AN ASSESSMENT OF ACCURACY OF THE 1992 NATIONAL RESOURCES INVENTORY AS PERFORMED BY THE SOIL CONSERVATION SERVICE IN OKLAHOMA

By

JEFFREY DEAN PHILLIPPI

Bachelor of Science

Oklahoma State University

Stillwater, Oklahoma

1985

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
MASTER OF SCIENCE
May, 1994

AN ASSESSMENT OF ACCURACY OF THE 1992
NATIONAL RESOURCES INVENTORY AS
PERFORMED BY THE SOIL
CONSERVATION
SERVICE IN
OKLAHOMA

Thesis Approved:

Thesis Adviser

Dean of the Graduate College

ACKNOWLEDGEMENTS

My sincerest thanks to Dr. Stadler for his time and continuous encouragement and assistance throughout my development and completion of this thesis. I thank Dr. Wikle for his advice, comments and support. I would also like to thank Dr. Engle and Dr. Miller for their support and advice.

I wish to thank Dr. Mark Payton for his assistance, patience, and advice during the statistical analysis and manipulations phase of this thesis.

I wish to thank my parents, Larry and Karen Phillippi, for their support, understanding, love and encouragement to continue my education for an advanced degree.

Finally I wish to thank my beautiful wife, Angelia Phillippi, for her love, support, encouragement, and sacrifice during my scholastic tenure. I can truly say she is my best friend. I wish to thank my to sons, Joshua and Collin Phillippi, who are two little miracles that have come into my life and have been a source of inspiration.

TABLE OF CONTENTS

Chapter		Page

LIST OF TABLES

LIST OF FIGURES

Introduction

The Rural Development Act of 1972 authorizes the resource activities within the Soil Conservation Service. The United States Department of Agriculture, Soil Conservation Service (SCS), has the responsibility to conduct the National Resources Inventory (NRI). This is an enormous task that takes place a minimum of every five years. Until recently, the use of remote sensing played a limited role in the gathering of NRI data. Remote sensing has been utilized in data acquisition of the 1987 NRI and almost exclusively in the 1992 NRI. The use of remote sensing has been judged to be of great value and has eased the burden of collecting NRI data for these efforts (NRI Summary Report, 1987).

Purpose

This research will attempt to assess the accuracy of the data collected during the 1992 NRI. If the data collected are highly accurate, a decrease in the number of sites to reference could be utilized. A decrease in referencing sites could save time, money, labor, and administrative costs involved with performing the NRI which could be utilized to carryout other duties set forth by law and the judicious use of funds.

Objective

The objective of this study was to assess the accuracy of the 1992 NRI in regard to "cover/use" data collected through remote sensing techniques and determine if a two percent or less reference check is adequate instead of the present five percent reference check being used at the present time. During this assessment, an attempt was made to identify land cover/use types for which remote sensing techniques may be deficient and identify techniques and procedures that might improve the accuracy of future inventories.

Research Problem/Question

A three to five percent reference check is the standard for the data gathered through the NRI. The research question or hypothesis is as follows. "Is the proportion of misclassified points at the five percent sampling level is equal to the proportion of misclassified points at the two percent level and is the proportion of misclassified points at the five percent sampling level is equal to the proportion of misclassified points at the one percent level of the 1992 NRI "cover/use" data collected?" Chapter III Methodology explains the manner in which this research problem/question is answered.

Background

The National Resources Inventory which is performed by the United States Department of Agriculture, Soil Conservation Service, is an inventory which today's leaders will use to develop policy and procedures affecting the future. The National Resources Inventory teams utilized remote sensing (aerial photo interpretation) as the primary tool in data collection efforts.

For fifty years, the Soil Conservation Service (SCS) has conducted periodic inventories of the nation's soil and water related resources. The earliest efforts were reconnaissance studies, including the Soil Erosion Inventory of 1934 and the Soil and Water Conservation Needs Inventory of 1945. The Potential Cropland Study of 1975, and the National Resources Inventories (NRI) of 1977 and 1982 were extensions and modifications of these earlier inventories.

The 1992 National Resources Inventory is the latest of these inventories conducted by the Soil Conservation Service, and is a vital part of the Soil Conservation Service mission. Data from the National Resources Inventory serve a variety of purposes. Soil Conservation Service technical and administrative personnel at all levels use the National Resources Inventory data to help determine staffing patterns and to focus on where conservation dollars can be utilized most efficiently and effectively. National Resources Inventory data are used to help formulate both national and state policy and priorities. The inventory was instrumental in developing the conservation provisions of the 1985 Farm Bill (Food Security Act) and was an important factor in the development of the 1990 Farm Bill known as the Food, Agriculture, Conservation and Trade Act (USDA Summary Report, 1987).

The National Resources Inventory is a multi-resource inventory based on soils and other types of resource data collected at randomly sampled sites. These sites were derived through a cooperative effort with the Iowa State University Statistical Laboratory and the Soil Conservation Service. The 1982 and 1987 National Resource Inventories collected data from 1,000,000 sample sites and 300,000 sample sites respectively (USDA Summary Report, 1987).

The Rural Development Act of 1972 authorized the resources inventory activities within the Soil Conservation Service. The Act directed the Secretary of Agriculture to carry out a land inventory and monitoring program and to issue a report which reflected soil, water, and related resource conditions at no less than five year intervals (USDA Summary Report, 1987). The past National Resource Inventories were conducted in 1977, 1982, and 1987, with the most recent completed December 31, 1992.

The 1992 National Resource Inventory instructions called for collection of many data elements. Primary Sampling Units (PSU's) include five separate categories with several attributes. The Primary Sampling Units data fields and attributes are outlined in table 1 (USDA Instructors Draft, 1991).

The Point Data of the 1992 National Resources Inventory include, but are not limited to the outline listed. The Point Data and Attributes are outlined in table 2 (USDA Instructors Draft, 1991).

TABLE 1

PSU DATA FIELDS AND ATTRIBUTES

PSU DATA

I. Data Gatherers
   - A. Name
   - B. Title
   - C. Field Visit? (Y/N)

II. General Information
   - A. Major Land Resource Area
   - B. Hydrologic Unit
   - C. Size of the Primary Sampling Unit
   - D. Entirely Federal Land? (Y/N)
   - E. "R"(Rainfall) Factor for the Universal Soil Loss Equation
   - F. "C"(Annual Climate) Factor for the Wind Erosion Equation

III. Imagery
   - A. Source or Type
   - B. Date of Imagery
   - C. Scale of Photography
   - D. Type of Photographic Film
   - E. Index Numbers

IV. Farmsteads and Built-Up Areas
   - A. Farmstead and Ranch Headquarters
   - B. Urban and Built-Up Areas
     1. Small Built-Up Areas 0.25 - 10 Acres
     2. Urban and Built-Up Areas At Least 10 Acres

V. *Windbreaks
   - A. Kind
   - B. Total Width
   - C. Width Within PSU
   - D. Total Length
   - E. Length Within PSU

SOURCE: NRI INSTRUCTORS DRAFT, 1991

TABLE 1 CONTINUED

PSU DATA FIELDS AND ATTRIBUTES

VI. Water Areas
- A. *Large Streams, At Least 1/8 Mile Wide (Census water)
  1. Area Within PSU
- B. *Small Streams Less Than 1/8 Mile Wide
  1. Width
  2. Length (Total Within PSU)
- C. *Census Water, Waterbodies At Least 40 Acres
  1. Kind
  2. Size Class, Total
  3. Size, Within PSU
- D. *Small Waterbodies (Less Than 40 Acres)
  1. Kind
  2. Total Size
  3. Size Within PSU

(*Indicates variable number of entries per PSU)

SOURCE: NRI INSTRUCTORS DRAFT, 1991

TABLE 2

PSU POINT DATA AND ATTRIBUTES

POINT DATA

I. Ownership

II. Soil Information
- A. Soil Mapping Unit Symbol
- B. SCS-SOI-5 Record Number
- C. Surface Texture
- D. Texture Modifier
- E. Slope Class, Low
- F. Slope Class, High
- G. Flooding Class
- H. Other Phase Determining Criteria
- I. Hydric? (Y/N)
- J. HEL? (Y/N)
- K. Prime Farmland? (Y/N)

III. (CRP) Conservation Reserve Program Information
- A. Under CRP Contract? (Y/N)
- B. Sign-Up Number
- C. Contracted Practice

IV. Earth Cover Determination
- Level I and II Categories, With %'s

V. Land Use
- A. Land Cover/Use
- B. Use of Land
  1. Primary
  2. Secondary
- C. Double Cropping?(Y/N)
- D. Second Crop, if Double Cropped
- E. Cropping History
  1. One Year Prior
  2. Two Years Prior
  3. Three Years Prior
- F. Forest Type, if Forest Land

VI. Distances to Habitats
- A. Cropland
- B. Forest Land (At Least 1 Acre)
- C. Water (Any Perennial Stream or Waterbody)
- D. Wetland (Type 1-20)
- E. Build-up Land, Farmstead, or Road
- F. Predominantly Grassy or Herbaceous Area (e.g., Pasture, Range, Roadsides, Grassy Fence Rows, and Odd or Idle Grassy Areas)

SOURCE: NRI INSTRUCTORS DRAFT, 1991

TABLE 2 CONTINUED

PSU POINT DATA AND ATTRIBUTES

VII. Irrigation
- A. Type
- B. Source of Water
- C. Field Delivery System

VIII. Erosion Data
- A. USLE
  1. C-factor
  2. P-factor
  3. Slope %
  4. Slope Length
- B. WEQ
  1. Knoll Erodibility
  2. K Factor (4yrs.)
  3. L Factor (4yrs.)
  4. V Factor (4yrs.)
  5. Length of Rotation

IX. Wetlands Data
- A. Wetland Type 1-20
- B. FSA Wetland Classification

X. Conservation Practices

XI. Conservation Treatment Needs
- A. Treatment Needed? (Y/N)
- B. Type of Treatment Needed
- C. Nonarable Due to Salinity

XII. Potential for Conversion to Cropland

XIII. Rangeland Data
- A. Range Site Number
- B. Total Woody Canopy, for Rangeland
- C. Range Data for Field-Visited Sites
  (only for a sub-sample of range sites##)
  1. Range Condition
  2. Apparent Trend
  3. Woody Canopy Cover, by Species
  4. Noxious Weeds
  5. Concentrated Flow Erosion

XIV. Conservation Tillage
  1. Type, if >30% cover or > 1000 lb. Residue
  2. Percentage, if <30% Cover

SOURCE: NRI INSTRUCTORS DRAFT, 1991

There was a vast amount of data collected in the 1992 National Resources Inventory. Effort and expense were utilized to accomplish the collection, storage, and analysis of the collected data. The data were gathered through field work and the interpretation of aerial slides.

Literature Review

Remote Sensing

Remote sensing was one tool used to collect data for the 1992 National Resources Inventory. 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, 1987). There are two basic processes involved with remote sensing. These processes are data acquisitions and data analysis.

The elements of the data acquisition process include energy sources, propagation of energy through the atmosphere, energy interactions with earth surface features, retransmission of energy through the atmosphere, airborne and/or spaceborne sensors, resulting in the generation of sensor data in pictorial and/or digital form. The data analysis process involves examining the data using various viewing and interpretation devices to analyze digital sensor data. Once the datum is extracted, the information is then compiled, generally in the form of hard copy maps and tables, or as computer files that can be merged with other "layers" of information in a Geographic Information System (GIS). Finally, the information is presented to the users, who use it in their decision-making process (Lillesand and Kiefer, 1987). The Soil Conservation Service has adopted this methodology, in performing the National Resources Inventory, through interpretation of aerial slides, gathering datum from this interpretation and then inputting this datum into a GIS.

Reference data is also known as "ground truth" or "ground truthing." This term is not meant literally since many forms of reference data are not collected on the ground and can only approximate actual ground conditions. Remote sensing is seldom used without the use of some type of reference data. Reference data might be used to serve any or all of the following purposes:

1) To aid in the analysis and interpretation of remotely sensed data;

2) To calibrate a sensor;

3) To verify information extracted from remotely sensed data (Lillesand and Kiefer, 1987); and

4) To train and calibrate information for ongoing surveys (Williams, 1977);

Purposes one and three were relevant to the 1992 National Resource Inventory in that the data gathered will aid in the validification, analysis, and interpretation of remotely sensed data.

One of the most vital phases of the evaluation of a remote sensing system is the collection of accurate and unbiased ground data. Little time is usually devoted to planning the collection of ground data. As a result, adequate remote sensing data may be collected, but adequate supporting data may not be available to permit its meaningful evaluation. Unless care is taken in the specification of ground data collection, the entire evaluation process becomes meaningless (Benson, 1972).

Reference data are an important part of any study or inventory. It is necessary to have a rapid per site coverage in order to achieve statistically significant sample sizes and make efficient use of personnel and equipment. In the 1992 NRI there was an average of 97 primary sampling units located within each county, with a total of 7436 within Oklahoma. The Soil Conservation Service uses a three to a five percent ground truth or "spot check" of the primary sampling units within each county. These spot check areas served as reference data to insure effectiveness and quality control of the remotely sensed (photo-interpretation) data.

The 1992 National Resources Inventory was completed primarily through remote sensing with a three to five percent reference data gathering. National Resources Inventory data gathering specialists utilized single date color slides with an approximate scale of 8"=1 mile or 1:7920.

The role of remote sensing has proven to be extremely valuable, in the National Resource Inventory. Remote

Sensing helped ease the burden of collecting the 1987 National Resources Inventory data in many states. Almost thirty percent of all sample sites did not require a field visit because information could be acquired from aerial photography. In an additional one-third of the samples, photography helped, but did not completely replace field visits (USDA Summary Report, 1987).

Many tools impact the effectiveness of remote sensing. Examples of some of these tools are crop calendars, multi-date imagery, and single-date imagery. Substantial knowledge of crops grown in an area is necessary in order to accurately identify crops from aerial photography. These data can be summarized by using a crop calendar and a detailed listing of the specific crops grown in an area along with rotational cycles. SCS analysts are thoroughly familiar with the local areas being inventoried, which helps aerial photo interpretations. Crop calendars in conjunction with multidate imagery available in various spectral bands greatly enhance crop identification up to, and at times exceeding 90 percent. Crop Calendars in combination with single-date imagery, show areas with comparable crops, such as wheat and alfalfa, crops with a crop identification accuracy rate rarely exceeding 55-65 percent. The reliability of crop identification on single-date photography can be improved by observing the following rules:

1) Schedule aerial coverage during the month when the most important crops are distinctly separable.

2) When a given crop exhibits no unique spectral signature during the growing season, obtain aerial coverage during the time when fewest other similar crops are present.

3) Use the critical bare soil months, or optimum crop discrimination periods to predict the occurrence of the next crop in the rotational cycle (Avery, 1977).

The accuracy of the NRI was performed on more than just crops. Other, natural areas such as rangeland, pastureland, and woodland were also assessed. The use of crop calendars on non cropland areas would have limited use. Lissesand and Kiefer (1987) state that a knowledge of land use and land cover/use is important for many planning and management activities concerned with the surface of the earth with regard to airphoto interpretation.

Geographic Information System

Information extracted through the process of airphoto interpretation is almost always "mapped in some sense." That is, the resource manager may normally wish to display and analyze the interpreted information in a spatial context (Lillesand and Kiefer, 1979). A Geographic Information System (GIS) is one way to both display and analyze spatial data.

A GIS is defined as a complete sequence of components for acquiring, processing, storing, and managing spatial data (Star and Estes, 1990). A GIS is both a database system with specific capabilities for spatially referenced data as well as set of operations for working with the data.

Raster and vector are two data structures common in geographic information systems. The simpler data structure is a raster or cellular organization of spatial data. In a raster structure a value for the parameter of interest is developed for every cell in a array over space. Vector data structures are based on elemental points where locations are known to arbitrary precision. Data gathered with the 1992 NRI will be placed and utilized within a raster based GIS.

Star and Estes (1990) have identified five essential elements that a GIS must contain:

1) Data acquisition- the process of identifying and gathering the data for your application.

2) Preprocessing- manipulation of data in several ways for entry into the GIS.

3) Data Management- functions govern the creation of, and access to, the database itself. These functions provide consistent methods for data entry, update, deletion, and retrieval.

4) Manipulation and Analysis- this portion of the system are the analytic operators that work the database contents to derive new information.

5) Product Generation- the phase where final outputs from the GIS are created. These might include statistical reports, maps, tables, graphics, etc. These products could be in soft copy and/or hard copy form.

This study is critical to the data acquisition phase of information for use in a GIS. If the information gathered is inaccurate, no amount of preprocessing will correct this error and product generation will be inaccurate and misleading. Some applications with regard to the product generating phase of GIS, by SCS, are soils mapping, land and/or crop classification, crop yield predictions, and an assessment of land use change.

Analysis of national resources issues will be enhanced using National Resources Inventory data sets as part of the input in a geographic information system study. With careful analysis, the National Resources Inventory database can be used to help guide thinking on many issues regarding the status and condition of the nation's resources. This information can be a valuable tool in helping protect America's natural resources and in using them wisely (NRI Summary Report, 1987). Without an accurate database which will be utilized by the GIS, SCS decisions and interpretations of NRI data might be flawed.

Analysis for Oklahoma will be performed using the Geographic Resources Analysis Support System (GRASS), a raster based geographic information system, developed by the U.S. Army Corps of Engineers, Construction Engineering Research Laboratory (Agee et al., 1989). The analyses will take place at the Soil Conservation Service State Office in Stillwater.

Two key tools that play a significant role in the National Resources Inventory are remote sensing (photo-interpretation) and geographic information systems. Data was acquired for the NRI through the use of RS techniques. The data collected will be analyzed through the use of a GIS. Remote Sensing and GIS have demonstrated their usefulness in today's high-tech society. The accuracy of data gathered and referenced is critical. If the proposed sampling decline indicates a decrease in quality and representativeness of the data gathered, this will decrease the integrity of the data being manipulated within the GIS and must be kept in mind when one knows that the integration will occur. When integration is performed using these two tools a synergistic effect is realized. Barker (1988) describes remote sensing as the unheralded component of a geographic information system. Ehlers (1989) describes the integration of remote sensing with GIS in a similar, but different sense as a necessary evolution.

The first step in integration of a GIS and remotely sensed data is the evaluation of data. With efficient data transfer, users can take advantage of the capabilities of two distinct systems by moving data from one to the other as necessary or convenient for a given processing task. This has been called the separate but equal approach to integration (Jordan, 1990). The task of data transfer from one system to another has not always been as easy as with today's technological advances. Barker (1988) stated the problems associated with gridded data (imagery) and vector data (map) continue to be an impeding factor in integrating image data into a GIS. Jordan (1990) seconds this by writing that both raster and vector GIS processing functions were developed to capture and store maps and to perform overlays of thematic maps. In most cases, the vector and raster technologies served different purposes, and their distinct data structures reflect the functions they need to fill. Data transfer between the two types of systems was difficult, if not impossible. Recently, this difficulty has been reduced significantly (Jordan, 1990).

An important aspect in the joint applications of remote sensing technology and GIS is to identify change (Star and Estes, 1990). These two processes are a very important in the National Resources Inventory. Remote sensing technology will provide a permanent record of the inventory as well as a system to record and identify change. The geographic information system will be used as an analytical tool to quantify the process of change in regard to previous National Resources Inventory results, such as NRI-1982 and NRI-1987.

Another area useful with the integration of geographic information systems and remote sensing is map updating. In some cases the classification accuracies achieved through standard automated image processing methods are inadequate. One way of improving accuracy is to incorporate geographic information systems data as ancillary information in the classification procedure. For instance, in a land cover/use mapping project, information on underlying soils and other physical characteristics was used to improve overall accuracy from 76 percent to 90 percent. In another study, vegetation types were classified with an accuracy rate of 88 percent by adding geographic information systems procedures to automated image classification (Jordan, 1990).

Other areas of application through integration of remote sensing and geographic informations include, but are not limited to, mapguided image interpretation, stratification, classifier modification (as mentioned earlier), and postclassifier sorting (Star and Estes, 1990). The use of remote sensing and geographic information systems separately and in combination is both exciting as well as overwhelming. Effective utilization of large amounts of remotely sensed data is dependent upon the existence of an efficient geographic handling and processing system that will transform the data into usable information for decision making activities (Zhou, 1989). Through the interfacing of geographic information systems technology with remote sensing, different management scenarios can be processed allowing the manager to analyze many management alternatives before selecting the alternatives that would be most suitable (Nellis et al., 1990). The data gathered through RS techniques will be utilized in the GIS environment, thus making the integrity of the sampled data critical, interrelated, and inseparable. Therefore, a discussion has been presented showing this relationship and the importance of geographic sampling procedure.

Geographic Sampling

The National Resources Inventory encompasses the result of a great deal of analyses through remote sensing and GIS. The success and integrity of the inventory will depend on geographic sampling, due to the impossibility of checking all locations in the United States each time an inventory is performed.

Past national resources inventories have used geographic sampling on approximately 1,000,000 sites in 1982 and almost 300,000 sites in 1987 in all counties of the United States, except those in Alaska, Puerto Rico and the Virgin Islands (USDA Summary Report, 1987). Though smaller than the survey carried out in 1982, the magnitude of geographic sampling that took place during the 1992 National Resources Inventory cannot be over-emphasized.

It is often desirable to base hypotheses on data that are not a complete set of the total population. Such a limited survey is termed a sample. This may be due to the inaccessibility of part of the population or perhaps the very large size of the whole population. The aims of the study should always be considered n developing a sampling plan. Some statistical models allow several variables to be considered, and the samples should be collected in sufficient numbers to allow bona fide results to be presented. The number in the sample collected depends on the degree of certainty that is required (Cole and King, 1969). In addition to the degree of certainty, the amount of time and resources will also have an effect on the size of the sample. It is important to realize that the larger the sample fraction, the more likely it is to give a true picture of the population being sampled, which in turn relates back to the degree of certainty (Lenon and Cleves, 1984). The sample size should be such that it is possible to infer from it sufficiently and accurately the character of the whole population (Cole and King, 1969). Perfect agreement should not be expected between sample estimates and the true population values. Estimates based on any given sampling procedure are distinguished by two properties: accuracy and precision. Given limited resources, sampling is the only way that the NRI can be accomplished.

Accuracy refers to correctness in estimating a population value (Berry and Baker, 1968). Fritscher and Gay (1979) define accuracy as the relationship between the measured and "true" value, or the closeness to an accepted standard. Star and Estes (1990) define accuracy as freedom from error, lack of bias, and closeness to true values. Although precision is not a factor to a spread of values in the case of wheat for example, however it is important to sampling and in the explanation of bias as demonstrated in the text following. Precision refers to the spread of estimates of the population value around the true value (Berry and Baker, 1968). Star and Estes (1990) define precision as the degrees of exactness with which a quantity is stated. This is directly related to the number of significant figures used in a description. Fitscher and Gay (1979) define precision the variability observed among numerous measurements of quality. A population value can be accurate but not very precise. A population value can be extremely precise but totally inaccurate. If there is a consistent over or underestimation of this value, the sample is said to be biased and inaccurate (Berry and Baker, 1968). Being biased can be looked upon as error, intentional or unintentional. Bias directly relates to the integrity or accuracy of the data gathered through sampling for the NRI by remote sensing techniques and the use of this data in the GIS environment.

The Soil Conservation Service has accomplished the task of gathering the geographic spatial data with the 1992 NRI. The main tools utilized in the completion of the National Resources Inventory have been presented. These tools are remote sensing, geographic information systems and geographic sampling. The Soil Conservation Service has performed a three to five percent ground truth or spot check. Davis and Dozier (1990) state that they tested the predictive values of classification using 300 samples identified by interpreting 1:24,000-scale aerial photos. Extensive ground reconnaissance in the study area confirmed the reliability of identified vegetation from the photographs. Keeping this in mind, the accuracy of the Soil Conservation Service National Resources Inventory surveys could be greater given the 1:7920-scale color aerial photography of the NRI as compared to the 1:24,000-scale photos. This could mean that a one or two percent ground truth could be sufficient for the NRI.

The 1992 National Resources Inventory photo interpretation phase of the NRI is complete. An evaluation is needed as to the level of sampling necessary to confirm the airphoto interpretations. If the interpretations are determined to be highly accurate then less time and money could be spent on the reference checking phase of future National Resources Inventories. The result could have an effect on future NRI's, future policy and procedures, as well as an effect on future generations.

CHAPTER II

STUDY AREA

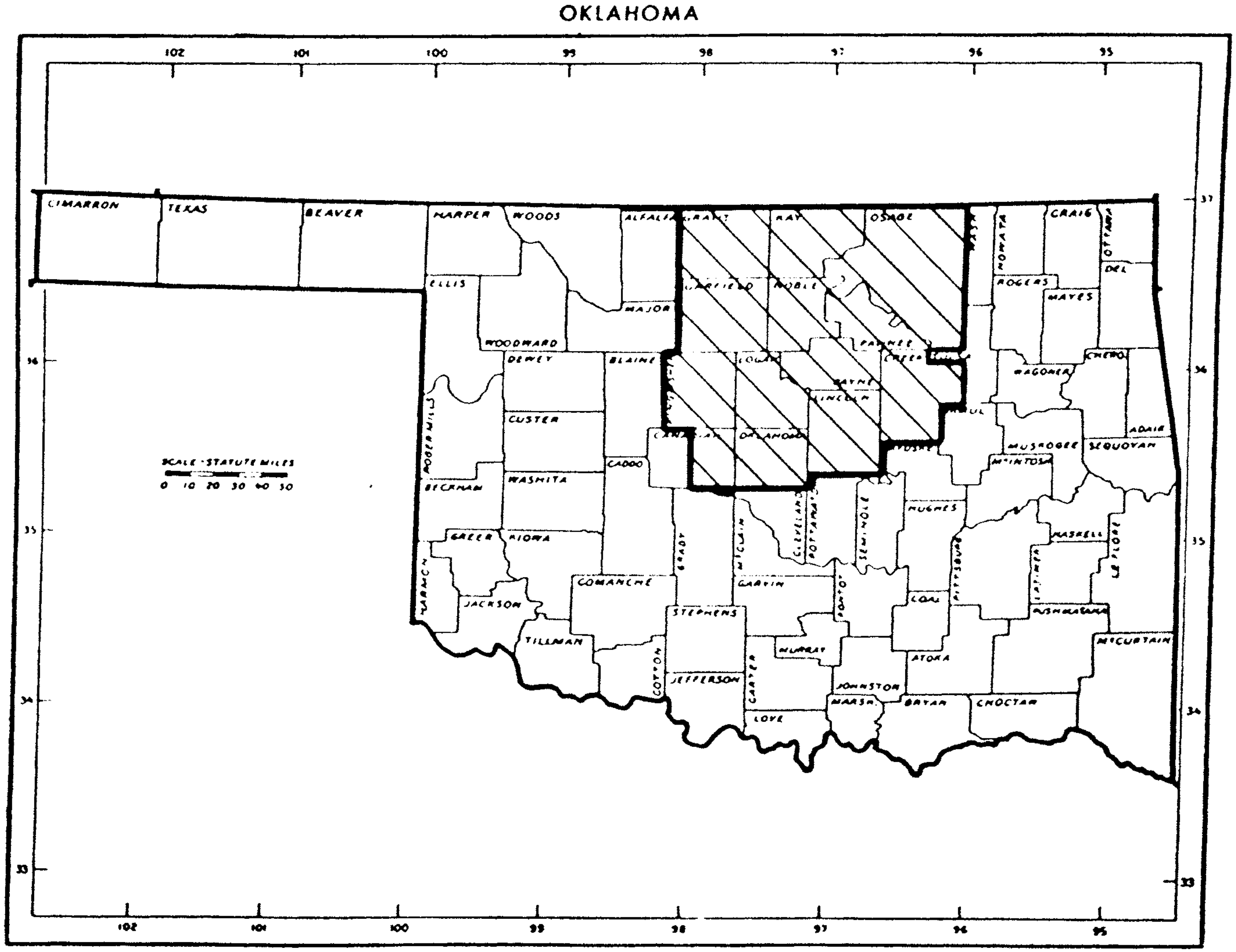
Introduction

The SCS has divided Oklahoma into a number of work areas (Figure 1). This thesis uses selected counties of Area II as a study area. The principal reason for using SCS Area II is due to the diversity of land cover/uses found within this geographical region. The land cover/uses of concern being cropland, rangeland, pastureland and woodland. Area II consists of the following counties: Grant, Kay, Osage, Garfield, Noble, Pawnee, Kingfisher, Logan, Payne, Creek, Lincon, Oklahoma, and Canadian. Time and resource considerations for this thesis dictated that work be limited to four of these counties. The four counties selected are Kay, Noble, Oklahoma, and Pawnee (Figure 2). These counties were chosen for their land cover/use diversity, centrality, and availability of PSU data sheets. Differences in climate and soils can affect the vegetation present and demonstrate the diversity of the study area. The following is a summary of the physical characteristics of the study area.

Climate

Oklahoma has a continental type climate with pronounced seasonal and geographic ranges in temperature and precipitation (Gray & Galloway, 1969). The average length of the growing season varies within the study area. The average

Figure 1. Area Boundary Map

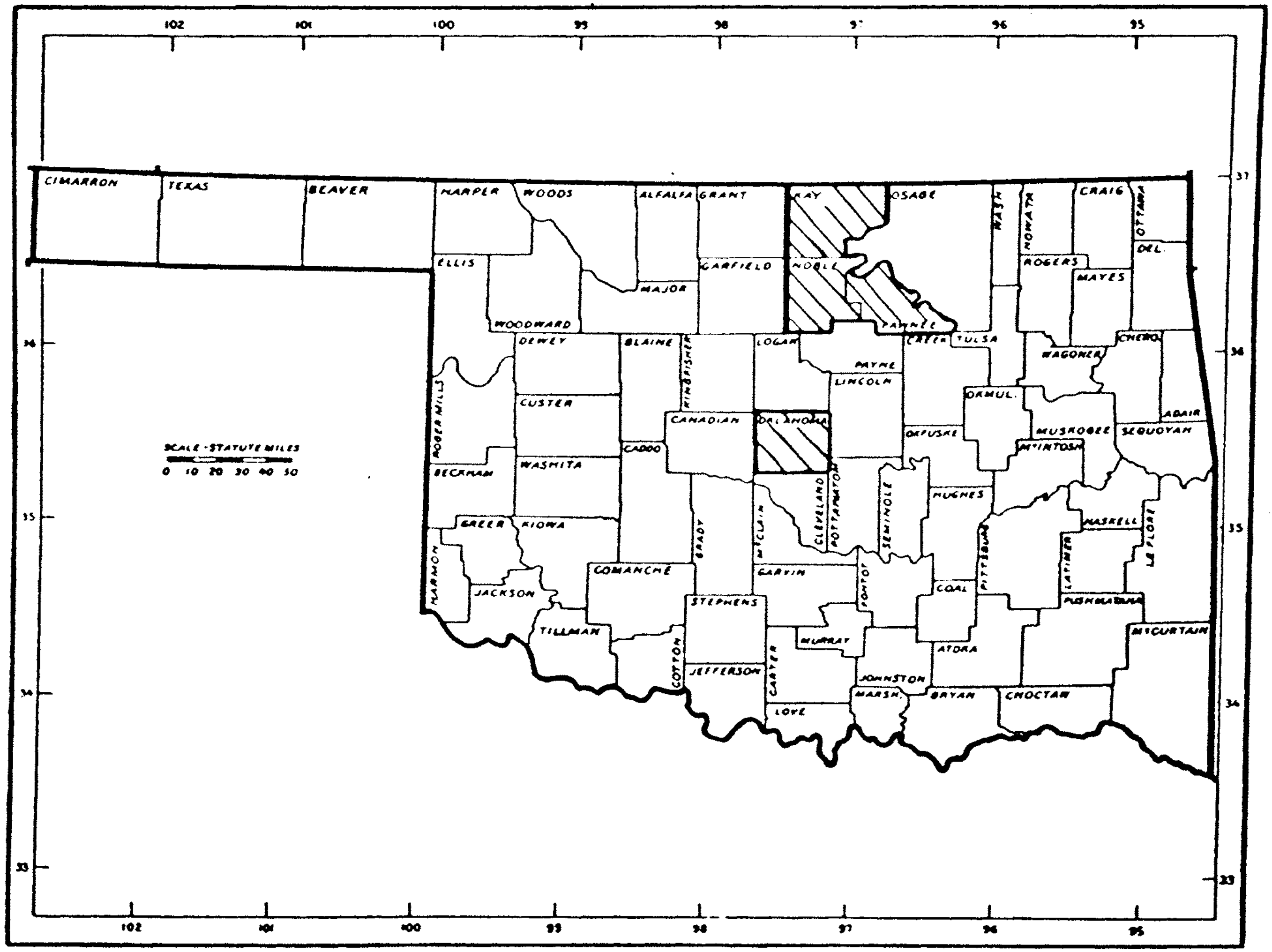


AD-400.34 11/58

Figure 1 Area 2 [hatched]

Figure 2. Study Area

OKLAHOMA



AD-480.34 11/58

Figure 2 Study Area

lengths of the frost free growing season for Oklahoma, Kay, Noble, and Pawnee are 221 days, 195 days, 206 days and 202 days respectively. This is the time period that extends from the last killing frost in the spring to the first frost in the fall.

The mean annual summer temperature for the study area is 78.5 degrees Fahrenheit. The mean annual summer temperature ranges from a low in Kay county at 72.9 degrees Fahrenheit to a high in Pawnee County at 80.8 degrees Fahrenheit. The summers are long with occasional periods of very high daytime temperatures. The winters are relatively short and mild, although minimum temperatures of zero or lower have been recorded at one or more stations in all except 3 of the 48 winters on record (Gray and Galloway, 1969).

Oklahoma's average annual precipitation varies across the state and within the study area. However, average annual precipitation fails to show the variation in rainfall from month to month. The mean annual precipitation, in inches, for Oklahoma, Kay, Noble, and Pawnee counties are 31.93, 32.11, 34.24, and 38.18 respectively. The precipitation received and soils present in a county will ultimately affect the water available for utilization by plant life and affect the plant community present. A space and time variability occurs within the study area with respect to the climate.

Soils

The following is a description of the soil associations found within the study area taken from the general soils map of Oklahoma. The soil associations present contribute to plant diversity and species locations, adding to differences among the counties selected for this study. The Bethany, Tabler, Kirkland soil complex and the Renfrom, Zaneis, Vernon soil complex are known as central reddish prairies, are dark soils with clayey subsoils developed under tall, grass mostly in clayey red beds. The Vaness, Mines, Yahola soil complex is within the central reddish prairies, with loamy soils and loamy subsoils developed under tall grass in loamy red beds or alluvium. The Sogn, Summit soil complex and the Parsons, Dennis, Bates soil complex are known as the Eastern (Cherokee) Prairies with dark colored soils mostly with clayey subsoils developed on shales, sandstones, and limestones under tall grasses. The Darnell, Stephenville soil complex and the Dougherty, Teller, Yahola soil complex are soils found within the cross timbers. they are light colored sandy soils with reddish subsoils on various sandy materials developed under oak-hickory forests with prairie openings (Savannah). The precipitation and soil type affect the amount of water received and the water holding potential of the soil, these in turn affect the amount and type of vegetation present.

Vegetation

The vegetation and land use/cover varies from county to county. This variation is one of the primary reasons that these four counties were selected. An overview of the vegetation present for each county follows.

The vegetation for the western part of Oklahoma County is dominantly cropland, with some grassland present. The savannah vegetation type is found in the central and eastern parts of the county. In the savannah vegetation blackjack oaks and grasses are dominant and the soils are generally medium to low in organic matter content. (Fisher, 1969).

The native vegetation in Kay County consists mainly of prairie grasses with a few small areas of trees in the uplands and along streams. The soils formed under prairie grasses have a dark-colored, friable, granular surface layer that holds moisture and plant nutrients well because the grasses, including their roots, contribute a large amount of organic matter to the soils. In addition, the fibrous roots of the grasses penetrate to a depth of 18 to 24 inches, and some of the smaller roots go much deeper. These roots absorb much of the rain that falls during the growing season and, therefore, lessen the leaching of plant nutrients. Also, the roots of the grasses bring nutrients, mainly calcium to the surface. These nutrients are returned to the surface layer in the organic residue of plants.

The soils of the uplands that formed under a cover consisting mostly of post oak and blackjack oak are less fertile than the dark colored, granular soils that formed under grasses (Culver, 1967).

The native vegetation of Noble County consisted principally of mixed tall and short grasses. There are three distinct associations in the county. The first association, the typical prairie type, occurs on the silt loam, loam, and clay loam soils of the uplands. When the county was first settled, the association was dominated by little bluestem. Associated with the little bluestem were big bluestem, sand bluestem, silver beardgrass, side-oats gramma. Bluestem grasses persist in meadows, along ungrazed roadsides, and on well-managed grazing land in the eastern and southern parts of the county.

The second plant association occupies deep and loose sandy soils in the northeastern part of the county. This area, surrounded on three sides by the Arkansas River, is locally called Big Bend Country. This second plant association is also on shallow sandy soils in the southern part of the county. Post and blackjack oaks are dominant. Associated with the oaks are little bluestem, sand bluestem, Indiangrass, various panicums, Johnsongrass, field sandbur, and hairy gramma.

The third plant association occurs along the streams on alluvial soils. This association consists of American elm, chinquapin, post, and blackjack oaks, hackberry, gum-elastic, willow, cottowood, green ash, and Chickasaw plum. Associated with the trees are several species of grass, principally bluestems. Common shrubs are the fragrant sumac, smooth sumac, roughleaf dogwood, poison ivy, and coralberry (Brenching, 1952).

Pawnee County is a part of the great grassland area of the United States known ecologically as the true prairie. The land cover/use is altered now. The normal cover for such prairie is bluestems and other medium tall grasses. The eastern third of the county is largely wooded and has grassy openings. The central third includes mostly grassland with post and blackjack oak tree openings, and the western third is nearly all grassland with an invasion of oaks on the sandy ridgetops. Bottom lands throughout the county are rather thickly forested, and much hardwood growth still remains. The native forest and grasses vary greatly on different types of soils (Galloway, 1959).

Summary

The study area shows variations with respect to climate, vegetation, and soils. The precipitation ranges from a low in Oklahoma county at 31.93 inches, to a high within the study area at 38.18 inches in Pawnee County. The vegetation types vary from the soils associated with the tall prairies and cropland such as the Renfrow, Zaneis, Vernon soil complex in the western part of Oklahoma county, to the soil complex associated with the oak-hickory forests with grass openings of the Darnell, Stephenville complexes of eastern Oklahoma county and the eastern part of Pawnee county. The soils associations of the land use of cropland are primarily the Bethany, Tabler, Kirkland soil complex and Renfrow, Zaneis, Vernon soil complex in western Kay, Noble, western Oklahoma and western Pawnee county. These four counties selected for the survey were chosen for the diversity demonstrated above. There is diversity shown within county boundaries and between counties, therefor the counties are unique and different enough making for a reasonable study. The variability and uniqueness present within and between counties ensures that all four land cover/use types are present and can be evaluated through this thesis.

CHAPTER III

METHODOLOGY

Introduction

There is a need for sampling when dealing with natural resource inventories. Most often it is not economically feasible to conduct a 100 percent survey of the population, and by the time an inventory is completed the data can be obsolete. The ultimate objective of all sampling is to obtain reliable data from the population sampled and to make certain inferences about that population (Avery, 1975). A sample is a part of a population (Steel and Torrie, 1980). Sampling is a very important technique. Avery (1975) states of all the techniques described in his book, the concept of sampling is perhaps the most important as applied to inventories of natural resources. A sampling is utilized as a reference check for the NRI. The objective of this study compares the accuracy of a 1, 2, and 5 percent samples of PSU's.

Study Period

The time period for this research was for the 1992 National Resources Inventory, beginning January 1, 1992 and completed December 31, 1992. This thesis is intended to be timely with regard to the inventory recently being completed and the data being placed within a GIS.

Study Emphasis

For this study, the population sampled are those Primary Sampling Unit (PSU's) points inventoried by the 1992 National Resources Inventory in Kay, Noble, Oklahoma, and Pawnee Counties. There are 69, 91, 185, and 72 PSU's surveyed within each county respectively in the 1992 NRI. There are three sampling points associated with each PSU. An assessment was performed to determine if the accuracy at a 1, 2, and 5 percent ground referencing of land cover/use points are equal.

Statistical Manipulation

Background Information

Four major land cover/use classes were referenced. These classes were chosen to parallel tables demonstrated in the 1987 USDA NRI summary which has several major land cover/use classes listed (USDA Summary Report, 1987). The classes are cropland, rangeland, pastureland, forest land, Minor land cover/uses and total rural land. The four land cover/use classes chosen are close grown crops (cropland), pasture and native pasture, rangeland, and forest land (woodland) which compose most of the non-urban land in these counties. The counties were selected based on the diversity of land cover/use present and the PSU points of these classes were pooled together. These classes were sampled using stratified random sampling techniques. In stratified sampling a population is divided into subpopulations of known size and a sample of at least two units are selected in each subpopulation (Avery, 1975). The four land cover/use classes were stratified within the sample area and the one, two, and five percent samples were taken randomly and independently of each other.

Fisher's Exact Test

In order to statistically test the probability of obtaining the same results in a reference check of a five percent sample of land cover/use points as compared to a two or one percent sample, it was necessary to understand the study design and variables present. The pattern present for the variables and the outcome of the reference check ultimately determined the selection of the statistical test utilized to analyze the data.

A comparison of the number of correct land cover/use identifications and incorrect land cover/use identifications of independent random samples of five percent level versus a two percent level, and a five percent level versus a one percent level were made. A 2 X 2 matrix was developed for each land cover/use and total under both sampling comparisons. Under the null hypothesis of independence, an exact distribution that is free of any known parameters results from conditioning on the marginal frequencies in both margins. When assuming independent multinomial sampling and then condition on the observed marginal totals a hypergeometric distribution is obtained.

This test for 2 X 2 tables is called Fishers exact test (Agresti, 1990). When it is required to test homogeneity, the probability of obtaining the observed distribution or a more extreme distribution is computed (Steel and Torri, 1980). The following tables are used to demonstrate what is meant by a more extreme distribution. The observed table shows the correct classifications and incorrect classifications of a five percent sample and a one percent sample, with a probability, of obtaining these tabled results, equal to 0.02105.

Fisher's Exact Test
Probability Calculations
Comparison of 5% sample and 1% sample
Sampling percent

	5%	1%	Total
Correct Classification	23	2	25
Incorrect Classification	2	3	5
Total	25	5	30

The table below is considered more extreme with a probability, of obtaining these tabled results, equal to 0.0008772.

Fisher's Exact Test
Probability Calculations
Comparison of 5% sample and 1% sample
Sampling percent

	5%	1%	Total
Correct Classification	24	1	25
Incorrect Classification	1	4	5
Total	25	5	30

These tables were not obtained from the present research study, they are shown only to demonstrate what is meant by the probability of obtaining an observed distribution or a more extreme distribution. This study has a small sampling size with data in a 2 X 2 matrix. The probabilities calculated are of obtaining the observed distribution or one more extreme as shown above.

Summary

The sample area is defined as Kay, Noble, Pawnee, and Oklahoma counties of Area II. The classes sampled were cropland, rangeland, pastureland, and woodland. An independent stratified random sample was taken at the 1, 2, and 5 percent levels of PSU points identified within the sampling land cover/use classes. The statistical manipulation was performed utilizing the Fisher's Exact Test in a 2 X 2 matrix. The Fisher's Exact Test was developed as a test appropriate for small sample sizes that are arranged in a 2 X 2 matrix.

CHAPTER IV

OBSERVATIONS

Introduction

An inventory of PSU points was performed within each county. All points that were identified within one of the four land cover/use categories were listed separately by land cover/use. All points were then sampled through a stratified random sampling procedure. The SCS has reference maps showing all PSU locations. These maps were used to locate those PSU points sampled. Each point selected for sampling was located on the PSU locator map and transferred to the appropriate soil survey aerial map and an onsite investigation performed. The actual land cover/use was then noted on the PSU sampling sheet and recorded on to the county sample sheet. Onsite investigations were performed on the following dates: Kay County December 29, 1993, Pawnee County January 3, 1994 and January 5, 1994, Noble County January 6, 1994, and Oklahoma County January 12, 1994 and January 13, 1994. Kay County has a total of 178 PSU points. 112 PSU points were found on cropland, 43 PSU points were found on rangeland, 21 PSU points were found on pastureland and 2 PSU points were found on woodland. Figure 3, shows the distribution of sampled PSU points by land cover/use that were randomly selected within Kay County for referencing and the accuracy of those onsite investigations. 11 PSU points were found on cropland, 4 PSU points were found on rangeland, 1 PSU point was found on pastureland and 0 PSU points were found on woodland within Kay County. Figure 4, shows a summary of PSU points within each percentage class by land cover/use within each county and the accuracy of those onsite investigations. 7 PSU points were randomly selected for the 5% sampling, 2 PSU points were randomly selected for the 2% sampling, 2 PSU points were randomly selected for the 1% sampling for cropland within Kay County. 1 PSU point was randomly selected for the 5% sampling, 1 PSU point was randomly selected for the 2% sampling, and 2 PSU points were randomly selected for the 1% sampling for rangeland within Kay County. 0 PSU points were randomly selected for the 5% sampling, 1 PSU point was randomly selected for the 2% sampling, and 0 PSU points were randomly selected for the 1% sampling for pastureland within Kay County. No PSU points were randomly selected for woodland within Kay County. The PSU points of each land cover/use class were pooled together. Figures 5 through 12 demonstrate the same relationships as listed above for, Noble County, Oklahoma County, Pawnee County, and the study area.

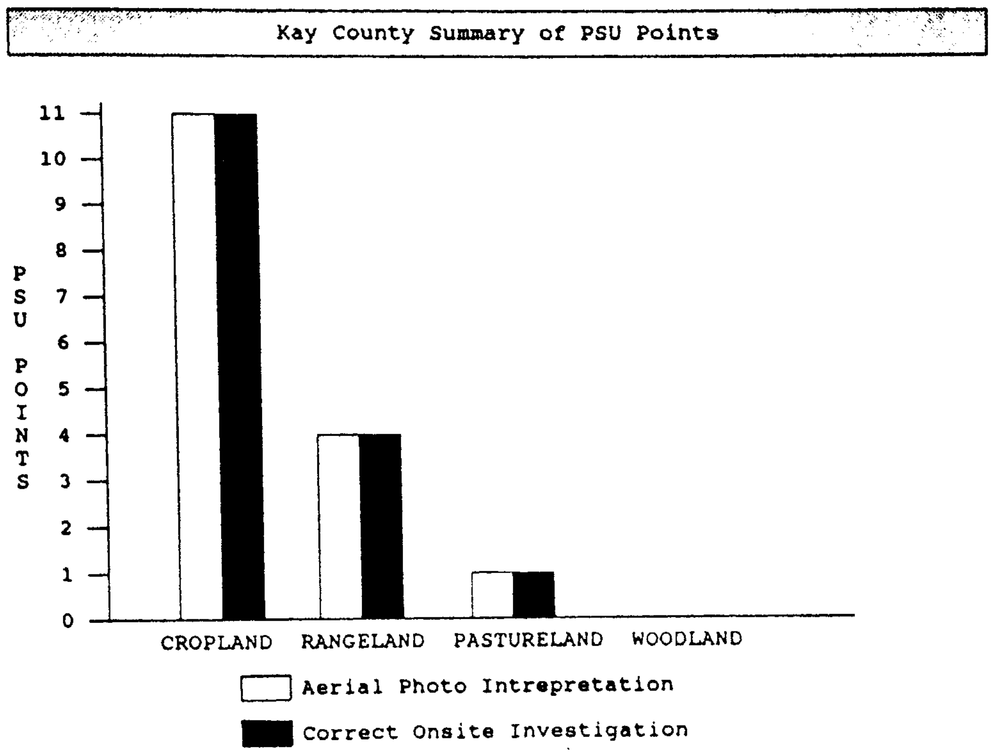


Figure 3. Kay County Summary of PSU Points

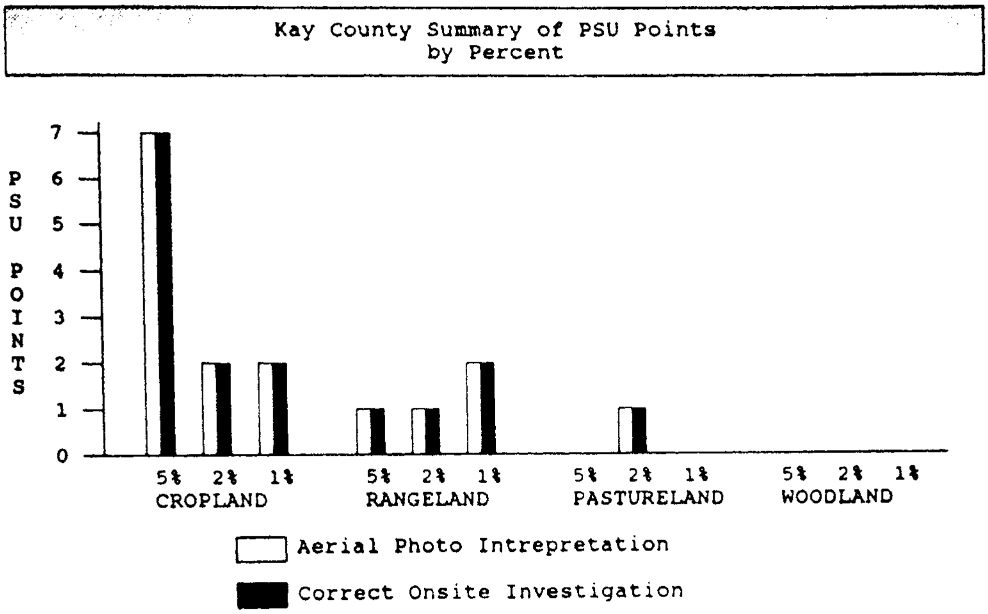


Figure 4. Kay County Summary of PSU Points by Percent

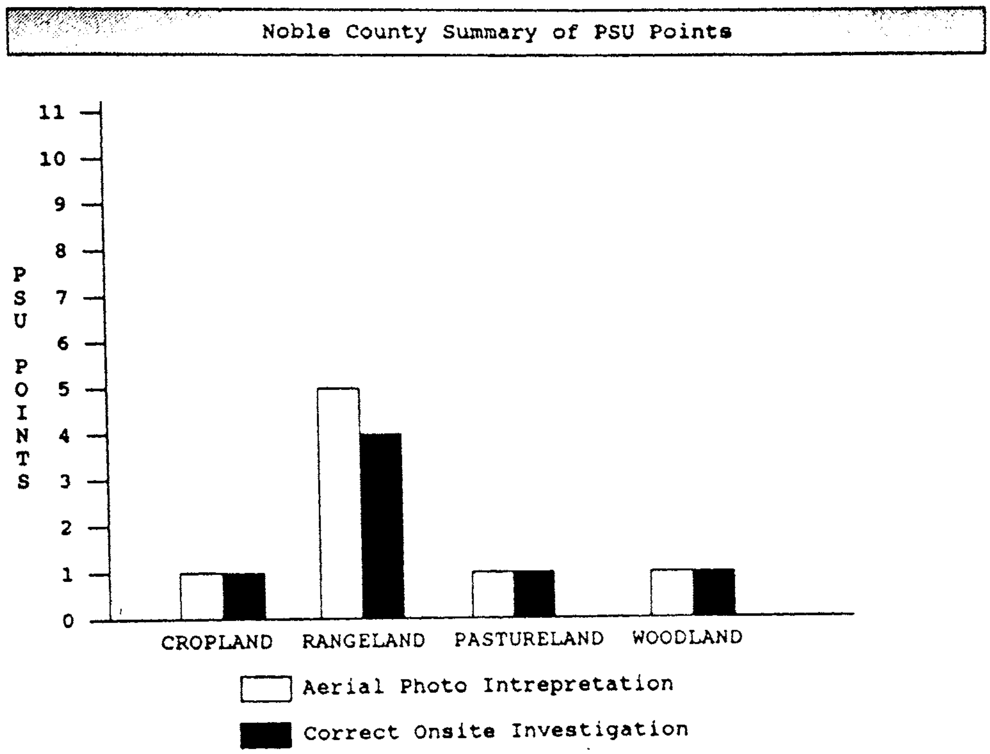


Figure 5. Noble County Summary of PSU Points

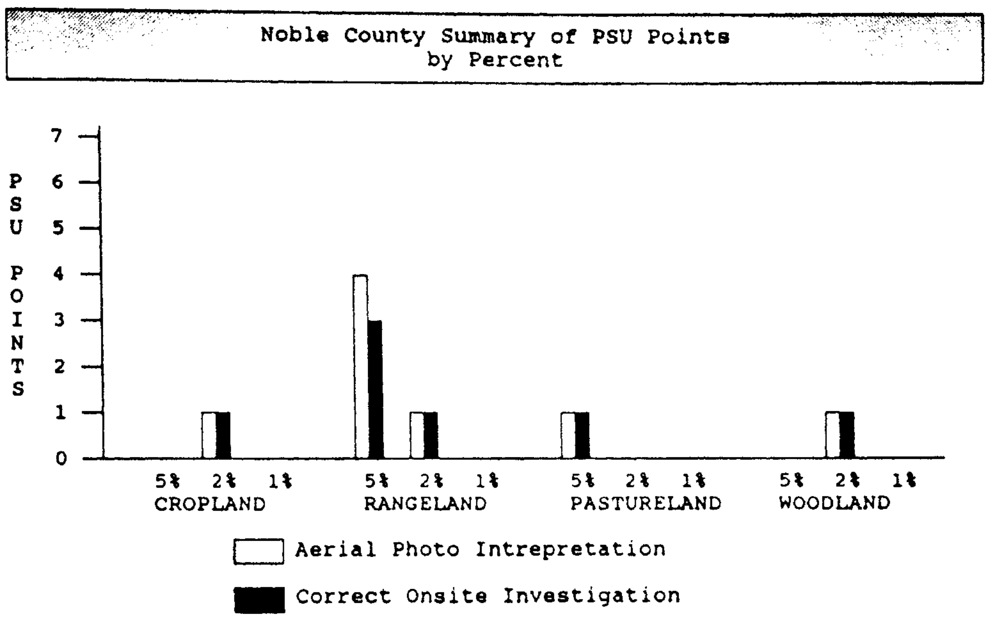


Figure 6. Noble County Summary of PSU Points by Percent

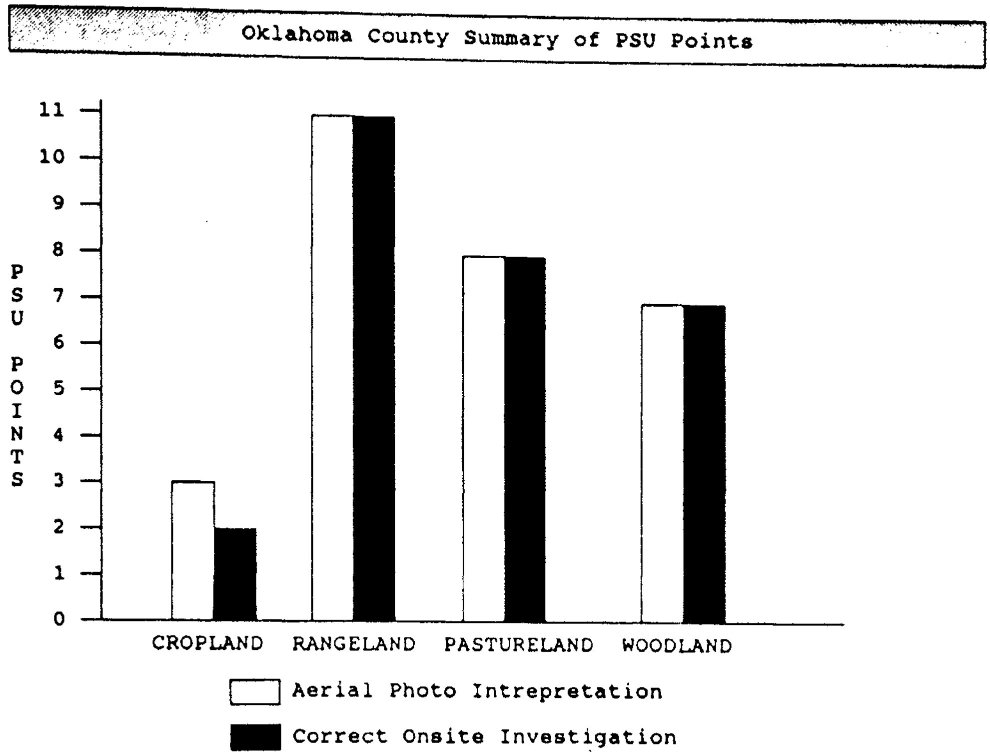


Figure 7. Oklahoma County Summary of PSU Points

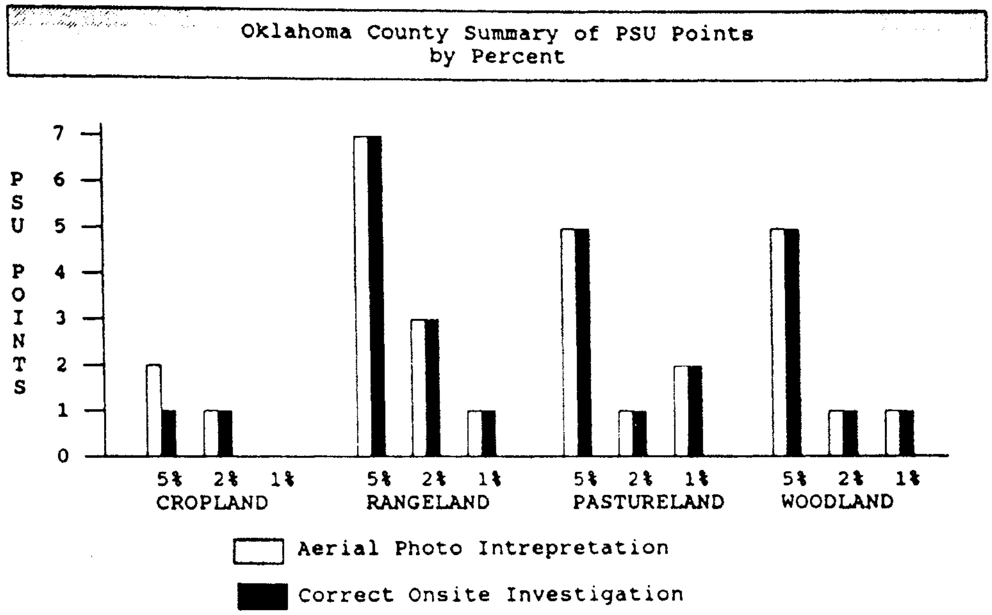


Figure 8. Oklahoma County Summary of PSU Points by Percent

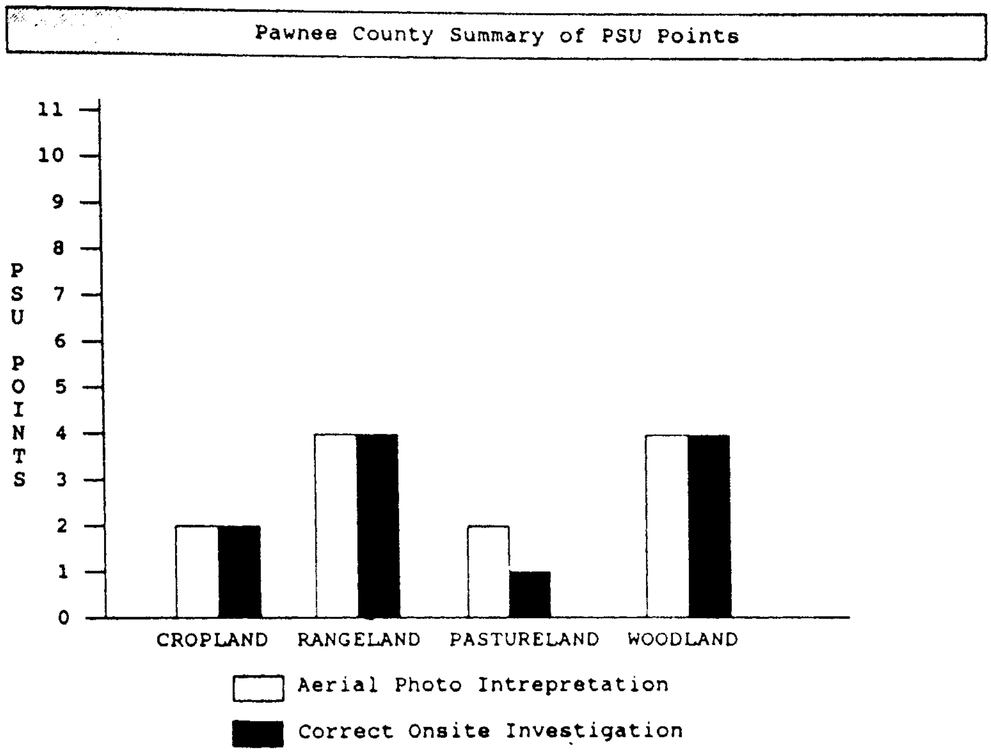


Figure 9. Pawnee County Summary of PSU Points

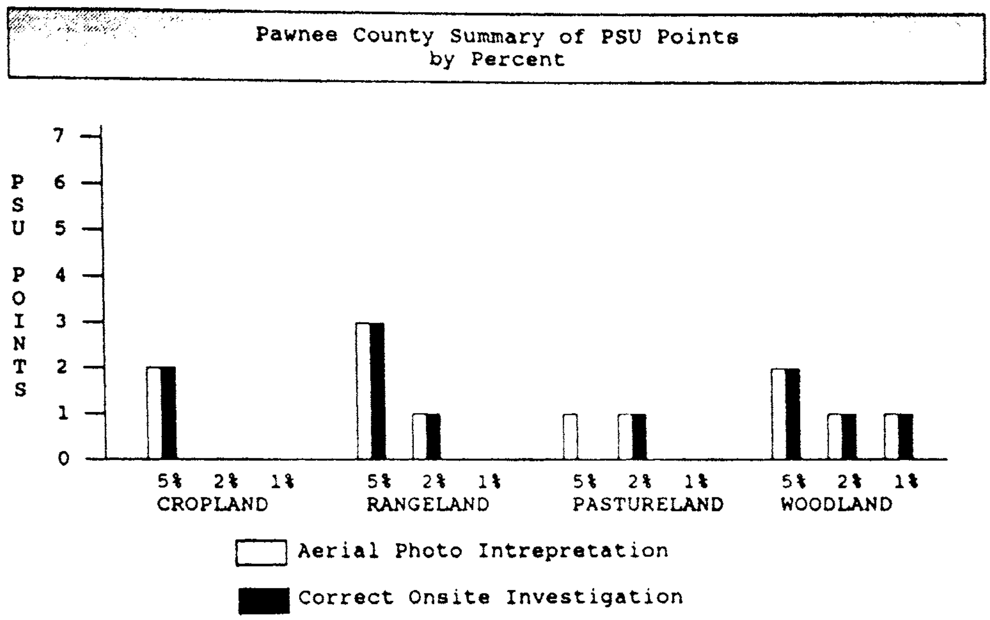


Figure 10. Pawnee County Summary of PSU Points by Percent

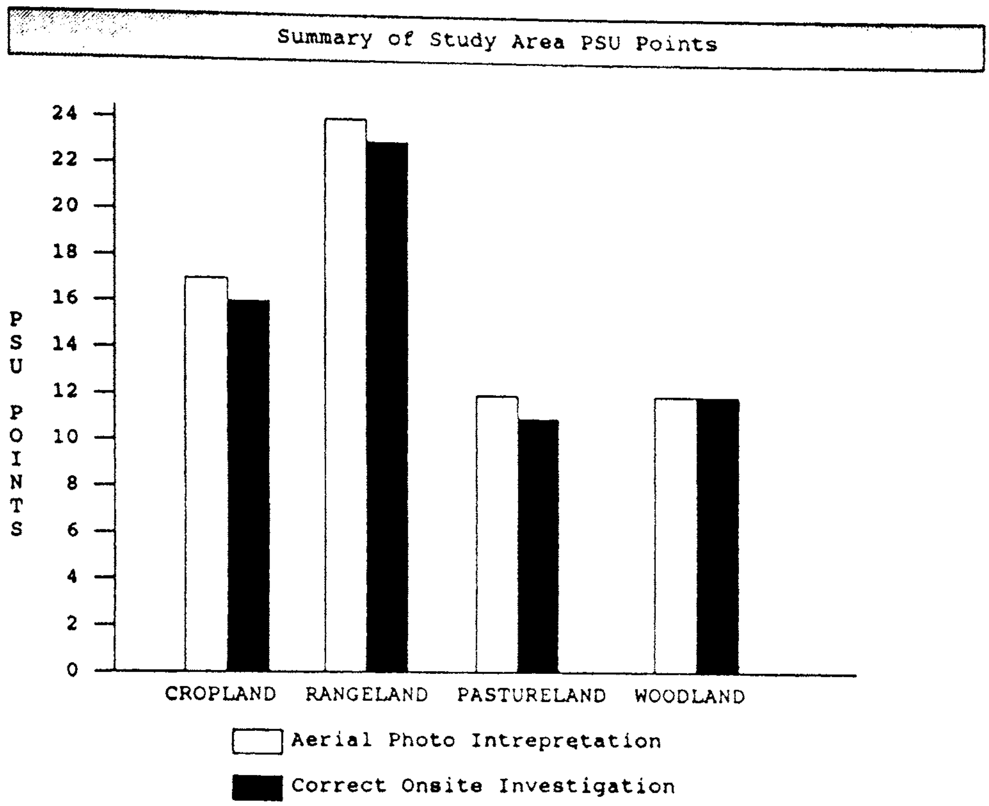


Figure 11. Study Area Summary of PSU Points

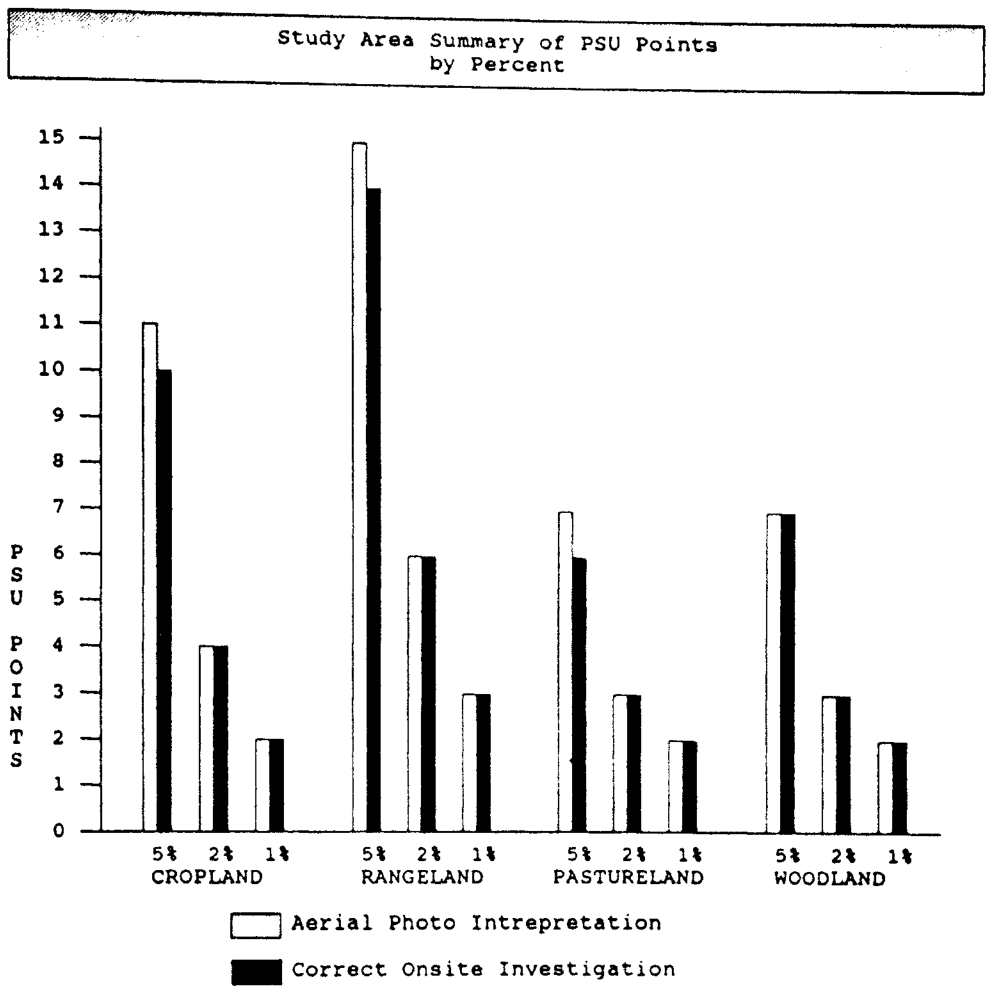


Figure 12. Study Area Summary of PSU Points by Percent

Table 3 demonstrates a total summary of PSU points inventoried and the percentage of the PSU points are shown for land cover/use within county and between counties. For example Oklahoma County has 41 PSU points of cropland, representing 14 percent of the total points within the county and 18 percent of the cropland points between the other three counties. Table 3 also shows that of the cropland points inventoried two were selected for sampling with the five percent sample, 1 was selected for the two percent sample and 0 were selected for the 1 percent sample. The table demonstrates this relationship for the four land cover/use classes and counties.

Statistical Results

The observations collected were placed in a 2 X 2 matrix and analyzed utilizing the Fisher's exact test. Table 4 and Table 5 demonstrates this 2 X 2 matrix for each land cover/use, sampling percent, and classification. The Probability value (P-value) is listed at the side of each matrix. The P-values obtained for the comparison of the five percent sample and the two percent sample were .73 for cropland, .71 for rangeland, .70 for pastureland, 1.0 for woodland, and .36 for the total of all the land cover/uses combined. The P-values obtained for the comparison of the five percent sample and the one percent sample were .85 for cropland, .83 for rangeland, .78 for pastureland, 1.0 for woodland, and .54 for the total of all the land cover/uses combined. The P-values obtained

TABLE 3
NRI INFORMATION
SUMMARY OF PSU POINTS INVENTORIED

LOCATION	OKC	5%	2%	1%	PAWNEE	5%	2%	1%	NOBLE	5%	2%	1%	KAY	5%	2%	1%	ALL	5%	2%	1%
CROP- LAND	18% 41 14%	2	1	0	10% 22 12%	2	0	0	23% 51 43%	0	1	0	49% 112 63%	7	2	2	100 226 29%	11	4	2
RANGE- LAND	32% 92 32%	7	3	1	35% 102 53%	3	1	0	18% 53 44%	4	1	0	15% 43 24%	1	1	2	100 290 37%	15	6	3
PASTURE- LAND	54% 72 24%	5	1	2	23% 31 16%	1	1	0	7% 10 8%	1	0	0	16% 21 12%	0	1	0	100 134 17%	7	3	2
WOOD- LAND	66% 89 30%	5	1	1	28% 37 19%	2	1	1	5% 6 5%	0	1	0	1% 2 1%	0	0	0	100 134 17%	7	3	2
TOTAL PERCENT WITHIN COUNTY	38% 294 100	19	6	4	24% 192 100	8	3	1	15% 120 100	5	3	0	23% 178 100	8	4	4	100 784	40	16	9

TABLE 4

FISHER'S EXACT TEST
PROBABLITY CALCULATIONS

CROPLAND — P-VALUE = .73

	SAMPLING PERCENT 5%	SAMPLING PERCENT 2%	
CORRECT CLASSIFICATION	10	4	14
INCORRECT CLASSIFICATION	1	0	1
	11	4	15

RANGELAND — P-VALUE = .7143

	SAMPLING PERCENT 5%	SAMPLING PERCENT 2%	
CORRECT CLASSIFICATION	14	6	20
INCORRECT CLASSIFICATION	1	0	1
	15	6	21

PASTURELAND — P-VALUE = .70

	SAMPLING PERCENT 5%	SAMPLING PERCENT 2%	
CORRECT CLASSIFICATION	6	3	9
INCORRECT CLASSIFICATION	1	0	1
	7	3	10

WOODLAND — P-VALUE = 1

	SAMPLING PERCENT 5%	SAMPLING PERCENT 2%	
CORRECT CLASSIFICATION	7	3	10
INCORRECT CLASSIFICATION	0	0	1
	7	3	10

TOTAL — P-VALUE = .36

	SAMPLING PERCENT 5%	SAMPLING PERCENT 2%	
CORRECT CLASSIFICATION	37	16	53
INCORRECT CLASSIFICATION	3	0	3
	40	16	56

TABLE 5

FISHER'S EXACT TEST
PROBABLITY CALCULATIONS

CROPLAND

	SAMPLING PERCENT 5%	1%	
CORRECT CLASSIFICATION	10	2	12
INCORRECT CLASSIFICATION	1	0	1
	11	2	13

P-VALUE = .85

RANGELAND

	SAMPLING PERCENT 5%	1%	
CORRECT CLASSIFICATION	14	3	17
INCORRECT CLASSIFICATION	1	0	1
	15	3	18

P-VALUE = .83

PASTURELAND

	SAMPLING PERCENT 5%	1%	
CORRECT CLASSIFICATION	6	2	8
INCORRECT CLASSIFICATION	1	0	1
	7	2	9

P-VALUE = .78

WOODLAND

	SAMPLING PERCENT 5%	1%	
CORRECT CLASSIFICATION	7	2	9
INCORRECT CLASSIFICATION	0	0	0
	7	2	9

P-VALUE = 1

TOTAL

	SAMPLING PERCENT 5%	1%	
CORRECT CLASSIFICATION	37	9	46
INCORRECT CLASSIFICATION	3	0	3
	40	9	49

P-VALUE = .54

support the null hypothesis Ho: the proportion of misclassified points found at the five percent sampling level is equal to the proportion of misclassified points found at the two percent and the one percent levels. The alternative hypothesis is Ha: the proportion of misclassified points at the five percent sampling level is greater than the proportion of misclassified points at the two or one percent sampling level. Since the samples are small and the number of misclassified points are also small, rejecting the null hypothesis in favor of the alternative hypothesis, using the Fisher's Exact Test, is unlikely. The NRI aerial photo interpretations were good and the SCS could perform the reference checks at a lower level than the three to five percent level being performed.

Fisher's Exact Test provides only a partial perspective of the relationship existing between these samples. In order to further examine the merit of taking larger samples, the probability of finding misclassified units were calculated (Table 6 and Table 7). For example, if N (sample size) equal to 400, q (percent sampled) equal to five percent, p (proportion of misclassified units) equal to .04, the probability of finding exactly one misclassified unit is 0.14022 of sampling units. While looking at this cell the probability at a q value of two percent is 0.03870 and the probability at a q value of one percent is 0.01255 of sampling units. Both are less than the probability at q= to the five percent sampling level. Therefore the probability of finding exactly one misclassified point is higher at the five percent sampling level than at the two or one percent sampling level. Table 7 lists the probabilities of finding exactly two misclassifications in the same manner as table 6. These tables demonstrate that as the sample size increases, the probability of identifying exactly one or exactly two, misclassified points increases depending on which table is being viewed. However, as the percentage of known misclassifications increase, along with an increase in the sample size, there is a noted decline in the probability of identifying only one or two (depending on which table is being viewed) misclassified PSU points, especially in the five percent sample. This decline in probability is an indication that there is a higher probability of finding more than exactly one or exactly two misclassifications, again depending on which table is being viewed. If SCS could specify the level of probability acceptable, these tables could be used to determine the sampling level percent to utilize.

Summary

PSU points in four land cover/use classes were sampled to check the accuracy at an one, two, and five percent level of sampling. The probabilities obtained comparing the five percent sample to the two percent sample of the four land cover/use classes in the four county area were 0.73, 0.71, 0.70, 1, and 0.36. These probabilities were for cropland, rangeland, pastureland, woodland and the total of all four classes, respectively. The probabilities obtained comparing the five percent sample to the one percent sample of the four land cover/use classes in the four county area were 0.85, 0.83, 0.78, 1, and 0.54. These probabilities were for cropland, rangeland, pastureland, woodland, and total of all four classes respectively. The probabilities support the null hypothesis.

TABLE 6
PROBABILITY OF FINDING EXACTLY ONE MISCLASSIFICATION

N	q %	p=	.01	.02	.03	.04	.05	.06	.07	.08	.09	.10
100	5		.05697	.10813	.15383	.19442	.23024	.26159	.28878	.31210	.33182	.34820
	2		.02934	.05758	.08460	.11045	.13516	.15874	.18121	.20260	.22291	.24218
	1		.01979	.03919	.05818	.07676	.09494	.11273	.13010	.14707	.16364	.17980
200	5		.09919	.17871	.24121	.28907	.32440	.34907	.36475	.37289	.37479	.37157
	2		.04800	.09214	.13259	.16951	.20308	.23347	.26082	.28529	.30704	.32620
	1		.02939	.05760	.08464	.11052	.13527	.15889	.18142	.20287	.22325	.24259
300	5		.13703	.23435	.30009	.34099	.36263	.36956	.36551	.35347	.33586	.31457
	2		.06585	.12384	.17454	.21852	.25632	.28843	.31533	.33747	.35525	.36908
	1		.03880	.07526	.10945	.14144	.17129	.19908	.22488	.24874	.27075	.29095
400	5		.17082	.27730	.33688	.36296	.36578	.35305	.33051	.30235	.27158	.24032
	2		.08297	.15285	.21101	.25871	.29711	.32725	.35012	.36659	.37748	.38351
	1		.04801	.092189	.13269	.16966	.20336	.23385	.26132	.28593	.30782	.32713
500	5		.20086	.30952	.35673	.36444	.34805	.31816	.28191	.24395	.20715	.17318
	2		.09937	.17935	.24251	.29117	.32738	.35296	.36955	.37857	.38129	.37822
	1		.05704	.10840	.15443	.19544	.23177	.26370	.29153	.31553	.33597	.35311
600	5		.22746	.33271	.36379	.35239	.31892	.27611	.23157	.18956	.15217	.12018
	2		.11508	.20348	.26950	.31686	.34880	.36811	.37716	.37801	.37241	.36182
	1		.06588	.12393	.17473	.21885	.25681	.28909	.31618	.33852	.35651	.37055
700	5		.25087	.34833	.36134	.33188	.28463	.23339	.18527	.14346	.10887	.08123
	2		.13011	.22539	.29240	.33666	.36284	.37481	.37582	.36852	.35512	.33740
	1		.07473	.13879	.19367	.24006	.27874	.31047	.33593	.35577	.37059	.38093
800	5		.27135	.35764	.35198	.30654	.24913	.19347	.14537	.10648	.076389	.05385
	2		.14450	.24523	.31160	.35134	.37073	.37486	.36782	.35288	.33262	.30903
	1		.08300	.15300	.21131	.25921	.29783	.32822	.35134	.36806	.37920	.38547
900	5		.28915	.36175	.33776	.27891	.21481	.15799	.11236	.07785	.05280	.03516
	2		.15825	.26311	.32747	.36158	.37355	.36971	.35501	.33323	.30723	.27914
	1		.09130	.16658	.22771	.27643	.31429	.34269	.36291	.37608	.38322	.38526
1000	5		.30447	.36158	.32029	.25078	.18303	.12749	.08582	056252	.03606	.02269
	2		.17139	.27918	.34034	.36800	.37221	.36060	.33885	.31119	.28063	.24934
	1		.09942	.17955	.24293	.29185	.32835	.35423	.37112	.38042	.38342	.38119

SOURCE: COMPUTED BY AUTHOR
N= TOTAL NUMBER OF SAMPLES UNITS q= PERCENT SAMPLED p= PROPORTION OF MISCLASSIFIED UNITS

TABLE 7
PROBABILITY OF FINDING EXACTLY TWO MISCLASSIFICATIONS

N	q %	p=	.01	.02	.03	.04	.05	.06	.07	.08	.09	.10
100	5		.00000	.00145	.00557	.01201	.02046	.03061	.04219	.05493	.06859	.08295
	2		.00000	.00029	.00118	.00264	.00465	.00718	.01024	.01378	.01781	.02229
	1		.00000	.00010	.00040	.00090	.00161	.00252	.00363	.00494	.00646	.00818
200	5		.00131	.01080	.02731	.04864	.07298	.09887	.12509	.15068	.17492	.19726
	2		.00024	.00215	.00581	.01104	.01768	.02558	.03459	.04457	.05539	.06690
	1		.00007	.00068	.00183	.00356	.00582	.00861	.01190	.01567	.019914	.02460
300	5		.00485	.02624	.05798	.09446	.13162	.16667	.19779	.22388	.24444	.25937
	2		.00090	.00537	.01306	.02344	.03601	.05033	.06600	.08264	.09993	.11757
	1		.00026	.00161	.00405	.00750	.01190	.01718	.02327	.03012	.03765	.04582
400	5		.01022	.04565	.09230	.14022	.18340	.21859	.24446	.26091	.26864	.26877
	2		.00192	.00975	.02239	.03870	.05766	.07840	.10017	.12233	.14432	.16571
	1		.00055	.00291	.00696	.01255	.01952	.02771	.03697	.04718	.05819	.06988
500	5		.01709	.06739	.12667	.18072	.22275	.25054	.26452	.26656	.25908	.24459
	2		.00327	.01513	.03335	.05591	.08110	.10755	.13413	.15995	.18431	.20671
	1		.00093	.00452	.01051	.01856	.02840	.03974	.05233	.06592	.08030	.09525
600	5		.02520	.09018	.15872	.21339	.24835	.26392	.26331	.25068	.23010	.20508
	2		.00494	.02137	.04554	.07430	.10514	.13615	.16588	.19329	.21766	.23854
	1		.00140	.00644	.01461	.02538	.03829	.05288	.06877	.08559	.10303	.12078
700	5		.03431	.11303	.18703	.23738	.26114	.26239	.24749	.22268	.19310	.16251
	2		.00689	.02834	.05862	.09324	.12886	.16306	.19419	.22119	.24346	.26077
	1		.00195	.00863	.01921	.03289	.04897	.06680	.08584	.10559	.12562	.14555
800	5		.04420	.13520	.21080	.25290	.26314	.25010	.22308	.18975	.15547	.12356
	2		.00912	.03592	.07229	.11225	.15158	.18751	.21834	.24316	.26166	.27396
	1		.00259	.01109	.02425	.04097	.06023	.28120	.10314	.12542	.14750	.16894
900	5		.05468	.15613	.22980	.26074	.25671	.23085	.19477	.15663	.12128	.09102
	2		.01160	.14400	.08630	.13089	.17280	.20904	.23804	.25923	.27275	.27971
	1		.00331	.01378	.02969	.04950	.07192	.09584	.12034	.14467	.16822	.19049
1000	5		.06559	.17544	.24399	.26198	.24415	.20778	.16584	.12610	.09227	.06541
	2		.01431	.05249	.10043	.14886	.19219	.22740	.25326	.26974	.27751	.27770
	1		.00410	.01670	.03547	.05840	.08386	.11049	.13718	.16307	.18747	.20989

SOURCE: COMPUTED BY AUTHOR
N= TOTAL NUMBER OF SAMPLES UNITS q= PERCENT SAMPLED p= PROPORTION OF MISCLASSIFIED UNITS

CHAPTER V

CONCLUSIONS AND RECOMMENDATIONS

Conclusions

The main objective of this study was to assess the accuracy of the 1992 NRI in regard to "cover/use" data collected through aerial photo interpretation. An effort was made to determine if a two percent or less reference check of PSU points is adequate in lieu of a five percent reference check of PSU points. A two percent or less reference check would save time, money, labor, and administrative costs involved with performing the NRI. The study area utilized for this study were four Oklahoma counties in SCS's Area II. This area was chosen due to the diversity of land forms and land cover/use found within this geographic region. Kay County is known for the large amount of cropland found in the western part of the county with prairie found in the eastern part of the county. Pawnee County has largely wooded and grassy openings in the eastern third. Its central third has grassland with oak openings, and its western portion is nearly all grassland with oaks found on sandy ridge tops. Noble County consists of mainly mixed tall and short grasses. The county also has a dominance of post oak and blackjack oak in the southern part of the county. Oklahoma County has mostly cropland found in the western part of the county and a savannah vegetation type in the central and eastern part of the county.

For this study the populations sampled were those PSU points inventoried by the 1992 NRI. Four major land cover/use classes were referenced. These were cropland, rangeland, pastureland, and woodland. The PSU points were pooled from the four county study area by each land cover/use class. These classes were sampled through stratified random sampling and independent of each sampling percent gathered. The data obtained was analyzed statistically utilizing Fisher's Exact Test in a 2 X 2 matrix. The P-values obtained for the comparison of the five percent sample and the two percent sample were .73 for cropland, .71 for rangeland, .70 for pastureland, 1.0 for woodland, and .36 for the total of all the land cover/use combined. The P-values obtained for the comparison of the five percent sample and the one percent sample were .85 for cropland, .83 for rangeland, .78 for pastureland, 1.0 for woodland, and .54 for the total of all the land cover/uses combined. These probabilities are high and support the null hypothesis that the proportion of misclassified points found at the five percent sampling level is equal to the proportion of misclassified points found at the two percent and the one percent levels. Based on the results of this study, the SCS could utilize a one or two percent reference check in place of the standard five percent reference check. While the results of this study, may not apply to all SCS Areas or counties, they can be used as a guide suggesting that a one percent or two percent reference check may be adequate for the NRI.

Recommendations

Fisher's exact test shows a high probability that the reference data collected at a five percent level is equal to the reference data collected at the two and one percent level of sampling of the four land cover/use classifications. This research indicates that the aerial photo interpretation techniques utilized in the 1992 NRI were relatively accurate with only 3 points found misclassified out of the 65 points referenced. Of these misclassified points, one point was found on rangeland, one point was found on cropland, and one point was found on pastureland. No misclassified points were found on woodland. This study was not able to identify aerial photo interpretation techniques common to a specific land cover/use misclassification. The misclassifications were spread relatively evenly across all of the four land cover/use referenced. Further studies, with regard to land cover/use reference checks, are recommended to identify reference levels at which the USDA SCS feels comfortable with the remote sensing techniques in place.

BIBLIOGRAPHY

Agee, James K.; Stitt, Susan C.F.; Nyquist, Maurice; and Root, Ralph. (1989). A Geographic Analysis of Historical Grizzly Bear Sightings in the North Cascades, Photogrammetric Engineering and Remote Sensing, 55(11), pp.1637-1642.

Agresti, Alan. (1990). Categorical Data Analysis. New York: John Wiley & Sons Inc.

Agterberg, F.P. (1981). Cell-Value distribution models in spatial pattern analysis, Future Trends in Geomathmatics. London England: Poin Limited. pp. 1-28.

Avery, Thomas Ergene. (1977). Interpretation of Aerial Photographs (3rd ed.). Minneapolis, Minnesota: Burgess Publishing Co.

Barker, G. Robinson. (1988). Remote Sensing: The Unheralded Component of a Geographic Information System, Photogrammetric Engineering and Remote Sensing, 54(2), pp.195-199.

Benson, Andrew S. (1977). A critique of Remote Sensing Evaluation Techniques, Remote Sensing of Earth Resources, Vol. 1, pp. 52-63.

Berry, Brian J.L.; Baker, Allen M. (1968). Spatial Analysis: A Reader In Statistical Geography. Englewood Cliffs, NJ: Prentice Hall.

Brenching, O.H.; Talley, E.C., Fitzpatrick, E.G.; Templin E. H.. (1952). Noble County Soil Survey. United States Department of Agriculture: Washington D.C.

Cole, John P.; King Cuchlaine A.M. (1969). Quantative Geography. New York: John Wiley & Sons Inc.

Culver, James R.; Bain, William R.; Baggett, Beryl G.. (1967). Kay County Soil Survey. United States Department of Agriculture: Washington D.C.

Davis, Frank W.; Dozier, Jeff. (1990). Information Analysis of a Spatial Database For Ecological Land Classification, Photogrammetric Engineering and Remote Sensing, 56(5), pp. 605-613.

Davis, J.C.; Harbaugh, J.W. (1981). Alternative Objectives for oil exploration by the United States Government: the dilemma of the National Petroleum Reserve in Alaska, Future Trends in Geomathmatics. London, England: Pion Limited.

Elders, Manfred; Edwards, Geoffrey, Bedard, Yvan. (1989). Integration of Remote Sensing With Geographic Information Systems: A Necessary Evolution, Photogrammetric Engineering and Remote Sensing, 55(11), pp. 1619-1627.

Fisher, Carl F.; Chelf, John V.. (1969). Oklahoma County Soil Survey. United States Department of Agriculture: Washington D.C.

Fritschen, Leo J.; Gay Lloyd W. (1979). Environmental Instrumentation. New York: Springer-Verlag.

Galloway, Harry M.; Templin, E.H., Oakes, Harvey. (1959). Pawnee County Soil Survey. United States Department of Agriculture: Washington D.C.

Gray, A.G.; Galloway, J.A.. (1959). Soils of Oklahoma. OSU Agricultural Experiment Station : Stillwater, Oklahoma.

Gould, Peter. (1985). The Geographer at Work. Routlege & Kegan Paul plc.

Jordon III, Lawrie E. (1990). Integration of GIS and Remote Sensing, Geographic Information Systems Supplementary Readings, Spring 1991. pp. 261-264.

Lenon, Barnaby J.; Cleves, Paul G. (1984). Techniques and Fieldwork in Geography. Slough: University Tutorial Press.

Lillesand, Thomas M.; Kiefer, Ralph W. (1987). Remote Sensing and Image Interpretation (2nd ed.). New York: John Wiley & Sons Inc.

Nellis, M. Duane; Lulla, Kamlesha, and Jensen , John. (1990). Interfacing Geographic Information Systems and Remote Sensing for Rural Land-Use Analysis, Photogrammetric Engineering and Remote Sensing, 55(5), pp. 329-331.

Star, Jeffrey; Estes, John. (1990). Geographic Information Systems: An Introduction. Englewood Cliffs, NJ: Pretice Hall.

2

Steel, Robert G.D.; Torrie, James H. (1980). Principles And Procedures of Statistics: A Biometrical Approach (2nd ed.). New York: McGraw-Hill Publishing.

United States Department of Agriculture, Soil Conservation Service. (1990). Fact Sheet: 1987 National Resources Inventory. Stillwater, Oklahoma: Soil Conservation Service.

United States Department of Agriculture, Soil Conservation Service. (1989). Summary Report 1987 National Resources Inventory. Statistical Bulletin Number 790.

United States Department of Agriculture, Soil Conservation Service. (1991). National Resources Inventory Instructors Draft. vol 11.

Williams, T.H. Lee. (1977). The Role of Ground Truth Data and An Approach to its Collection, Remote Sensing of Earth Resources, Vol. 6, pp. 39-49.

Zhou, Qiming. (1989). A Method for integrating Remote Sensing and Geographic Information Systems, Photogrammetric Engineering and Remote Sensing, 55(5), pp. 591-596.