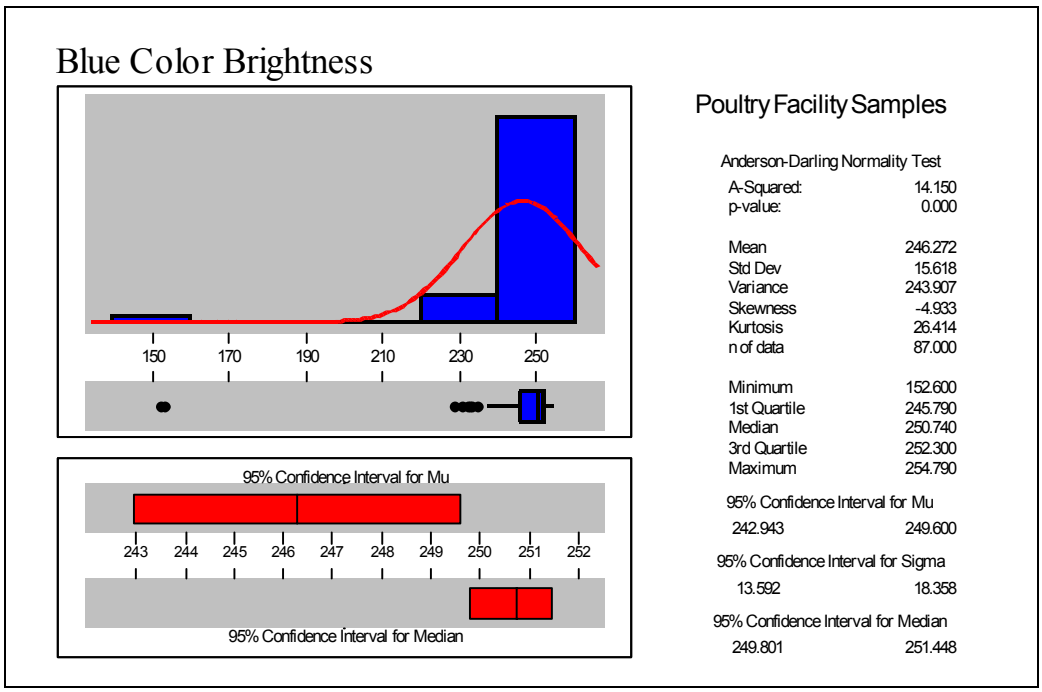


**Figure 4-9. The statistical analysis of mean intensity of blue color band for poultry facility samples.**

## 4.4 Threshold Analysis

In the thresholding procedure, statistical techniques are used to group spectral characteristics of the 35-mm image that relate to poultry facilities while eliminating other features. The first step in threshold analysis was to generate histograms of brightness across all color bands for a slide image (see example in Figure 4-10). The objective was to determine which region of the histogram included poultry facilities. For illustration purposes, in Figure 4-10 brightness ranges from 19 (bottom left of Figure 4-10) to 255 (bottom right of Figure 4-10). The average brightness for poultry facilities for the red color band, which is located in the right of Figure 4-10, was 219 (Table 4-10). For the red color band, brightness ranged from a minimum of 130 to a maximum of 252. The minimum value from the descriptive statistical analysis was used for threshold analysis on each color band.

Thresholding analysis was performed on 30 slide images for each color band. The process is illustrated in Figure 4-11a through Figure 11c. The upper left image (Figure 4-11a) is the original image. The lower left image (Figure 4-11b) shows the image after the thresholding analysis. Note that this image includes the results from thresholding over the three-color bands and only includes those objects with brightness characteristics similar to poultry houses within the poultry facility samples. The upper right image (Figure 4-11c) is an enlargement of two individual poultry houses after the thresholding procedure.

The thresholding procedure significantly reduced the information not related to confined poultry facilities as shown in Figure 4-11b and 4-11c. The "*Calculate Statistics*" function of ER Mapper was used to determine how much information was retained. For a typical 35-mm slide image, the red color band contained 3,486,800 non-null cells (i.e. recorded data cells). After the threshold analysis, the image red color channel contained only 20,800 non-null cells, which represents over 99 percent reduction in non-relevant information. Similar reductions were found for the blue and green bands. The remaining information, however, contained all features that had spectral characteristics similar to the poultry houses. However, note that the poultry facilities contain some information

from each color band (i.e. the brightness intensity values are not homogeneous neither within nor over the three bands). The next task was to utilize procedures that were developed to create a homogeneous spectral characteristic for each object created from the thresholding analysis.

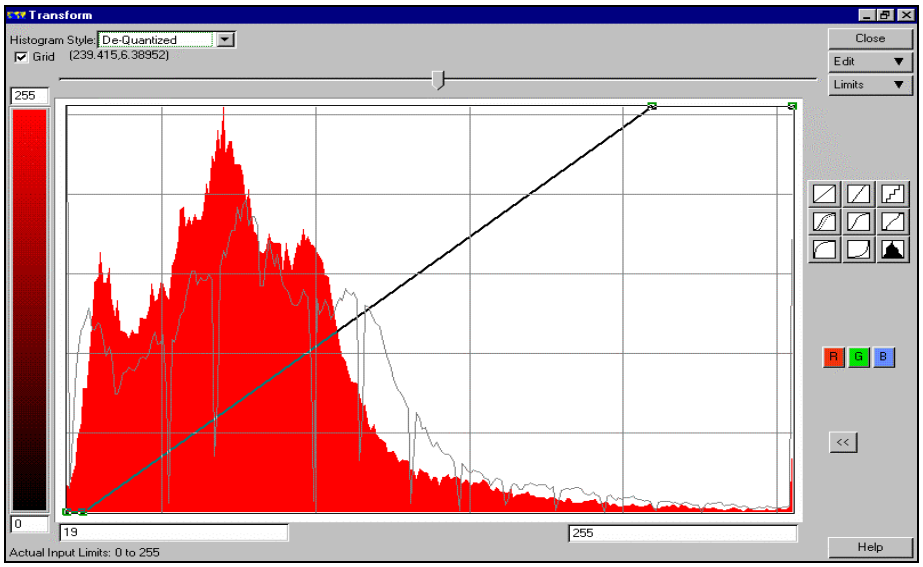


**Figure 4-10. Example histogram of brightness with multiple peaks and valleys.**

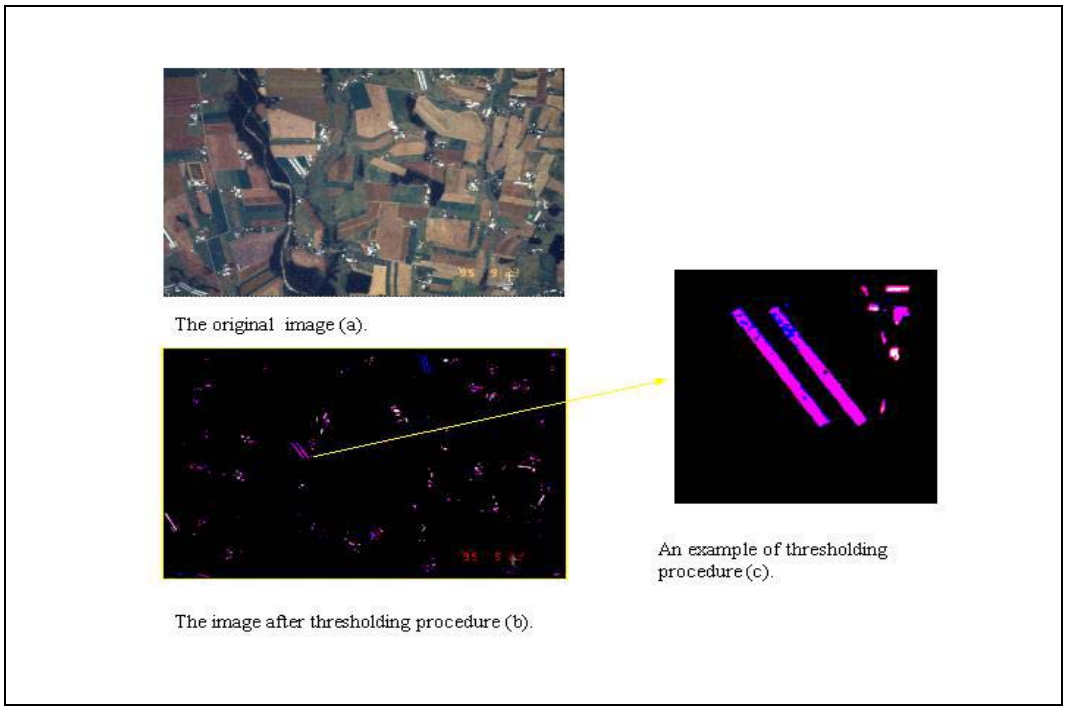


**Figure 4-11. Removal of spectral characteristics not similar to poultry facilities by thresholding analysis.**

## 4.5 Image Processing and Enhancement

The thresholding analysis removed all objects with spectral characteristics dissimilar to those from poultry facilities. The purpose of these analyses was to remove noise in the spectral image from poultry facilities due to such factors as sun angle when image was taken, line of sight of camera used, variation in roof reflectance due to rust, and shading by trees. Additionally, the date of the flight, which was placed on each slide image with a permanent marker, resulted in noise that needed to be removed. Four procedures were developed to enhance (remove) the above spectral signatures, from features with spectral characteristics similar to poultry facilities. These included PCA on the red color band, RGB to Hue on the red and green color bands, and RGB to Saturation on the red color band. These steps are illustrated in Figure 4.12a through 4.12f. These analyses were selected based on the procedures developed in Chapter 3, Section 3.2.3. The intensities for each primary color band had different spectral characteristics with the red and green color bands having the widest distribution, which implies that these bands have significant noise. Because all bands need to be used in an attempt to identify a broader spectrum of poultry facilities, procedures were needed to remove the noise from the red and green color bands. The first step was to use PCA analysis on the red color band, which removed all spectral characteristics not related to the red color band. This step is illustrated with Figures 4.12a, 12b, and 12c where Figure 12a is the image from thresholding analysis, Figure 12b shows the image after PCA analysis, and Figure 12c is an enlargement of two poultry facilities from Figure 12b.

The next step was to use the RGB to Hue procedure on the red color band to process the image resulting from PCA analysis (Figure 12b). This step is illustrated in Figures 12c and 12d using only the view of the two poultry facilities. The purpose of this step was to separate the red color from green and blue. The next step was to remove the red color band from the image created by RGB to Hue on the red band using the procedure developed for RGB to Saturation on the red color band. The result from this step is illustrated in Figure 12e, which contains only green and blue color bands. The last step is to remove the green color band from the image, represented by Figure 12e, using the

procedure for RGB to Saturation on the green color band (Figure 12f). The final image represented by Figure 12f includes only one color band in which all objects will have a uniform color. These steps were important in removing noise such as the flight date and noise within and immediately adjacent to all objects with spectral characteristics similar to poultry houses. The uniform color was required in order to retain the physical shape of the object when the image was converted to an ArcView shape file format for shape analysis procedures discussed in the next section.

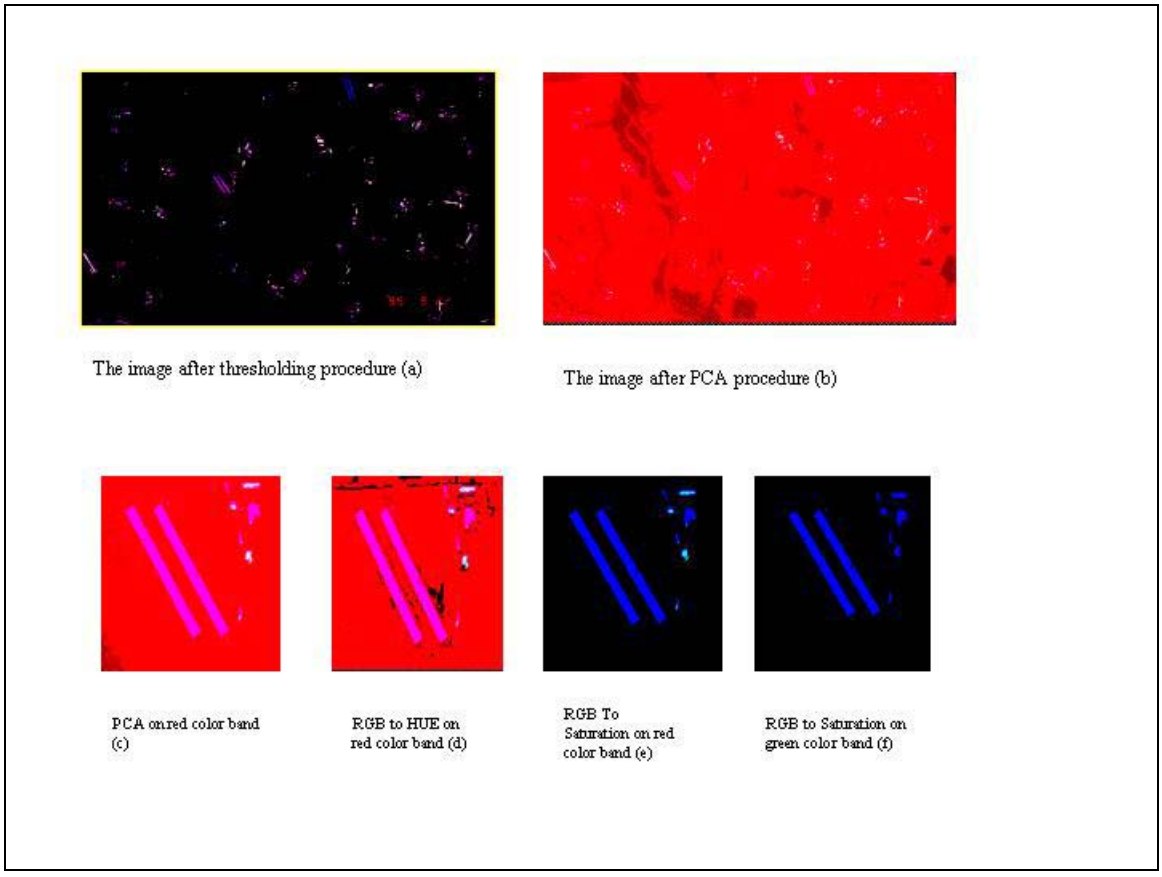


**Figure 4-12. Threshold Analysis Image (Figure 4-12a); PCA analysis from Figure 4-12a to Figure 4-12b. Then zoom of two poultry houses from PCA image (Figure 4-12c); Then zoom of two poultry houses from RGB to HUE, RGB to Saturation (Figures 4-12d to 4-12i).**

## 4.6 Shape Analysis of Training Data

The purpose of these procedures was to develop an understanding of the physical characteristics of poultry houses represented in the digital images. A total of 60 poultry facilities were selected form the original 83 facility samples used for spectral analysis of training data. When poultry houses are grouped in clusters of two or three, the shape characteristics often are nearly identical although the spectral characteristics may vary significantly. For these situations, only one of the groups was selected for shape analysis resulting in a final total of 60 facilities. The procedures discussed in Chapter 3 were used to calculate shape parameters for these sixty poultry facility samples. The parameters utilized to describe shape included width, length, compactness and length-width ratio. The purpose of these analysis were to determine the statistical characteristics of the different shape parameters for poultry facility samples, which would then be used to identify and extract poultry facilities from the image file resulting from spectral analysis discussed in previous sections.

The descriptive statistics for width of poultry facility samples are presented in Figure 4-13. The average width was 8.92 pixels or approximately 12.6 m. The width of modern poultry houses in Rockingham County ranges from 9.7 m to 15.2 m with an average of 12.5 m (Fitzgerald, 1999). The difference between the computer generated width from shape analysis was only 0.5 m. The Anderson-Darling Normality Test also showed that the width of poultry facility samples followed a normal distribution with a standard deviation of 1.52 pixels or approximately 2.1 m (p value=0.7311 from Figure 4-13), which indicates that the width of all poultry houses in the facility sample were nearly the same. Some small variation was expected because of distortion in the building footprint due to camera angle, weather conditions, etc. The descriptive statistics for the length of poultry facilities samples are also given in Figure 4-14. The minimum, maximum and average lengths were 44.34 (63 m), 184 (261 m), and 105 (149 m) pixels, respectively. The length of poultry houses can vary significantly. The length of modern poultry facilities in Rockingham County, Virginia ranged from 78.7 m to 246.9 m ((Fitzgerald, 1999) with an average of 162.8 m. The large variation in length was primarily due to the

size of the poultry operation. That is, the capacity of the building is increasing with the length while the width is unchanged. Some facilities use the smaller house (i.e. shorter building) to house poultry during the first two to three weeks then move the poultry to a larger facility for finishing. Because the variation in length was significant, this shape parameter alone would result in the selection of many non-poultry house features. An attempt was made to define a shape parameter that included both length and width because the width is generally a fixed parameter. This was accomplished by normalizing the length by the width to define a length-width ratio. Descriptive statistics for this shape parameter are given in Figure 4-15. The minimum, maximum, and average length-width ratios were 6, 23, and 11.9, respectively. The length of confined poultry facilities averaged approximately 12 times the building width.

An additional shape parameter, which is related to the length and width parameters, used in this study was the compactness parameter, defined as the square of the perimeter of the feature footprint divided by the area. This parameter has been demonstrated as an effective parameter for identifying shapes (Davis, 1986). The compactness parameter was determined for each facility sample and descriptive statistics were calculated using Minitab software. The results are summarized in Figure 4-16. The Compactness parameter ranged from 30 to 130 with a mean value approximately equal to 57, which is approximately equal to the median value of 55. This represents significant variation, which is primarily attributed to the variation in length of poultry facilities. The analysis also indicates that a much narrower range (i.e. 53 to 61) could be used if only a 95% confidence of including all poultry facilities is desired. Some additional discussion on the use of a narrower band to extract poultry shapes from the images derived from spectral analysis will be given in a later section.

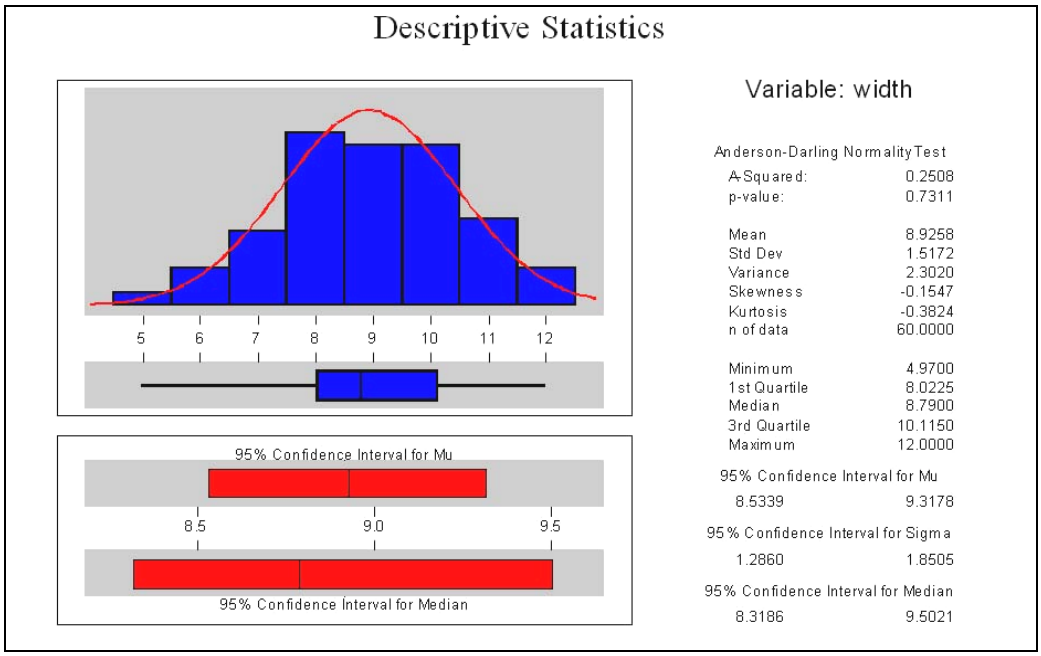


**Figure 4-13. A summary of descriptive statistics for width of poultry facility samples.**

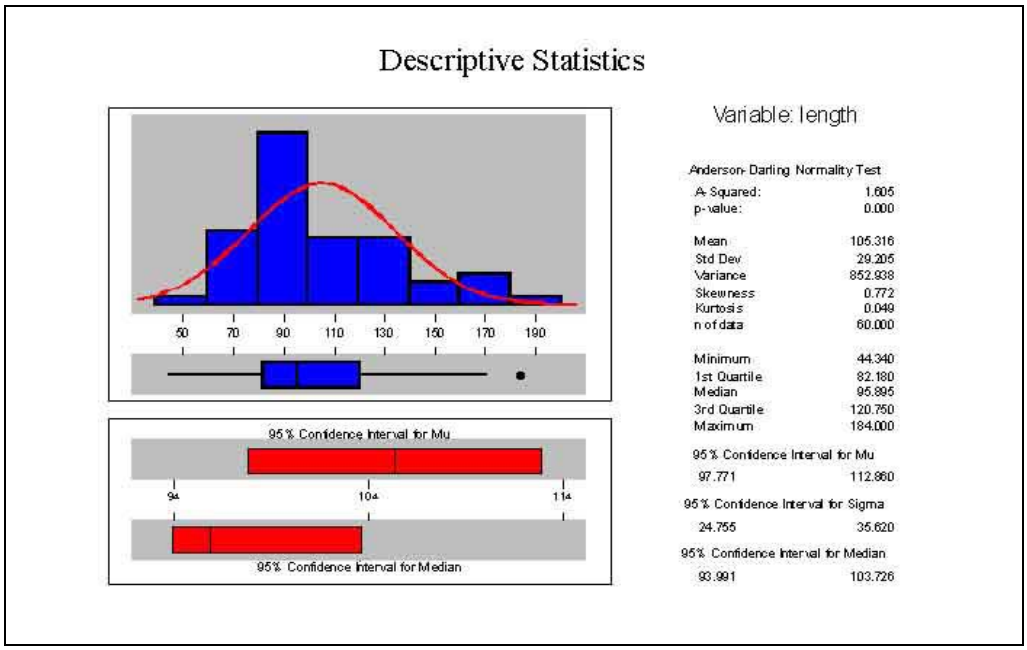


**Figure 4-14. Summary of descriptive statistics for length of poultry facility samples.**

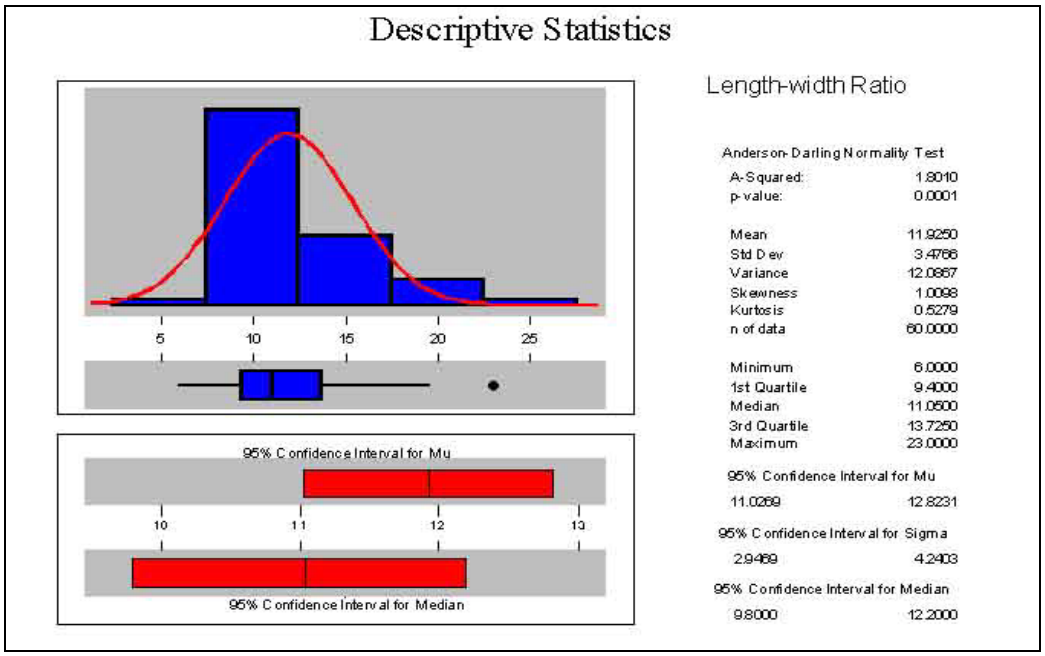


**Figure 4-15. Summary of descriptive statistics length-width ratio of poultry facility samples.**

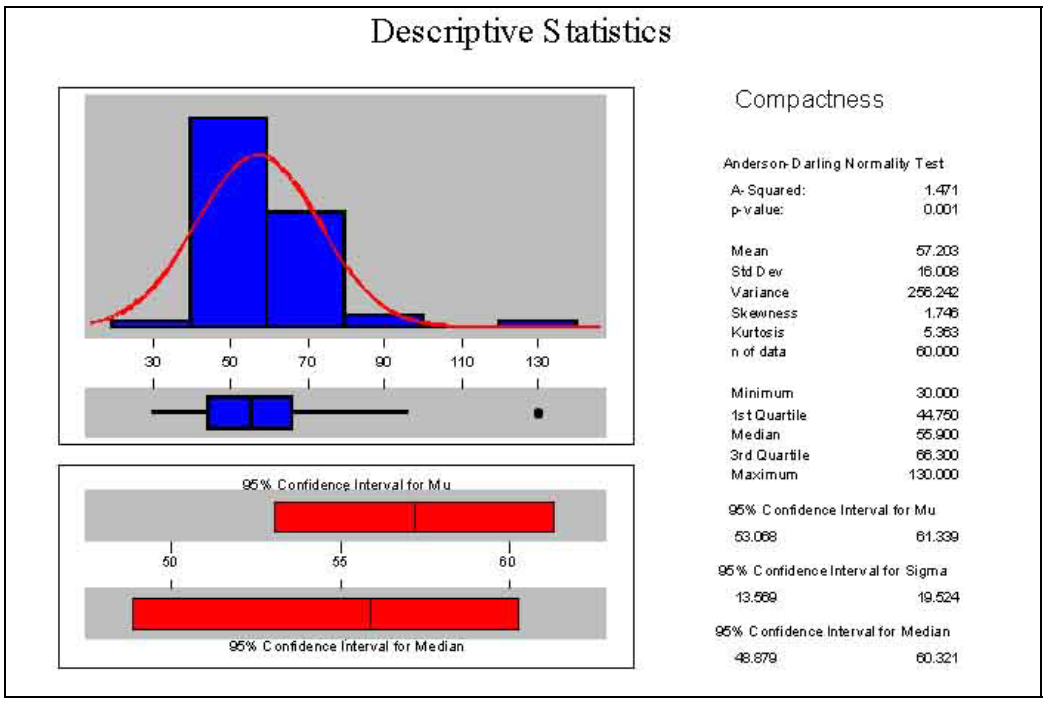


**Figure 4-16. Summary of descriptive statistics for compactness shape parameter for poultry facility samples.**

## 4.7 Identification of Poultry Buildings in Images

The purpose of this task was to identify poultry building from the digital image resulting from the RGB to Saturation on the green color band and converted to an ArcView shape file format (Section 4.5). This is the final step in feature identification. The procedure involves using the compactness parameter to identify all features that fall within a specified range and then use the length-width ratio to fine-tune the selection. ArcView GIS was used to convert the objects (i.e. polygons) from the slide image file generated from previous spectral analysis. Each polygon was converted from objects in a raster image to a vector format. Functions were programmed for ArcView GIS to calculate area, perimeter, compactness, length, width, and length-width ratio for each polygon. The ArcView query function was then used to cluster polygons based on descriptive statistics determined for compactness and length-width ratio from the results of training data for facility samples. As noted previously, compactness has been found to be a good identifier of objects. The length-width ratio was used to refine the selection because of the unique characteristics of poultry houses and the fact that it retains the geometric characteristics of the poultry facility. The length-width ratio that would include all poultry houses with this shape characteristics found in the training dataset was 6 to 23 (Figure 4-15). The compactness range that would include all poultry houses with shape characteristics found in the training dataset was 30 to 130 (Figure 4-16). Objects were selected based on the above criteria. Figure 4-17 shows an example of polygons selected based on compactness followed by length-width ratio. The polygons selected based on compactness are highlighted with green color. Those selected based on length-width ratio are highlighted in red. From Figure 4-17, it can be seen that the polygons selected based on compactness not only included the poultry facilities but also non-poultry buildings with similar compactness values. Note that compactness value includes all poultry facilities that fall within the range of those found for the training data set, but also includes any object that has an area-perimeter configuration that gives a compactness value within the range that identifies poultry houses. These objects may be machine sheds, a cluster of buildings located close together, etc. The length-width ratio reduces the number of non-poultry facilities but still will not guarantee exclusiveness. Since a constant value of width is used

to calculate the length of a polygon, the length-width ratio preserves the overall geometric characteristics of the facility.

## 4.8 Evaluation of Procedures on Random Dataset

A dataset that included thirty 35-mm slides randomly selected from the Rockingham County database was used for the purpose of an independent evaluation of the automated procedure developed from training data (Chapter 3) and discussed in Sections 4.1 through 4.7. These slides contained 182 poultry houses previously identified by manual procedures (Mostaghimi, et. al., 1998). The results are summarized in Table 4-3 as interpreted, missed, and misinterpreted. The interpreted category included all poultry houses that were correctly identified. The missed category referred to all poultry houses that were not identified. The misinterpreted category referred to objects in a slide image that were incorrectly identified as poultry houses. The total facilities category includes the total number of poultry houses existing on the slide image, i.e. the sum of the interpreted and the missed categories. The number included in the parenthesis is the total number of poultry facilities identified on the slide plus the number of features misinterpreted.

A review of Table 4-3 shows that all poultry houses were correctly identified in 28 of the 30 slides analyzed. A total of five poultry facilities, located on or across slide boundaries, were not identified in image slide B2539r1008 and two poultry slides were not identified in slide image B1938q3003. Sixteen slide images had no objects misclassified as a poultry house. Three slide images had two objects misclassified as poultry houses while eleven slide images had one object misclassified as a poultry house. The objects misclassified as poultry houses were dairy operations based on manual classifications (Mostaghimi, et al, 1998).


**Figure 4-17. An example of polygons selected based on compactness and length-width.**

**Table 4-3. Summary of independent dataset analysis to identify poultry houses.**

| Slide Name | Interpreted | Missed | Misinterpreted | Total facilities[1] |
|---|---|---|---|---|
| B1737q05002 | 4 | 0 | 1 | 4 (5) |
| B1738q01005 | 2 | 0 | 0 | 2 (2) |
| B1938q03003 | 5 | 2 | 0 | 7 (7) |
| B1938q03009 | 7 | 0 | 2 | 7 (9) |
| B1938q04002 | 3 | 0 | 1 | 3 (4) |
| B1939q01007 | 7 | 0 | 1 | 7 (8) |
| B2138r04001 | 6 | 0 | 1 | 6 (7) |
| B2138r05001 | 9 | 0 | 0 | 9 (9) |
| B2239r01001 | 9 | 0 | 0 | 9 (9) |
| B2539q02001 | 8 | 0 | 0 | 8 (8) |
| B2539q02004 | 6 | 0 | 1 | 6 (7) |
| B2539q04001 | 5 | 0 | 0 | 5 (5) |
| B2539r01005 | 5 | 0 | 0 | 5 (5) |
| B2539r01008 | 7 | 5 | 1 | 12 (13) |
| B2539r01010 | 14 | 0 | 0 | 14 (14) |
| B2539r02002 | 12 | 0 | 1 | 12 (13) |
| B2539r02008 | 7 | 0 | 2 | 7 (9) |
| B2539r02013 | 5 | 0 | 0 | 5 (5) |
| B2539r02016 | 10 | 0 | 2 | 10 (12) |
| B3541q01005 | 1 | 0 | 0 | 1 (1) |
| B3541r06002 | 0 | 0 | 1 | 0 (1) |
| B3542r01006 | 3 | 0 | 0 | 3 (3) |
| B3542r02002 | 7 | 0 | 0 | 7 (7) |
| B3542r02007 | 2 | 0 | 0 | 2 (2) |
| B4740r01011 | 6 | 0 | 0 | 6 (6) |
| B4740r06001 | 2 | 0 | 1 | 2 (3) |
| B4740s04007 | 3 | 0 | 1 | 3 (4) |
| B4740s06004 | 13 | 0 | 1 | 13 (14) |
| B4741s01006 | 3 | 0 | 0 | 3 (3) |
| B4741t04005 | 4 | 0 | 0 | 4 (4) |
| **Total Facilities** | 175 | 7 | 17 | 182 (199) |
| **% Error** | 89 | 3 | 8 | |

[1] The number inside the parenthesis refers to the number of facilities identified as poultry facilties. The number outside the parenthesis is the number of poultry facilities in the slide.

Slide image B2539r01005 (Figure 4-18) is an example of an image with all (5) poultry houses correctly identified. A shape file for poultry houses overlaid onto the original slide image demonstrates that the shapes of the poultry houses were well preserved after image processing and extraction from the original digital image. Another slide image (B4740s4004) is displayed in Figure 4-18 with thirteen poultry houses correctly identified and one object (dairy) misidentified as a poultry house. The remaining slides that were correctly identified are listed in Appendix A.

Of the two slide images with poultry houses not identified (i.e. they were missed), the least accurate identification was for slide B2539r01008 (Figure 4-19). For this slide, 12 poultry houses had been previously manually identified. Of the twelve, only seven (7) were automatically identified. One object (dairy facility) was misinterpreted as a poultry house. The unidentified poultry houses were located near or crossed the slide boundary. As discussed in the Chapter 3, the poultry facility samples for the training dataset were not selected from the edge of the slide image. The objects along the boundary of a slide were generally severely distorted which significantly affects geometric values. In some situations, the poultry house extended across the boundary making identification difficult. The distortion of brightness resulted in missing five poultry houses. The intensities of brightness of the five poultry houses were not in the range described in Figures 4-7 through 4-9, which resulted in them being excluded by thresholding analysis.


**Figure 4-18. Image slide B2539r01005 with all (5) poultry houses correctly identified.**


**Figure 4-19. Slide image (B4740s06004) with thirteen poultry houses correctly identified and one object (dairy) misidentified.**


**Figure 4-20. Slide image B2539r01008 with least accurate poultry house classification. Total of 12 poultry houses with five not identified.**

Seven poultry houses were located in slide B1938q03003 with five poultry houses correctly identified and two not identified. The two poultry houses that were not identified (i.e. missed) are displayed in Figure 4-21. In this situation, the two poultry houses were located very close to each other and spectral characteristics of objects between the buildings resulted in the buildings to appear joined at several points. The computed compactness for the two polygons was 197, which exceeded the range 30 to 130 that was determined from training samples (Figure 4-16).

Ninety seven percent of the poultry houses located within the 30 slide images were correctly identified by this procedure, which demonstrates that the sampling strategy represented the characteristics of confined poultry houses with reasonable accuracy. With the exception of two houses, the other problems were the result of houses located at the

edges or across the slide boundary. These houses were not included in the facility (house) training dataset described in Chapter 3 because of distortion. Distortion significantly


**Figure 4-21. Slide image (B1938r03003) with five poultry houses correctly identified but with the two displayed not identified.**

impacts spectral characteristics and geometric properties used to define shape. To include the boundary areas will require significant additional preprocessing of the scanned images to improve geo-referencing and distortion correction. Additionally, to correctly identify houses that cross the boundary would in many instances require joining adjacent slides and addressing the various issues stated there (e.g. distortion, scale, etc.).

A total of seventeen objects in 13 slide images were misinterpreted when only the length-width ratio was based on an average or constant width assumption. Careful examination of the misinterpreted objects revealed that 16 of the 17 objects were dairy farms. The other object was unknown. The materials used for roofs of these dairy farm buildings were the same as poultry. The compactness and length-width ratio of the dairy facilities were within the range defined by the training data for poultry houses. A slide image with examples of two objects that were misinterpreted as poultry houses is displayed in Figure 4-22. Although the two objects did not have a normal rectangle shape, when the defined width for a poultry houses was assigned to the two polygons, the length-width ratio was within the range of a poultry house. The image slides containing the remaining objects that were misinterpreted are listed in Appendix A.

After reviewing the results presented in Table 4-3, an additional shape parameter was developed and applied to the dataset resulting from the previous shape analysis. For the results presented in Table 4-3, the new procedure were applied to the dataset containing 199 polygons. The procedure involved the calculation of the centroid of each polygon, the distance from the centroid to the vertices of each polygon, and a determination of the minimum and maximum distances. The maximum distance was assumed to be a measure of the length and the shortest distances a measure of the polygon width. The measures of length, width and length-width ratio based on the polygon centroid were calculated for each polygon with the results shown only for the misinterpreted objects from Table 4-3 in Table 4-4. Two additional queries were developed in ArcView GIS to filter all 199 objects (Table 4-3). The 95% of confidence interval for the mean of width (8.53 to 9.32, Figure 4-13) and the 95% confidence interval of the mean of length-width ratio (11.02 to 12.82, Figure 4-15) were used to determine if a polygon had similar shape characteristics

as a poultry object. The ratio of the length and width of poultry houses in Rockingham County ranged from 8 to 16 (Fitzgerald, 1999). This procedure resulted in the removal of the mis-interpreted objects shown in Table 4-3 as shown in Table 4.5.

As previously discussed in the literature review, the accuracy of the manual interpretation of 35-mm slide images for poultry houses was approximately 97% (Mostaghimi, et. al., 1999). The results from this test case suggest an accuracy of 97% with 3% of the poultry facilities missed. Little quantitative information was needed to identify the poultry buildings using the manual procedures. The facilities located at the edge of slides were identified by manual interpretation because the effect of distortion had little impact on visual interpretations. Poultry houses with significantly different spectral patterns also were usually identified by manual interpretation techniques.

The manual procedure, however, is very time consuming. An experienced interpreter requires approximately 15 minutes to manually identify a poultry building and enter the center coordinates into a database. It requires approximately an additional 30 minutes to digitize each building boundary and to perform post-processing tasks such as correcting digitizing errors and geo-referencing. Geo-referencing to a known coordinate system is required to determine the size of the building from which calculations to obtain the number of birds and waste load determinations could be made. The manual procedure, furthermore, requires extensive GIS and agricultural knowledge. Time consuming training is usually required. The 35-mm slides are often updated, sometime once a year by USDA/FSA, which may contain new houses. Thus, frequent update of the information base becomes difficult with manual interpretation. The automated procedure requires approximately two minutes to process one slide through all components of the method. If software is written to link all components to run sequentially in batch mode, the process time is expected to be reduced by a factor of four or 30 seconds per slide.

A comparison of the two procedures is made based on processing 1000 slides, with an average of 10 poultry facilities per slide, and pre-processing completed for both procedures. For this situation, 750 hours would be required to process 1000 slides

manually and approximately 34 hours using the automated procedure. These estimates are also based on each procedure identifying the facility and providing location and building geometry. Customizing the software would be expected to significantly reduce computer-processing time for the automated procedure.

**Figure 4-22. Slide image with example of two objects mis-interpreted as poultry houses.**

**Table 4-4. A summary of the results from application of centroid shape parameter. This summary only includes the results for the misinterpreted objects given in Table 4-3.**

| Slide Name | Length | Width | Length/Width Ratio |
|---|---|---|---|
| B1737q05002 | 66 | 0 | Out range |
| B1938q03009 (dairy 1) | 51.8 | 11.4 | 4.5 |
| B1938q03009 (dairy 2) | 64.4 | 11.3 | 5.7 |
| B1948q04002 | 32.5 | 8.34 | 3.89 |
| B1939q01007 | 65.0 | 16.7 | 3.9 |
| B2138r04001 | 51.9 | 3.6 | 14.4 |
| B2539q02004 | 48.3 | 5.1 | 9.47 |
| B2539r01008 | 51.9 | 12.2 | 4.3 |
| B2539r02002 | 43.1 | 6.7 | 6.4 |
| B2539r02008 (dairy 1) | 48.9 | 4 | 12.3 |
| B2539r02008 (dairy 2) | 63.3 | 3.2 | 20.0 |
| B2539r02016 (dairy 1) | 40.2 | 4.24 | 9.5 |
| B2539r02016 (dairy 2) | 93.6 | 22.2 | 4.2 |
| B3541r06002 | 51.9 | 25.9 | 1.99 |
| B4740r06001 | 71.3 | 12.4 | 5.76 |
| B4740s04007 | 49.0 | 7.8 | 6.3 |
| B4740s06004 | 46.7 | 7.1 | 6.6 |

**Table 4-5. Final summary of independent dataset analysis to identify poultry houses.**

| Slide Name | Interpreted | Missed | Misinterpreted | Total facilities |
|---|---|---|---|---|
| B1737q05002 | 4 | 0 | 0 | 4 |
| B1738q01005 | 2 | 0 | 0 | 2 |
| B1938q03003 | 5 | 2 | 0 | 7 |
| B1938q03009 | 7 | 0 | 0 | 7 |
| B1938q04002 | 3 | 0 | 0 | 3 |
| B1939q01007 | 7 | 0 | 0 | 7 |
| B2138r04001 | 6 | 0 | 0 | 6 |
| B2138r05001 | 9 | 0 | 0 | 9 |
| B2239r01001 | 9 | 0 | 0 | 9 |
| B2539q02001 | 8 | 0 | 0 | 8 |
| B2539q02004 | 6 | 0 | 0 | 6 |
| B2539q04001 | 5 | 0 | 0 | 5 |
| B2539r01005 | 5 | 0 | 0 | 5 |
| B2539r01008 | 7 | 5 | 0 | 12 |
| B2539r01010 | 14 | 0 | 0 | 14 |
| B2539r02002 | 12 | 0 | 0 | 12 |
| B2539r02008 | 7 | 0 | 0 | 7 |
| B2539r02013 | 5 | 0 | 0 | 5 |
| B2539r02016 | 10 | 0 | 0 | 10 |
| B3541q01005 | 1 | 0 | 0 | 1 |
| B3541r06002 | 0 | 0 | 0 | 0 |
| B3542r01006 | 3 | 0 | 0 | 3 |
| B3542r02002 | 7 | 0 | 0 | 7 |
| B3542r02007 | 2 | 0 | 0 | 2 |
| B4740r01011 | 6 | 0 | 0 | 6 |
| B4740r06001 | 2 | 0 | 0 | 2 |
| B4740s04007 | 3 | 0 | 0 | 3 |
| B4740s06004 | 13 | 0 | 0 | 13 |
| B4741s01006 | 3 | 0 | 0 | 3 |
| B4741t04005 | 4 | 0 | 0 | 4 |
| **Total Facilities** | *175* | *7* | 0 | *182* |
| **% Error** | **97** | **3** | 0 | |

A reliable automated identification system can significantly reduce costs and can be an effective tool for periodic updating of poultry facility database. Geometric parameters such as area, length, width and the coordinates of the centers of the confined facility are automatically determined. The number of birds and waste generated by each facility also is automatically estimated. The automated procedure can be set up and executed in batch mode so that hundreds of slides could be processed without operator intervention.

## 4.9 Summary

The automated procedures developed to identify poultry houses were applied to thirty 35-mm slides selected randomly from over several hundred slides, for Rockingham County, Virginia. The dataset contained 182 poultry facilities previously identified by manual procedures (Mostaghimi, et. al., 1998). The basic procedures and purpose of each are summarized in Table 4-6.

**Table 4-6. Automated procedures and purpose.**

| Procedure | Purpose |
|---|---|
| Thresholding Analysis | Procedure to remove all features with spectral characteristics different then poultry houses |
| Image Processing/PCA | Removes all spectral characteristics not related to red color band |
| Image Processing/RGB to Hue on Red band | Enhances the red, green, and blue color bands |
| Image Processing/RGB to Saturation on red band | Removes the red from the red color band leaving green and blue |
| Image Processing/RGB to Saturation on Green band | Removes the green from green color band leaving only blue. |
| Conversion of raster to vector | Creates polygon shapes that can be used in shape analysis |
| Calculate shape parameters | Gives values for compactness and two length-width ratio |
| Extract polygons based on compactness | Gives a map containing polygons with given range of compactness values |

| | |
|---|---|
| Extract polygons based on length-width ratio based on width assumed equal to average width | Gives a map containing polygons with given range of length-width ratios |
| Extract polygons based on length-width ratio based on calculated width from centroid analysis | Gives a map containing poultry houses |

As indicated in Table 4-6, the automated procedure involved describing the spectral characteristics of the slide image in relation to the spectral characteristics of training data developed for known poultry facilities, performing image enhancement to sharpen the image, conducting shape analysis to identify the poultry facilities and using GIS functions to extract desired facility characteristics. The spectral analysis of training data (i.e. poultry facilities) indicated that poultry facilities were distinctly distinguishable from neighboring features. The spectral signature for poultry facilities had a mean intensity of 78 for color red, 87 for color green, and 103 for color blue. The standard deviation of brightness for color red was highest among the three primary colors. This can be significantly impacted by the condition of the roof as the brightness of metal materials tends to degrade with time of exposure to weather elements.

The next step in the automated process involved defining shape parameters that could be utilized to extract poultry facilities from an image that contained other features with similar spectral characteristics. The shape analysis of training data indicated that the length-width ratio and the compactness coefficient gave an excellent identification of poultry houses after appropriate spectral analysis, threshold analysis and image enhancement were completed. The average width of poultry houses based on the descriptive statistics was approximately 12.6 m. with a standard deviation of approximately 2.1 m, which is very close to the width of a modern poultry house. The average length of poultry houses in the training sample was approximately 149 m. The compactness parameter ranged from 30 to 130 with the range 53 to 61 representing a

95% confidence interval about the mean of 57.

The final step in the automated procedure was to import the image into ArcView environment, to convert objects from raster to a vector format and then to use procedures developed to extract polygons based on compactness and length-width ratio to extract polygons that represent poultry houses. The procedure was applied to thirty 35-mm slide images randomly selected from the slide database for Rockingham County, Virginia. For this test, all poultry houses were correctly identified in 28 of the 30 slides analyzed. A total of five poultry houses located on or across slide boundaries were not identified in one slide and two poultry houses were not identified in another slide. A total of eighteen features were misinterpreted as poultry houses. The results from this test suggest an accuracy of 97%. The accuracy of manual interpretation was approximately 97% (Mostaghimi, et. al., 1999). The results suggest that the automated procedure can provide reasonable identification of poultry houses and the geometric characteristics of each facility.

The manual procedure described by Mostaghimi, et al (1999) only gives the center coordinates for each poultry facility. The automated procedure not only gives the center coordinate for each poultry building but also gives estimates for geometric parameters area, length, width along with the capacity of building (i.e. number of birds), and waste load generated by birds including nutrient and bacteria content. The nutrient and bacteria load generated by each poultry facility is important information for conducting TMDL studies currently being developed for all impaired Virginia streams. The information is expected to be very helpful to consultants and state agencies conducting the studies. Other agricultural support agencies such as USDA/NRCS and USDA/FSA, Extension Service, etc. will find the information very helpful in the development of implementation plans designed to meet TMDL target water quality goals. The data also should be useful to Water Authorities for selection of appropriate treatment of water supplies and to county and local government jurisdictions for developing policies to minimize the degradation of water supplies.

Some of the difficulties encountered with the proposed procedures included houses that crossed the slide boundary, distortion at the boundary, and rectangular features with spectral and shape characteristics similar to poultry houses. One solution is to tile the individual slides into larger composites and remove distortion. The problem with this option is that the boundary problem will always exist, but could be minimized for the region of interest. Another possible option would be to develop a procedure that would import sections of adjacent slide images to include all features that had spectral signatures similar to poultry features. This would result in an image with irregular boundaries but otherwise may resolve the problem of poultry facilities located at the boundary. Additionally, the size of the data files (i.e. image files) would generally be much smaller than tiling into a larger unit. Other problems that exist in the USDA/FSA 35-mm slide imagery that impact image processing include cloud cover, inconsistency in the flight patterns and height of the flight. These issues are considered pre-existing and can only be resolved by changing how future imagery is collected.

## 4.10 Data for NPS Pollution Control Applications

This section describes the databases that were generated to support NPS pollution control studies (e.g. TMDL, NPS pollution potential, etc.) and an is illustration of how the data might be used in non-point source pollution control studies. An example of a shape file containing seven poultry houses identified from a slide image is given in Figure 4-23. Each poultry house was assigned a numerical code beginning with 1. The shape parameters and estimate of the number of poultry (birds) per house per cycle are given in Table 4-7. The largest house for this slide was house number 1 with a length of approximately 201 m. It had a capacity of approximately 13,130 market age birds. To obtain annual capacity, the above numbers would be increased by a factor of 4 to 5 to represent the number of cycles of birds raised per year. The smallest buildings were houses 2 and 3 with market bird capacity of approximately 6,989.

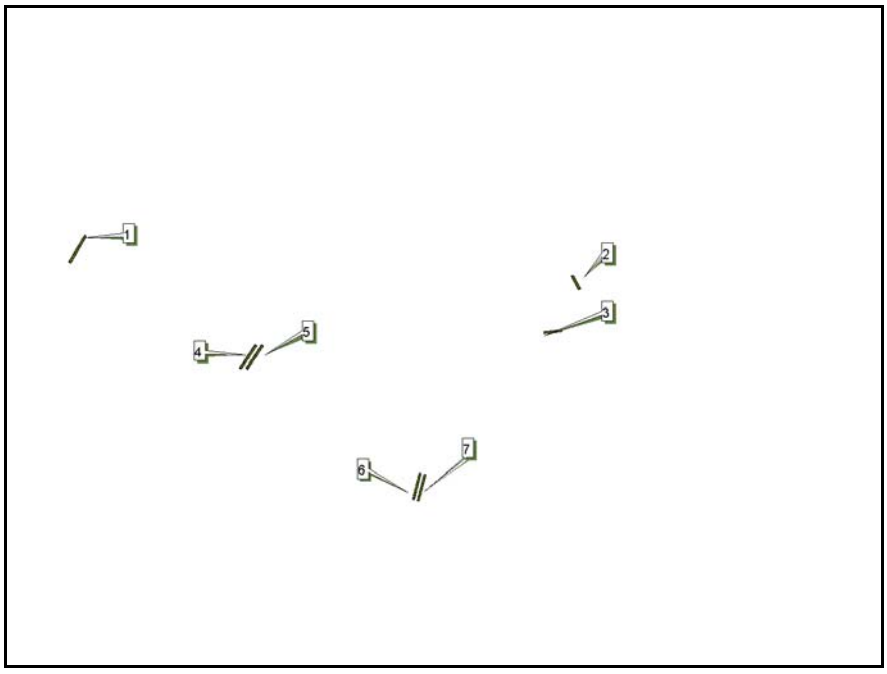


**Figure 4-23. Shape file of confined poultry facilities.**

**Table 4-7. Poultry house shape parameters with estimates of number of birds.**

| Facility ID | Area | Perimeter | Compactness | Width | Length | L/W | Market Bird |
|---|---|---|---|---|---|---|---|
| | (Acre) | (Meter) | | (Meter) | (Meter) | | (Number) |
| 1 | 0.62 | 435.03 | 73.00 | 12.82 | 201.23 | 18.04 | 13130 |
| 2 | 0.33 | 233.54 | 40.00 | 12.82 | 105.82 | 9.49 | 6989 |
| 3 | 0.33 | 250.47 | 46.00 | 12.82 | 105.66 | 9.47 | 6989 |
| 4 | 0.55 | 410.70 | 74.00 | 12.82 | 177.66 | 15.93 | 11648 |
| 5 | 0.55 | 409.93 | 74.00 | 12.82 | 176.95 | 15.86 | 11648 |
| 6 | 0.48 | 401.83 | 82.00 | 12.82 | 153.85 | 13.79 | 10165 |
| 7 | 0.50 | 387.25 | 72.00 | 12.82 | 162.55 | 14.57 | 10589 |

Area is the polygon area from the shape file.

Perimeter is the distance around the object calculated from the shape file

Compactness is the square of the perimeter divided by the area.

Width is the average width of a poultry house as determined from the training dataset.

Length is calculated from area and width.

L/W is the length-width ratio.

Chickens 10 week to market require about 2 $ft^2$/bird (Midwest Plan Service, 1987, page 515.1

The location coordinates (Universal Transverse Mercator (UTM)) for each poultry house, (Figure 4-23), waste generated, and estimates for two HSPF parameters used to model fecal coliform (ACQCP and SQOLIM) are presented in Table 4-8. Of these data, only the global coordinates are a direct result of the classification procedure developed during the course of this study.

The data presented in Table 4-6 can be utilized to obtain information useful in many types of applications e.g. determining the non-point source pollution potential of poultry facilities, pollutant source assessment studies required in development of TMDLs, etc. Because the development of bacterial TMDLs is a major issue in many states, a brief discussion is presented to illustrate how the information in Table 4-8 can be used to support TMDL development in Virginia.

A large number of water bodies in Virginia have been listed as impaired because of elevated levels of fecal coliform. Fecal coliform is an indicator bacteria but research has

shown that when it is present there is a very high probability that more dangerous bacteria will be present. Virginia, like many states, is under a consent decree to develop TMDLs for the impaired water bodies over the next eight to ten years. Most fecal coliform TMDLs, which are developed to-date in Virginia, have used the HSPF continuous simulation model or an enhanced derivative such as WinHSP or the BASINS/NPSM system. These model frameworks are used to link sources within a watershed to the water quality end-point or for the bacteria concentrations. The TMDL process requires extensive pollutant source assessment to identify potential sources of fecal coliform, location in the watershed, potential load delivered to the watershed, and pathways to water bodies (e.g. applied to pasture or cropland with the potential to enter the stream from storm runoff, or direct deposition by livestock in stream). Poultry waste may be stored and spread on cropland or pastureland at a later date or be removed from the facility and immediately applied to nearby cropland or pastureland. The fecal coliform density for poultry waste in storage can be significantly less because of die-off than poultry waste applied directly from the facility to the field. This procedure can not determine the various management options that the operator may use, but the geographic coordinates and the estimates of physical parameters (length, projected area and width) can be very helpful in identifying sources and the spatial distribution, e.g. by sub-watershed.

The geographic coordinates are used to position the facility within the watershed or sub-watershed, which provides an estimate of the location of waste application to the land. The physical dimensions of the facility as determined from the ArcView shape file can be used to approximate the number of poultry housed in each facility (Midwest Plan Service, 1987). With research that is available on the fecal coliform content of fresh poultry waste and the expected die-off from storage (Mostaghimi, et al 2001), expected waste production and fecal coliform production can be estimated for each poultry facility. GIS can be used to link the geographic coordinates with sub-watersheds and appropriate spreadsheets can then be used to calculate loads to be applied to each sub-watershed, by land use category, based on the number of poultry per facility, waste generated, fecal coliform in waste, and die-off. These data are utilized to determine HSPF model

parameters ACQOP, the fecal coliform accumulation rate and the maximum rate of storage SQOLUM. Parameters ACQOP and SQOLUM are used within HSPF to simulate a time series of in-stream pollutant concentration, i.e. fecal coliform concentration from poultry) to determine existing conditions within a given watershed. Further analysis then can be conducted to determine the TMDL or the assimilative capacity of the water body.

To illustrate the calculation process, the land use/cover associated with each facility was arbitrarily set to 1,000 acres and the waste produced from each poultry house was spread on this area. Poultry house one (1) generated approximately 1,221 kg per day of poultry waste. Based on the fecal coliform densities for poultry, the waste contained approximately 1.44E+12 counts #/day of fecal coliform (Metcalf and Eddy, 1991). Assuming the waste was removed from the poultry facility and immediately applied to the pastureland, the accumulation rate (ACQOP) for fecal coliform would be 4.3E+08 #/acre/day and the maximum rate of storage (SQOLIM) would be 6.45E+08 (#/acre/day). The slide image was projected to Universal Transverse Mercator (UTM) zone 17 coordinates. The location coordinate represents the center point. For building 1, the coordinates were 674349.88 m easting and 4250306.99 m) northing (see Table 4-7).

**Table 4-8. Estimates of HSPF model parameters for fecal coliform.**

| Facility Id | Two-week Bird | Chicken litter | FC | Pasture | ACQOP | SQOLIM | UTM | |
|---|---|---|---|---|---|---|---|---|
| | | | | | Pasture | Pasture | Centroid | Centriod |
| | | Lb/day | #/day | acre | #/acre/day | #/acre/day | X (meter) | Y (meter) |
| **1** | 38529 | 1221 | 1.44E+12 | 1000 | 4.3E+08 | 6.45E+08 | 674349.88 | 4250306.99 |
| **2** | 20263 | 642 | 7.58E+11 | 1000 | 3.7E+08 | 5.54E+08 | 676440.71 | 4250154.14 |
| **3** | 20231 | 641 | 7.57E+11 | 1000 | 2.3E+08 | 3.39E+08 | 676349.78 | 4249947.22 |
| **4** | 34017 | 1078 | 1.27E+12 | 1000 | 3.8E+08 | 5.70E+08 | 675026.87 | 4249843.75 |
| **5** | 33880 | 1073 | 1.27E+12 | 1000 | 3.8E+08 | 5.68E+08 | 675119.82 | 4249873.85 |
| **6** | 29457 | 933 | 1.1E+12 | 1000 | 3.3E+08 | 4.94E+08 | 675779.84 | 4249284.11 |
| **7** | 31123 | 986 | 1.16E+12 | 1000 | 3.5E+08 | 5.21E+08 | 675801.78 | 4249286.81 |

# 5. SUMMARY AND CONCLUSIONS

In 1998, streams in 187 watersheds encompassing over ten million acres of land in Virginia were identified by the Virginia Department of Environmental Quality (VADEQ), as impaired. Fecal coliform was identified as a leading cause of the impairments and agricultural activities were thought to be the primary source. Section 303(d) of the Clean Water Act requires that Total Maximum Daily Loads (TMDLs) be developed for all streams that do not meet standards (i.e. one or more pollutants impair them) and the USEPA is currently under a consent decree to develop TMDLs for the impaired water bodies over the next 10-years. Confined livestock facilities have been identified as a significant potential pollution source for water bodies impaired by fecal coliform. The identification of these sources is very important for the development of implementable strategies to meet the state's water quality standards. This study was initiated to investigate the use of USDA/FSA 35-mm aerial photography for developing an automated procedure to identify poultry facilities.

The overall goal of this study was to utilize pattern recognition and GIS technologies to identify and extract confined animal facilities from 35-mm slide imagery. The extracted information can be utilized to estimate potential stream pollutant loads from confined animal facilities. This study is expected to provide important data for Total Maximum Daily Load (TMDL) analysis. The results of this research are expected to reduce the cost for data preparation and to improve decision makers' abilities to better understand potential impacts of pollutants on receiving water systems and to develop effective alternative management strategies. The following specific objectives were undertaken: (a) Develop spectral and geometric signatures to identify poultry houses; (b) Develop pattern recognition algorithms to extract confined poultry facilities; and (c) Evaluate the accuracy of the procedure on an independent data set.

An automated procedure was developed which utilizes image processing and GIS techniques to identify and extract facilities. The procedure involved converting 35-mm slides (i.e. photographs) to a digital image and developing signatures based on spectral

characteristics and shape of the facilities. Adobe and Minitab packages were used to define spectral characteristics of training data. Histogram analysis was used to describe spectral characteristics and to identify the range of spectral characteristics that included poultry facilities. Descriptive statistics were used to define the thresholds or boundaries between spectral characteristics of poultry building and other image features.

Thresholding analyses were utilized to eliminate all image features with spectral characteristics outside of this range. Additional analyses were made to remove noise in the spectral image due to the sun angle, line of sight of camera, variation in roof reflectance due to rust and/or aging, shading by trees, etc. A primary objective in these analyses was to enhance the spectral characteristics for the poultry facility while at the same time retaining physical characteristics, i.e. the spectral characteristic is represented by a single blue color with a high brightness value. The techniques developed to achieve a single blue color involved the use of Principal Component Analysis (PCA) on the red color band from histogram analysis followed by RGB to Hue and RGB to Saturation analyses on the red and green color bands, respectively, of the image resulting from the PCA analysis. The features remaining from this series of analyses were converted into polygons (shape file) using ArcView GIS, which was then used to calculate the area and perimeter of each polygon.

The geometric characteristics of poultry shapes were analyzed based on poultry facilities found in the training data set. The width of poultry houses were found to follow a normal distribution with a low standard deviation indicating that the building widths were nearly the same with an average width of approximately 12.6 m (48 feet). The width of modern poultry houses in the area average approximately 50 feet. The length of poultry facilities was found to vary significantly ranging from 63 (240 ft) m to 261 (994 ft) m with an average length of 149 (567 ft) m. These findings were consistent with poultry house construction with the width fixed and additional length added to increase capacity. The ratio of width to length was identified as a useful parameter for identifying poultry house shapes. The compactness ratio, which is defined as the square of the perimeter divided by area, has also been identified as a good indicator for identifying shape. This parameter

was found to range from 30 to 130 with a mean value of approximately 57. It was found that the order for identifying shapes was first to select polygons based on compactness. The length-width ratio was then used to select poultry facilities from this dataset. An additional shape parameter based on the polygon centroid was calculated and used as the final selection criteria.

Thirty 35-mm slide photographs containing 182 confined poultry houses were selected randomly from over 2000 digital slides data set for Rockingham County and used to evaluate the effectiveness of the above automated procedures.

All poultry houses were identified on 28 of the 30 slides. The principal difficulty involved identifying poultry houses located at the edge of slides, particularly those that crossed the slide boundary. On two slides, seven poultry houses out of a total of 19 were not identified (i.e. they were missed). With the exception of two houses, five were located near or across the slide boundary. Two houses were parallel and very close to each other. Spectral characteristics of the feature(s) between the two poultry houses resulted in the two buildings appearing as joined at several locations. This resulted in a compactness value outside of the range defined by the training dataset developed for the study. With the first two steps of the shape analysis, i.e. selection based on compactness followed by selection based on the length-width ratio, some polygons will be selected that are not poultry houses. For the current randomly selected database, 17 non-poultry houses were selected. All but one of these facilities was identified as dairy facilities. Three slides had two feature misidentified and eleven slides had one feature misidentified. Applying a third shape parameter based on polygon centroid eliminated all non-poultry house objects. Ninety seven percent of existing poultry facilities were correctly identified which compares well with manual procedures that resulted in an accuracy of 97% and was considered acceptable for the type of imagery being utilized for the analysis.

The manual procedure described by Mostaghimi, et al (1999) only gives the center coordinates for each poultry facility. The automated procedure not only gives the center coordinate for each poultry building but also gives estimates for geometric parameters

area, length, width along with the capacity of building (i.e. number of birds), and waste load generated by birds including nutrient and bacteria content. The nutrient and bacteria load generated by each poultry facility is important information for conducting TMDL studies currently being developed for all impaired Virginia streams. The information is expected to be very helpful to consultants and state agencies conducting the studies. Other agricultural support agencies such as USDA/NRCS and USDA/FSA, Extension Service, etc. will find the information very helpful in the development of implementation plans designed to meet TMDL target water quality goals. The data also should be useful to Water Authorities for selection of appropriate treatment of water supplies and to county and local government jurisdictions for developing policies to minimize the degradation of water supplies.

Some of the difficulties encountered with the proposed procedures included houses that crossed the slide boundary, distortion at the boundary, and rectangular features with spectral and shape characteristics similar to poultry houses. One solution is to tile the individual slides into larger composites and remove distortion. The problem with this option is that the boundary problem will always exist, but could be minimized for the region of interest. Another possible option would be to develop a procedure that would import sections of adjacent slide images to include all features that had spectral signatures similar to poultry features. This would result in an image with irregular boundaries but otherwise may resolve the problem of poultry facilities located at the boundary. Additionally, the size of the data files (i.e. image files) would generally be much smaller than tiling into a larger unit. Other problems that exist in the USDA/FSA 35-mm slide imagery that impact image processing include cloud cover, inconsistency in the flight patterns and height of the flight. These issues are considered pre-existing and can only be resolved by changing how future imagery is collected.

This study was the first attempt to merge GIS, remote sensing and pattern recognition technologies to extract information from 35-mm aerial slides important to NPS pollution studies and/or applications. The following conclusions could be drawn from the study:

- The resolution of the 35-mm aerial slide photography was adequate for identifying poultry facilities. The focal length of the camera used to obtain the 1995 Rockingham County 35-mm imagery was 24-mm and the flight height was approximately 9500 feet, which gives a scale of approximately 1:120065. With this scale and a scanning resolution of 2150 ppi one pixel on the slide film is approximately equal to 0.14 m ground distance. Widths of poultry facilities ranged from approximately 8.4 m (older buildings) to 13.1 m (newer buildings), the slide photography was adequate for this study.
- Poultry facilities had well defined spectral characteristics. The mean brightness for all color bands of poultry facility samples was 222 compared to a mean brightness of 87 for the background or all other objects.
- The shapes of the confined poultry facilities were well preserved after image processing and shape analysis. Principle Component Analysis, RGB to Hue and RGB to Saturation procedures were utilized to remove the noise from objects while maintaining the geometric shape of the object. This preservation was important for identifying poultry facilities and for estimating the size of each facility.
- The width of confined poultry facilities followed a normal distribution with a standard deviation of 2.1 m which indicates that the width of the buildings were nearly the same.
- Compactness and length-width ratio were good indicators of building shape.
- The automated procedure was able to identify 97 percent of 182 poultry houses in an independent evaluation of the method. This accuracy was identical to the 97 % achieved by manual procedures for the same sample.
- The automated procedure can save significant time and resources and appears to be a useful tool for periodic updating of the poultry facilities database.
- The automated procedure provides data important for TMDL development and for ranking the NPS pollution potential of poultry facilities. The estimate of building size (i.e. area) was used to estimate the number of birds being produced. These estimates allow calculation of bacteria densities (e.g. fecal coliform) and nutrient availability,

which are both important parameters in TMDL analysis and for other non-point source studies.

# 6. RECOMMENDATIONS

The following recommendations are suggested for future research.

1. This study used 35-mm slides flown with only limited quality control. Future research with this application should focus on achieving better flight control and improved techniques for distortion correction and joining the digital images to minimize boundary problems. The proposed procedures should be applicable to large-format images collected at resolutions sufficient for the purposes of this study.
2. Development of model interface to link the system directly to NPS pollution models will greatly enhance the usefulness of the procedures developed in this study.
3. Extensive neighborhood analysis to identify possible commonalities between poultry houses and other features is needed. This would involve statistical analyses of features surrounding known poultry facilities at some predefined distance. The sample should be fairly large, possibly several thousand, and should be based on sampling theory to support the type(s) of statistical analysis perceived to be required. This may be the only alternative to achieve reasonable accuracies for regions where a large majority of the poultry facilities are relatively narrow and much shorter then typically found in Virginia. For these situations, other facilities such as machine sheds, feed lots, etc. likely would have footprints similar to poultry facilities.
4. Higher resolution images and extensive neighborhood analysis are expected to be necessary to develop a useful automated procedure to identify other confined animal facilities. To identify features typically associated with dairy operations, e.g. silos and waste storage pits will require imagery collected at a higher resolution than current 35-mm slide imagery used in this study.

   Some the issues that must be considered in developing an automated procedure to identify dairy operations from 35-mm slide imagery are summarized in the following paragraphs.

Like poultry operations, dairy operations are highly confined and often involve relatively large herds. Unlike a poultry operation, a dairy facility normally consists of several buildings that are designed to support the operation. A dairy farm normally has a milk facility, loafing lot, a waste collector, silos and other support facilities (Figure 6-1). These buildings are usually separated but some may be interconnected under one roof. The milk herd facility can include free-stall housing, loose housing, confined free-stall housing or stanchion-tie (comfort stall housing) for milking the cows. The facility is normally located in a relative large separate building to orient sanitation and cows for comfort.

Types of waste storage facilities typically include underground tank(s), roofed storage, holding pond, slurry, and lagoon storage or a reception pit. The receptor pit is the most common type used in Virginia. The storage pit is normally located in the center or on the end of a free-stall barn but some distance from water supplies and the milking center. A dairy farm normally has several silos located besides the milking facility building, which are used to store feed. Upright, trench and ground-level silos, which are common features of dairy farms but are very difficult to identify on 35-mm aerial slide photography.

Loafing lots characteristically are relatively small areas of pasture that are well grazed and show evidence of heavy animal traffic. These areas are usually located adjacent to the milking facility.


**Figure 6-1. A dairy farmer in Rockingham County, Virginia is shown in a 35-mm slide.**

# 7. BIBLIOGRAPHY


ADOBE. 1996. Photoshop 5.0. http://www.adobe.com/.

Baston. 1974. Report on Results of A Statewide Field Test of 35 mm Aerial Photography System USDI-Bureau of Land Management, Billings, Montana.

Bonham-Carter. G.F. 1994, Geographic Information Systems for Geoscientists: Modeling with GIS,: Kidlington, Elsevier, 398 p.

Campbell, J. B. 1987. Introduction to Remote Sensing. The Guilford Press. ISBN 0-89862-776-1.

Chesapeake Bay Foundation. 1998. "The need to regulate poultry: what science and the experts say" Fact sheet. http:// www.eieio.org/poultry/needpoulregs.htm.

Cleanwater. 1999. Clean Water Action Plan, The first year. The future. http://www.cleanwater.gov/anniv.

Clements, H. B., D. E. Ward, and C. W. Adkins. 1983. Measuring fire behavior with photography. Photogrammetric Engineering and Remote Sensing, 49, 213-217.

Cipra, B. 1993. Wavelet Applications Come to the Fore. SIAM News, Volume 26, Number, 7.

Davis, J. C. 1986. Statistics and Data Analysis in Geology. John Wiley &and Son, Inc.

Dejesus, E. X. 1995. Faster and more sophisticated algorithms are helping computerized facial-recognition systems come of age. http://www.byte.com/art/9502/sec10/art2.htm.

Duda, R. O. and P. E. Hart. 1973. Pattern Classification and Scene Analysis. New York, NY: Wiley and Sons.

Eastman, J. R. 1997. The IDRISI User Manual. The IDRISI Project Clark Labs for Cartographic Technology and Geographic Analysis Clark University 950 Main Street, Worcester, MA USA 01610-1477.

ER Mapper. 1995. ER Mapper Tutorial. Earth Resource Mapping 4370 La Jolla Village Drive, Suite 900, San Diego, CA.

ESRI.. 1997. Arc/View GIS 3.0.

Farmer, L.D. and R.Q. Robe. 1977. Photogrammetric determination of iceberg volumes. Photogrammetric Engineering and Remote Sensing, 43, 183-189.

FDA/CFSAN. 1998. Analysis of Comments to the Proposed Guide: Guide to Minimize Microbial Food Safety Hazards for Fresh Fruits and Vegetables. http://vm.cfsan.fda.gov/~dms/prodrpt.html.

Fitzgerald, R. 1999. Personal Communication. Farm Service Agency, USDA, Virginia. Tel: 540-332-9992.

FSA. 1999. http://www.fsa.usda.gov/pas/fsanews/html/1999/aug99/gis.htm

Fu, Y. T. 1994. Comparison of two hydrological models on a Virginia piedmont watershed. M. S. Thesis, Virginia Tech. Blacksburg, Virginia.

Georgia Tech. 1997. Learning from experience: new pattern recognition & detection technique may help radiologists analyze digital mammograms. Research News. http://gtresearchnews.gatech.edu/newsrelease/MAMMOG.htm.

Goodchild, M. F. 1999. "GIS and transportation: status and challenges" Keynote address, International Workshop on GIS-T and ITS, Chinese University of Hong Kong.

Goolsby D. A. and W.A., Battaglin. 2000. Nitrogen in the Mississippi Basin-estimating sources and predicting flux to the Gulf of Mexico. USGS Fact Sheet 135-00 December.

Hall, E., 1979. Computer Image Processing and Recognition. Academic Press, Inc., 111 Fifth Avenue, New York, New York 10003.

Hall, R. J. and A. H., Aldred. 1992. Forest regeneration appraisal with large-scale aerial photographs. The Forestry Chronicle, 68, 142-150.

Hall, W. B. and T. H. Walsh. 1977. Color oblique stereo aerial photographs for planning rapid geological reconnaissance and other types of geological programs. Wyoming Geological Association Guidebook, 741-760.

Hiby, A. R. and D. Thompson and A. J.Ward. 1988. Census of gray seals by aerial photography. Phtogrammetric record, 12, 589-594.

Hixson, M., D. Scholz, and N. Fuhs. 1980. Evaluation of several schemes for classification of remotely sensed data. Programmatic Engineering and Remote Sensing, Vol. 46, pp. 1547-1533.

Horsley & Whitten, Inc.. 1996. Identification and evaluation of nutrient bacterial loadings to Maquiot Bay, Brunswick and Freeport, Maine. Casco Bay Estuary Project.

Hu, M. K. 1962. Visual problem recognition by moment invariant. IRE Trans. In Form Theory, Vol. IT-8, pp179-187, Feb.

Jaggard, K. and Clark, C. J. 1990. Remote sensing to predict the yield of sugar beet in England. In: M.D. Steven and J.A. Clark (Editors), Applications of Remote Sensing in Agricultural. Butterworths, London, pp.201-208.

Ketting, R. L. and D. A. Landgrebe. 1975. Classification on Multispectral Image Data by Extraction and Classification of Homogeneous Objects. In Proceedings, Symposium on Machine Classification of Remotely Sensed Data. West Lafayette. IN: Laboratory for Applications in Remote Sensing, pp. 2A-1-2A-11.

Kirby, C. L. 1980. A Camera and Interpretation Systems for Assessment in Forest Regeneration. Inf. Report Nor-X-221. Canadian Forestry Service.

Kodak. 1998. http://www.kodak.com/productInfo/productInfo.shtml.

Kratz, G. P.1998. Will Corporate Farms Hog Market? Deseret News September 16.

Lee, Y. J. and R. W., McKelvey. 1984. Digitized small format aerial photography as a tool for measuring food consumption by trumper swans. Photogrametric Engineering and Remote Sensing, 50, 215-219.

Liang, J. J. and V. Clarson. 1989. A new approach to classification of brainwaves. Pattern Recognition, 22:767-774

Lipton, E. 1997. A Growth Industry Spurs Concern. Poultry Farming Employs Many, but It May Threaten Water For Many More. The Washington Post. Sunday, June 1, 1997.

Lillesand, T. M. and R. W. Kiefer. 1994; Remote Sensing and Image Interpretation; 3rd ed. John Wiley and Sons, Inc.

Meyer, M., F. Batson and D. Whitmer. 1982. Helicopter-born 35-mm Aerial Photography Applications to Range and Riparian Studies (IAFHE No.82-1) Univ. Minn, College of Forestry.

Midwest Plan Service (MWPS-1). 1987. Structures and Environment Handbook. Eleven Editions.

Minitab Inc. 1995. http://www.minitab.com/.

Mok, E. C. and A. L. Boyer. 1986. Encoding patient contours using Fourier descriptors for computer treatment planning. Med. Phys., 13:413-415.

Mostaghimi, S. and P.W. McClellan. 1997. Personal communication. Dept. of Biological Systems Engineering, Virginia Tech, Blacksburg, Virginia.

Mostaghimi, S., McClellan, J. Miller and Y. Fu. 1998. Identification of confined animal sites for selected watershed in Virginia using Remote Sensing: Phase I. Report No. CAS-P1-0298. Prepared for Virginia Division of Soil and Water Conservation. Biological Systems Engineering Department, Blacksburg, VA 24061-0303.

Mostaghimi, S., McClellan, J. Miller and Y. Fu. 1999 Identification of confined animal sites for selected watershed in Virginia using Remote Sensing: Phase I. Report No. CAS-P1-0299 Prepared for Virginia Division of Soil and Water Conservation. Biological Systems Engineering Department, Blacksburg, VA 24061-0303.

Mostaghimi, S., K. DeGuise and P. W. McCllan. 2001. Field investigation of bacteria die-off rates in cow manure and poultry litter. Report No. 0301BAC. The Virginia Department of Conservation and Recreation Division of Soil and Water Conservation, Richmond, Virginia.

Norton-Griffiths, M. 1973. Counting the Serengeti migratory wildebeest using two-stage sampling. East African Wildlife Journal, 11, 35-49.

Norton-Griffiths, M. 1988. Aerial point sampling for land use survey. Journal of Biogeography, 15, 149-156.

Novotny, V., and B. Chesters. 1981. Handbook on non-point pollution: Sources and management. Van Nostrand and Reinhold Company, New York, NY

Nikon Corporation. 1996. Electronic Image Division 4-25, Nishi-Ohio 1-chome Shinagawa-ku, Tokyo 140, Japan

Otterloo, P. J. Van. 1991. A Contour-oriented approach to shape analysis. Prentice Hall International (UK) Ltd 66 Wood Lane End, Hemel Jempstead Hertfrshire HP2 4RG.

Pane, T. 1997. Personal Communication. Dept. of Conservation and Recreation, Division of Soil and Water Conservation, Richmond, Virginia.

Pavlidis, T. 1978. A review of algorithms for shape analysis. Computer Graphics and Image Processing 1978. Selected Papers on Industrial Machine Vision Systems**,** B.G. Batchelor and P.F. Whelan (Eds)**.**SPIE Milestone Series MS 97, SPIE Optical Engineering Press, 654 pages; 46 papers, ISBN 0-8194-1580-4 (hardbound); ISBN 0-8194-1579-0 (softbound), 1994.

Rozum, M. A and F. Huminek. 1998. Animal waste management, an initiative for the cooperative extension System. http://www.escop.msstate.edu/minutes/min2-99.htm.

Scholz, D., N. Fuhs, and M. Hixon. 1979. An evaluation of seven different classifications schemes, their parameters, and performance. In Proceedings of the Thirteenth International Symposium on remote sensing of the Environment. Ann Arbor: University of Michigan, pp.1143-1149.

Stafford, J.V.; B. Ambler and H. C. Wheeler. 1994. Mapping yield variation to aid crop management. Proc. XII World Congress on Agric. Eng., Milan, Aug 29-Sept 1, 1994, vol. 1, p 88-96.

Steven, M. D. 1993. Satellite remote sensing for agricultural management: opportunities and logistic constraints. Volume 48, number 4, 1993. Journal of Photogrammetr and Remote Sensing.

Stevens, A. A. 1972. Application of small formats color and color infrared aerial photography to Dutch elm disease detection. In 38th Annual Meeting of American Society of Photogrammetry, 349-357. Washington: ASP.

Tello, R., R. W. Mann, and D. Rowell. 1986. Fourier descriptors for biomedical computer graphics. In Proceedings of the Eighth Annual Conference of the IEEE/Engineering in Medicine and Biology Society, pages 1062-1064, IEEE.

The Muddy Creek TMDL Establishment Workgroup. 1999. Virginia Department of Environmental Quality, Virginia Department of Conservation and Recreation.

The New Shorter Oxford English Dictionary. 1993. Oxford: Oxford University Press, 1993, pp.3836.

Tueller, P. T., P. C., Lent, R. D., Stager, E. A. Jacobsen, and K. A., Platou. 1988. Rangeland vegetation changes measured from helicopter-born 35 mm aerial photography. Photogrammetric Engineering and Remote Sensing, 54, 609-614.

Tomita, F. and S. Tsuji. 1990. Computer analysis of visual textures. Kluwer Academy Publishers, Norwell, Massachusetts.

USDA/NRCS. 1995. Animal Manure Management. http://www.nrcs.usda.gov/technical/land/pubs/ib7text.html

USEPA. 1991. Guidance for Water Quality-based Decisions: The TMDL process. U.S.EPA office of Water Regulations and Standards. Washington, D.C. 65 pages.

USEPA. 1996. National Water Quality Inventory: 1996 Report to Congress.

USEPA. 1997. http://www.epa.gov/OWOW/NPS/qa.html.

USEPA. 1997. The Total Maximum Daily Loading (TMDL) program. http://www.epa.gov/owow/tmdl/advisory.html.

USEPA. 1999. BASINS 2.0 Fact Sheet. http://www.epa.gov/ost/basins/basinsv2.htm.

USGS. 1998. Fish Health, Fungal Infections, and Pfiesteria: The Role of the U.S. Geological Survey. http://biology.usgs.gov/pr/newsrelease/1998/9-23d.html.

Ullmann, J. 1973, Pattern Recognition Techniques, Crane, Russak, New York.

VADCR. 1997. http://www.dcr.state.va.us/sw/swcds.htm

VADCR. 1998. http://www.dcr.state.va.us/sw/swcds.htm

Virginia Department of Conservation and Recreation. 1999. http:// www.dcr.state.va.us.

Wang, J. F. and P.J, Howarth. 1987. Automatic road network extraction from Landsat TM imagery. ASPRS-ACSM Annual Convention, Baltimore, Vol.1, pp429-438.

Warner, W. S., R. W. Graham and R. R. Read. 1996. Small Format Aerial Photography. American Society for Photogrammetry and Remote Sensing, 5410 Grosvenor Lane, Suite 210, Bethesda, Maryland 20814-2160, USA. ISBN 1-57083-0347

Weng, J. 1993. Image Matching Using the Windowed Fourier Phase, Int. J. of Comp. Vis. 3, 211-236

Wracher, D. A. 1973. Small Format Aerial Photography. Mining Engineering, 27, 47-48.

Willett, A. M. and B. K. Ward. 1978. Forest insect and disease survey by remote sensing. Bulletin of the Remote Sensing Association of Australia, 11-12.

# APPENDIX A: SLIDE IMAGES

This appendix includes displays for the following slide images used to validate this study.


**Figure A-1 Automated identification results for slide B1737q05002 (4 poultry houses identified, 0 missed, 1 misinterpreted).**

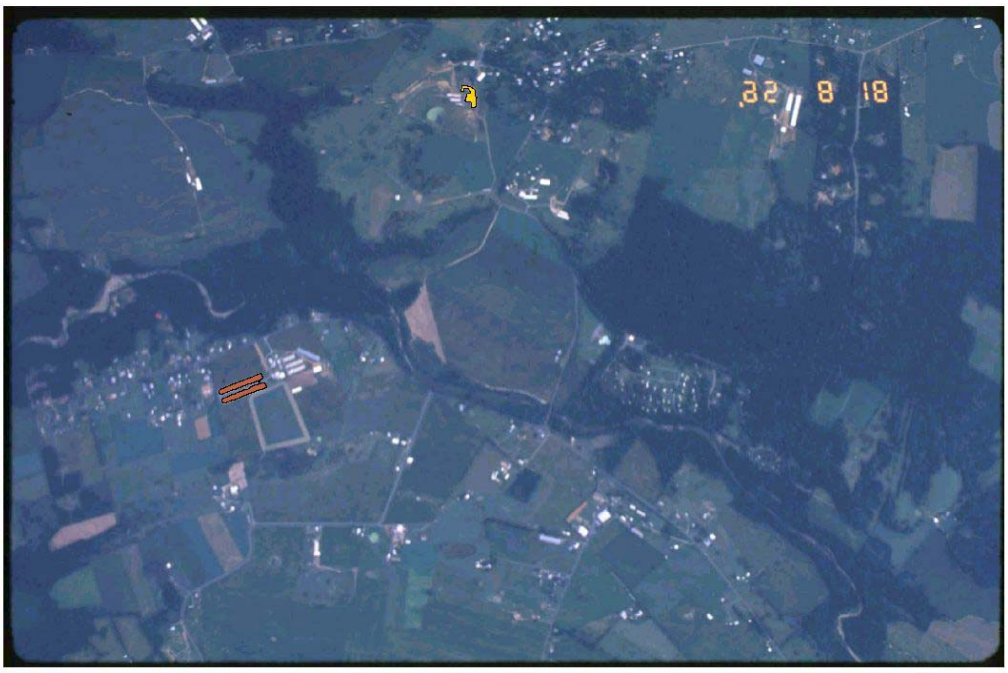


**Figure A-2. Automated identification results for slide B1738q01005 (2 poultry houses identified, 0 missed, 0 misinterpreted).**

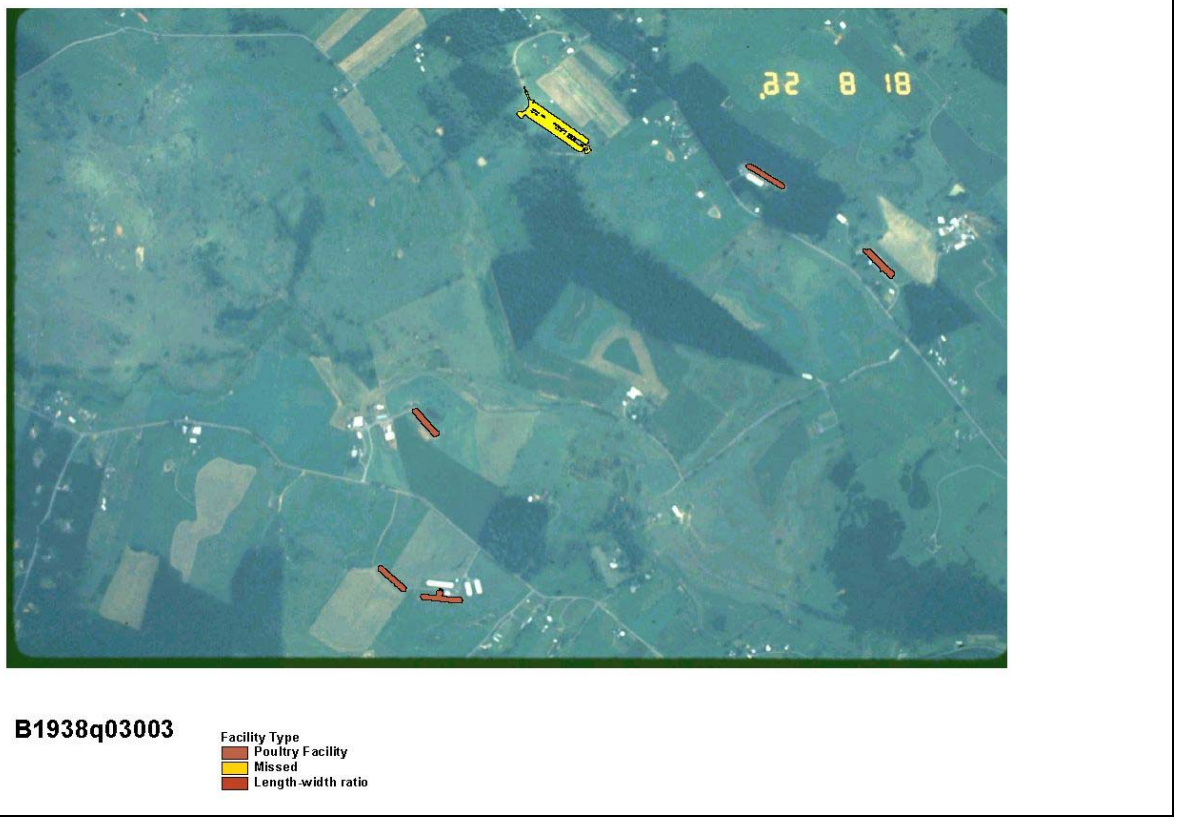


**Figure A-3. Automatic identification results on slide of B1938q03003 (5 poultry houses identified, 2 missed, 0 misinterpreted).**


**Figure A-4. Automated identification results for slide B1938q03009 (7 poultry houses identified, 0 missed, 2 misinterpreted).**


**Figure A-5. Automated identification results for slide B1938q04002 (3 poultry houses identified, 0 missed, 1 misinterpreted).**

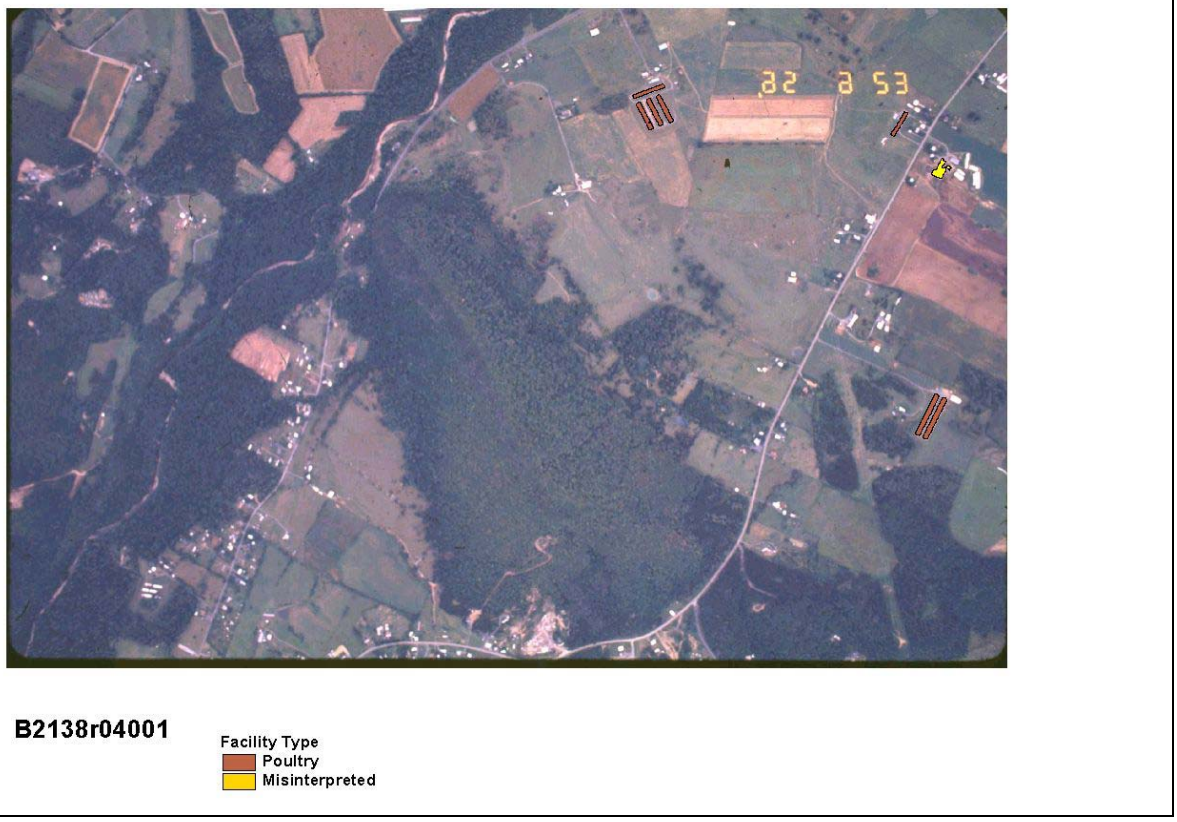


**Figure A-6. Automated identification results for slide B2138r04001 (6 poultry houses identified, 0 missed, 1 misinterpreted).**


**Figure A-7. Automated identification results for slide B2138r05001 (9 poultry houses identified, 0 missed, 0 misinterpreted).**


**Figure A-8. Automated identification results for slide B2239r01001 (9 poultry houses identified, 0 missed, 0 misinterpreted).**


**Figure A-9. Automated identification results for slide B2539q02001 (8 poultry houses identified, 0 missed, 0 misinterpreted).**


**Figure A-10 Automated identification results for slide B2539q02004 (6 poultry houses identified, 0 missed, 1 misinterpreted).**


**Figure A-11. Automated identification results for slide B1939q04001 (2 poultry houses identified, 0 missed, 0 misinterpreted).**


**Figure A-12. Automated identification results for slide B2539q01010 (14 poultry houses identified, 0 missed, 0 misinterpreted).**


**Figure A-13. Automated identification results for slide B2539r02002 (12 poultry houses identified, 0 missed, 1 misinterpreted).**


**Figure A-14. Automated identification results for slide B2539r02008 (7 poultry houses identified, 0 missed, 2 misinterpreted).**

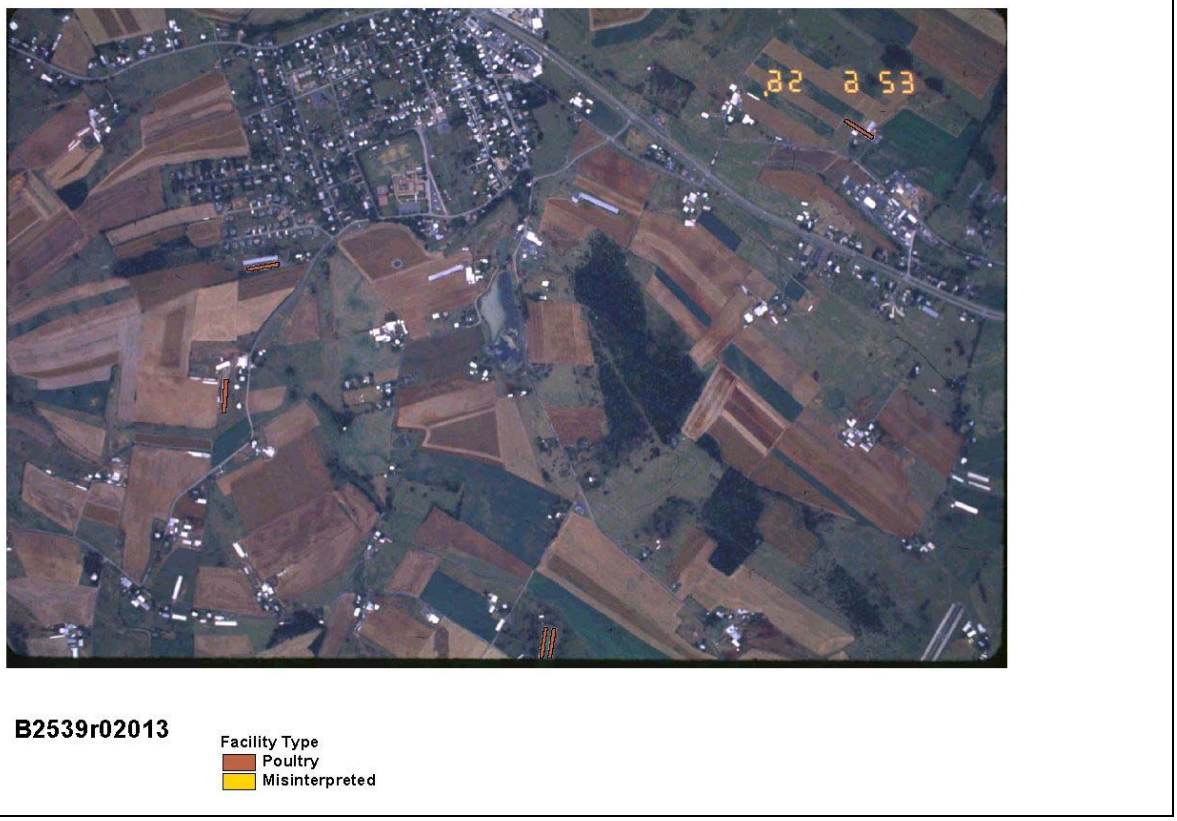


**Figure A-15. Automated identification results for slide B2539r02013 (5 poultry houses identified, 0 missed, 0 misinterpreted).**

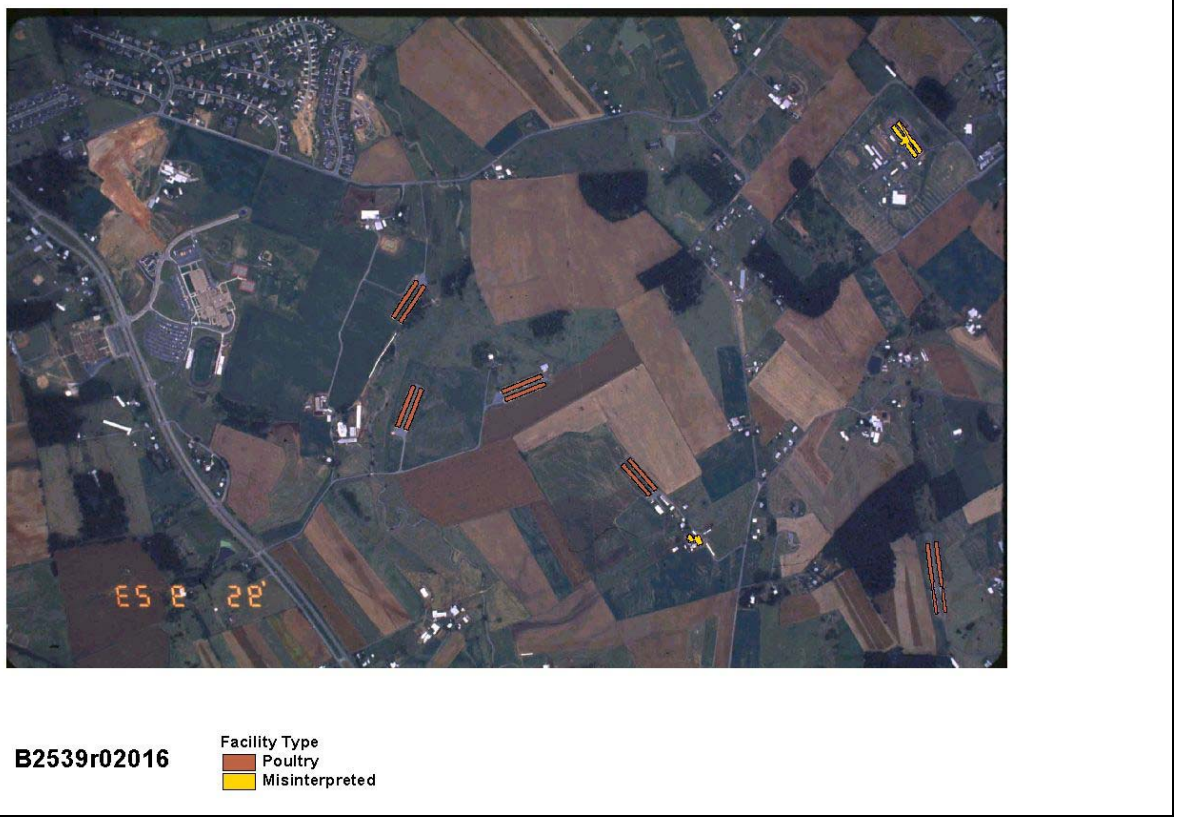


**Figure A-16. Automated identification results for slide B2539r02016 (10 poultry houses identified, 0 missed, 2 misinterpreted).**


**Figure A-17. Automated identification results for slide B3541q01005 (1 poultry houses identified, 0 missed, 0 misinterpreted).**


**Figure A-18. Automated identification results for slide B3541r06002 (0 poultry houses identified, 0 missed, 1 misinterpreted).**


**Figure A-19. Automated identification results for slide B3542r01006 (3 poultry houses identified, 0 missed, 0 misinterpreted).**


**Figure A-20. Automated identification results for slide B3542r02002 (7 poultry houses identified, 0 missed, 0 misinterpreted).**


**Figure A-21. Automated identification results for slide B2542r02007 (2 poultry houses identified, 0 missed, 0 misinterpreted).**


**Figure A-22. Automated identification results for slide B4740r01011 (6 poultry houses identified, 0 missed, 0 misinterpreted).**


**Figure A-23. Automated identification results for slide B4740r06001 (2 poultry houses identified, 0 missed, 1 misinterpreted).**


**Figure A-24. Automated identification results for slide B4740s04007 (3 poultry houses identified, 0 missed, 1 misinterpreted).**


**Figure A-25. Automated identification results for slide B4741s01006 (3 poultry houses identified, 0 missed, 0 misinterpreted).**


**Figure A-26. Automated identification results for slide B2539q01010 (4 poultry houses identified, 0 missed, 0 misinterpreted).**


**Figure A-27. Automated identification results for slide B1939q01007 (7 poultry houses identified, 0 missed, 1 misinterpreted).**


**Figure A-28. Automated identification results for slide B2539r01005(5 poultry houses identified, 0 missed, 0 misinterpreted).**


**Figure A-29. Automated identification results for slide B2539r01008 (7 poultry houses identified, 5 missed, 1 misinterpreted).**


**Figure A-30. Automated identification results for slide B4740s06004(13 poultry houses identified, 0 missed, 1 misinterpreted).**

# APPENDIX B. STATISTICS FROM PRINCIPAL COMPONENT ANALYSIS FOR THIRTY SLIDE IMAGES FROM VALIDATION STUDY AREA

**Table B-1. Statistics for dataset: b1738q01005.**

| | Band1 | Band2 | Band3 |
|---|---|---|---|
| Non-Null Cells | 3465504 | 3465504 | 3465504 |
| Area In Hectares | 346.550 | 346.550 | 346.550 |
| Area In Acres | 856.345 | 856.345 | 856.345 |
| Minimum | 0.000 | 4.000 | 0.000 |
| Maximum | 255.000 | 255.000 | 255.000 |
| Mean | 64.755 | 89.836 | 122.448 |
| Median | 64.000 | 91.000 | 125.000 |
| Std. Dev. | 24.244 | 23.431 | 26.360 |
| Std. Dev. (n-1) | 24.244 | 23.431 | 26.360 |
| Corr. Eigenval. | 2.750 | 0.218 | 0.032 |
| Cov. Eigenval. | 1672.860 | 140.192 | 18.548 |

| Correlation Matrix | Band1 | Band2 | Band3 |
|---|---|---|---|
| Band1 | 1.000 | 0.938 | 0.785 |
| Band2 | 0.938 | 1.000 | 0.900 |
| Band3 | 0.785 | 0.900 | 1.000 |
| Determinant | 0.019 | | |

| Corr. Eigenvectors | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.572 | -0.650 | -0.500 |
| Band2 | 0.596 | -0.089 | 0.798 |
| Band3 | 0.563 | 0.755 | -0.337 |

| Inv. of Corr. Ev. | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.572 | 0.596 | 0.563 |
| Band2 | -0.650 | -0.089 | 0.755 |
| Band3 | -0.500 | 0.798 | -0.337 |

| Covariance Matrix | Band1 | Band2 | Band3 |
|---|---|---|---|
| Band1 | 587.763 | 532.627 | 501.559 |
| Band2 | 532.627 | 549.007 | 556.007 |
| Band3 | 501.559 | 556.007 | 694.831 |
| Determinant | 4349996.794 | | |

| Cov. Eigenvectors | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.558 | -0.665 | -0.496 |
| Band2 | 0.565 | -0.133 | 0.814 |
| Band3 | 0.608 | 0.735 | -0.301 |

| Inv. of Cov. Ev. | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.558 | 0.565 | 0.608 |
| Band2 | -0.665 | -0.133 | 0.735 |
| Band3 | -0.496 | 0.814 | -0.301 |

**Table B-2. Statistics for dataset: b1737q05002.**

| | Band1 | Band2 | Band3 |
|---|---|---|---|
| Non-Null Cells | 3465504 | 3465504 | 3465504 |
| Area In Hectares | 346.550 | 346.550 | 346.550 |
| Area In Acres | 856.345 | 856.345 | 856.345 |
| Minimum | 0.000 | 0.000 | 0.000 |
| Maximum | 255.000 | 255.000 | 255.000 |
| Mean | 59.408 | 84.714 | 111.786 |
| Median | 54.000 | 82.000 | 113.000 |
| Std. Dev. | 24.174 | 24.367 | 25.546 |
| Std. Dev. (n-1) | 24.174 | 24.367 | 25.546 |
| Corr. Eigenval. | 2.780 | 0.190 | 0.029 |
| Cov. Eigenval. | 1694.604 | 118.590 | 17.570 |

| Correlation Matrix | Band1 | Band2 | Band3 |
|---|---|---|---|
| Band1 | 1.000 | 0.950 | 0.814 |
| Band2 | 0.950 | 1.000 | 0.905 |
| Band3 | 0.814 | 0.905 | 1.000 |
| Determinant | 0.016 | | |

| Corr. Eigenvectors | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.574 | -0.625 | -0.529 |
| Band2 | 0.593 | -0.127 | 0.795 |
| Band3 | 0.564 | 0.770 | -0.298 |

| Inv. of Corr. Ev. | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.574 | 0.593 | 0.564 |
| Band2 | -0.625 | -0.127 | 0.770 |
| Band3 | -0.529 | 0.795 | -0.298 |

| Covariance Matrix | Band1 | Band2 | Band3 |
|---|---|---|---|
| Band1 | 584.397 | 559.682 | 502.691 |
| Band2 | 559.682 | 593.753 | 563.119 |
| Band3 | 502.691 | 563.119 | 652.614 |
| Determinant | 3530876.491 | | |

| Cov. Eigenvectors | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.560 | -0.631 | -0.537 |
| Band2 | 0.585 | -0.157 | 0.796 |
| Band3 | 0.586 | 0.760 | -0.281 |

| Inv. of Cov. Ev. | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.560 | 0.585 | 0.586 |
| Band2 | -0.631 | -0.157 | 0.760 |
| Band3 | -0.537 | 0.796 | -0.281 |

**Table B-3. Statistics for dataset: b1938q01007.**

| | Band1 | Band2 | Band3 |
|---|---|---|---|
| Non-Null Cells | 3450304 | 3450304 | 3450304 |
| Area In Hectares | 345.030 | 345.030 | 345.030 |
| Area In Acres | 852.589 | 852.589 | 852.589 |
| Minimum | 0.000 | 0.000 | 1.000 |
| Maximum | 255.000 | 255.000 | 255.000 |
| Mean | 53.526 | 63.882 | 116.166 |
| Median | 51.000 | 66.000 | 118.000 |
| Std. Dev. | 21.643 | 19.493 | 22.415 |
| Std. Dev. (n-1) | 21.643 | 19.493 | 22.415 |
| Corr. Eigenval. | 2.675 | 0.285 | 0.040 |
| Cov. Eigenval. | 1195.869 | 138.674 | 16.292 |

| Correlation Matrix | Band1 | Band2 | Band3 |
|---|---|---|---|
| Band1 | 1.000 | 0.938 | 0.726 |
| Band2 | 0.938 | 1.000 | 0.845 |
| Band3 | 0.726 | 0.845 | 1.000 |
| Determinant | 0.030 | | |

| Corr. Eigenvectors | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.576 | -0.587 | -0.568 |
| Band2 | 0.602 | -0.166 | 0.781 |
| Band3 | 0.553 | 0.792 | -0.258 |

| Inv. of Corr. Ev. | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.576 | 0.602 | 0.553 |
| Band2 | -0.587 | -0.166 | 0.792 |
| Band3 | -0.568 | 0.781 | -0.258 |

| Covariance Matrix | Band1 | Band2 | Band3 |
|---|---|---|---|
| Band1 | 468.399 | 395.617 | 352.112 |
| Band2 | 395.617 | 379.985 | 369.028 |
| Band3 | 352.112 | 369.028 | 502.451 |
| Determinant | 2701876.932 | | |

| Cov. Eigenvectors | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.587 | -0.612 | -0.531 |
| Band2 | 0.552 | -0.177 | 0.815 |
| Band3 | 0.592 | 0.771 | -0.234 |

| Inv. of Cov. Ev. | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.587 | 0.552 | 0.592 |
| Band2 | -0.612 | -0.177 | 0.771 |
| Band3 | -0.531 | 0.815 | -0.234 |

**Table B-4. Statistics for dataset: b1938q03003**

| | Band1 | Band2 | Band3 |
|---|---|---|---|
| Non-Null Cells | 3450304 | 3450304 | 3450304 |
| Area In Hectares | 345.030 | 345.030 | 345.030 |
| Area In Acres | 852.589 | 852.589 | 852.589 |
| Minimum | 0.000 | 0.000 | 2.000 |
| Maximum | 255.000 | 255.000 | 255.000 |
| Mean | 59.884 | 72.294 | 123.139 |
| Median | 58.000 | 75.000 | 126.000 |
| Std. Dev. | 21.644 | 19.527 | 20.975 |
| Std. Dev. (n-1) | 21.644 | 19.527 | 20.975 |
| Corr. Eigenval. | 2.704 | 0.259 | 0.037 |
| Cov. Eigenval. | 1158.551 | 116.063 | 15.125 |

| Correlation Matrix | Band1 | Band2 | Band3 |
|---|---|---|---|
| Band1 | 1.000 | 0.942 | 0.751 |
| Band2 | 0.942 | 1.000 | 0.860 |
| Band3 | 0.751 | 0.860 | 1.000 |
| Determinant | 0.026 | | |

| Corr. Eigenvectors | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.576 | -0.592 | -0.564 |
| Band2 | 0.599 | -0.163 | 0.784 |
| Band3 | 0.556 | 0.789 | -0.261 |

| Inv. of Corr. Ev. | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.576 | 0.599 | 0.556 |
| Band2 | -0.592 | -0.163 | 0.789 |
| Band3 | -0.564 | 0.784 | -0.261 |

| Covariance Matrix | Band1 | Band2 | Band3 |
|---|---|---|---|
| Band1 | 468.479 | 398.110 | 340.889 |
| Band2 | 398.110 | 381.315 | 352.222 |
| Band3 | 340.889 | 352.222 | 439.945 |
| Determinant | 2033761.122 | | |

| Cov. Eigenvectors | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.604 | -0.602 | -0.522 |
| Band2 | 0.564 | -0.140 | 0.814 |
| Band3 | 0.563 | 0.786 | -0.255 |

| Inv. of Cov. Ev. | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.604 | 0.564 | 0.563 |
| Band2 | -0.602 | -0.140 | 0.786 |
| Band3 | -0.522 | 0.814 | -0.255 |

**Table B-5. Statistics for dataset: b1938r03009.**

| | Band1 | Band2 | Band3 |
|---|---|---|---|
| Non-Null Cells | 3450304 | 3450304 | 3450304 |
| Area In Hectares | 345.030 | 345.030 | 345.030 |
| Area In Acres | 852.589 | 852.589 | 852.589 |
| Minimum | 0.000 | 0.000 | 1.000 |
| Maximum | 255.000 | 255.000 | 255.000 |
| Mean | 89.903 | 93.657 | 147.358 |
| Median | 90.000 | 100.000 | 151.000 |
| Std. Dev. | 40.499 | 31.336 | 28.334 |
| Std. Dev. (n-1) | 40.499 | 31.336 | 28.334 |
| Corr. Eigenval. | 2.718 | 0.231 | 0.050 |
| Cov. Eigenval. | 3127.085 | 244.312 | 53.476 |

| Correlation Matrix | Band1 | Band2 | Band3 |
|---|---|---|---|
| Band1 | 1.000 | 0.925 | 0.773 |
| Band2 | 0.925 | 1.000 | 0.878 |
| Band3 | 0.773 | 0.878 | 1.000 |
| Determinant | 0.032 | | |

| Corr. Eigenvectors | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.574 | -0.631 | -0.523 |
| Band2 | 0.596 | -0.116 | 0.795 |
| Band3 | 0.562 | 0.767 | -0.309 |

| Inv. of Corr. Ev. | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.574 | 0.596 | 0.562 |
| Band2 | -0.631 | -0.116 | 0.767 |
| Band3 | -0.523 | 0.795 | -0.309 |

| Covariance Matrix | Band1 | Band2 | Band3 |
|---|---|---|---|
| Band1 | 1640.134 | 1174.042 | 886.465 |
| Band2 | 1174.042 | 981.942 | 779.129 |
| Band3 | 886.465 | 779.129 | 802.797 |
| Determinant | 40855237.943 | | |

| Cov. Eigenvectors | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.703 | -0.595 | -0.389 |
| Band2 | 0.549 | 0.107 | 0.829 |
| Band3 | 0.452 | 0.796 | -0.402 |

| Inv. of Cov. Ev. | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.703 | 0.549 | 0.452 |
| Band2 | -0.595 | 0.107 | 0.796 |
| Band3 | -0.389 | 0.829 | -0.402 |

**Table B-6. Statistics for dataset: b1938q04002.**

| | Band1 | Band2 | Band3 |
|---|---|---|---|
| Non-Null Cells | 3450304 | 3450304 | 3450304 |
| Area In Hectares | 345.030 | 345.030 | 345.030 |
| Area In Acres | 852.589 | 852.589 | 852.589 |
| Minimum | 0.000 | 0.000 | 0.000 |
| Maximum | 255.000 | 255.000 | 255.000 |
| Mean | 76.206 | 72.820 | 114.477 |
| Median | 73.000 | 74.000 | 115.000 |
| Std. Dev. | 33.256 | 25.250 | 25.125 |
| Std. Dev. (n-1) | 33.256 | 25.250 | 25.125 |
| Corr. Eigenval. | 2.782 | 0.180 | 0.038 |
| Cov. Eigenval. | 2204.886 | 143.312 | 26.571 |

| Correlation Matrix | Band1 | Band2 | Band3 |
|---|---|---|---|
| Band1 | 1.000 | 0.939 | 0.822 |
| Band2 | 0.939 | 1.000 | 0.911 |
| Band3 | 0.822 | 0.911 | 1.000 |
| Determinant | 0.019 | | |

| Corr. Eigenvectors | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.573 | -0.653 | -0.495 |
| Band2 | 0.592 | -0.088 | 0.801 |
| Band3 | 0.567 | 0.752 | -0.336 |

| Inv. of Corr. Ev. | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.573 | 0.592 | 0.567 |
| Band2 | -0.653 | -0.088 | 0.752 |
| Band3 | -0.495 | 0.801 | -0.336 |

| Covariance Matrix | Band1 | Band2 | Band3 |
|---|---|---|---|
| Band1 | 1105.964 | 788.654 | 686.489 |
| Band2 | 788.654 | 637.543 | 578.003 |
| Band3 | 686.489 | 578.003 | 631.261 |
| Determinant | 8395942.004 | | |

| Cov. Eigenvectors | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.689 | -0.621 | -0.373 |
| Band2 | 0.529 | 0.080 | 0.845 |
| Band3 | 0.495 | 0.779 | -0.384 |

| Inv. of Cov. Ev. | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.689 | 0.529 | 0.495 |
| Band2 | -0.621 | 0.080 | 0.779 |
| Band3 | -0.373 | 0.845 | -0.384 |

**Table B-7. Statistics for dataset: b2138r04001.**

| | Band1 | Band2 | Band3 |
|---|---|---|---|
| Non-Null Cells | 3450304 | 3450304 | 3450304 |
| Area In Hectares | 345.030 | 345.030 | 345.030 |
| Area In Acres | 852.589 | 852.589 | 852.589 |
| Minimum | 0.000 | 0.000 | 4.000 |
| Maximum | 255.000 | 255.000 | 255.000 |
| Mean | 77.192 | 80.621 | 131.057 |
| Median | 70.000 | 79.000 | 130.000 |
| Std. Dev. | 38.203 | 30.206 | 26.262 |
| Std. Dev. (n-1) | 38.203 | 30.206 | 26.262 |
| Corr. Eigenval. | 2.776 | 0.190 | 0.034 |
| Cov. Eigenval. | 2861.560 | 164.473 | 35.472 |

| Correlation Matrix | Band1 | Band2 | Band3 |
|---|---|---|---|
| Band1 | 1.000 | 0.957 | 0.822 |
| Band2 | 0.957 | 1.000 | 0.883 |
| Band3 | 0.822 | 0.883 | 1.000 |
| Determinant | 0.018 | | |

| Corr. Eigenvectors | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.579 | -0.555 | -0.598 |
| Band2 | 0.591 | -0.220 | 0.776 |
| Band3 | 0.562 | 0.803 | -0.201 |

| Inv. of Corr. Ev. | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.579 | 0.591 | 0.562 |
| Band2 | -0.555 | -0.220 | 0.803 |
| Band3 | -0.598 | 0.776 | -0.201 |

| Covariance Matrix | Band1 | Band2 | Band3 |
|---|---|---|---|
| Band1 | 1459.435 | 1103.847 | 824.718 |
| Band2 | 1103.847 | 912.393 | 700.533 |
| Band3 | 824.718 | 700.533 | 689.677 |
| Determinant | 16694954.671 | | |

| Cov. Eigenvectors | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.701 | -0.531 | -0.477 |
| Band2 | 0.557 | -0.011 | 0.830 |
| Band3 | 0.446 | 0.848 | -0.288 |

| Inv. of Cov. Ev. | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.701 | 0.557 | 0.446 |
| Band2 | -0.531 | -0.011 | 0.848 |
| Band3 | -0.477 | 0.830 | -0.288 |

**Table B-8. Statistics for dataset: b2138r05001.**

| | Band1 | Band2 | Band3 |
|---|---|---|---|
| Non-Null Cells | 3450304 | 3450304 | 3450304 |
| Area In Hectares | 345.030 | 345.030 | 345.030 |
| Area In Acres | 852.589 | 852.589 | 852.589 |
| | | | |
| Minimum | 0.000 | 0.000 | 0.000 |
| Maximum | 255.000 | 255.000 | 255.000 |
| Mean | 74.402 | 66.300 | 103.922 |
| Median | 69.000 | 65.000 | 102.000 |
| Std. Dev. | 37.789 | 27.753 | 27.149 |
| Std. Dev. (n-1) | 37.789 | 27.753 | 27.149 |
| Corr. Eigenval. | 2.786 | 0.177 | 0.037 |
| Cov. Eigenval. | 2728.587 | 175.986 | 30.727 |
| Correlation Matrix | Band1 | Band2 | Band3 |
| Band1 | 1.000 | 0.931 | 0.823 |
| Band2 | 0.931 | 1.000 | 0.924 |
| Band3 | 0.823 | 0.924 | 1.000 |
| Determinant | 0.018 | | |
| | | | |
| Corr. Eigenvectors | PC1 | PC2 | PC3 |
| Band1 | 0.571 | -0.694 | -0.438 |
| Band2 | 0.592 | -0.022 | 0.806 |
| Band3 | 0.569 | 0.719 | -0.398 |
| | | | |
| Inv. of Corr. Ev. | PC1 | PC2 | PC3 |
| Band1 | 0.571 | 0.592 | 0.569 |
| Band2 | -0.694 | -0.022 | 0.719 |
| Band3 | -0.438 | 0.806 | -0.398 |
| | | | |
| Covariance Matrix | Band1 | Band2 | Band3 |
| Band1 | 1428.025 | 976.198 | 844.471 |
| Band2 | 976.198 | 770.229 | 696.122 |
| Band3 | 844.471 | 696.122 | 737.046 |
| Determinant | 14754727.148 | | |
| | | | |
| Cov. Eigenvectors | PC1 | PC2 | PC3 |
| Band1 | 0.704 | -0.639 | -0.310 |
| Band2 | 0.522 | 0.170 | 0.836 |
| Band3 | 0.481 | 0.751 | -0.453 |
| | | | |
| Inv. of Cov. Ev. | PC1 | PC2 | PC3 |
| Band1 | 0.704 | 0.522 | 0.481 |
| Band2 | -0.639 | 0.170 | 0.751 |
| Band3 | -0.310 | 0.836 | -0.453 |

**Table B-9. Statistics for dataset: b2539q02001.**

| | Band1 | Band2 | Band3 |
|---|---|---|---|
| Non-Null Cells | 3486800 | 3486800 | 3486800 |
| Area In Hectares | 348.680 | 348.680 | 348.680 |
| Area In Acres | 861.607 | 861.607 | 861.607 |
| | | | |
| Minimum | 0.000 | 3.000 | 0.000 |
| Maximum | 255.000 | 255.000 | 255.000 |
| Mean | 86.441 | 85.687 | 100.793 |
| Median | 85.000 | 85.000 | 99.000 |
| Std. Dev. | 36.987 | 27.729 | 27.002 |
| Std. Dev. (n-1) | 36.987 | 27.729 | 27.002 |
| Corr. Eigenval. | 2.787 | 0.186 | 0.028 |
| Cov. Eigenval. | 2663.671 | 179.428 | 22.981 |

| Correlation Matrix | Band1 | Band2 | Band3 |
|---|---|---|---|
| Band1 | 1.000 | 0.937 | 0.815 |
| Band2 | 0.937 | 1.000 | 0.927 |
| Band3 | 0.815 | 0.927 | 1.000 |
| Determinant | 0.014 | | |

| Corr. Eigenvectors | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.570 | -0.692 | -0.443 |
| Band2 | 0.594 | -0.026 | 0.804 |
| Band3 | 0.568 | 0.722 | -0.396 |

| Inv. of Corr. Ev. | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.570 | 0.594 | 0.568 |
| Band2 | -0.692 | -0.026 | 0.722 |
| Band3 | -0.443 | 0.804 | -0.396 |

| Covariance Matrix | Band1 | Band2 | Band3 |
|---|---|---|---|
| Band1 | 1368.051 | 960.555 | 813.628 |
| Band2 | 960.555 | 768.915 | 694.316 |
| Band3 | 813.628 | 694.316 | 729.115 |
| Determinant | 10983625.105 | | |

| Cov. Eigenvectors | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.697 | -0.638 | -0.329 |
| Band2 | 0.530 | 0.149 | 0.835 |
| Band3 | 0.483 | 0.756 | -0.442 |

| Inv. of Cov. Ev. | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.697 | 0.530 | 0.483 |
| Band2 | -0.638 | 0.149 | 0.756 |
| Band3 | -0.329 | 0.835 | -0.442 |

**Table B-10. Statistics for dataset: b2539q02004.**

| | Band1 | Band2 | Band3 |
|---|---|---|---|
| Non-Null Cells | 3486800 | 3486800 | 3486800 |
| Area In Hectares | 348.680 | 348.680 | 348.680 |
| Area In Acres | 861.607 | 861.607 | 861.607 |
| | | | |
| Minimum | 2.000 | 7.000 | 0.000 |
| Maximum | 255.000 | 255.000 | 255.000 |
| Mean | 86.673 | 91.100 | 110.924 |
| Median | 85.000 | 92.000 | 110.000 |
| Std. Dev. | 39.683 | 28.950 | 27.034 |
| Std. Dev. (n-1) | 39.683 | 28.950 | 27.034 |
| Corr. Eigenval. | 2.780 | 0.186 | 0.034 |
| Cov. Eigenval. | 2928.485 | 183.679 | 31.575 |

| Correlation Matrix | Band1 | Band2 | Band3 |
|---|---|---|---|
| Band1 | 1.000 | 0.943 | 0.817 |
| Band2 | 0.943 | 1.000 | 0.910 |
| Band3 | 0.817 | 0.910 | 1.000 |
| Determinant | 0.018 | | |

| Corr. Eigenvectors | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.573 | -0.646 | -0.504 |
| Band2 | 0.593 | -0.097 | 0.800 |
| Band3 | 0.566 | 0.757 | -0.327 |

| Inv. of Corr. Ev. | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.573 | 0.593 | 0.566 |
| Band2 | -0.646 | -0.097 | 0.757 |
| Band3 | -0.504 | 0.800 | -0.327 |

| Covariance Matrix | Band1 | Band2 | Band3 |
|---|---|---|---|
| Band1 | 1574.756 | 1082.828 | 876.220 |
| Band2 | 1082.828 | 838.126 | 711.948 |
| Band3 | 876.220 | 711.948 | 730.857 |
| Determinant | 16984057.936 | | |

| Cov. Eigenvectors | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.717 | -0.597 | -0.360 |
| Band2 | 0.527 | 0.126 | 0.841 |
| Band3 | 0.457 | 0.792 | -0.405 |

| Inv. of Cov. Ev. | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.717 | 0.527 | 0.457 |
| Band2 | -0.597 | 0.126 | 0.792 |
| Band3 | -0.360 | 0.841 | -0.405 |

**Table B-11. Statistics for dataset: b2539q04001.**

| | Band1 | Band2 | Band3 |
|---|---|---|---|
| Non-Null Cells | 3486800 | 3486800 | 3486800 |
| Area In Hectares | 348.680 | 348.680 | 348.680 |
| Area In Acres | 861.607 | 861.607 | 861.607 |
| Minimum | 1.000 | 5.000 | 0.000 |
| Maximum | 255.000 | 255.000 | 255.000 |
| Mean | 83.606 | 81.979 | 91.961 |
| Median | 81.000 | 81.000 | 91.000 |
| Std. Dev. | 35.285 | 26.156 | 25.238 |
| Std. Dev. (n-1) | 35.285 | 26.156 | 25.238 |
| Corr. Eigenval. | 2.789 | 0.180 | 0.031 |
| Cov. Eigenval. | 2389.388 | 153.922 | 22.798 |

| Correlation Matrix | Band1 | Band2 | Band3 |
|---|---|---|---|
| Band1 | 1.000 | 0.938 | 0.820 |
| Band2 | 0.938 | 1.000 | 0.924 |
| Band3 | 0.820 | 0.924 | 1.000 |
| Determinant | 0.015 | | |

| Corr. Eigenvectors | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.571 | -0.682 | -0.456 |
| Band2 | 0.593 | -0.042 | 0.804 |
| Band3 | 0.568 | 0.730 | -0.381 |

| Inv. of Corr. Ev. | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.571 | 0.593 | 0.568 |
| Band2 | -0.682 | -0.042 | 0.730 |
| Band3 | -0.456 | 0.804 | -0.381 |

| Covariance Matrix | Band1 | Band2 | Band3 |
|---|---|---|---|
| Band1 | 1245.041 | 865.625 | 730.404 |
| Band2 | 865.625 | 684.124 | 609.895 |
| Band3 | 730.404 | 609.895 | 636.942 |
| Determinant | 8384476.453 | | |

| Cov. Eigenvectors | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.703 | -0.629 | -0.332 |
| Band2 | 0.527 | 0.148 | 0.837 |
| Band3 | 0.477 | 0.763 | -0.436 |

| Inv. of Cov. Ev. | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.703 | 0.527 | 0.477 |
| Band2 | -0.629 | 0.148 | 0.763 |
| Band3 | -0.332 | 0.837 | -0.436 |

**Table B-12. Statistics for dataset: b2539r01005.**

| | Band1 | Band2 | Band3 |
|---|---|---|---|
| Non-Null Cells | 3486800 | 3486800 | 3486800 |
| Area In Hectares | 348.680 | 348.680 | 348.680 |
| Area In Acres | 861.607 | 861.607 | 861.607 |
| | | | |
| Minimum | 1.000 | 2.000 | 0.000 |
| Maximum | 255.000 | 255.000 | 255.000 |
| Mean | 76.282 | 78.549 | 93.968 |
| Median | 69.000 | 79.000 | 93.000 |
| Std. Dev. | 42.090 | 31.537 | 29.516 |
| Std. Dev. (n-1) | 42.090 | 31.537 | 29.516 |
| Corr. Eigenval. | 2.821 | 0.153 | 0.026 |
| Cov. Eigenval. | 3428.700 | 180.572 | 28.039 |

| Correlation Matrix | Band1 | Band2 | Band3 |
|---|---|---|---|
| Band1 | 1.000 | 0.949 | 0.848 |
| Band2 | 0.949 | 1.000 | 0.934 |
| Band3 | 0.848 | 0.934 | 1.000 |
| Determinant | 0.011 | | |

| Corr. Eigenvectors | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.572 | -0.677 | -0.463 |
| Band2 | 0.590 | -0.052 | 0.806 |
| Band3 | 0.569 | 0.735 | -0.369 |

| Inv. of Corr. Ev. | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.572 | 0.590 | 0.569 |
| Band2 | -0.677 | -0.052 | 0.735 |
| Band3 | -0.463 | 0.806 | -0.369 |

| Covariance Matrix | Band1 | Band2 | Band3 |
|---|---|---|---|
| Band1 | 1771.534 | 1259.265 | 1053.458 |
| Band2 | 1259.265 | 994.597 | 869.432 |
| Band3 | 1053.458 | 869.432 | 871.181 |
| Determinant | 17359718.512 | | |

| Cov. Eigenvectors | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.704 | -0.625 | -0.337 |
| Band2 | 0.532 | 0.149 | 0.833 |
| Band3 | 0.471 | 0.766 | -0.438 |

| Inv. of Cov. Ev. | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.704 | 0.532 | 0.471 |
| Band2 | -0.625 | 0.149 | 0.766 |
| Band3 | -0.337 | 0.833 | -0.438 |

**Table B-13. Statistics for dataset: b2539r01008.**

| | Band1 | Band2 | Band3 |
|---|---|---|---|
| Non-Null Cells | 3486800 | 3486800 | 3486800 |
| Area In Hectares | 348.680 | 348.680 | 348.680 |
| Area In Acres | 861.607 | 861.607 | 861.607 |
| Minimum | 3.000 | 6.000 | 0.000 |
| Maximum | 255.000 | 255.000 | 255.000 |
| Mean | 80.860 | 80.906 | 90.602 |
| Median | 78.000 | 78.000 | 88.000 |
| Std. Dev. | 37.232 | 27.813 | 26.950 |
| Std. Dev. (n-1) | 37.232 | 27.813 | 26.950 |
| Corr. Eigenval. | 2.796 | 0.174 | 0.030 |
| Cov. Eigenval. | 2691.200 | 170.203 | 24.651 |

| Correlation Matrix | Band1 | Band2 | Band3 |
|---|---|---|---|
| Band1 | 1.000 | 0.935 | 0.826 |
| Band2 | 0.935 | 1.000 | 0.932 |
| Band3 | 0.826 | 0.932 | 1.000 |
| Determinant | 0.014 | | |

| Corr. Eigenvectors | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.570 | -0.702 | -0.427 |
| Band2 | 0.592 | -0.009 | 0.806 |
| Band3 | 0.569 | 0.712 | -0.411 |

| Inv. of Corr. Ev. | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.570 | 0.592 | 0.569 |
| Band2 | -0.702 | -0.009 | 0.712 |
| Band3 | -0.427 | 0.806 | -0.411 |

| Covariance Matrix | Band1 | Band2 | Band3 |
|---|---|---|---|
| Band1 | 1386.209 | 968.084 | 828.903 |
| Band2 | 968.084 | 773.568 | 698.620 |
| Band3 | 828.903 | 698.620 | 726.277 |
| Determinant | 11291207.246 | | |

| Cov. Eigenvectors | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.699 | -0.645 | -0.310 |
| Band2 | 0.528 | 0.172 | 0.831 |
| Band3 | 0.483 | 0.745 | -0.461 |

| Inv. of Cov. Ev. | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.699 | 0.528 | 0.483 |
| Band2 | -0.645 | 0.172 | 0.745 |
| Band3 | -0.310 | 0.831 | -0.461 |

**Table B-14. Statistics for dataset: b2539r01010.**

| | Band1 | Band2 | Band3 |
|---|---|---|---|
| Non-Null Cells | 3486800 | 3486800 | 3486800 |
| Area In Hectares | 348.680 | 348.680 | 348.680 |
| Area In Acres | 861.607 | 861.607 | 861.607 |
| | | | |
| Minimum | 3.000 | 7.000 | 0.000 |
| Maximum | 255.000 | 255.000 | 255.000 |
| Mean | 77.843 | 77.531 | 85.024 |
| Median | 75.000 | 76.000 | 82.000 |
| Std. Dev. | 36.095 | 27.460 | 26.430 |
| Std. Dev. (n-1) | 36.095 | 27.460 | 26.430 |
| Corr. Eigenval. | 2.794 | 0.179 | 0.027 |
| Cov. Eigenval. | 2565.478 | 168.731 | 21.216 |

| Correlation Matrix | Band1 | Band2 | Band3 |
|---|---|---|---|
| Band1 | 1.000 | 0.931 | 0.821 |
| Band2 | 0.931 | 1.000 | 0.938 |
| Band3 | 0.821 | 0.938 | 1.000 |
| Determinant | 0.013 | | |

| Corr. Eigenvectors | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.569 | -0.718 | -0.401 |
| Band2 | 0.593 | 0.020 | 0.805 |
| Band3 | 0.570 | 0.696 | -0.437 |

| Inv. of Corr. Ev. | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.569 | 0.593 | 0.570 |
| Band2 | -0.718 | 0.020 | 0.696 |
| Band3 | -0.401 | 0.805 | -0.437 |

| Covariance Matrix | Band1 | Band2 | Band3 |
|---|---|---|---|
| Band1 | 1302.826 | 922.778 | 783.344 |
| Band2 | 922.778 | 754.073 | 680.583 |
| Band3 | 783.344 | 680.583 | 698.527 |
| Determinant | 9184002.058 | | |

| Cov. Eigenvectors | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.692 | -0.658 | -0.297 |
| Band2 | 0.535 | 0.190 | 0.823 |
| Band3 | 0.485 | 0.729 | -0.483 |

| Inv. of Cov. Ev. | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.692 | 0.535 | 0.485 |
| Band2 | -0.658 | 0.190 | 0.729 |
| Band3 | -0.297 | 0.823 | -0.483 |

**Table B-15. Statistics for dataset: b2539r02008.**

| | Band1 | Band2 | Band3 |
|---|---|---|---|
| Non-Null Cells | 3486800 | 3486800 | 3486800 |
| Area In Hectares | 348.680 | 348.680 | 348.680 |
| Area In Acres | 861.607 | 861.607 | 861.607 |
| | | | |
| Minimum | 3.000 | 2.000 | 0.000 |
| Maximum | 255.000 | 255.000 | 255.000 |
| Mean | 84.176 | 85.159 | 99.971 |
| Median | 80.000 | 85.000 | 99.000 |
| Std. Dev. | 39.037 | 28.189 | 26.836 |
| Std. Dev. (n-1) | 39.037 | 28.189 | 26.836 |
| Corr. Eigenval. | 2.778 | 0.189 | 0.033 |
| Cov. Eigenval. | 2822.459 | 187.936 | 28.329 |

| Correlation Matrix | Band1 | Band2 | Band3 |
|---|---|---|---|
| Band1 | 1.000 | 0.936 | 0.811 |
| Band2 | 0.936 | 1.000 | 0.918 |
| Band3 | 0.811 | 0.918 | 1.000 |
| Determinant | 0.017 | | |

| Corr. Eigenvectors | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.571 | -0.678 | -0.463 |
| Band2 | 0.593 | -0.049 | 0.803 |
| Band3 | 0.567 | 0.734 | -0.374 |

| Inv. of Corr. Ev. | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.571 | 0.593 | 0.567 |
| Band2 | -0.678 | -0.049 | 0.734 |
| Band3 | -0.463 | 0.803 | -0.374 |

| Covariance Matrix | Band1 | Band2 | Band3 |
|---|---|---|---|
| Band1 | 1523.906 | 1029.591 | 849.903 |
| Band2 | 1029.591 | 794.640 | 694.699 |
| Band3 | 849.903 | 694.699 | 720.178 |
| Determinant | 15026990.126 | | |

| Cov. Eigenvectors | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.717 | -0.616 | -0.327 |
| Band2 | 0.522 | 0.164 | 0.837 |
| Band3 | 0.462 | 0.770 | -0.439 |

| Inv. of Cov. Ev. | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.717 | 0.522 | 0.462 |
| Band2 | -0.616 | 0.164 | 0.770 |
| Band3 | -0.327 | 0.837 | -0.439 |

**Table B-16. Statistics for dataset: b2539r02013.**

| | Band1 | Band2 | Band3 |
|---|---|---|---|
| Non-Null Cells | 3486800 | 3486800 | 3486800 |
| Area In Hectares | 348.680 | 348.680 | 348.680 |
| Area In Acres | 861.607 | 861.607 | 861.607 |
| | | | |
| Minimum | 2.000 | 2.000 | 0.000 |
| Maximum | 255.000 | 255.000 | 255.000 |
| Mean | 86.615 | 85.022 | 97.403 |
| Median | 84.000 | 82.000 | 95.000 |
| Std. Dev. | 34.858 | 26.029 | 27.462 |
| Std. Dev. (n-1) | 34.858 | 26.029 | 27.462 |
| Corr. Eigenval. | 2.760 | 0.204 | 0.035 |
| Cov. Eigenval. | 2427.131 | 193.275 | 26.326 |

| Correlation Matrix | Band1 | Band2 | Band3 |
|---|---|---|---|
| Band1 | 1.000 | 0.921 | 0.796 |
| Band2 | 0.921 | 1.000 | 0.922 |
| Band3 | 0.796 | 0.922 | 1.000 |
| Determinant | 0.020 | | |

| Corr. Eigenvectors | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.568 | -0.710 | -0.416 |
| Band2 | 0.595 | 0.005 | 0.804 |
| Band3 | 0.569 | 0.704 | -0.425 |

| Inv. of Corr. Ev. | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.568 | 0.595 | 0.569 |
| Band2 | -0.710 | 0.005 | 0.704 |
| Band3 | -0.416 | 0.804 | -0.425 |

| Covariance Matrix | Band1 | Band2 | Band3 |
|---|---|---|---|
| Band1 | 1215.073 | 835.245 | 761.623 |
| Band2 | 835.245 | 677.497 | 659.334 |
| Band3 | 761.623 | 659.334 | 754.162 |
| Determinant | 12349424.809 | | |

| Cov. Eigenvectors | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.682 | -0.661 | -0.313 |
| Band2 | 0.520 | 0.137 | 0.843 |
| Band3 | 0.515 | 0.738 | -0.437 |

| Inv. of Cov. Ev. | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.682 | 0.520 | 0.515 |
| Band2 | -0.661 | 0.137 | 0.738 |
| Band3 | -0.313 | 0.843 | -0.437 |

**Table B-17. Statistics for dataset: b2539r02016.**

| | Band1 | Band2 | Band3 |
|---|---|---|---|
| Non-Null Cells | 3486800 | 3486800 | 3486800 |
| Area In Hectares | 348.680 | 348.680 | 348.680 |
| Area In Acres | 861.607 | 861.607 | 861.607 |
| | | | |
| Minimum | 2.000 | 5.000 | 0.000 |
| Maximum | 255.000 | 255.000 | 255.000 |
| Mean | 74.995 | 75.441 | 84.964 |
| Median | 69.000 | 73.000 | 83.000 |
| Std. Dev. | 34.033 | 25.634 | 25.356 |
| Std. Dev. (n-1) | 34.033 | 25.634 | 25.356 |
| Corr. Eigenval. | 2.774 | 0.195 | 0.031 |
| Cov. Eigenval. | 2271.747 | 164.411 | 22.087 |

| Correlation Matrix | Band1 | Band2 | Band3 |
|---|---|---|---|
| Band1 | 1.000 | 0.929 | 0.805 |
| Band2 | 0.929 | 1.000 | 0.925 |
| Band3 | 0.805 | 0.925 | 1.000 |
| Determinant | 0.017 | | |

| Corr. Eigenvectors | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.569 | -0.700 | -0.432 |
| Band2 | 0.594 | -0.013 | 0.804 |
| Band3 | 0.568 | 0.714 | -0.409 |

| Inv. of Corr. Ev. | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.569 | 0.594 | 0.568 |
| Band2 | -0.700 | -0.013 | 0.714 |
| Band3 | -0.432 | 0.804 | -0.409 |

| Covariance Matrix | Band1 | Band2 | Band3 |
|---|---|---|---|
| Band1 | 1158.213 | 810.692 | 694.620 |
| Band2 | 810.692 | 657.103 | 600.964 |
| Band3 | 694.620 | 600.964 | 642.930 |
| Determinant | 8249695.297 | | |

| Cov. Eigenvectors | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.692 | -0.647 | -0.321 |
| Band2 | 0.530 | 0.152 | 0.834 |
| Band3 | 0.491 | 0.747 | -0.448 |

| Inv. of Cov. Ev. | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.692 | 0.530 | 0.491 |
| Band2 | -0.647 | 0.152 | 0.747 |
| Band3 | -0.321 | 0.834 | -0.448 |

**Table B-18. Statistics for dataset: b3541r06002.**

| | Band1 | Band2 | Band3 |
|---|---|---|---|
| Non-Null Cells | 3450304 | 3450304 | 3450304 |
| Area In Hectares | 345.030 | 345.030 | 345.030 |
| Area In Acres | 852.589 | 852.589 | 852.589 |
| | | | |
| Minimum | 0.000 | 0.000 | 2.000 |
| Maximum | 255.000 | 255.000 | 255.000 |
| Mean | 68.700 | 82.595 | 131.349 |
| Median | 59.000 | 78.000 | 129.000 |
| Std. Dev. | 36.955 | 32.681 | 28.971 |
| Std. Dev. (n-1) | 36.955 | 32.681 | 28.971 |
| Corr. Eigenval. | 2.823 | 0.150 | 0.027 |
| Cov. Eigenval. | 3092.968 | 149.555 | 30.572 |

| Correlation Matrix | Band1 | Band2 | Band3 |
|---|---|---|---|
| Band1 | 1.000 | 0.961 | 0.856 |
| Band2 | 0.961 | 1.000 | 0.917 |
| Band3 | 0.856 | 0.917 | 1.000 |
| Determinant | 0.012 | | |

| Corr. Eigenvectors | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.576 | -0.599 | -0.556 |
| Band2 | 0.589 | -0.168 | 0.791 |
| Band3 | 0.567 | 0.783 | -0.256 |

| Inv. of Corr. Ev. | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.576 | 0.589 | 0.567 |
| Band2 | -0.599 | -0.168 | 0.783 |
| Band3 | -0.556 | 0.791 | -0.256 |

| Covariance Matrix | Band1 | Band2 | Band3 |
|---|---|---|---|
| Band1 | 1365.687 | 1160.749 | 916.049 |
| Band2 | 1160.749 | 1068.061 | 867.879 |
| Band3 | 916.049 | 867.879 | 839.347 |
| Determinant | 14141686.794 | | |

| Cov. Eigenvectors | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.650 | -0.584 | -0.485 |
| Band2 | 0.582 | -0.028 | 0.813 |
| Band3 | 0.488 | 0.811 | -0.322 |

| Inv. of Cov. Ev. | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.650 | 0.582 | 0.488 |
| Band2 | -0.584 | -0.028 | 0.811 |
| Band3 | -0.485 | 0.813 | -0.322 |

**Table B-19. Statistics for dataset: b3542r01006.**

| | Band1 | Band2 | Band3 |
|---|---|---|---|
| Non-Null Cells | 3450304 | 3450304 | 3450304 |
| Area In Hectares | 345.030 | 345.030 | 345.030 |
| Area In Acres | 852.589 | 852.589 | 852.589 |
| | | | |
| Minimum | 0.000 | 0.000 | 1.000 |
| Maximum | 255.000 | 255.000 | 255.000 |
| Mean | 73.343 | 78.593 | 118.979 |
| Median | 71.000 | 79.000 | 119.000 |
| Std. Dev. | 34.735 | 29.471 | 27.234 |
| Std. Dev. (n-1) | 34.735 | 29.471 | 27.234 |
| Corr. Eigenval. | 2.821 | 0.151 | 0.027 |
| Cov. Eigenval. | 2657.796 | 133.438 | 25.493 |

| Correlation Matrix | Band1 | Band2 | Band3 |
|---|---|---|---|
| Band1 | 1.000 | 0.961 | 0.854 |
| Band2 | 0.961 | 1.000 | 0.915 |
| Band3 | 0.854 | 0.915 | 1.000 |
| Determinant | 0.012 | | |

| Corr. Eigenvectors | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.576 | -0.596 | -0.559 |
| Band2 | 0.589 | -0.171 | 0.790 |
| Band3 | 0.566 | 0.784 | -0.253 |

| Inv. of Corr. Ev. | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.576 | 0.589 | 0.566 |
| Band2 | -0.596 | -0.171 | 0.784 |
| Band3 | -0.559 | 0.790 | -0.253 |

| Covariance Matrix | Band1 | Band2 | Band3 |
|---|---|---|---|
| Band1 | 1206.515 | 984.066 | 808.049 |
| Band2 | 984.066 | 868.536 | 734.697 |
| Band3 | 808.049 | 734.697 | 741.675 |
| Determinant | 9041141.030 | | |

| Cov. Eigenvectors | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.659 | -0.583 | -0.475 |
| Band2 | 0.566 | -0.031 | 0.824 |
| Band3 | 0.495 | 0.812 | -0.310 |

| Inv. of Cov. Ev. | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.659 | 0.566 | 0.495 |
| Band2 | -0.583 | -0.031 | 0.812 |
| Band3 | -0.475 | 0.824 | -0.310 |

**Table B-20. Statistics for dataset: b3542r02002.**

| | Band1 | Band2 | Band3 |
|---|---|---|---|
| Non-Null Cells | 3450304 | 3450304 | 3450304 |
| Area In Hectares | 345.030 | 345.030 | 345.030 |
| Area In Acres | 852.589 | 852.589 | 852.589 |
| | | | |
| Minimum | 0.000 | 0.000 | 2.000 |
| Maximum | 255.000 | 255.000 | 255.000 |
| Mean | 81.804 | 87.248 | 128.194 |
| Median | 76.000 | 86.000 | 129.000 |
| Std. Dev. | 38.150 | 32.393 | 28.446 |
| Std. Dev. (n-1) | 38.150 | 32.393 | 28.446 |
| Corr. Eigenval. | 2.815 | 0.160 | 0.025 |
| Cov. Eigenval. | 3128.660 | 157.219 | 28.022 |

| Correlation Matrix | Band1 | Band2 | Band3 |
|---|---|---|---|
| Band1 | 1.000 | 0.965 | 0.848 |
| Band2 | 0.965 | 1.000 | 0.909 |
| Band3 | 0.848 | 0.909 | 1.000 |
| Determinant | 0.011 | | |

| Corr. Eigenvectors | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.577 | -0.578 | -0.577 |
| Band2 | 0.590 | -0.194 | 0.784 |
| Band3 | 0.565 | 0.793 | -0.229 |

| Inv. of Corr. Ev. | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.577 | 0.590 | 0.565 |
| Band2 | -0.578 | -0.194 | 0.793 |
| Band3 | -0.577 | 0.784 | -0.229 |

| Covariance Matrix | Band1 | Band2 | Band3 |
|---|---|---|---|
| Band1 | 1455.412 | 1192.641 | 920.205 |
| Band2 | 1192.641 | 1049.330 | 837.144 |
| Band3 | 920.205 | 837.144 | 809.160 |
| Determinant | 13783717.363 | | |

| Cov. Eigenvectors | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.669 | -0.557 | -0.492 |
| Band2 | 0.574 | -0.034 | 0.818 |
| Band3 | 0.472 | 0.830 | -0.297 |

| Inv. of Cov. Ev. | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.669 | 0.574 | 0.472 |
| Band2 | -0.557 | -0.034 | 0.830 |
| Band3 | -0.492 | 0.818 | -0.297 |

**Table B-21. Statistics for dataset: B4740r01011.**

| | Band1 | Band2 | Band3 |
|---|---|---|---|
| Non-Null Cells | 3465504 | 3465504 | 3465504 |
| Area In Hectares | 346.550 | 346.550 | 346.550 |
| Area In Acres | 856.345 | 856.345 | 856.345 |
| Minimum | 0.000 | 0.000 | 1.000 |
| Maximum | 255.000 | 255.000 | 255.000 |
| Mean | 67.664 | 73.407 | 128.015 |
| Median | 56.000 | 67.000 | 126.000 |
| Std. Dev. | 33.574 | 27.184 | 28.581 |
| Std. Dev. (n-1) | 33.574 | 27.184 | 28.581 |
| Corr. Eigenval. | 2.767 | 0.214 | 0.019 |
| Cov. Eigenval. | 2473.152 | 194.079 | 15.822 |

| Correlation Matrix | Band1 | Band2 | Band3 |
|---|---|---|---|
| Band1 | 1.000 | 0.967 | 0.798 |
| Band2 | 0.967 | 1.000 | 0.883 |
| Band3 | 0.798 | 0.883 | 1.000 |
| Determinant | 0.011 | | |

| Corr. Eigenvectors | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.578 | -0.570 | -0.584 |
| Band2 | 0.595 | -0.195 | 0.780 |
| Band3 | 0.558 | 0.798 | -0.227 |

| Inv. of Corr. Ev. | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.578 | 0.595 | 0.558 |
| Band2 | -0.570 | -0.195 | 0.798 |
| Band3 | -0.584 | 0.780 | -0.227 |

| Covariance Matrix | Band1 | Band2 | Band3 |
|---|---|---|---|
| Band1 | 1127.217 | 882.544 | 765.352 |
| Band2 | 882.544 | 738.972 | 685.981 |
| Band3 | 765.352 | 685.981 | 816.865 |
| Determinant | 7594232.182 | | |

| Cov. Eigenvectors | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.655 | -0.568 | -0.498 |
| Band2 | 0.542 | -0.107 | 0.834 |
| Band3 | 0.527 | 0.816 | -0.238 |

| Inv. of Cov. Ev. | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.655 | 0.542 | 0.527 |
| Band2 | -0.568 | -0.107 | 0.816 |
| Band3 | -0.498 | 0.834 | -0.238 |

**Table B-22. Statistics for dataset: B4740r06001**

| | Band1 | Band2 | Band3 |
|---|---|---|---|
| Non-Null Cells | 3450304 | 3450304 | 3450304 |
| Area In Hectares | 345.030 | 345.030 | 345.030 |
| Area In Acres | 852.589 | 852.589 | 852.589 |
| Minimum | 0.000 | 0.000 | 3.000 |
| Maximum | 255.000 | 255.000 | 255.000 |
| Mean | 61.952 | 73.204 | 120.798 |
| Median | 57.000 | 72.000 | 125.000 |
| Std. Dev. | 32.453 | 33.101 | 33.395 |
| Std. Dev. (n-1) | 32.453 | 33.101 | 33.395 |
| Corr. Eigenval. | 2.844 | 0.137 | 0.019 |
| Cov. Eigenval. | 3094.085 | 148.987 | 21.083 |

| Correlation Matrix | Band1 | Band2 | Band3 |
|---|---|---|---|
| Band1 | 1.000 | 0.966 | 0.867 |
| Band2 | 0.966 | 1.000 | 0.932 |
| Band3 | 0.867 | 0.932 | 1.000 |
| Determinant | 0.008 | | |

| Corr. Eigenvectors | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.575 | -0.623 | -0.530 |
| Band2 | 0.589 | -0.135 | 0.797 |
| Band3 | 0.568 | 0.771 | -0.289 |

| Inv. of Corr. Ev. | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.575 | 0.589 | 0.568 |
| Band2 | -0.623 | -0.135 | 0.771 |
| Band3 | -0.530 | 0.797 | -0.289 |

| Covariance Matrix | Band1 | Band2 | Band3 |
|---|---|---|---|
| Band1 | 1053.196 | 1038.147 | 939.327 |
| Band2 | 1038.147 | 1095.699 | 1030.111 |
| Band3 | 939.327 | 1030.111 | 1115.259 |
| Determinant | 9718594.913 | | |

| Cov. Eigenvectors | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.565 | -0.623 | -0.540 |
| Band2 | 0.591 | -0.151 | 0.793 |
| Band3 | 0.576 | 0.767 | -0.283 |

| Inv. of Cov. Ev. | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.565 | 0.591 | 0.576 |
| Band2 | -0.623 | -0.151 | 0.767 |
| Band3 | -0.540 | 0.793 | -0.283 |

**Table B-23. Statistics for dataset: b4740s04007**

| | Band1 | Band2 | Band3 |
|---|---|---|---|
| Non-Null Cells | 3450304 | 3450304 | 3450304 |
| Area In Hectares | 345.030 | 345.030 | 345.030 |
| Area In Acres | 852.589 | 852.589 | 852.589 |
| Minimum | 0.000 | 0.000 | 1.000 |
| Maximum | 255.000 | 255.000 | 255.000 |
| Mean | 72.062 | 70.360 | 112.396 |
| Median | 73.000 | 72.000 | 113.000 |
| Std. Dev. | 29.499 | 23.348 | 22.834 |
| Std. Dev. (n-1) | 29.499 | 23.348 | 22.834 |
| Corr. Eigenval. | 2.783 | 0.181 | 0.036 |
| Cov. Eigenval. | 1801.790 | 112.898 | 22.008 |

| Correlation Matrix | Band1 | Band2 | Band3 |
|---|---|---|---|
| Band1 | 1.000 | 0.951 | 0.826 |
| Band2 | 0.951 | 1.000 | 0.896 |
| Band3 | 0.826 | 0.896 | 1.000 |
| Determinant | 0.018 | | |

| Corr. Eigenvectors | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.577 | -0.592 | -0.563 |
| Band2 | 0.591 | -0.173 | 0.788 |
| Band3 | 0.564 | 0.787 | -0.250 |

| Inv. of Corr. Ev. | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.577 | 0.591 | 0.564 |
| Band2 | -0.592 | -0.173 | 0.787 |
| Band3 | -0.563 | 0.788 | -0.250 |

| Covariance Matrix | Band1 | Band2 | Band3 |
|---|---|---|---|
| Band1 | 870.217 | 655.018 | 556.246 |
| Band2 | 655.018 | 545.106 | 477.763 |
| Band3 | 556.246 | 477.763 | 521.373 |
| Determinant | 4476914.868 | | |

| Cov. Eigenvectors | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.678 | -0.580 | -0.451 |
| Band2 | 0.542 | -0.020 | 0.840 |
| Band3 | 0.497 | 0.814 | -0.301 |

| Inv. of Cov. Ev. | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.678 | 0.542 | 0.497 |
| Band2 | -0.580 | -0.020 | 0.814 |
| Band3 | -0.451 | 0.840 | -0.301 |

**Table B-24. Statistics for dataset: b4740s06004**

| | Band1 | Band2 | Band3 |
|---|---|---|---|
| Non-Null Cells | 3450304 | 3450304 | 3450304 |
| Area In Hectares | 345.030 | 345.030 | 345.030 |
| Area In Acres | 852.589 | 852.589 | 852.589 |
| Minimum | 0.000 | 0.000 | 1.000 |
| Maximum | 255.000 | 255.000 | 255.000 |
| Mean | 66.220 | 76.892 | 117.823 |
| Median | 64.000 | 80.000 | 120.000 |
| Std. Dev. | 33.724 | 33.188 | 29.452 |
| Std. Dev. (n-1) | 33.724 | 33.188 | 29.452 |
| Corr. Eigenval. | 2.862 | 0.115 | 0.023 |
| Cov. Eigenval. | 2971.045 | 110.429 | 24.685 |

| Correlation Matrix | Band1 | Band2 | Band3 |
|---|---|---|---|
| Band1 | 1.000 | 0.969 | 0.889 |
| Band2 | 0.969 | 1.000 | 0.935 |
| Band3 | 0.889 | 0.935 | 1.000 |
| Determinant | 0.007 | | |

| Corr. Eigenvectors | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.577 | -0.600 | -0.555 |
| Band2 | 0.586 | -0.170 | 0.792 |
| Band3 | 0.569 | 0.782 | -0.254 |

| Inv. of Corr. Ev. | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.577 | 0.586 | 0.569 |
| Band2 | -0.600 | -0.170 | 0.782 |
| Band3 | -0.555 | 0.792 | -0.254 |

| Covariance Matrix | Band1 | Band2 | Band3 |
|---|---|---|---|
| Band1 | 1137.291 | 1084.208 | 882.898 |
| Band2 | 1084.208 | 1101.476 | 914.012 |
| Band3 | 882.898 | 914.012 | 867.392 |
| Determinant | 8098875.127 | | |

| Cov. Eigenvectors | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.606 | -0.589 | -0.534 |
| Band2 | 0.604 | -0.095 | 0.791 |
| Band3 | 0.517 | 0.802 | -0.299 |

| Inv. of Cov. Ev. | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.606 | 0.604 | 0.517 |
| Band2 | -0.589 | -0.095 | 0.802 |
| Band3 | -0.534 | 0.791 | -0.299 |

**Table B-25. Statistics for dataset: b4741s01006**

| | Band1 | Band2 | Band3 |
|---|---|---|---|
| Non-Null Cells | 3450304 | 3450304 | 3450304 |
| Area In Hectares | 345.030 | 345.030 | 345.030 |
| Area In Acres | 852.589 | 852.589 | 852.589 |
| Minimum | 0.000 | 0.000 | 3.000 |
| Maximum | 255.000 | 255.000 | 255.000 |
| Mean | 76.592 | 82.198 | 122.716 |
| Median | 75.000 | 84.000 | 124.000 |
| Std. Dev. | 31.705 | 26.949 | 27.133 |
| Std. Dev. (n-1) | 31.705 | 26.949 | 27.133 |
| Corr. Eigenval. | 2.816 | 0.158 | 0.026 |
| Cov. Eigenval. | 2316.985 | 130.125 | 20.572 |

| Correlation Matrix | Band1 | Band2 | Band3 |
|---|---|---|---|
| Band1 | 1.000 | 0.961 | 0.848 |
| Band2 | 0.961 | 1.000 | 0.914 |
| Band3 | 0.848 | 0.914 | 1.000 |
| Determinant | 0.012 | | |

| Corr. Eigenvectors | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.576 | -0.599 | -0.556 |
| Band2 | 0.590 | -0.166 | 0.790 |
| Band3 | 0.566 | 0.783 | -0.258 |

| Inv. of Corr. Ev. | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.576 | 0.590 | 0.566 |
| Band2 | -0.599 | -0.166 | 0.783 |
| Band3 | -0.556 | 0.790 | -0.258 |

| Covariance Matrix | Band1 | Band2 | Band3 |
|---|---|---|---|
| Band1 | 1005.225 | 821.496 | 729.182 |
| Band2 | 821.496 | 726.275 | 668.107 |
| Band3 | 729.182 | 668.107 | 736.181 |
| Determinant | 6202298.851 | | |

| Cov. Eigenvectors | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.642 | -0.596 | -0.483 |
| Band2 | 0.554 | -0.075 | 0.829 |
| Band3 | 0.530 | 0.800 | -0.282 |

| Inv. of Cov. Ev. | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.642 | 0.554 | 0.530 |
| Band2 | -0.596 | -0.075 | 0.800 |
| Band3 | -0.483 | 0.829 | -0.282 |

**Table B-26. Statistics for dataset: b4741t04005**

| | Band1 | Band2 | Band3 |
|---|---|---|---|
| Non-Null Cells | 3465504 | 3465504 | 3465504 |
| Area In Hectares | 346.550 | 346.550 | 346.550 |
| Area In Acres | 856.345 | 856.345 | 856.345 |
| Minimum | 0.000 | 0.000 | 0.000 |
| Maximum | 255.000 | 255.000 | 255.000 |
| Mean | 66.056 | 78.034 | 125.515 |
| Median | 63.000 | 80.000 | 128.000 |
| Std. Dev. | 27.785 | 25.158 | 27.942 |
| Std. Dev. (n-1) | 27.785 | 25.158 | 27.942 |
| Corr. Eigenval. | 2.709 | 0.244 | 0.047 |
| Cov. Eigenval. | 1965.380 | 188.462 | 31.875 |

| Correlation Matrix | Band1 | Band2 | Band3 |
|---|---|---|---|
| Band1 | 1.000 | 0.938 | 0.768 |
| Band2 | 0.938 | 1.000 | 0.855 |
| Band3 | 0.768 | 0.855 | 1.000 |
| Determinant | 0.031 | | |

| Corr. Eigenvectors | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.578 | -0.572 | -0.582 |
| Band2 | 0.596 | -0.191 | 0.780 |
| Band3 | 0.558 | 0.798 | -0.230 |

| Inv. of Corr. Ev. | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.578 | 0.596 | 0.558 |
| Band2 | -0.572 | -0.191 | 0.798 |
| Band3 | -0.582 | 0.780 | -0.230 |

| Covariance Matrix | Band1 | Band2 | Band3 |
|---|---|---|---|
| Band1 | 772.014 | 656.003 | 595.944 |
| Band2 | 656.003 | 632.928 | 600.858 |
| Band3 | 595.944 | 600.858 | 780.775 |
| Determinant | 11806445.223 | | |

| Cov. Eigenvectors | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.595 | -0.593 | -0.542 |
| Band2 | 0.555 | -0.185 | 0.811 |
| Band3 | 0.581 | 0.784 | -0.219 |

| Inv. of Cov. Ev. | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.595 | 0.555 | 0.581 |
| Band2 | -0.593 | -0.185 | 0.784 |
| Band3 | -0.542 | 0.811 | -0.219 |

**Table B-27. Statistics for dataset: b2239r01001.**

| | Band1 | Band2 | Band3 |
|---|---|---|---|
| Non-Null Cells | 3486800 | 3486800 | 3486800 |
| Area In Hectares | 348.680 | 348.680 | 348.680 |
| Area In Acres | 861.607 | 861.607 | 861.607 |
| Minimum | 1.000 | 0.000 | 3.000 |
| Maximum | 255.000 | 255.000 | 255.000 |
| Mean | 80.489 | 74.983 | 112.516 |
| Median | 75.000 | 73.000 | 110.000 |
| Std. Dev. | 35.010 | 25.893 | 25.163 |
| Std. Dev. (n-1) | 35.010 | 25.893 | 25.163 |
| Corr. Eigenval. | 2.756 | 0.201 | 0.042 |
| Cov. Eigenval. | 2326.804 | 172.182 | 30.359 |

| Correlation Matrix | Band1 | Band2 | Band3 |
|---|---|---|---|
| Band1 | 1.000 | 0.921 | 0.799 |
| Band2 | 0.921 | 1.000 | 0.914 |
| Band3 | 0.799 | 0.914 | 1.000 |
| Determinant | 0.023 | | |

| Corr. Eigenvectors | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.570 | -0.696 | -0.437 |
| Band2 | 0.594 | -0.019 | 0.804 |
| Band3 | 0.568 | 0.718 | -0.403 |

| Inv. of Corr. Ev. | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.570 | 0.594 | 0.568 |
| Band2 | -0.696 | -0.019 | 0.718 |
| Band3 | -0.437 | 0.804 | -0.403 |

| Covariance Matrix | Band1 | Band2 | Band3 |
|---|---|---|---|
| Band1 | 1225.730 | 834.540 | 703.572 |
| Band2 | 834.540 | 670.455 | 595.328 |
| Band3 | 703.572 | 595.328 | 633.160 |
| Determinant | 12162731.970 | | |

| Cov. Eigenvectors | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.704 | -0.639 | -0.311 |
| Band2 | 0.526 | 0.175 | 0.832 |
| Band3 | 0.477 | 0.749 | -0.459 |

| Inv. of Cov. Ev. | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.704 | 0.526 | 0.477 |
| Band2 | -0.639 | 0.175 | 0.749 |
| Band3 | -0.311 | 0.832 | -0.459 |

**Table B-28. Statistics for dataset: b2539r02002.**

| | Band1 | Band2 | Band3 |
|---|---|---|---|
| Non-Null Cells | 3486800 | 3486800 | 3486800 |
| Area In Hectares | 348.680 | 348.680 | 348.680 |
| Area In Acres | 861.607 | 861.607 | 861.607 |
| Minimum | 0.000 | 2.000 | 0.000 |
| Maximum | 255.000 | 255.000 | 255.000 |
| Mean | 74.835 | 83.121 | 97.305 |
| Median | 73.000 | 84.000 | 98.000 |
| Std. Dev. | 35.389 | 28.249 | 27.045 |
| Std. Dev. (n-1) | 35.389 | 28.249 | 27.045 |
| Corr. Eigenval. | 2.807 | 0.157 | 0.036 |
| Cov. Eigenval. | 2606.228 | 144.670 | 30.923 |

| Correlation Matrix | Band1 | Band2 | Band3 |
|---|---|---|---|
| Band1 | 1.000 | 0.942 | 0.844 |
| Band2 | 0.942 | 1.000 | 0.923 |
| Band3 | 0.844 | 0.923 | 1.000 |
| Determinant | 0.016 | | |

| Corr. Eigenvectors | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.573 | -0.665 | -0.479 |
| Band2 | 0.590 | -0.071 | 0.805 |
| Band3 | 0.569 | 0.744 | -0.351 |

| Inv. of Corr. Ev. | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.573 | 0.590 | 0.569 |
| Band2 | -0.665 | -0.071 | 0.744 |
| Band3 | -0.479 | 0.805 | -0.351 |

| Covariance Matrix | Band1 | Band2 | Band3 |
|---|---|---|---|
| Band1 | 1252.390 | 941.891 | 807.785 |
| Band2 | 941.891 | 797.983 | 705.225 |
| Band3 | 807.785 | 705.225 | 731.448 |
| Determinant | 11659163.318 | | |

| Cov. Eigenvectors | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.676 | -0.638 | -0.370 |
| Band2 | 0.545 | 0.095 | 0.833 |
| Band3 | 0.496 | 0.764 | -0.412 |

| Inv. of Cov. Ev. | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.676 | 0.545 | 0.496 |
| Band2 | -0.638 | 0.095 | 0.764 |
| Band3 | -0.370 | 0.833 | -0.412 |

**Table B-29. Statistics for dataset: b3542r02007.**

| | Band1 | Band2 | Band3 |
|---|---|---|---|
| Non-Null Cells | 3450304 | 3450304 | 3450304 |
| Area In Hectares | 345.030 | 345.030 | 345.030 |
| Area In Acres | 852.589 | 852.589 | 852.589 |
| Minimum | 0.000 | 0.000 | 1.000 |
| Maximum | 255.000 | 255.000 | 255.000 |
| Mean | 67.843 | 78.080 | 126.434 |
| Median | 58.000 | 74.000 | 125.000 |
| Std. Dev. | 38.978 | 33.803 | 30.206 |
| Std. Dev. (n-1) | 38.978 | 33.803 | 30.206 |
| Corr. Eigenval. | 2.828 | 0.150 | 0.022 |
| Cov. Eigenval. | 3383.121 | 164.550 | 26.616 |

| Correlation Matrix | Band1 | Band2 | Band3 |
|---|---|---|---|
| Band1 | 1.000 | 0.966 | 0.855 |
| Band2 | 0.966 | 1.000 | 0.920 |
| Band3 | 0.855 | 0.920 | 1.000 |
| Determinant | 0.009 | | |

| Corr. Eigenvectors | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.576 | -0.600 | -0.555 |
| Band2 | 0.589 | -0.165 | 0.791 |
| Band3 | 0.566 | 0.783 | -0.259 |

| Inv. of Corr. Ev. | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.576 | 0.589 | 0.566 |
| Band2 | -0.600 | -0.165 | 0.783 |
| Band3 | -0.555 | 0.791 | -0.259 |

| Covariance Matrix | Band1 | Band2 | Band3 |
|---|---|---|---|
| Band1 | 1519.288 | 1272.224 | 1006.666 |
| Band2 | 1272.224 | 1142.614 | 938.930 |
| Band3 | 1006.666 | 938.930 | 912.383 |
| Determinant | 14816774.760 | | |

| Cov. Eigenvectors | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.656 | -0.582 | -0.481 |
| Band2 | 0.577 | -0.024 | 0.817 |
| Band3 | 0.487 | 0.813 | -0.320 |

| Inv. of Cov. Ev. | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.656 | 0.577 | 0.487 |
| Band2 | -0.582 | -0.024 | 0.813 |
| Band3 | -0.481 | 0.817 | -0.320 |

**Table B-30. Statistics for dataset: b4740S06004.**

| | Band1 | Band2 | Band3 |
|---|---|---|---|
| Non-Null Cells | 3450304 | 3450304 | 3450304 |
| Area In Hectares | 345.030 | 345.030 | 345.030 |
| Area In Acres | 852.589 | 852.589 | 852.589 |
| Minimum | 0.000 | 0.000 | 1.000 |
| Maximum | 255.000 | 255.000 | 255.000 |
| Mean | 66.220 | 76.892 | 117.823 |
| Median | 64.000 | 80.000 | 120.000 |
| Std. Dev. | 33.724 | 33.188 | 29.452 |
| Std. Dev. (n-1) | 33.724 | 33.188 | 29.452 |
| Corr. Eigenval. | 2.862 | 0.115 | 0.023 |
| Cov. Eigenval. | 2971.045 | 110.429 | 24.685 |

| Correlation Matrix | Band1 | Band2 | Band3 |
|---|---|---|---|
| Band1 | 1.000 | 0.969 | 0.889 |
| Band2 | 0.969 | 1.000 | 0.935 |
| Band3 | 0.889 | 0.935 | 1.000 |
| Determinant | 0.007 | | |

| Corr. Eigenvectors | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.577 | -0.600 | -0.555 |
| Band2 | 0.586 | -0.170 | 0.792 |
| Band3 | 0.569 | 0.782 | -0.254 |

| Inv. of Corr. Ev. | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.577 | 0.586 | 0.569 |
| Band2 | -0.600 | -0.170 | 0.782 |
| Band3 | -0.555 | 0.792 | -0.254 |

| Covariance Matrix | Band1 | Band2 | Band3 |
|---|---|---|---|
| Band1 | 1137.291 | 1084.208 | 882.898 |
| Band2 | 1084.208 | 1101.476 | 914.012 |
| Band3 | 882.898 | 914.012 | 867.392 |
| Determinant | 8098875.127 | | |

| Cov. Eigenvectors | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.606 | -0.589 | -0.534 |
| Band2 | 0.604 | -0.095 | 0.791 |
| Band3 | 0.517 | 0.802 | -0.299 |

| Inv. of Cov. Ev. | PC1 | PC2 | PC3 |
|---|---|---|---|
| Band1 | 0.606 | 0.604 | 0.517 |
| Band2 | -0.589 | -0.095 | 0.802 |
| Band3 | -0.534 | 0.791 | -0.299 |

# VITA

**NAME: YOUTONG FU**

**BIRTH PLACE: BEIJING, P.R. CHINA**

**BIRTH DATE: APRIL 11, 1967**

**MARRIED: QUN ZHUANG (WIFE): VIRGINIA YU FU (DAUGHTER)**

**PARENTS: BIJIA FU (FATHTER) AND AIXUE WEI (MOTHER)**

**EDUCATION:**

**Ph.D.**, **Biological Systems Engineering** -- Land and Water Resource Engineering Emphasis, Virginia Polytechnic Institute and State University, 2002

**M.S**., **Agricultural Engineering** -- Watershed Engineering and GIS Emphasis**,** Virginia Polytechnic Institute and State University, 1994

**B.S**., **Geo and Ocean Sciences**--GIS and Computer Cartography Emphasis**,** Nanjing University, P. R. China, 1989

**AREAS OF SPECIALIZATION:**

- TMDL Development
- Geographic information systems (GIS) and Image Processing
- Hydrologic modeling/water quality modeling

**PROFESSIONAL EXPERIENCE:**

**Vice President of International Division**, MapTech, Inc., Virginia Tech Corporate Research Center, Blacksburg, Virginia (August 2002 – Present)

**Environmental Engineer***,* Tetra Tech, Inc. Fairfax, Virginia (August 1999----June 2002)

**GIS and Remote Sensing Project Supervisor**, Department of Biological Systems Engineering, Virginia Polytechnic Institute and State University (Virginia Tech), Blacksburg, Virginia (January 1997 - August 1999)

**GIS Specialist and Environment Engineer**, MapTech, Inc. Virginia Tech Corporate Research Center, Blacksburg, Virginia (January 1995-December 1996)

**Research Associate, Information Support Systems Laboratory** (VirGIS Project), Department of Agricultural Engineering, Virginia Tech, Blacksburg, Virginia (May 1994-December 1994)

**Graduate Research Assistant**, Department of Agricultural Engineering, Virginia Tech, Blacksburg, Virginia (August, 1991- April, 1994)

**Environmental Engineer**, Research Center for Eco-Environmental Studies, The Chinese Academy of Sciences, Beijing, China, (August 1989-July 1991)

**PUBLICATIONS:**

**Fu Y.,** S. Mostaghimi, P McClellan, J. Miller, V. Shanholtz, 1999. Use of Small Format Aerial Photography in NPS Pollution Control Applications. American Society of Agricultural Engineering (ASAE) /Canadian Society of Agricultural Engineering (CSAE), Paper No. 993125, Toronto, Canada.

**Fu, Y.** Environmental Impacts of Precision Farming, 1997. One chapter of the final reports (Mostaghimi et. al. 1997) to Commonwealth of Virginia Department of Environmental Quality, Virginia Coastal Resources Management Program.

Mostaghimi, S., P .W. McClellan, K. Brannan and **Y. Fu**, 1997. Precision Farming Technology for Reducing Agricultural Chemical Inputs in Virginia's Coastal Resources Management Area. Report No. PF-0298. Virginia Polytechnic Institute and State University, Biological Systems Engineering Department Blacksburg, VA 24061-0303.

Mostaghimi, S., P .W .McClellan, J .L Miller and **Y. Fu**, 1997. Identification of Confined Animal Sites for Selected Watershed in Virginia Using Remote Sensing: Phase I. Report No. CAS-P1-0298. Virginia Polytechnic Institute and State University, Biological Systems Engineering Department Blacksburg, VA 24061-0303.

**Fu, Y.**, 1994. Comparison of Two Hydrological Models on a Virginia Piedmont Watershed. Unpublished Master Sciences' thesis. Virginia Polytechnic Institute

and State University, Biological Systems Engineering Department, Blacksburg, VA 24061-0303.

**Fu, Y**., V. O. Shanholtz, 1993. Comparison of Two Hydrological Models on a Virginia Piedmont Watershed. American Society of Agricultural Engineer (ASAE) Conference Paper No. 932508. Chicago, IL.