THE FLORIDA STATE UNIVERSITY

COLLEGE OF ARTS AND SCIENCES

A QUESTION OF BIAS IN THE NORTH AMERICAN FLUTED-POINT SAMPLE

By

Christopher A. Schaefer

A Thesis submitted to the
Department of Anthropology
in partial fulfillment of the
requirements for the degree of
Master of Science

Degree Awarded:
Spring Semester, 2006

The members of the Committee approve the thesis Christopher A. Schaefer defended on November 16, 2005.

Glenn H. Doran
Professor Directing Thesis

Rochelle A. Marrinan
Committee Member

William Parkinson
Committee Member

Approved:

Dean Falk, Chairperson, Department of Anthropology

The Office of Graduate Studies has verified and approved the above named committee members

TABLE OF CONTENTS

List of Tables iv
List of Figures v
Abstract vi

1. INTRODUCTION...1

2. STATISTICS AND ANTHROPOLOGY 11

3. METHODOLOGY 13

4. RESULTS 21

5. DISCUSSION 42

CONCLUSION 49

APPENDIX A: Detailed Results for the one-way ANOVA and Tukey Multiple Comparison Tests…………………………………….52

APPENDIX B: Thesis Research Database……………………………………… 69

BIBLIOGRAPHY 70

BIOGRAPHICAL SKETCH 73

LIST OF TABLES

Table 1.1. Possible Biases of Fluted-Points Found and Recorded....................................2

Table 3.1 Ordinal Classes for People Per Square Kilometer 17

Table 3.2 Ordinal Classes for Cultivated Square Kilometers 17

Table 3.3 Region Composition..20

Table 4.1 Non-Transformed Descriptive Statistic...22

Table 4.2 Transformed Descriptive Statistics 27

Table 4.3 Pearson Association Matrix on the Untransformed Variables.........................31

Table 4.4 Spearman's Correlation Matrix on the Untransformed Variables 32

Table 4.5 Statistical Significance Summary: Population Versus Fluted-Points 40

Table 4.6 Statistical Significance Summary: Cultivation Versus Fluted-Points..............41

LIST OF FIGURES

Figure 1.1 Contours of Total-Fluted Points Per County 5

Figure 1.2 Contours of Total-Fluted Points Divided by County Area x1000....................6

Figure 1.3 Contours of Total Population 7

Figure 1.4 Contours of Population Per Square Kilometer 8

Figure 1.5 Contours of Cultivation Per Square Kilometer..9

Figure 3.1 Regional Groupings..19

Figure 4.1 Non-Transformed Frequency Distribution for Total Fluted-points per County..23

Figure 4.2 Non-Transformed Frequency Distribution for Modern Population 24

Figure 4.3 Non-Transformed Frequency Distribution for Cultivation 25

Figure 4.4 Non-Transformed Frequency Distribution for Total Fluted-points Divided by County Area 26

Figure 4.5 Log Transformed Frequency Distribution for People Per Square Kilometer 28

Figure 4.6 Log Transformed Frequency Distribution for Total Fluted-points per County..29

Figure 4.7 Log Transformed Frequency Distribution for Cultivation 30

ABSTRACT

A variety of statistical tests were used to analyze data from the fluted-point database compiled by Anderson and Faught (2000b) to determine if modern population density and/or cultivated square kilometers of land bias the fluted-point sample. Results for each statistical test are presented to show how the different tests can influence interpretation of the data. Ultimately, it was determined that only one test, the non-parametric Kruskal-Wallis, was appropriate for analysis of the data compiled in the fluted-point sample.

When the lower forty-eight states are conflated, statistical analysis showed that the fluted-point database was biased by modern population density and amount of cultivation. Counties with higher population density have more fluted-points recorded than those with lower population density. As with population density, counties with more cultivation per square kilometer have more fluted-points recorded than those with less cultivation. The same is true for the northeast, south, and midwest regions of the United States, as well as those states east and west of the Mississippi. There was no statistically significant relationship, however, between cultivation and modern population density and fluted-points recorded in the Western region.

CHAPTER 1

INTRODUCTION

This thesis examines the possibility that the North American Fluted-Point Database compiled by Anderson and Faught (1998, 2000a, 2000b) is subject to biases. Specifically, this research focuses on the possibility that modern population density and/or area of land under cultivation bias the fluted-point sample.

There are 12,851 fluted-points recorded by county level in the North American fluted-point database. The database is an aggregate of many samples that differed greatly in design, scale, and purpose, depending on who conducted them, how, when, and where (Shott 2002; Morrow and Morrow 1999). The database shows fluted-points occur in distinct concentrations in North America and with greater frequency in Eastern North America. Anderson and Faught (2000b) indicate that areas of distinct concentrations yield important insights into Paleoindian settlement pattern and history. For example, Anderson and Gillam (2000) propose the uneven distribution of fluted-points across North America reflects a leap frog pattern of technology dispersement rather than the traditionally held wave of adance model.

Distributional studies of fluted-points at various levels can help determine Paleoindian land use and adaptations (Shott 2002). Fluted-points are a useful tool in distributional studies because they are diagnostic, recognizable, and recorded nationwide.

At the state level, Seeman and Prufer (1982) conducted intense Paleoindian distributional studies in Ohio while Largent et al. (1991) documented the distribution and characteristics of Folsom Points in Texas. Regionally, Ronald Mason was one of the first people to note the distribution, varieties, and frequencies of fluted-points in Eastern North America (Shott 2002) Mason and Quimby led distributional studies in the Midwest (Shott 2002) and Blackmar (2001) analyzed the distribution and variability of Clovis, Folsom, and Cody points in Kansas, Oklahoma, and Texas. On a larger scale, Anderson and Faught (1998, 2000a, 2000b) compiled the first nationwide fluted-point sample documenting Paleoindian fluted-point distributions in the lower forty-eight states and Morrow and Morrow (1999) examined the geographic variation of projectile point morphology in North and South America.

It has been suggested that when we find many fluted-points, it is an indication of past Paleoindian activity (Shott 2002). The distribution of projectile points over broad geographic areas might yield important insights about Paleoindian settlement patterns and history (Faught and Anderson 2000). As noted by other researchers, (Shott 2002; Anderson and Faught 2000; Morrow and Morrow 1999; Lepper 1983) these insights into Paleoindian settlement patterns and history include mode and tempo of migration, human-animal interaction, types of resources exploited and which regions were preferred by Paleoindians.

Because distributional studies of fluted-points found and recorded are often used to make broad generalizations about past Paleoindian behavior, it is important for anthropologists to acknowledge the possibility that biases in these studies may exist and could therefore provide a skewed picture past Paleoindian behavior. Potential factors that may bias the number of fluted-points found and recorded are indicated in Table 1.1.

Table 1.1. Possible Biases of Fluted-Points Found and Recorded

Potential biases
1. Modern population density
2. Cultivation (plow zones)
3. Geologic visibility
4. Collector efforts
5. Survey Effort
6. Urbanization
7. Amount of ground cover

Anderson and Faught (2002) noted that surface visibility, amount of prior collection, amount of prior research, and differential erosion may bias the fluted-point sample. Specifically, Faught (1996) noted that Eastern states exhibit more exposure to relevant Pleistocene surfaces and Western states are more concealed by alluviation in drainage channel systems where Paleoindians may have left evidence of their activities.

Blackmar (2001) examined projectile point distrbibutions in Kansas, Oklahoma, and Texas to help determine Paleoindian land use. She provides the examples of the Aubrey, Domebo, and Mclean sites as those covered by thick alluvial fills while extensive sand dunes around playa basins in the high plains may also decrease archaeological visibility. She also addresses the notion that if distributional studies are biased, they may tell us more about archaeological sampling and reporting, amount of ground cover, or modern land use rather than actual patterns of Paleoindian land use. Anderson and Faught (2002) note that while these factors influenced the fluted-point sample, it appears that Paleoindians still utilized some areas and avoided others. However, they did not demonstrate or test this assertion.

Shott (2002) explored the possibility that modern population density and cultivated acreage could bias the fluted-point sample. He looked at these factors in seven Midwestern states using the premise that more people per area and more landscape exposure from cultivation, specifically plowing, would increase the chances of finding fluted-points. His research indicated a high association between modern population density and the fluted-point count and a negative association between cultivated acreage and fluted-point count (Shott 2002).

Shott (2002) statistically determined that the fluted-point sample was biased by modern population density but not cultivated acreage in seven Midwestern states. Lepper (1983) also used statistical analysis to determine if population density and/or area of land under cultivation influenced the amount of fluted-points found and recorded in Indiana, Kentucky, Michigan, Ohio, Tennessee, and West Virginia. Through his research, Lepper (1983) found a statistically significant correlation between cultivation and fluted points recorded and an even stronger correlation between modern population density and fluted-points recorded.

The purpose of this research is to explore through statistical analysis the possibility that the fluted-point sample compiled by Anderson and Faught (2000b) is biased by both factors: modern population density and/or cultivation. It was the explicit expectation at the outset that where higher modern population densities occur, more fluted-points have been found and recorded. Likewise, I also expected that the greater the amount of land under cultivation, the more fluted-points have been found and recorded. The reason for choosing modern population density and cultivation is quite simple. These are variables that others have tested on smaller samples and it seemed reasonable to test the newly compiled database holistically.

The following maps (Figures 1.1-1.5) illustrate densities of fluted-points, modern population, and area of land under cultivation in the lower forty-eight states.

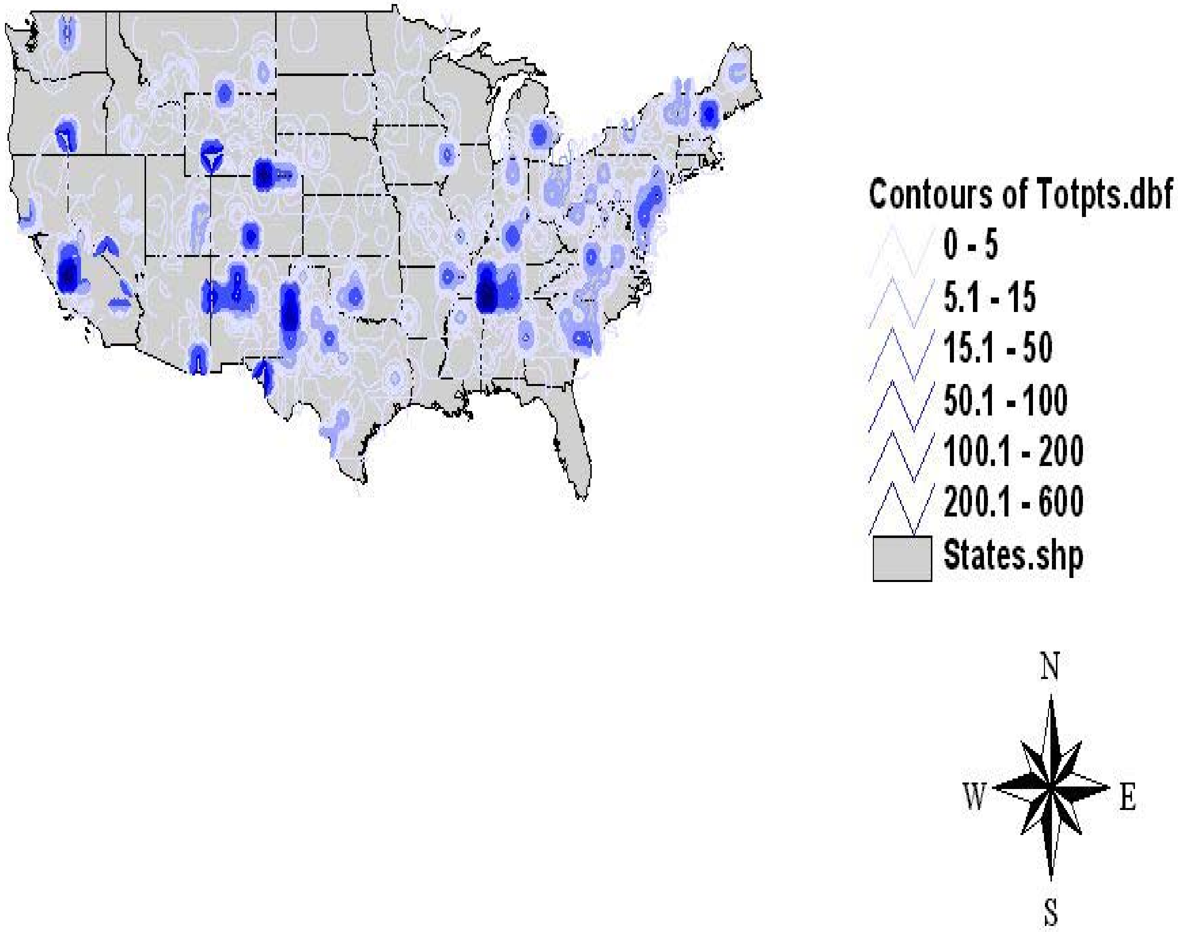

Figure 1.1 Contours of Total-Fluted Points Per County

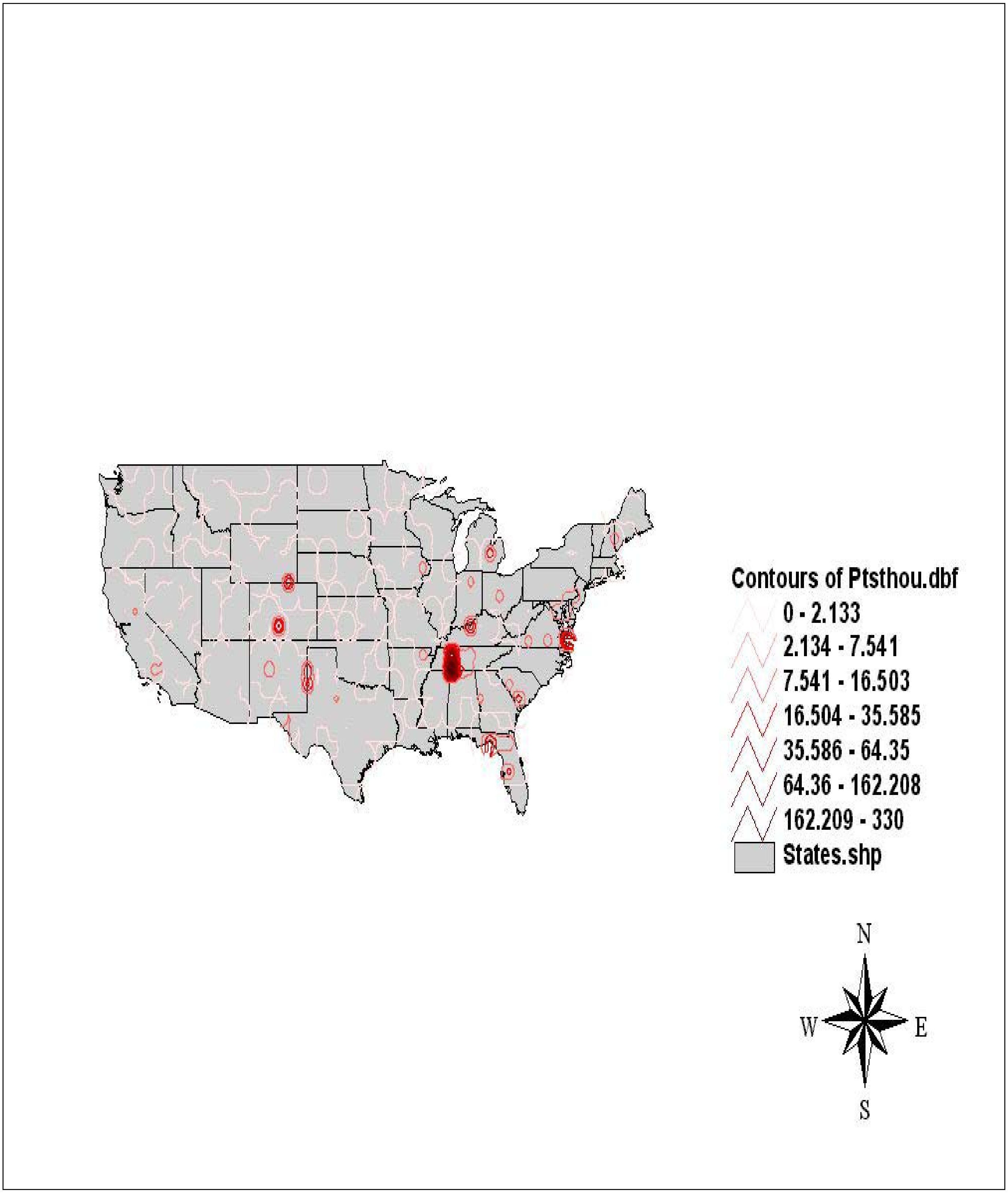

Figure 1.2 Contours of Total-Fluted Points Divided by County Area x1000.

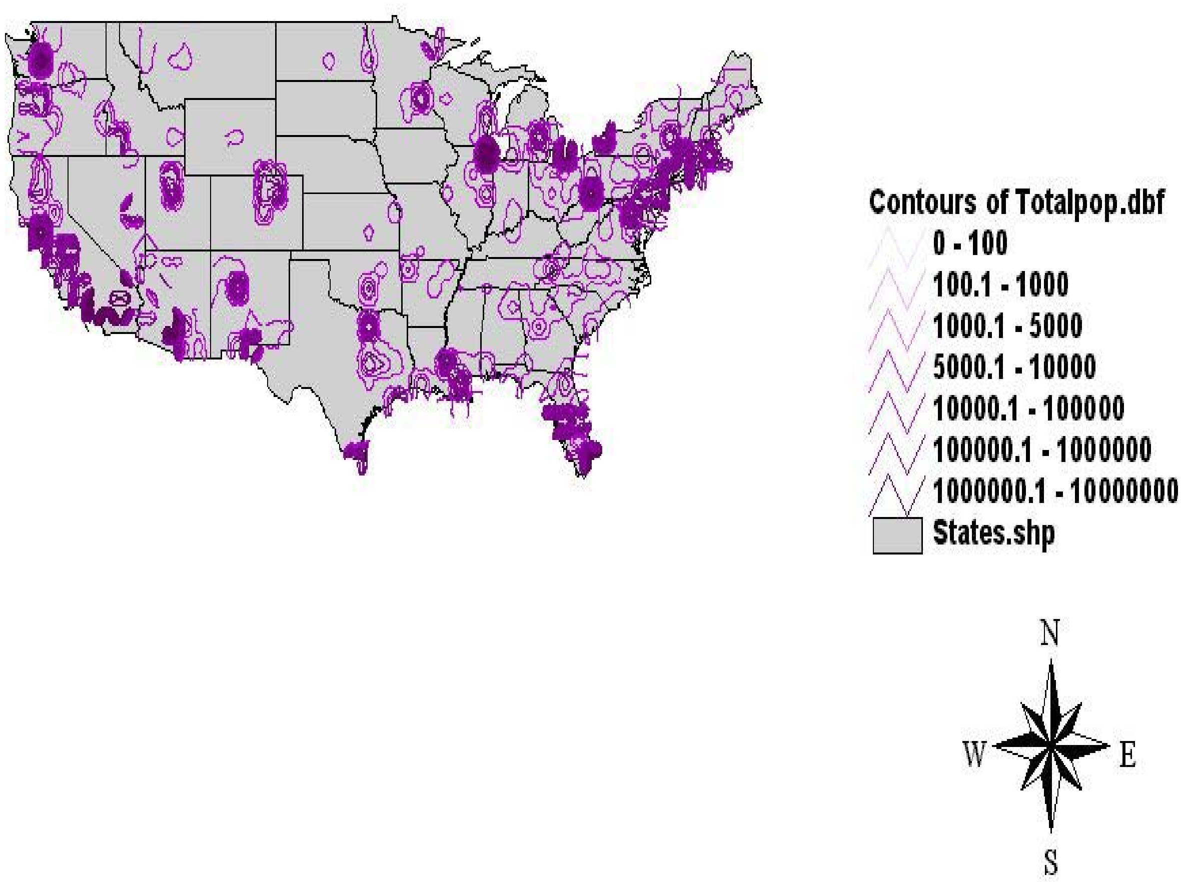

Figure 1.3 Contours of Total Population

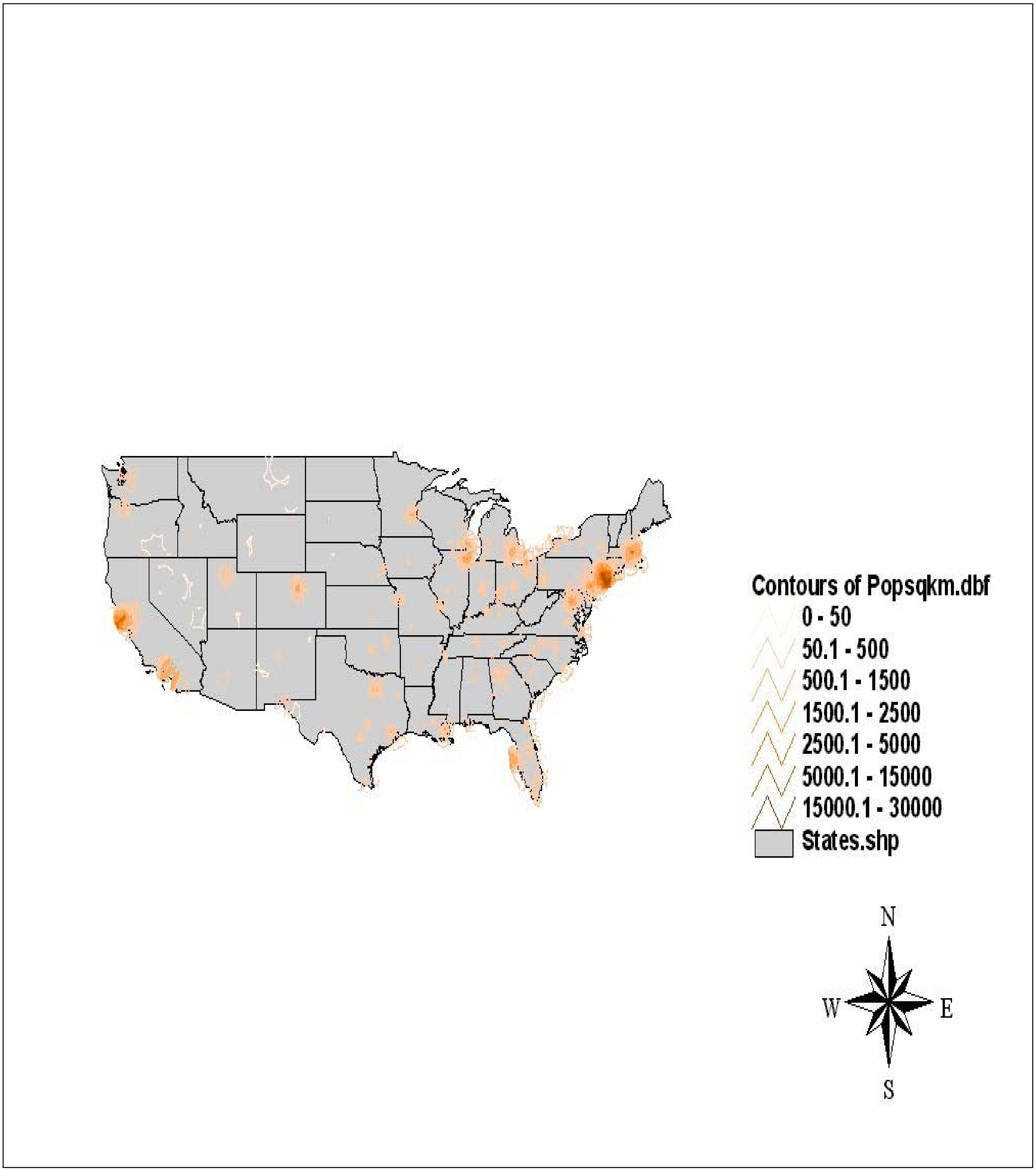

Figure 1.4 Contours of Population Per Square Kilometer

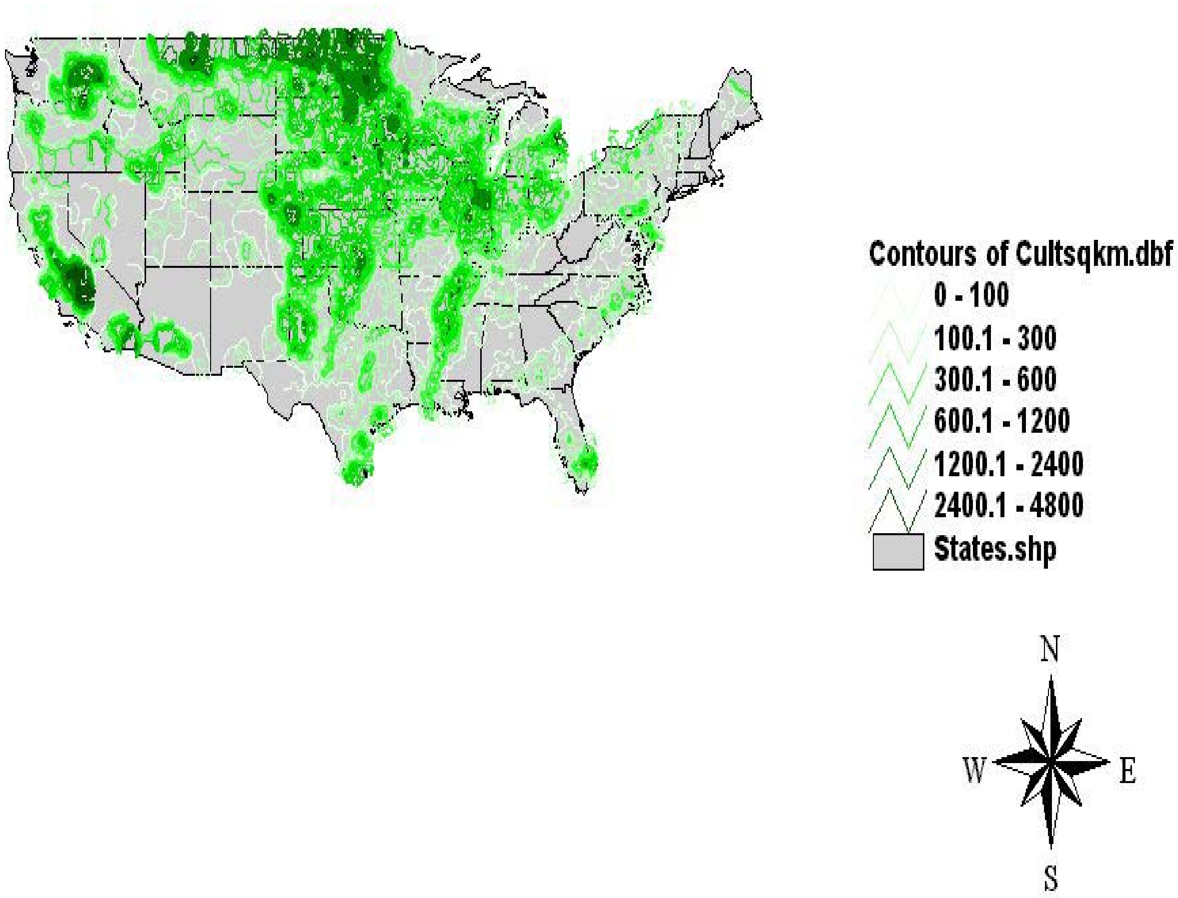

Figure 1.5 Contours of Cultivation Per Square Kilometer

In order to determine if modern population density and/or cultivation bias the North American fluted-point sample, a variety of statistical analyses were employed. As demonstrated by previous research (Shott 2002; Blackmar 2001; Lepper 1983), statistical analysis can be a useful tool in the evaluation of bias in fluted-point distributional studies. Within the past two decades there has been a notable increase in the application of statistical analyses to anthropological problems (Thomas 1986).

Chapter 2 discusses the increased use of statistics in anthropology. It includes discussion of other researchers who have applied statistical analysis to distributional studies to determine if they are subject to biases.

Chapter 3 discusses the tools and methods used to acquire the data pertinent to this research. In addition, it discusses the methods used to construct the database and the statistical tests used to analyze the data.

Chapter 4 presents the results of the statistical analysis used in the research. This includes both descriptive and inferential statistics. The inferential statistics consisted of both parametric and non-parametric methods.

Chapter 5 discusses the conclusions reached by the author as a result of the statistical analysis.

CHAPTER 2

STATISTICS AND ANTHROPOLOGY

A working knowledge of statistical inference has become virtually essential for anthropologists and particularly valuable for archaeologists working with large databases. Anthropologists such as Shott (2002), Blackmar (2001), Largent et al. (1991) and Lepper (1983) have used statistical analysis to determine meaningful patterns of Paleoindian occupation and land use.

Professional journals provide an excellent measure of significant trends and patterns within a discipline (Thomas 1986). To illustrate this point, Thomas (1986) conducted a survey in which a major journal was selected from each sub field of Anthropology (*American Journal of Physical Anthropology, American Antiquity, Language, and American Anthropologist*). The results of the survey show that between 1900 and 1970, the percentage of articles in major journals dealing with inferential statistics steadily increased in each of anthropology's subfields (Thomas 1986).

Thomas'(1986) findings illustrate the increasing importance of statistical thinking in anthropology and the relative application of quantitative procedures by anthropology's sub-fields. Both inferential and descriptive statistics are presented in this study.

Data can hardly be destroyed once they are generated. A society may become extinct, an archaeological site may be bulldozed, or a skeletal series may be lost or destroyed but all these things occur to people, objects, or things, not to data (Thomas 1986). This is why it is essential that databases, such as those compiled by Anderson and Faught are created. Even if the artifacts, in this case, fluted-points, do become unavailable for analysis, at least they have been documented and preserved in the data.

The underlying structure of data is subject to certain explicit assumptions when statistical tests are applied. Also, data generally require some minimal level of measurement. The relationship between scale of measurement and statistical test is a complex topic (Thomas 1986). For example, Shott (2002) used counties as his scale of measurement and Analysis of Variance (ANOVA) as his statistical test. Blackmar (2000) used physiographic provinces as well as counties as her scale of measurement and used ubiquity analysis. Lepper (1983) used only physiographic provinces as his scale of measurement and the Spearman's Rank-Order association as his statistical test.

A variety of statistical analyses can be applied in attempts to solve similar anthropological problems, in this case, the identification of past Paleoindian cultures through the analysis of fluted-point distributions. There are many factors that effect the decision of what statistical test is appropriate for a given situation. In the process of determining if modern population density and/or cultivated square kilometers bias the fluted-point sample, I used a variety of statistical analyses including, Pearson's Correlation Coefficient, Spearman's rho, one-way Analysis of Variance (ANOVA), and Kruskal-Wallis. The reasons for choosing these particular tests and their results will be discussed later.

One characteristic of most statistical populations is that they are incompletely observable. Populations of variates must usually be estimated from a small subset of the actual statistical population (Thomas 1986.)

We will never know the true number of fluted-points produced by Paleoindians. For the purposes of this thesis, the data compiled by Anderson and Faught (2000b) is all those fluted-points that have been found and recorded. Faught (1996) notes that distributional maps probably represent a statistical estimate of "real" fluted-point numbers. We cannot measure, photograph, and observe all projectile points. In fact, we cannot even collect all the information on those projectile points that have been discovered either by collectors or professionals. There are a variety of reasons for this but most notably Shott (2002) points out that many fluted-points are reported in unpublished records, many in obscure publications, and many never reported at all. In this case, the data provided by Anderson and Faught (2000b) represents the best sample of those fluted-points found and recorded within the lower forty-eight states.

CHAPTER 3

METHODOLOGY

The null hypothesis (H0) of this research is: “Fluted-points found and recorded are not affected by modern population density and/or cultivation.” The alternative hypothesis (H1) proposes that fluted-points found and recorded are affected by modern population density/and or cultivation.

This chapter outlines the methodology used either to accept or reject the null hypothesis. First, I discuss the methods, tools, and sources used to acquire the data relative to this research. Second, I present the methods used to construct a database of variables relevant to this research. Last, I provide a description of each statistical test employed, including strengths and weaknesses, assumptions, and the reason each was deemed appropriate to analyze the acquired data.

The fluted-point dataset in this study consists of point counts by county compiled by Anderson and Faught (2000b). Population is measured directly from the most current census records (U.S. Census Bureau 2002) and land use is measured as cultivated square kilometers as of 1997 (U.S. Department of Agriculture 1999). Modern population is measured as people per square kilometer for each county in the lower forty-eight states. Cultivation is measured as cultivated area in square kilometers for each county. Fluted-point counts were measured as both total numbers of fluted-points per county, as well as total number of fluted-points per county divided by the county area (points/area).

Anderson and Faught’s (2000b) data is the most robust and recent sample. I am in agreement with Shott (2002) that these data are easily accessable, available to all, and are reasonably accurate and a reliable source for this research. Census data used are nearest to the time of this publication, because, as Shott (2002) has observed, population differs between censuses, but values are highly correlated between censuses 20 and more years apart. The agricultural data are those produced nearest to the time of this publication.

Most researchers (Shott 2002; Blackmar 2001; Lepper 1983) have used one of three units of analysis either independently or in combination to count fluted-points: physiographic provinces, counties, or a combination of physiographic provinces and counties. There are various reasons to correlate fluted-point distribution using physiographic provinces as a unit of analysis, however the small number and large size of provinces tends to hamper statistical analysis (Shott 2002).

Counties on the other hand are a good unit of analysis and county level data are presented in this research. Counties are small in size relative to the study areas (the lower forty-eight states, individual regions, individual states, and those states east and west of the Mississippi); counties are many in number; and counties are arbitrary in location. (Shott 2002). Hence, counties were chosen as the unit of analysis in this research.

To determine if the numbers of fluted-points found and recorded correlate with modern population density and/or the amount of exposed landscape, in this case cultivation, I compiled a database in Microsoft Excel® 2000 that included all counties in the United States except Hawaii and Alaska.

From Anderson and Faught (2000b), each county was matched with its fluted-point count. In addition, I also divided the number of points in each county by the county area. Dividing the number of points in each county by the county area controls for difference in county size and creates the variable "points per unit area." For example, county sizes in states west of the Mississippi tend to be much larger and less numerous than counties in states east of the Mississippi.

To make all variable measurements uniform, the units of measurement were converted to square kilometers. For example, the area under cultivation in each county is in square kilometers, population is measured as people per square kilometer, and the area of each county is in square kilometers.

After compiling the database in Microsoft Excel® 2000, I transferred the data to Systat® 10.0. With the Systat® 10.0 program, I was able to analyze the data with both descriptive and inferential statistical methods. The inferential statistical methods used in this research include the Pearson Correlation Coefficient, Spearman's rho, one-way Analysis of variance (ANOVA), and the Kruskal-Wallis test.

In statistics, the probability of rejecting the null hypothesis is known as the power. The level of statistical significance used to reject the null hypothesis in each of the aforementioned statistical tests performed was 0.05. This rejection level is known as the alpha (α).

Pearson's Correlation Coefficient is a parametric test statistic and is the most common measure of association (refer to McClave and Dietrich II 1994 for the Pearson's Correlation Coefficient computational formula). It reflects the degree of linear relationship between two variables (Agresti and Findlay 1997). The non-parametric Spearman's rho is also used to assess the linear relationship between two variables (Agrestin and Findlay 1997). However, Spearman's rho uses ranks rather than raw data. Rank correlations are used to measure the correlation between any pair of variables (Agresti nd Findlay).

If two variables are measured in an experiment, one would rank the measurements for each variable seperately. For example, consider the variables: total fluted points per county and cultivated square kilometers per county. To determine the relationship between the variables, the number of fluted-points per county are ranked, assigining a 1 to the smallest number (0) and a 30 to the highest number (489). Likewise, the number of cultivated square kilometers in the same county are ranked. The rank measurement for fluted-points per county is subtracted from the rank measurement for cultivated square kilometers to determine the correlation between the variables.

Parametric statistics such as the Pearson Correlation Coefficient, make specific assumptions with regard to one or more population parameters that characterize the underlying distribution for which the test is employed (McClave and Dietrich II 1994). It is a test that assumes the population has a normal distribution and involves hypotheses about population parameters (McClave and Dietrich 1994). If we have a basic knowledge of the underlying distribution of a variable, we can make predictions about how, in repeated samples of equal size, this particular statistic will be distributed (McClave and Dietrich II 1994).

Non-parametric statistical tests such as the Spearman's rho do not make assumptions about underlying population distributions (refer to McClave and Dietrich II 1994 for the Sparman's rho computational formula). Non-parametric methods are used in cases when the researcher knows nothing about the parameters of the variable of interest in the population(Agresti and Findlay 1997). Non-parametric methods do not rely on the estimation of parameters (such as the mean or standard deviation) describing the distribution of the variable at interest in the population (Agresti and Findlay 1997).

Usually, non-parametric methods are good for small samples. However, in large samples, if the population follows a normal distribution, the P value will be nearly identical to the P value that would be obtained from a parametric test. However, non-parametric statistics are not as statistically as powerful as parametric statistics. Non-parametric statistics can be used when underlying assumptions for a parametric method have been violated. Parametric interval /ratio scale data can be easily transformed into non-parametric ordinal scale data. When the data are transformed, the counterpart non-parametric method can be used (Rice 1995). For example, if the Pearson Correlation Coefficient violates an assumption (the population does not have a normal distribution), the data can be transformed from interval/ratio in scale into ordinal data and the counterpart Spearman's rho can be used.

It is important to note here the difference between parametric and non-parametric statistical tests. As pointed out previously, parametric tests are not robust measures of correlation in skewed data. Parametric tests assume that a variable is defined on an interval or ratio scale and that the sample population distribution exhibits a normal curve (Agresti and Findlay 1997). A statistic is considered non-parametric if the variable is ordinal or nominal in scale and if the distributions for the sample populations do not assume normality.

Although parametric statistics such as Pearson's Correlation Coefficient are effective for measuring variables that are interval in scale, parametric statistics are not robust measures of association in skewed data (Shott 2002). As suggested by Shott (2002), another way to seek patterning is to reduce the variables from ratio or interval to ordinal to measure association, not correlation. In an attempt to find patterning, modern population density and cultivated square kilometers were placed in ordinal classes.

Fluted-points were not placed in ordinal classes for statistical analysis, however they were transformed. Population per square kilometer and cultivation per square kilometer were also transformed prior to being placed in ordinal classes. Logarithmic transformation changes raw data to a percentage of the absolute number in order to reduce error. Because log transformations of fluted-points, modern population density, and cultivated square kilometers are roughly normal in distribution, intervals were defined by the standard deviation of log-transformed variables.

In Shott's (2002) analysis, log population was roughly normal in distribution so his intervals were also defined by the standard deviation of the log-transformed population. However, cultivated acreage per county was roughly normal in distribution, so the standard deviation of the untransformed variable was used.

Following Shott (2002), I created four classes for modern population density and cultivation. This made the variables ordinal. For each variable, cases whose value is less than or equal to the mean minus one standard deviation (≤ mean-1 standard deviation) represents the first class, cases whose value is from the mean minus one standard deviation (mean-1 standard deviation) represent the second class, those from the mean to mean plus one standard deviation (mean to mean+1 standard deviation) represent the third class, and those greater or equal to the mean plus one standard deviation (≥ mean+1 standard deviation) represent the fourth. Tables 3.1 and 3.2 show each variable, the range for each class, and the number of cases (counties) in each class.

Table 3.1 Ordinal Classes for People Per Square Kilometer

Class	Range	Number
1	≤1.069 people per sq km	462
2	≥1.070 and ≤2.728 people per sq km	1031
3	≥2.729 and ≤4.387 people per sq km	1167
4	≥4.388 people per sq km	430

Table 3.2 Ordinal Classes for Cultivated Square Kilometers

Class	Range	Number
1	≤3.747 cultivated sq km	462
2	≥3.748 and ≤5.247 cultivated sq km	908
3	≥5.248 and ≤6.747 cultivated sq km	1179
4	≥6.748 cultivated sq km	538

After the ordinal classes for each variable were created, a one-way Analysis of Variance (ANOVA) that included the Tukey post hoc test was performed (refer to Rice 1995 for the one-way ANOVA computational formula). One-way ANOVA compares the values of more than two groups. The one-way ANOVA was used because there was only one dependent variable (fluted-points) with two categories (modern population density and cultivated square kilometers).

Post hoc tests determine which pairs of means differ significantly. The Tukey post hoc test is effective when testing a large number of pairs of means (Systat 2000). It looks at differences in means located in the ordinal classes (1-4) table. The Tukey pairwise comparison is a post hoc test used for samples that differ significantly in size.

One-way ANOVA as well as other parametric statistical tests assume that one has sampled data from populations that follow a Gaussian (bell shaped) distribution. One-way ANOVA works well even if the distribution is only approximately Gaussian (especially with large samples such as this one).

An alternative approach does not assume that data follow a Gaussian distribution. With this approach, values are ranked and the analyses are based on the distribution of the ranks. These approaches are non-parametric tests. The non-parametric equivalent of the one-way ANOVA is the Kruskal-Wallis one-way Analysis of Variance (refer to Rice 1995 for the Kruskal-Wallis one-way ANOVA computational formula).

Although the Kruskal-Wallis test does not assume normality of the distributions for the sample populations, it does assume that the populations have the same distribution, except for a possible difference in population medians. The Kruskal-Wallis test is also effective in analyzing data when outliers are suspected, even if the underlying distributions are close to normal. In the data analyzed here, outliers may be contributing to the skewness of the data.

An outlier is an observation that falls far from the rest of the data and can highly influence the mean (Agresti and Finlay 1997). However, the median is highly resistant to outliers (Drennan 1996). As the one-way ANOVA tests for differences in means and can be affected by outliers, the Kruskal-Wallis tests for differences in population medians and should be less effected by outliers.

As noted previously, the North American fluted-point database served as the primary data source for this research. The database consists of fluted-point counts per county and was subjected to statistical analyses as such. All counties were included in the analyses regardless if they had zero fluted-points or a many fluted-points. For example, in Shott's (2002) analysis, northern Minnesota, northern Michigan, and northern Wisconsin were excluded because few fluted-points have been found and recorded in those areas. In this research, no counties were excluded. This study simply attempts to find the best statistical method to analyze the data presented, with no exclusions.

In an effort to better understand the effects of modern population density and cultivated square kilometers on fluted-points found and recorded in the United States, the U.S. was divided into four separate regions: Northeast, Midwest, South, and West (see Figure 3.1 and Table 3.3). The regions and the states within each are defined by the U.S. Census Bureau (2002).

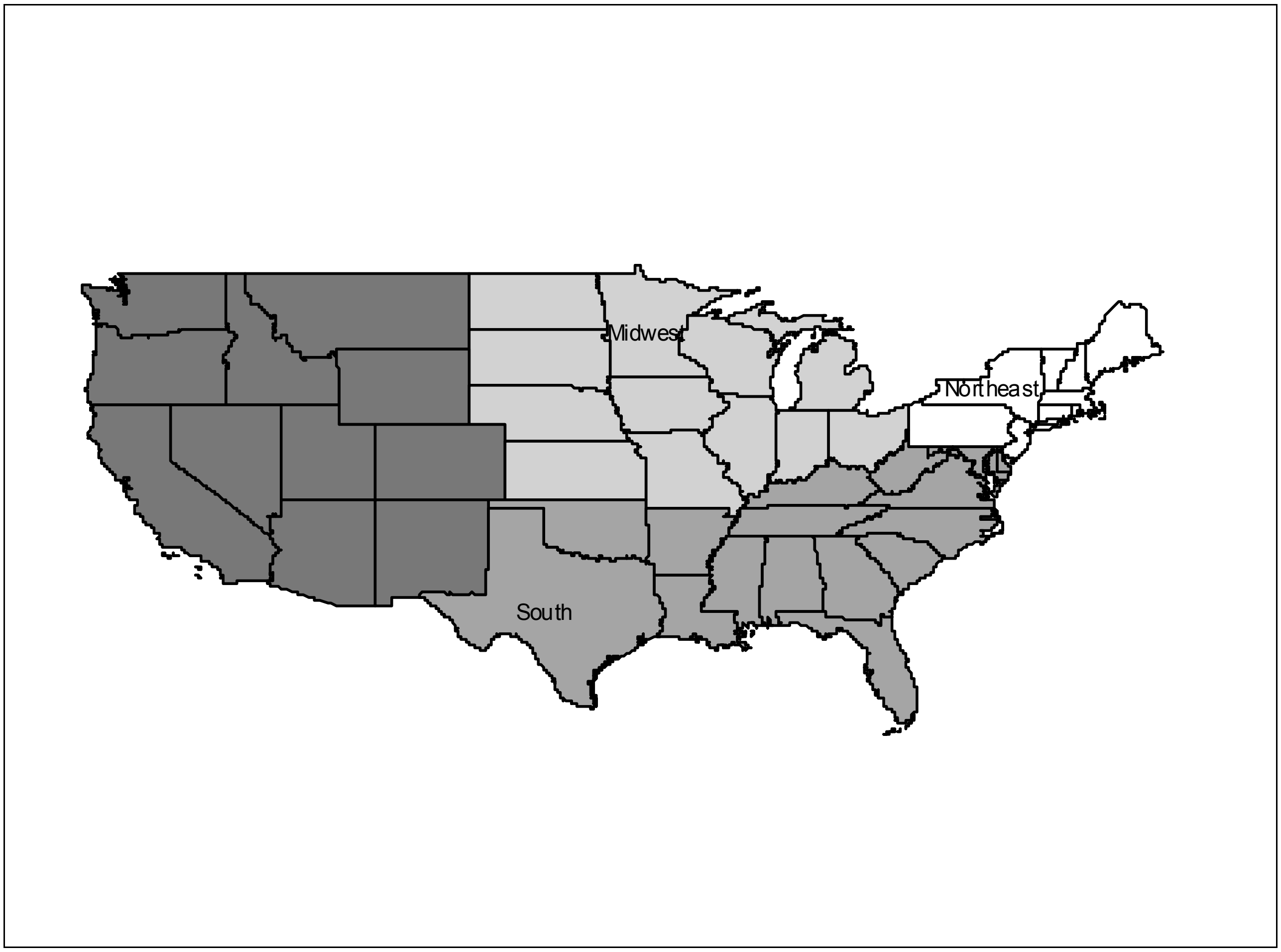

Figure 3.1 Regional Groupings

Table 3.3 **Region Composition**

Northeast	Midwest	South	West
Connecticut	Illinois	Alabama	Arizona
Maine	Indiana	Arkansas	California
Massachusetts	Iowa	Delaware	Colorado
New Hampshire	Kansas	District of Columbia	Hawaii
New Jersey	Michigan	Florida	Idaho
New York	Minnesota	Georgia	Montana
Pennsylvania	Missouri	Kentucky	Nevada
Rhode Island	Nebraska	Louisiana	New Mexico
Vermont	North Dakota	Maryland	Oregon
	Ohio	Mississippi	Utah
	South Dakota	North Carolina	Washington
	Wisconsin	Oklahoma	Wyoming
		South Carolina	
		Tennessee	
		Texas	
		Virginia	
		West Virginia	

This chapter has discussed the methodology by which the question: "Do fluted-points found and recorded depend on modern population density and/or area of land under cultivation?" will be addressed It is my intention to make other researchers aware of the possible bias or biases in distributional studies. The following chapter reports the results outlined by this methodology.

CHAPTER 4

RESULTS

In this chapter, the following combinations of data were analyzed in an attempt to find statistically significant associations: modern population and total fluted-point counts, modern population and fluted-point counts divided by county area (points/area), cultivation and total fluted-point counts, and cultivation and fluted-point counts divided by county area (points/area).

Both parametric and non-parametric tests were applied to the United States overall, each region individually, and states east and west of the Mississippi. By applying statistical analysis to individual regions, a clearer picture of the fluted-point distribution in the United States has emerged.

Descriptive Statistics

Descriptive statistics describe a sample by summarizing raw data. This includes measures of central tendency (the value around which much of the sample is distributed) and dispersion (how the sample is distributed around the central tendency value) such as the sample mean and standard deviation. Table (4.1) illustrates the descriptive statistics for the raw data compiled in this research.

Table 4.1 Non-Transformed Descriptive Statistics

Population per Square Kilometer	
Number of Counties	3089
Minimum Number	.020 people per square kilometer
Maximum Number	26977 people per square kilometer
Median	16 people per square kilometer
Mean	82 people per square kilometer
Cultivation per Square Kilometer	
Number of Counties	3055
Minimum Number	0 cultivated square kilometers
Maximum Number	3848 cultivated square kilometers
Median	231cultivated square kilometers
Mean	460 cultivated square kilometers
Total Points per County	
Number of Counties	3091
Minimum Number	0 points
Maximum Number	489 points
Median	0 points
Mean	4 points
Total Points Divided by County Area	
Number of Counties	3091
Minimum Number	0 points
Maximum Number	9.8 points
Median	0 points
Mean	.05 points

These frequency distributions indicate that the number of fluted-points per county is highly skewed (Figure 4.1). As shown by Shott (2000), many more counties than expected have zero fluted-points or only one, and many fewer counties have more than one. This makes statistical analysis difficult.

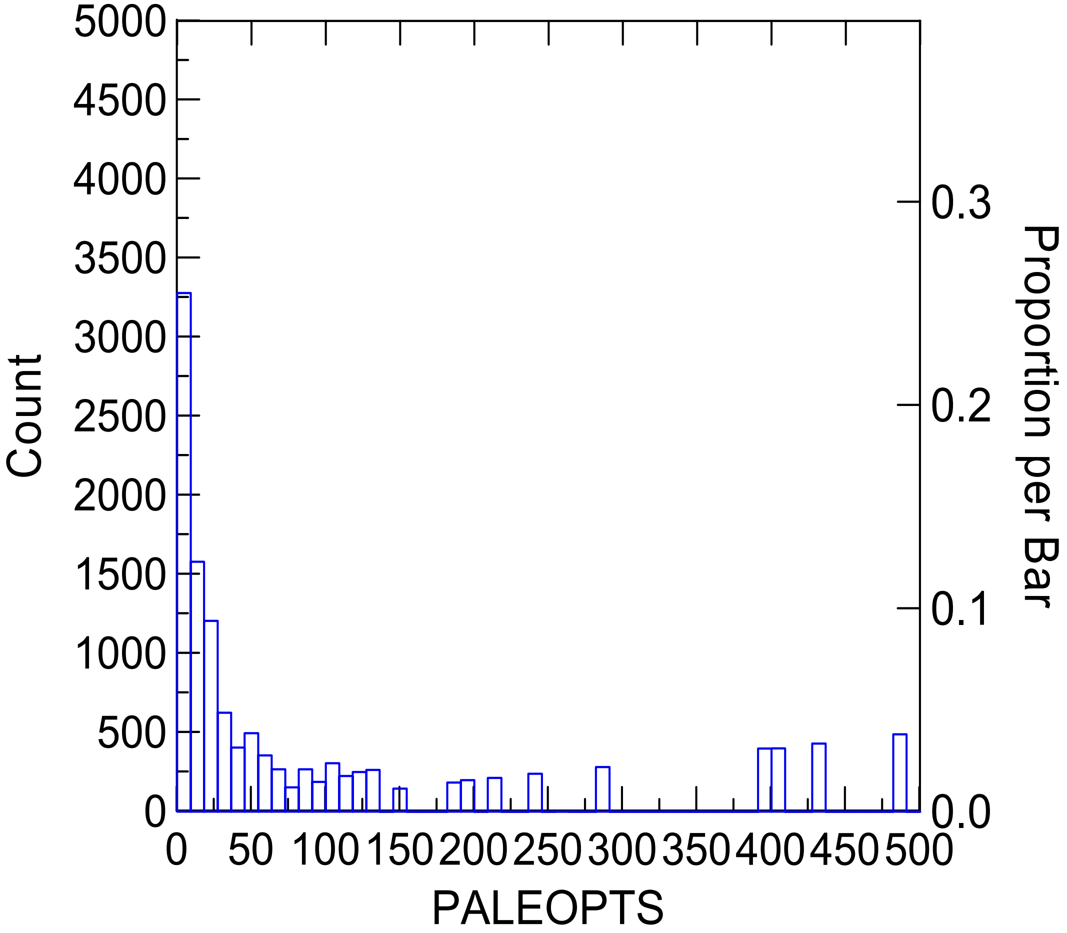

Figure 4.1 Non-transformed Frequency Distribution for Total Fluted-points per County

Modern population density, cultivated square kilometers per county, and fluted-points divided by county area are also skewed as shown in Figures 4.2-4.4.

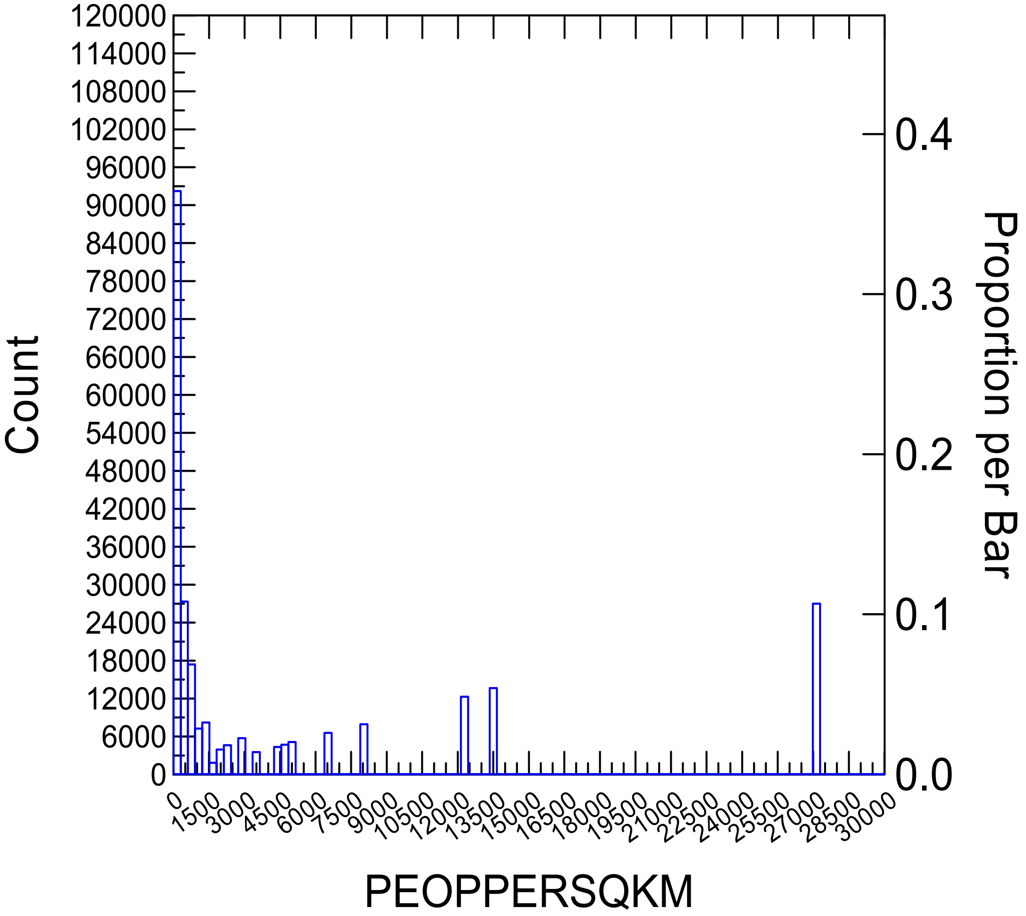

Figure 4.2 Non-transformed Frequency Distribution for Modern Population

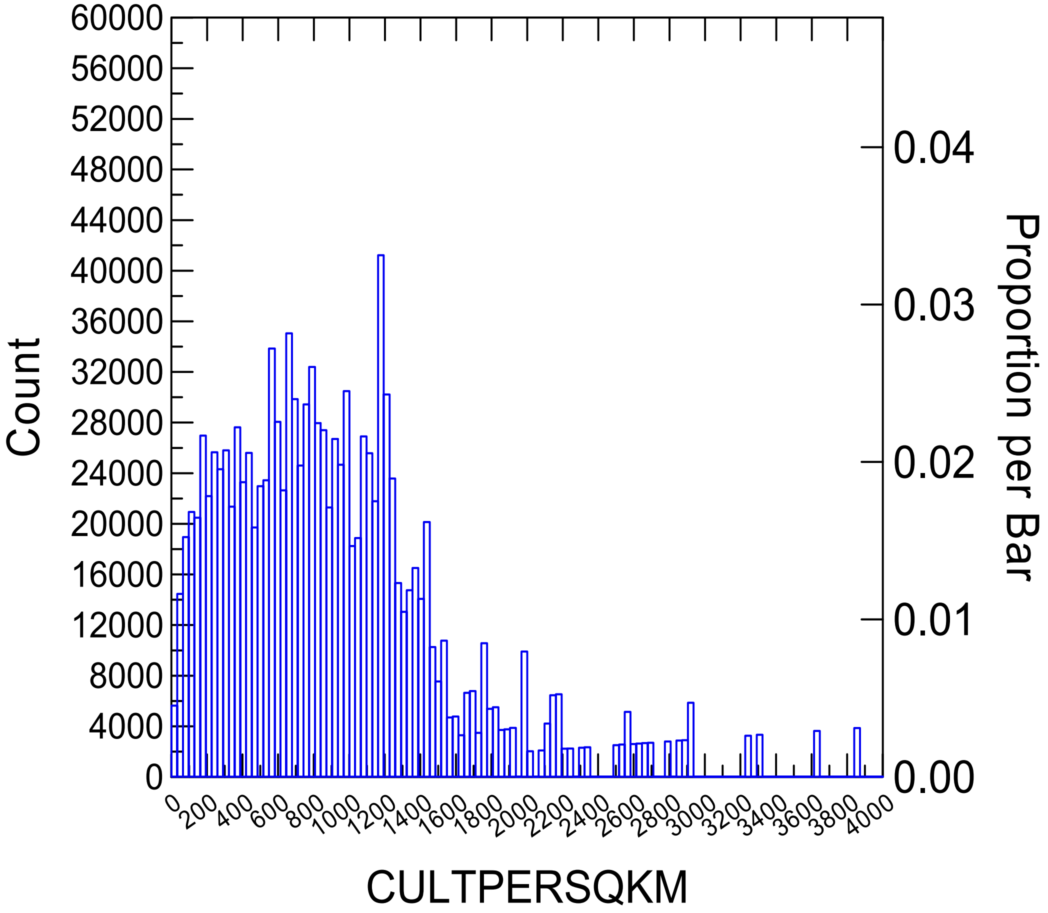

Figure 4.3 Non-transformed Frequency Distribution for Cultivation

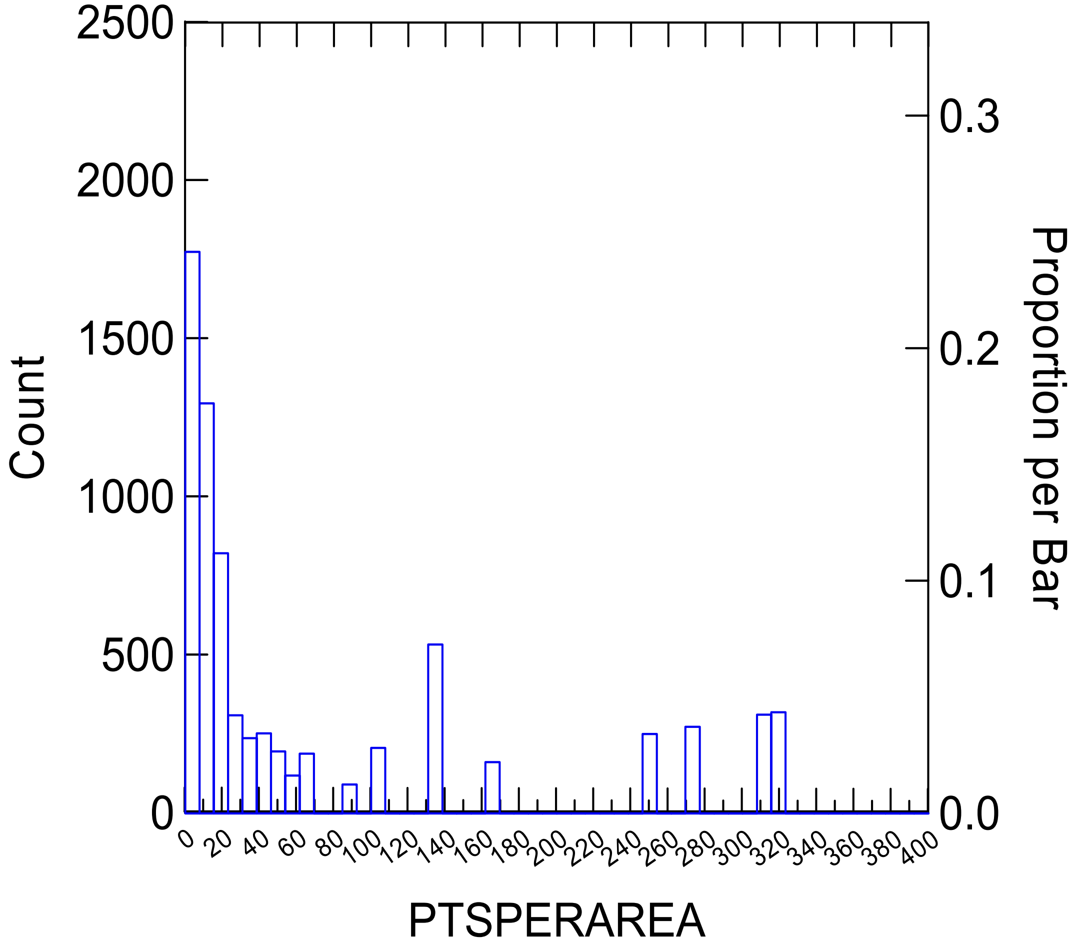

Figure 4.4 Non-transformed Frequency Distribution for Total Fluted-points Divided by County Area

Because the raw data were highly skewed, they were log transformed before any further analysis. Table 4.2 summarizes the log transformed statistics. Log transformation rescales the raw data and makes it relative rather than absolute (Buchanan 2003). In log transformations, nonlinear relationships are mathematically converted into linear proportions (Thomas 1986). The non-transformed variables of total fluted points, modern population density, and cultivated square kilometers per county were skewed to the right. However, as shown in Figure 4.5, the log-transformed variable of people per sqaure kilometer represents a normal distribution. Total fluted-points per county also shows a more normal distribtuion when log transformed (Table 4.6). The log-transformed variable of cultivated sqaure kilometers per county is represented in Table 4.7.

Table 4.2 Transformed Descriptive Statistics

Population per Square Kilometer	
Number of Counties	3089
Minimum Number	3.9 people per square kilometer
Maximum Number	10.2 people per square kilometer
Median	2.8 people per square kilometer
Mean	2.7 people per square kilometer
Cultivation per Square Kilometer	
Number of Counties	3026
Minimum Number	3.4 cultivated square kilometers
Maximum Number	8.5 cultivated square kilometers
Median	5.5cultivated square kilometers
Mean	5.2 cultivated square kilometers
Total Points per County	
Number of Counties	1394
Minimum Number	0 points
Maximum Number	6.2 points
Median	1.1 points
Mean	1.2 points

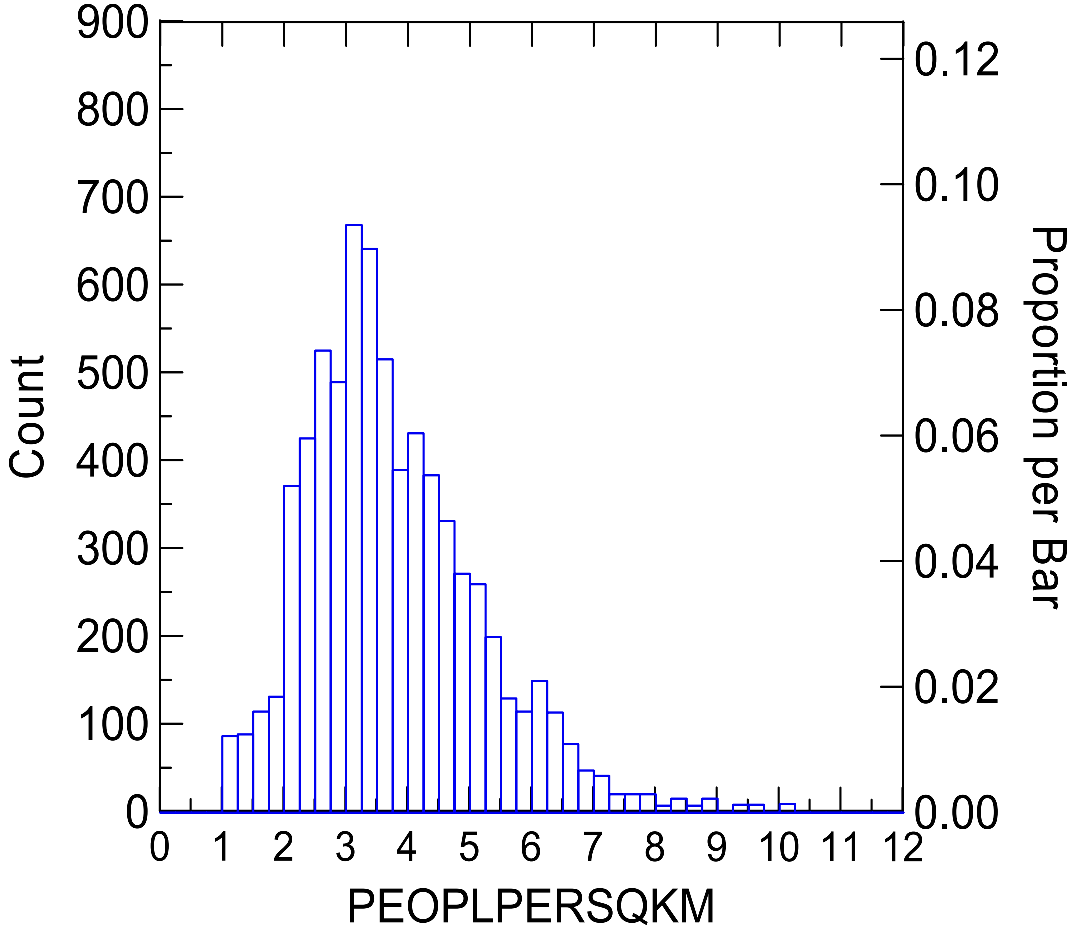

Figure 4.5 Log Transformed Frequency Distribution for People Per Square Kilometer

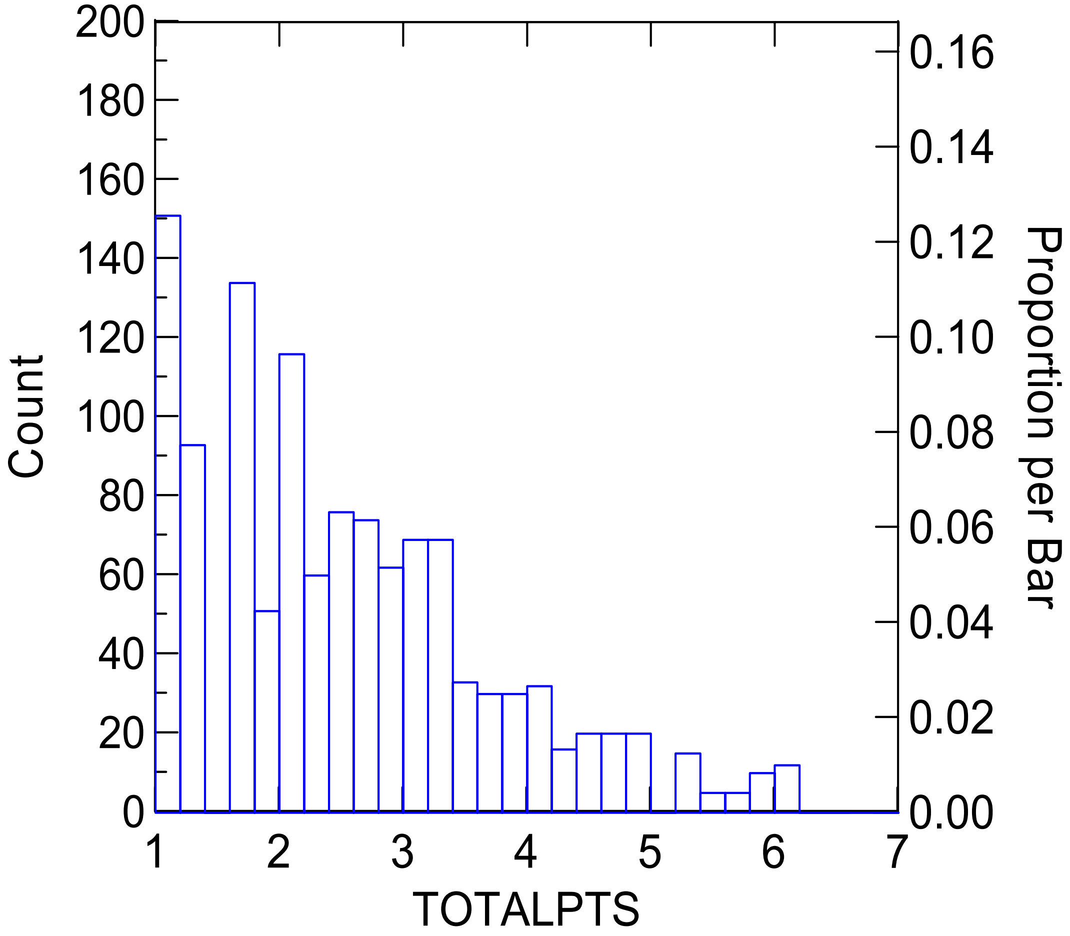

Figure 4.6 Log Transformed Frequency Distribution for Total Fluted-points per County

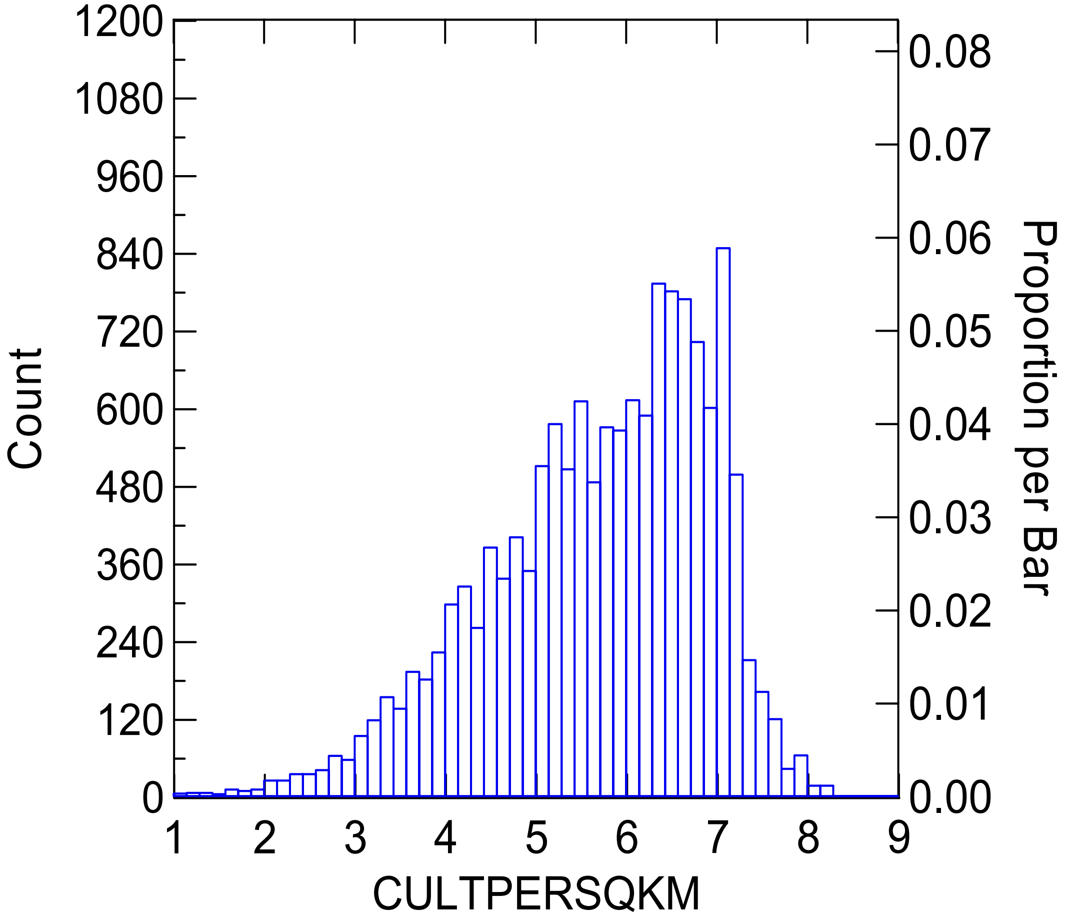

Figure 4.7 Log Transformed Frequency Distribution for Cultivation

Shott (2002) also found that modern population density was non-parametric. However, in contrast to the test results reported here, Shott (2002) found that cultivated land per county was roughly normal in distribution. It has been noted by Shott (2002) and Lepper (1983) that in the Eastern United States, archaeological visibility is primarily a function of landscape exposure, specifically cultivation. This seems to support a normal distribution of cultivation per county in Shott's study.

Because the raw data were highly skewed, the data were log transformed in an effort to establish normal distribution. Inferential statistics were applied to the log transformed data to make inferences about the population of total fluted-points from which the sample was drawn.

Inferential Statistics

As stated in the previous section, inferential statistics use sample data to make inferences about the population from which the sample is drawn. Specifically, in this research, it is expected that inferential statistics will provide better insight into the relationship between fluted-points and modern population density as well as fluted-points and cultivation.

Pearson's Correlation Coefficient

When evaluating the Pearson Correlation table, Thomas (1986) notes that the correlation coefficient has a number of desirable properties. First, a value of r =0 indicates that no linear relationship exists between the two variables. Second, the magnitude of r denotes the strength of linear relationship. Large values of r indicate a close relationship while smaller values indicate a weak relationship. Third, the sign of r (+/-) indicates the direction of the relationship. For example, if r =-1 or close to it, then as variable (x) increases, variable (y) decreases. If r =1 or close to it, then as variable (x) increases, variable y increases. Finally, the maximum value of r = +1.00 indicates a perfect positive correlation.

Pearson's Correlation Coefficient on the total raw data consisted of the conflated lower forty-eight states (Table 4.3). The results in the table below indicate a negative correlation between cultivation per square kilometer and fluted-points (r =-.008). Because of the weak correlation, there is a low r value. Although there is a positive correlation between population per square kilometer and fluted-points (r =.010), the low r value reflects a weak correlation.

Table 4.3 Pearson Association Matrix on the Untransformed Variables

	Fluted Points	Cultivation per square kilometer	People per square kilometer
Fluted Points	1.000		
Cultivation per square kilometer	-0.008	1.000	
People per square kilometer	0.010	-0.008	1.000

Spearman's rho

The non-parametric Spearman's rho was also used to assess the relationship between the variables (fluted points and population/fluted-points and cultivation) in the conflated lower forty-eight states (Table 4.4). However, the Spearman's rho uses ranked data rather than raw data. The Spearman's rho test indicated a negative relationship between cultivation and fluted points per county (r_s =-.016). The relationship is weak as indicated by the small value of r_s. There is a positive relationship between modern population density and fluted-points recorded per county (r_s=.256), however, the low value of r reflects a low correlation.

Table 4.4 Spearman's Correlation Matrix on the Untransformed Variables

	Fluted Points	Cultivation per square kilometer	People per square kilometer
Fluted Points	1.000		
Cultivation per square kilometer	-0.016	1.000	
People per square kilometer	0.256	-0.158	1.000

One-way Analysis of Variance

In the following discussion of the one-way ANOVA and Tukey multiple comparison, Tables A.1-A.38 are discussed independently, however, for deatiled results refer to Appendix A.

The one-way ANOVA compares average fluted-point counts for the four different population classes (refer to Tables 3.1 and 3.2). It performs a pairwise comparison of the means to find where the significant difference occurs. Although the one-way ANOVA indicates whether one population class (1-4) differs from another, you cannot determine if there is a statistically significant association from only the one-way ANOVA results. This is where the Tukey post hoc test is used.

The Tukey post hoc test examines specific group mean differences. For example, on the one-way ANOVA table, a difference between population class 1 and population class 2 can be observed. However, one would have to look at the Tukey pairwise comparison to determine if the difference between population class 1 and population class 2 was statistically significant.

One-way ANOVA applied to modern population density and total fluted-points in the lower forty-eight states. The parametric one-way ANOVA indicates there is a statistically significant relationship between total fluted-points recorded and population per square kilometer in the conflated lower forty-eight states, p=0.000, F=10.688 (Table A.1).

The results of the Tukey pairwise comparison analysis are presented in Table A.2. As indicated, only two tests returned significant results. Population classes 3 and 4 had significantly more points than population class 2, but not more than population class 1. In other words, population classes 3 and 4, counties with higher population density, have more fluted-points found and recorded than counties in population class 2, which has a lower population density. However, as shown, the means of population classes 3 and 1, as well as 4 and 1 do not differ significantly. This means one is unable to determine if population density influences the amount of fluted-points found recorded in these population classes.

By region, one-way ANOVA indicates a statistically significant relationship between fluted-points recorded and people per square kilometer in the Southern region, p=0.011, F=3.733 (Table A.3) and Midwestern region, p=0.000, F=17.013 (Table A.5).

The results of the Tukey pairwise comparison analysis for the Southern region are presented in Table A.4. As is indicated, only two tests returned significant results. Population classes 3 and 4 had significantly more points than population class 2, but not more than population class 1. Population classes 3 and 4, counties with higher population density, have more points recorded than counties in population class 2, which has a lower population density. However, as shown, the means of population classes 3 and 1, as well as 4 and 1 do not differ significantly. Again, if the means for each population class do not differ significantly, you can not determine if the independent variable (population density) influences the dependent variable (fluted-points recorded).

Table A.6 presents the results of the Tukey pairwise comparison analysis in the Midwestern region. As indicated, four tests returned significant results. Population classes 3 and 4 had significantly more points than population classes 2 and 1. Population classes 3 and 4, counties with higher population density, have more fluted-points recorded than counties in population classes 1 and 2, which have a lower population density.

There is no statistically significant relationship between fluted-points recorded and people per square kilometer in the Northeastern region, p=.428, F=19.046 (Table A.7) or in the Western region p=0.087, F=2.22 (Table A.8).

In states east of the Mississippi, one-way ANOVA shows a statistically significant relationship between total fluted-points recorded and population per square kilometer, p=0.000, F=6.772 (Table A.9).

The results of the Tukey pairwise comparison analysis are presented in table A.10. As is indicated, two tests returned significant results. Population classes 3 and 4 had significantly more points than population classes 2, but not more than in population class 1. Population classes 3 and 4, counties with higher population density, have more points recorded than counties in population class 2, which has a smaller population density. However, as shown, the means of population classes 3 and 1, as well as 4 and 1 do not differ significantly.

One-way ANOVA shows no statistically significant relationship between total fluted-points recorded and population per square kilometer in those states west of the Mississippi, p=0.283, F=1.273 (Table A.11).

One-way ANOVA applied to cultivation and total fluted-points in the lower forty-eight states. The parametric one-way ANOVA indicates there is no statistically significant relationship between total fluted-points recorded and cultivated square kilometers in the conflated lower forty-eight states, p=0.387, F=1.011 (Table A.12).

There is also no statistically significant relationship between fluted-points recorded and cultivated square kilometers in the Northern region, Southern region, or Western region (Tables A.13-A.15). However, one-way ANOVA did indicate a statistically significant relationship between fluted-points recorded and cultivated square kilometers in the Midwestern region, p=0.010, F=3.869 (Table A.16)

The results of the Tukey pairwise comparison analysis for the Midwestern region are presented in Table A.17. As is indicated, one test returned a significant result. Cultivation class 4 had significantly more points than cultivation class 3 but not more than 2 or 1. Population class 4, counties with higher population density, have more points recorded than counties in population class 3, which have a smaller population density. However, as shown, the means of the other population classes do not differ significantly.

Interestingly, when combining all states east of the Mississippi, one-way ANOVA showed a statistically significant relationship, p=0.002, F=4.852 (Table A.18). The same was true when all states west of the Mississippi were combined, p=0.036, F=2.867 (Table A.20).

The results of the Tukey pairwise comparison for all states east of the Mississippi are presented in Table A.19. As is indicated, two tests returned significant results. Cultivation class 3 had significantly more points than both cultivation classes 2 and 1. Counties in cultivation class 3 have more cultivation per square kilometer than those in cultivation classes 2 or 1.

The results of the Tukey pairwise comparison analysis for all states west of the Mississippi are presented in Table A.21. As is indicated, one test returned a significant result. Cultivation class 3 had significantly more points than cultivation class 1.

As stated previously, differences in county land area bias the fluted-point distribution, therefore one-way ANOVA was repeated using fluted-points divided by county area as the dependent variable.

One-way ANOVA applied to population and total-points divided by county area. The parametric one-way ANOVA indicates there is a statistically significant relationship between total fluted-points divided by county area and population per square kilometer in the conflated lower forty-eight states p=0.000, F=14.443 (Table A.22).

The results of the Tukey pairwise comparison analysis are presented in Table A.23. As is indicated, three tests returned significant results. Although significant results were returned, the relationship was inverse. Population class 1 had significantly more points than population classes 2, 3, and 4. As population per square kilometer increases, the amount of fluted-points recorded decreases.

One-way ANOVA indicated only the Southern region of the United States had a statistically significant relationship between total fluted-points divided by county area and population per square kilometer in the Southern region, p=0.007, F=4.023 (Table A.24).

The results of the Tukey pairwise comparison analysis for the Southern region are presented in Table A.25. As is indicated, only one test returned a significant result. Population class 3 had significantly more points than population class 2.

There were no statistically significant associations between people per square kilometer and fluted-points divided by county area in the Northeastern (p=0.120, F=1.959), Midwestern (p=0.632, F=0.574), or Western regions (p=0.238, F=1.415) (Tables A.26-28).

In states east of the Mississippi, one-way ANOVA also showed no statistically significant relationship, p=0.500, F=0.789 (Table A.29). However, states west of the Mississippi did indicate a statistically significant relationship between people per square kilometer and fluted-points divided by county area, p=0.002, F=5.137 (Table A.30).

The results of the Tukey pairwise comparison analysis are presented in table A.31. As is indicated, only two tests returned significant results. Population classes 2 and 3 had significantly more points than population class 1.

One-way ANOVA applied to cultivation and total points divided by county area. The parametric one-way ANOVA indicates there is no statistically significant relationship between total fluted-points divided by county area and cultivation in the conflated lower forty-eight states, p=0.559, F=0.688 (Table A.32) In fact, there are no significant relationships in any region between cultivation and total points divided by county area when the one-way ANOVA is applied (Tables A.33-38).

One-way ANOVA makes specific assumptions regarding population parameters that characterize underlying distributions. If data from one or more samples come from a population whose distribution violates the assumption of normality (bell shaped distribution) or outliers are suspected to be present, one-way ANOVA may provide misleading results by influencing the mean.

Outliers were encountered in the log-transformed data presented here. For example, when considering total points per county, the maximum number of fluted-points is 6.2 with the mean being 1.1 points. The maximum number of people per square kilometer is 10.2, while the mean is 2.7 people per sqaure kilometer. The mean of cultivation per square kilometer is 5.2, with the maximum number being 8.5. Furthermore, as stated previously, many of the counties in the database have either zero or one fluted-point, making the data highly skewed.

The Kruskal-Wallis test does not assume normality of distributions for sample populations and the mean is less effected by outliers.

Kruskal-Wallis

Kruskal-Wallis applied to modern population density and total fluted-points in the lower forty-eight states. Overall, in the lower forty-eight states, the non-parametric Kruskal-Wallis indicated a significant relationship (p=0.000, H=38.047) between modern population density and fluted-points recorded. Median points per square kilometer increase as people per square kilometer increases. All of the population classes (1-4) indicate a significant relationship. Counties with greater population density have significantly more points than counties with lower population density.

There is also a significant relationship between people per square kilometer and fluted-points recorded in the Southern (p=0.005, H=13.050) Western (p=0.046, H=8.005) and Midwestern regions (p=0.000, H=44.653) as well as those states east of the Mississippi (p=0.000, H=24.233).

In the Southern, Midwestern, and Western regions where a relationship between modern population and fluted-points did occur, counties in population class 1 had a higher density of fluted-points than counties in population class 2.

In the states east of the Mississippi, counties in population class 1 had a higher density fluted-points than counties in population class 3.

Kruskal-Wallis applied to total points and cultivation. Overall, in the lower 48 states, there is no significant relationship between cultivated square kilometers per county and total fluted-points recorded (p=0.246, H=4.143).

There is also no significant relationship between fluted-points recorded and cultivation in the Southern (p=0.238, H=4.230), Western (p=0.385, H=3.043) or Northeastern (p=0.127, H=5.701) regions of the United States.

However, the Midwestern region does exhibit a significant relationship between fluted-points and cultivation (p=.011, H=11.087). Counties with a higher density of cultivation have significantly more fluted-points than counties with a lower cultivation density.

Even though there is a significant relationship in the Midwestern region, counties in cultivation class 1 have a higher density of fluted-points than both counties in cultivation classes 2 and counties in cultivation class 3. Also, counties in cultivation class 2 have a higher density of fluted-points than counties in cultivation class 3.

There is also a significant relationship between cultivation and fluted-points in those states located east of the Mississippi (p=.002, H=15.361).

However, counties in cultivation class 1 have a higher density of fluted-points than counties in cultivation class 2 and counties in cultivation class 4. Counties in cultivation class 2 also have a higher density of fluted-points than counties in cultivation class 4. There was no significant association between cultivation and fluted-points recorded in those states west of the Mississippi (p=0.234, H=4.265).

Kruskal-Wallis test applied to population and total points divided by county area. In the lower forty-eight states, when the number of points within a county are divided by the area of that county (points/area) and then compared to people per square kilometer by the Kruskal-Wallis statistical test, there is a significant association (p=.000, H=239.011). Counties that have a higher population density have significantly more fluted-points than counties with a lower population density. All of the population classes (1-4) indicated significant relationships except counties in population class 1 versus counties population class 3.

There is also a significant relationship between points/area and population per square kilometer in the Southern region (p=0.000, H=52.751), Northern region (p=0.000, H=37.580), Midwestern region (p=0.000, H=162.258), states east of the Mississippi (p=0.000, H=110.811), and states west of the Mississippi (p=0.008, H=11.983). Counties with lower population densities have significantly less points than counties with higher population densities.

In the Southern region, all the population classes indicated significant relationships, except class 1 versus class 2. In the Northern region, all the classes except class 1 versus class 3 indicated a significant relationship. In the Midwestern region, all of the classes indicated significant relationships.

In states east of the Mississippi, there was no significant relationship between counties in population class 1 and counties in population class 2. There was also no significant relationship between counties in population class 2 and counties in population class 3.

There was no significant relationship between median points per square kilometer and people per square kilometer in the Western region.

Kruskal-Wallis applied to cultivation and total points divided by county area. When the number of points within a county are divided by the area of that county and then compared to cultivated square kilometers by the Kruskal-Wallis statistical test in the lower forty-eight states, there is a significant association (p=.000, H=105.091). All of the cultivation classes (1-4) indicated significant relationships. Counties with a higher density of cultivated land have significantly more fluted-points than counties with a lower density of cultivated land.

There is also a significant association between points/area and cultivation per square kilometer in the Southern region (p=0.000, H=39.097), Northern region (p=0.000, H=37.749), Midwestern region (p=0.000, H=45.407), and states east (p=0.000) and west (p=0.008, H=11.816) of the Mississippi.

In the Southern region, all the cultivation classes indicated a significant association. As cultivation increased, so did the amount of points found and recorded.

In the Northern region, all of the cultivation classes indicated a significant relationship except counties in cultivation class 2 versus counties in cultivation class1.

In the Midwestern region, all classes indicated a significant relationship except counties in cultivation class 1 verses counties in cultivation class 2, counties in cultivation class 1 verses counties in cultivation class 3, and counties in cultivation class 2 verses counties in cultivation class 3.

In states east of the Mississippi, all classes show a significant association except counties in cultivation class 1 versus counties in cultivation class 4.

In states west of the Mississippi, all classes show a significant association except counties in cultivation class 1 versus counties in cultivation classes 2 and 3.

The Kruskal-Wallis test results have shown that modern population density and cultivated square kilometers bias the fluted-point sample in the lower forty-eight states overall, and all in all regions except the Western region.

Table 4.5 Statistical Significance Summary: Population Versus Fluted-Points

	ANOVA Population and Total Points	**Kruskal-Wallis Population and Total Points**	**ANOVA Population and Total Points/County Area**	**Kruskal-Wallis Population and Total Points/County Area**
Lower 48 States	Significant Association	Significant Association	Significant Association	Significant Association
Northern Region	No Significant Association	No Significant Association	No Significant Association	Significant Association
Southern Region	Significant Association	Significant Association	Significant Association	Significant Association
Midwestern Region	Significant Association	Significant Association	No Significant Association	Significant Association
Western Region	No Significant Association	Significant Association	No Significant Association	No Significant Association
States East of the Mississippi	Significant Association	Significant Association	No Significant Association	Significant Association
States West of the Mississippi	No Significant Association	No Significant Association	No Significant Association	Significant Association

Table 4.6 Statistical Significance Summary: Cultivation Versus Fluted-Points

	ANOVA Cultivation and Total Points	**Kruskal Wallis Cultivation and Total Points**	**ANOVA Cultivation and Total Points/County Area**	**Kruskal Wallis Cultivation and Total Points/County Area**
Lower 48 States	No Significant Association	No Significant Association	No Significant Association	Significant Association
Northern Region	No Significant Association	No Significant Association	No Significant Association	Significant Association
Southern Region	No Significant Association	No Significant Association	No Significant Association	Significant Association
Midwestern Region	Significant Association	Significant Association	No Significant Association	Significant Association
Western Region	No Significant Association	No Significant Association	No Significant Association	No Significant Association
East of the Mississippi	Significant Association	Significant Association	No Significant Association	Significant Association
West of the Mississippi	Significant Association	No Significant Association	No Significant Association	Significant Association

CHAPTER 5

DISCUSSION

Statistical analyses of correlation, both parametric (one-way ANOVA) and non-parametric (Kruskal-Wallis) concluded that modern population is significantly associated with fluted-points when the entire United States is conflated. This conclusion was reached when total fluted-points per county and when fluted-points divided by county area were treated as the dependent variable for each test. Both tests were also applied to individual regions in the United States. Although the tests indicated a statistically significant relationship in the U.S. overall, individually, some regions showed no relationship while others did. However, in both the conflated United States and individual regions, the Kruskal-Wallis test with fluted-points divided by county area as the dependent variable is the most accurate measure of bias in the fluted-point sample.

With non-parametric tests, values are ranked and the analyses are based on the distribution of the ranks. The non-parametric equivalent of the one-way ANOVA is the Kruskal-Wallis One-Way Analysis of Variance. Although the Kruskal-Wallis test does not assume normality of the distributions for the sample populations, it does assume that the populations have the same distribution, except for a possible difference in population medians.

Also, the Kruskal-Wallis test is more effective than the one-way ANOVA when analyzing data when outliers are suspected even if the underlying distributions are close to normal. In the case of the data presented here, outliers were encountered. As pointed out earlier, outliers are observations that fall far from the rest of the data. As a result, they can influence the mean. However, the median is highly resistant to outliers. The one-way ANOVA tests for differences in means and can be affected by the outliers, while the Kruskal-Wallis test analyzes differences in population medians and should be less affected by outliers.

Dividing the number of points in each county by the county area controls for difference in county size and creates the variable “points per unit area.” It is important to control for county because as pointed out earlier, county sizes west of the Mississippi tend to be larger than those east of the Mississippi. When using total points per county as the dependent variable, the density of point distributions are skewed.

As with the analysis of modern population and recorded fluted-points, when analyzing the relationship between cultivated square kilometers and fluted-points, the same variety of statistical tests were applied. Again, the non-parametric Kruskal-Wallis test, with points divided by county as the dependent variable was determined to be the best test for the data presented. In the conflated United States, the Kruskal-Wallis test showed a significant relationship. In fact there was a statistically significant relationship in all regions except the Western region.

Other authors (Buchanan 2003;Shott 2002; Lepper 1983) who have conducted similar studies have also found correlation between modern population and fluted-points recorded. Specifically, as modern population increases within an area (in this case counties), so does the amount of fluted points recorded within that area. However, it is important to note that Shott (2002) did not control for county area when taking into account fluted- points, but Buchanan (2003) did. Lepper also controlled for area, however his units of analysis were physiographic provinces (points/area of physiographic province).

There are a variety of other possibilities that may cause an increase in the amount of fluted-points recovered in certain areas and these possibilities should be considered. Some of these possibilities include: more intensive archaeological survey efforts, more archaeologists, differential collecting activities, geologic visibility, and intensive site excavation.

Lepper (1983) proposes more fluted-points are found and recorded in the Eastern United States because the East has a higher population density and more farming as opposed the Western United States. Archaeological visibility is primarily a function of cultivation in the Eastern United States according to Lepper (1983).

Faught (1996) also proposes that archaeological visibility in the Eastern United States might be due to a denser population and longer farming history, but also points out that there is more fine grained cryptocrystalline chert available in the east. Cryptocrystalline chert is a resource that could possibly have drawn Paleoindians to the region. Anderson and Faught (2000), point out that over 70% of the total fluted point sample occurs east of the Mississippi.

Buchanan (2003) suggests the lack of visibility of points in the West is due to the shrub lands in the western United States. He equates shrub lands in the west during the Late Pleistocene with lower densities of game animals.

It has also been suggested (Buchanan 2003) that settlement patterns of early Paleoindian cultures resemble the same patterns used by modern populations, specifically, locating around water sources and waterways, as well as in forested areas. Anderson (1990) suggests that Paleoindians would have followed major river valleys, but, Shott (2002) disagrees. He proposes that Paleoindians did not inhabit the same areas attractive to modern populations. Hence, both populations should be treated independently. He argues that modern populations are primarily sedentary and Paleoindians would have been highly mobile and occupied many places in their annual rounds.

There is no doubt that there is a greater modern population density in the eastern United States. This may lead one to assume more people would find and record more fluted-points. However, it also must be considered that the reason for more fluted-point finds is because of the greater density of Paleoindian populations in the east. As reported in Buchanan's (2003) study, there is a high correlation between fluted-point finds and woodlands, as well as wetland areas.

Through a series of computer simulations using a standard demic expansion model, Steele et al. (1998) also found a greater density of fluted-points in the east. Their model attributed the density of fluted-points to a greater density of Paleoindian populations in the eastern woodlands.

The point data used by Steele et al. (1998) was taken from the database compiled by Anderson and Faught (2000b). Admittedly, they assume that the fluted-point densities are the result of discard rates by Paleoindians and not due to current recovery rates. Steele et al. (1998) assume the sample is not biased.

Storck (1982) proposes that the larger number of points in the east is due to different processes of settlement in the eastern and western United States. He suggests that the true antiquity of Paleoindian tradition can only be determined by radiocarbon dating and he points out that few eastern sites withstand this (dating). Even today, although there are more fluted-points recorded in the east, there are fewer early sites in the east that have unequivocal radiocarbon dates than in the west.

Storck (1983) notes that it had been assumed that the large number of artifacts at places like Bull Brook are the result of numerous occupations by small groups. In actuality, however, they may be higher populations of people. In fact, the higher density of populations may represent an evolution from a mobile to a more sedentary lifestyle. This evolution would have developed in the east first rather than the west because environmental conditions are more conducive to population growth in the east. This is in contrast to Shott (2002) who believed that Paleoindians were highly mobile and occupied territories annually.

Lepper (1983) also suggests discovery of archaeological data are not equal in all areas due to geomorphic processes and modern land use patterns. In addition, he notes that artifact collecting may also not be equal in all regions. Open fields may be revisited often and woodland areas not very often, if at all by collectors.

Lepper (1983) found a correlation between cultivated area and points recorded. Specifically, as the amount of cultivated area increases, so does point counts. As indicated earlier, Lepper (1983) also accounted for area when considering the amount of points recorded in that area. Interestingly, Buchanan (2003) also accounted for area but did not find a correlation between cultivated area and fluted-points.

Lepper (1983) examined data from Indiana, Kentucky, Michigan, Ohio, Tennessee, and West Virginia. He suggested there is no reason to believe that Paleoindian hunter-gatherers would have inhabited the same habitats as modern farmers because Paleoindian settlement pattern was not sensitive to the same environmental constraints as modern farmers. Because Paleoindian settlement pattern was not sensitive to the same environmental variables that constrain modern farming practices, a bias in fluted-point counts would exist if there were an association between cultivation and points found.

However, as Buchanan (2003) and Steele et al. (1998) demonstrated through their studies, Paleoindian populations seem to have had the same type of settlement patterns as modern populations.

Although Lepper (1983) found a correlation between population and fluted-points recorded as well as cultivation and points, there was a higher correlation between population and points than cultivation and points. In the study presented here, when the lower forty-eight states are conflated, modern population and amount of land under cultivation have an equal association with fluted-points ($p=.000$).

This research has relied on data collected in the North American fluted-point database to evaluate potential biases in the conflated lower forty-eight states, as well as individual regions within the United States. Admittedly, this coarse level of analysis could be refined to determine potential biases on a finer scale. As shown in regional studies, similar biases affect the amount of fluted-points found and recorded and support the general trends described in this research (Lepper 1983; Shott 2002).

Although there was no correlation between area of land under cultivation and fluted-points found and recorded in Shott's (2002) analysis, he did not take into account all states in the Midwest. Illinois, Kansas, Nebraska, North Dakota, and South Dakota were excluded, as well as northern Minnesota, northern Michigan, northern Wisconsin, and southwestern Wisconsin. Shott (2002) does not provide a reason for excluding Kansas, Nebraska, North Dakota, and South Dakota in his research. He excludes Illinois and southwestern Wisconsin, citing lack of available data. Northern Minnesota, northern Michigan, and northern Wisconsin were excluded because few fluted-points have been found and recorded in these areas and, ''it is unlikely these areas were occupied extensively by Paleoindians (Shott 2002).'' Northern Minnesota was also excluded because,''its counties are more irregular in shape and variable in size than elsewhere in the Midwest (Shott 2002).''

I disagree with the methodology of Shott's (2002) statistical analysis in regards to omitting portions of states. The North American fluted-point database served as the primary data source for this research as well as Shott's (2002). The database consists of fluted-point counts per county and should be subjected to statistical analysis as such. Because there are few or no fluted-points in a county, the county can not be excluded. One can not eliminate areas because there may or may not have been extensive Paleoindian occupation in a particular area. Also if Shott (2002) accounted for the number of fluted-points found and recorded divided by county area, the large size of counties in Minnesota, as compared to elsewhere in the Midwest would not matter. Furthermore, the omission of Kansas, Nebraska, North Dakota, and South Dakota from the Midwest could explain why the statistical analysis presented in this research indicates a correlation bewteen cultivation and fluted-points but not in Shott's (2002). Kansas, Nebraska, North Dakota. and South Dakota have a very high density of cultivation, specifically, northwest Minnesota and North Dakota.

Lepper's (1983) statistical anlysis showed a correlation between the number of fluted-points found and recorded and modern population density, as well as the number of fluted-points found and recorded and area of land under cultivation. Lepper (1983) used physiographic provinces as his unit of measurement. However, the large area of physiographic regions and small number tends to hamper statistical analysis. For this reason, it is the opinion of this author that counties are a better unit of measurement. Counties are arbitrary in location with respect to the archaeological record.

One problem with Lepper's study is that he converted county based point frequencies into points per physiographic province, however, some provinces had few counties while others overlapped physiographic areas. Counties that overlapped physiographic provinces were eliminated from analysis unless the county had more than 50% in one physiograpic province or the either. As the a result, a total of 35 counties were excluded from the analysis.

Considering units of analysis, Blackmar (2001) has put forth perhaps the most efficient unit of measurement for statistical analysis. Blackmar (2001) combined physiograpic regions and counties. Her study area consisted of Kansas, Texas, and Oklahoma. The counties of each state were grouped into environmental regions based on elevation, topography, vegetation, paleoecological maps, and physiographic maps. Statistical analysis was applied to determine the frequency of fluted-points divided by the area of each physiogrpahic region for Clovis, Folsom, and Cody cultural complexes.

Because the fluted-point sample is incredibly large, it was difficult to determine which particular statistical test was appropriate for analyzing the data in this research. Large samples equal more power, which leads to more statistically significant results. It must be acknowledged that this fact could influence the statistical results to some degree. However, in this study, every attempt was made to account for the robust nature of the data. Hence, a variety of statistical analyses was attempted before ultimately deciding that the Kruskal-Wallis test was the most appropriate (the reasons for choosing the Kruskal-Wallis test have been discussed in this chapter, as well as in chapter 3).

Also, because a large number of counties in the fluted-point database have zero or one fluted-point, the data were heavily skewed to the right. In addition, there were also quite a few outliers encountered. Both of these factors hampered statistical analysis. Again, in both cases, the Kruskal-Wallis test was chosen as the most appropriate test.

Using the Kruskal-Wallis test, it was possible to draw positive conclusions regarding biases in the fluted-point sample.

CONCLUSION

Through a variety of statistical analyses, it has been shown that both modern population density and cultivated square kilometers, at minimum, influence the fluted-point sample depending on the area tested. The non-parametric Kruskal-Wallis test indicated a statistically significant relationship between cultivated square kilometers and fluted-points divided by county area, as well as, modern population and fluted-points divided by county area.

To some extent, the fluted-point sample is determined by both people per square kilometer and amount of cultivation per square kilometer. In agreement with Shott (2002), we cannot attribute the distribution and abundance of points only to Paleoindian population distribution, colonization routes, land use, etc.

Although cultivation and modern population do to some extent bias the fluted-point sample, they do not determine it. The data set used here, the fluted-point sample compiled by Anderson and Faught (2000b), is a tremendously large sample. In statistics, larger samples equal more power, which leads to more statistically significant results. In this case, power is the probability of rejecting the null hypothesis when it is false. As acknowledged in the previous chapter, the robust size of the sample makes statistical analysis difficult and may influence the statistical results to some degree. In an attempt to counter the effects of the robustness of the sample, the data were subjected to a variety of statistical analyses. Ultimately, the non-parametric Kruskal-Wallis test was chosen as the most appropriate for analyzing the data at hand.

Other factors either independent of, or in combination with population and cultivation should also be considered as contributing to fluted-point finds. One factor is the intensity of archaeological survey at the county, state, and regional level. Regionally, Shott (2002) has already shown that survey effort in seven Midwestern states bias the fluted-point sample. Buchanan (2003) also notes that survey effort may explain the large densities in Ohio and Virginia.

Impact of collectors on an area could also be a source inflated fluted point numbers in an area. Lepper (1983) used population of an area as a proxy for the amount of collectors in that area. However, Shott (2002) did not think that population of an area was a good measure of collector activities. On this point I have to agree with Shott, the majority of people in the United States are not collecting artifacts. For that matter, I would assume the majority of people within a county are not collecting.

Lepper (1983) believes people collect in their "own back yards." In contrast, Seeman and Prufer (1982) thought collectors may travel long distances. In order to measure the true effects of collecting, a survey of collectors would have to occur. This would be a daunting task. There would have to documentation of collector's assemblages, how long they had been collecting, and areas they collect. Most collectors, unfortunately, are skeptical of answering any questions of the type proposed here (Shott 2002). The true impact of collectors will most likely never be known, even if an intensive survey is conducted. This is due to the collector's unwillingness to reveal their assemblages or locations of collecting.

Another factor leading to the discrepancy between point distributions in the eastern and western United States is geologic visibility. In the United States, it is readily apparent that different regions have quite different geomorphological occurrences happening. These processes may lead to the discovery of sites or vice versa.

Billeck's (1998) study is an excellent example of how geomorphic processes can determine site discovery. In the Loess Hills of Southwestern Iowa, thirty-three fluted-points were found. He showed that the fluted-point distribution was not only due to Paleoindian exploitation of the land, but also to geomorphological processes. The settings include hilltops, hillsides, foot slopes, and creek beds. Large portions of the Paleoindian landscape have been destroyed by erosion or are deeply buried.

In addition, Blackmar (2001) provides the examples of the Aubrey, Domebo, and Mclean sites as those covered by thick alluvial fills while extensive sand dunes around playa basins in the high plains may also decrease archaeological visibility.

All of the factors presented here could work independently of, or in combination with cultivation and modern population to bias the fluted-point sample.

In summary, the data in this research was compiled to determine if modern population density and/or area of land under cultivation bias the fluted-point sample compiled by Anderson and Faught (2000b). After a variety of statistical analyses, it was ultimately determined that the non-parametric Kruskal-Wallis test was the appropriate test for the data analyzed. The results showed that there is a significant relationship between cultivation and fluted-points as well as modern population and fluted-points.

Although these variables indicate a bias in the fluted-point sample, they do not determine it. In fact, more databases on various geographical levels should be compiled and existing databases should be expanded upon. The importance of this study is to at least warn those compiling or analyzing other databases of this type that biases do exist and should be addressed.

APPENDIX A

DETAILED RESULTS FOR THE ONE-WAY ANOVA AND TUKEY MULTIPLE COMPARISON TESTS

Table A.1 One-way ANOVA Results for Total-Fluted Points and Population in the Lower Forty-Eight States

Source	**Sum of Squares**	**df**	**Mean-Square**	**F-Ratio**	**P**
Population Classes 1-4	43.586	3	14.529	10.688	0.000
Error	1892.959	1390	1.362		

Table A.2 Tukey Multiple Comparison-Lower Forty-Eight States

	Population class 1 (≤1.069 people per sq km)	**Population Class 2** (≥1.070 and ≤ 2.728 people per sq km)	**Population Class 3** (≥2.729 and ≤ 4.387 people per sq km)	**Population Class 4** (≥4.388 people per sq km)
Population class 1 (≤1.069)	0.000			
Population Class 2 (≥1.070 and ≤ 2.728 people per sq km)	.601 (.205)			
Population Class 3 (≥2.729 and ≤ 4.387 people per sq km)	1.349 (.611)	2.249 (.000)		
Population Class 4 (≥ 4.388 people per sq km)	1.828 (.119)	3.048 (.000)	1.355 (.405)	

Table A.3 One-way ANOVA Results for Total-Fluted Points and Population in the Southern Region

Source	**Sum of Squares**	**df**	**Mean-Square**	**F-Ratio**	**P**
Population Classes 1-4	14.376	3	4.792	3.733	0.011
Error	899.812	701	1.284		

Table A.4 Tukey Multiple Comparison-Southern Region

	Population class 1 (≤1.069 people per sq km)	**Population Class 2** (≥1.070 and ≤ 2.728 people per sq km)	**Population Class 3** (≥2.729 and ≤ 4.387 people per sq km)	**Population Class 4** (≥4.388 people per sq km)
Population class 1 (≤1.069)				
Population Class 2 (≥1.070 and ≤ 2.728 people per sq km)	.838 (.982)			
Population Class 3 (≥2.729 and ≤ 4.387 people per sq km)	1.556 (.762)	1.854 (.036)		
Population Class 4 (≥ 4.388 people per sq km)	2.032 (.474)	2.427 (.018)	1.309 (.773)	

Table A.5 One-way ANOVA Results for Total-Fluted Points and Population in Midwestern Region

Source	**Sum of Squares**	**df**	**Mean-Square**	**F-Ratio**	**P**
Population Classes 1-4	56.409	3	18.803	17.013	0.000
Error	389.040	352	1.105		

Table A.6 Tukey Multiple Comparison-Midwestern Region

	Population class 1 (≤1.069 people per sq km)	**Population Class 2** (≥1.070 and ≤ 2.728 people per sq km)	**Population Class 3** (≥2.729 and ≤ 4.387 people per sq km)	**Population Class 4** (≥4.388 people per sq km)
Population class 1 (≤1.069)				
Population Class 2 (≥1.070 and ≤ 2.728 people per sq km)	1.216 (.981)			
Population Class 3 (≥2.729 and ≤ 4.387 people per sq km)	4.093 (.021)	4.977 (.000)		
Population Class 4 (≥ 4.388 people per sq km)	9.036 (.000)	10.990 (.000)	2.208 (.109)	

Table A.7 One-way ANOVA results for total-fluted points and population in Northeastern Region

Source	Sum of Squares	df	Mean-Square	F-Ratio	P
Population Classes 1-4	60.027	3	20.320	19.046	.428
Error	400.078	376	1.106		

Table A.8 One-Way ANOVA Results for Total-Fluted Points and Population in the Western Region

Source	Sum of Squares	df	Mean-Square	F-Ratio	P
Population Classes 1-4	14.169	3	4.723	2.228	0.087
Error	339.193	160	2.120		

Table A.9 One-way ANOVA Results for Total-Fluted Points and Population in Those States East of the Mississippi

Source	Sum of Squares	df	Mean-Square	F-Ratio	P
Population Classes 1-4	27.361	3	9.120	6.772	0.000
Error	1146.076	851	1.347		

Table A.10 Tukey Multiple Comparison-States East of the Mississippi

	Population class 1 (≤1.069 people per sq km)	**Population Class 2** (≥1.070 and ≤ 2.728 people per sq km)	**Population Class 3** (≥2.729 and ≤ 4.387 people per sq km)	**Population Class 4** (≥ 4.388 people per sq km)
Population class 1 (≤1.069)				
Population Class 2 (≥1.070 and ≤ 2.728 people per sq km)	.521 (.986)			
Population Class 3 (≥2.729 and ≤ 4.387 people per sq km)	1.303 (.999)	2.500 (.002)		
Population Class 4 (≥ 4.388 people per sq km)	1.910 (.986)	3.664 (.000)	1.466 (.284)	

Table A.11 One-way ANOVA Results for Total-Fluted Points and Population in Those States West of the Mississippi

Source	**Sum of Squares**	**df**	**Mean-Square**	**F-Ratio**	**P**
Population Classes 1-4	5.147	3	1.716	1.273	0.283
Error	720.762	535	1.347		

Table A.12 One-way ANOVA Results for Total-Fluted Points and Cultivation in the Lower Forty-Eight States

Source	Sum of Squares	df	Mean-Square	F-Ratio	P
Cultivation Classes 1-4	4.216	3	1.405	1.011	0.387
Error	1930.017	1388	1.391		

Table A.13 One-way ANOVA Results for Total-Fluted Points and Cultivation in the Northern Region

Source	Sum of Squares	df	Mean-Square	F-Ratio	P
Cultivation Classes 1-4	3.303	3	1.101	0.827	0.481
Error	219.654	165	1.331		

Table A.14 One-way ANOVA Results for Total-Fluted Points and Cultivation in the Southern Region

Source	Sum of Squares	df	Mean-Square	F-Ratio	P
Cultivation Classes 1-4	4.608	3	1.536	1.182	0.315
Error	909.321	700	1.299		

Table A.15 One-way ANOVA Results for Total-Fluted Points and Cultivation in the Western Region

Source	**Sum of Squares**	**df**	**Mean-Square**	**F-Ratio**	**P**
Cultivation Classes 1-4	9.244	3	3.081	1.433	0.235
Error	344.118	160	2.151		

Table A.16 One-way ANOVA Results for Total-Fluted Points and Cultivation in the Midwestern Region

Source	**Sum of Squares**	**df**	**Mean-Square**	**F-Ratio**	**P**
Cultivation Classes 1-4	14.192	3	4.731	3.869	0.010
Error	429.162	351	1.223		

Table A.17 Tukey Multiple Comparison-Midwestern Region

	Cultivation class 1 (≤3.747 cultivated sq km)	**Cultivation Class 2** (≥3.748 and ≤ 5.247 cultivated sq km)	**Cultivation Class 3** (≥5.248 and ≤ 6.747 cultivated sq km)	**Cultivation Class 4** (≥6.748 cultivated sq km)
Cultivation class 1 (≤3.747 cultivated sq km)				
Cultivation Class 2 (≥3.748 and ≤ 5.247 cultivated sq km)	.824 (.998)			
Cultivation Class 3 (≥5.248 and ≤ 6.747 cultivated sq km)	.877 (.999)	1.064 (.999)		
Cultivation Class 4 (≥6.748 cultivated sq km)	.296 (.617)	.359 (.160)	.337 (.005)	

Table A.18 One-way ANOVA Results for Total-Fluted Points and Cultivation in Those States East of the Mississippi

Source	**Sum of Squares**	**df**	**Mean-Square**	**F-Ratio**	**P**
Cultivation Classes 1-4	19.744	3	6.581	4.852	0.002
Error	1151.585	849	1.356		

Table A.19 Tukey Multiple Comparison-States East of the Mississippi

	Cultivation class 1 (≤3.747 cultivated sq km)	**Cultivation Class 2** (≥3.748 and ≤ 5.247 cultivated sq km)	**Cultivation Class 3** (≥5.248 and ≤ 6.747 cultivated sq km)	**Cultivation Class 4** (≥6.748 cultivated sq km)
Cultivation class 1 (≤3.747 cultivated sq km)				
Cultivation Class 2 (≥3.748 and ≤ 5.247 cultivated sq km)	1.442 (.599)			
Cultivation Class 3 (≥5.248 and ≤ 6.747 cultivated sq km)	2.523 (.009)	1.750 (.029)		
Cultivation Class 4 (≥6.748 cultivated sq km)	1.050 (1.000)	.728 (.887)	.416 (.186)	

Table A.20 One-way ANOVA Results for Total-Fluted Points and Cultivation in Those States West of the Mississippi

Source	**Sum of Squares**	**df**	**Mean-Square**	**F-Ratio**	**P**
Cultivation Classes 1-4	11.847	3	3.829	2.867	0.036
Error	714.422	535	1.335		

Table A.21 Tukey Multiple Comparison-States West of the Mississippi

	Cultivation class 1 (≤3.747 cultivated sq km)	**Cultivation Class 2** (≥3.748 and ≤ 5.247 cultivated sq km)	**Cultivation Class 3** (≥5.248 and ≤ 6.747 cultivated sq km)	**Cultivation Class 4** (≥6.748 cultivated sq km)
Cultivation class 1 (≤3.747 cultivated sq km)				
Cultivation Class 2 (≥3.748 and ≤ 5.247 cultivated sq km)	-.490 (.067)			
Cultivation Class 3 (≥5.248 and ≤ 6.747 cultivated sq km)	-.511 (.027)	-.021 (.999)		
Cultivation Class 4 (≥6.748 cultivated sq km)	-.354 (.282)	.136 (.799)	.157 (.607)	

Table A.22 One-way ANOVA Results for Fluted-Points Divided by County Area and Population in the Lower Forty-Eight States

Source	**Sum of Squares**	**df**	**Mean-Square**	**F-Ratio**	**P**
Population Classes 1-4	16.070	3	5.357	14.443	0.000
Error	1144.555	3086	0.371		

Table A.23 Tukey Multiple Comparison-Lower Forty-Eight States

	Population class 1 (≤1.069 people per sq km)	**Population Class 2** (≥1.070 and ≤ 2.728 people per sq km)	**Population Class 3** (≥2.729 and ≤ 4.387 people per sq km)	**Population Class 4** (≥4.388 people per sq km)
Population class 1 (≤1.069)				
Population Class 2 (≥1.070 and ≤ 2.728 people per sq km)	-0.181(.000)			
Population Class 3 (≥2.729 and ≤ 4.387 people per sq km)	-0.211(.000)	-0.030 (.667)		
Population Class 4 (≥ 4.388 people per sq km)	-0.208 (.000)	-0.027 (.868)	0.003 (1.000)	

Table 4.24 One-way ANOVA Results for Total-Fluted Points Divided by County Area and Population in the Southern Region

Source	**Sum of Squares**	**df**	**Mean-Square**	**F-Ratio**	**P**
Population Classes 1-4	0.003	3	0.001	4.023	0.007
Error	0.340	1382	0.000		

Table A.25 Tukey Multiple Comparison-Southern Region

	Population class 1 (≤1.069 people per sq km)	**Population Class 2** (≥1.070 and ≤ 2.728 people per sq km)	**Population Class 3** (≥2.729 and ≤ 4.387 people per sq km)	**Population Class 4** (≥ 4.388 people per sq km)
Population class 1 (≤1.069)				
Population Class 2 (≥1.070 and ≤ 2.728 people per sq km)	0.001(.988)			
Population Class 3 (≥2.729 and ≤ 4.387 people per sq km)	0.004 (.202)	0.003 (.008)		
Population Class 4 (≥ 4.388 people per sq km)	0.003 (.547)	0.002 (.387)	-0.001 (.916)	

Table A.26 One-way ANOVA Results for Total-Fluted Points Divided by County Area and Population in the Northeastern Region

Source	**Sum of Squares**	**df**	**Mean-Square**	**F-Ratio**	**P**
Population Classes 1-4	1.904	3	0.635	1.959	0.120
Error	94.258	291	0.324		

Table A.27 One-way ANOVA Results for Total-Fluted Points divided by County Area and Population in the Midwestern Region

Source	Sum of Squares	Df	Mean-Square	F-Ratio	P
Population Classes 1-4	0.145	3	0.048	0.574	0.632
Error	81.369	966	0.084		

Table A.28 One-way ANOVA Results for Total-Fluted Points Divided by County Area and Population in the Western Region

Source	Sum of Squares	df	Mean-Square	F-Ratio	P
Population Classes 1-4	9.162	3	3.054	1.415	0.238
Error	938.678	435	2.158		

Table A.29 One-way ANOVA Results for Total-Fluted Points Divided by County Area and Population in Those States East of the Mississippi

Source	Sum of Squares	df	Mean-Square	F-Ratio	P
Population Classes 1-4	0.147	3	0.049	0.789	0.500
Error	96.657	1558	0.062		

Table A.30 One-way ANOVA Results for Total-Fluted Points Divided by County Area and Population in Those States West of the Mississippi

Source	**Sum of Squares**	**df**	**Mean-Square**	**F-Ratio**	**P**
Population Classes 1-4	10.596	3	3.532	5.137	0.002
Error	1047.174	1523	0.688		

Table A.31 Tukey Multiple Comparison-States West of the Mississippi

	Population class 1 (≤1.069 people per sq km)	**Population Class 2** (≥1.070 and ≤ 2.728 people per sq km)	**Population Class 3** (≥2.729 and ≤ 4.387 people per sq km	**Population Class 4** (≥ 4.388 people per sq km)
Population class 1 (≤1.069)				
Population Class 2 (≥1.070 and ≤ 2.728 people per sq km)	.667 (.003)			
Population Class 3 (≥2.729 and ≤ 4.387 people per sq km)	.643 (.009)	.964 (.993)		
Population Class 4 (≥ 4.388 people per sq km)	.653 (.159)	.979 (1.000)	1.016 (1.000)	

Table A.32 One-way ANOVA Results for Fluted-Points Divided by County Area and Cultivation in the Lower Forty-Eight States

Source	**Sum of Squares**	**df**	**Mean-Square**	**F-Ratio**	**P**
Cultivation Classes 1-4	0.777	3	0.259	0.688	0.559
Error	1159.841	3083	0.376		

Table A.33 One-way ANOVA Results for Fluted-Points Divided by County Area and Cultivation in the Midwestern Region

Source	**Sum of Squares**	**df**	**Mean-Square**	**F-Ratio**	**P**
Cultivation Classes 1-4	0.023	3	0.008	0.089	0.966
Error	81.491	965	0.084		

Table A.34 One-way ANOVA Results for Fluted-Points Divided by County Area and Cultivation in the Northeastern Region

Source	**Sum of Squares**	**df**	**Mean-Square**	**F-Ratio**	**P**
Cultivation Classes 1-4	0.531	3	0.177	0.539	0.656
Error	95.631	291	0.329		

Table A.35 One-way ANOVA Results for Fluted-Points Divided by County Area and Cultivation in the Southern Region

Source	**Sum of Squares**	**df**	**Mean-Square**	**F-Ratio**	**P**
Cultivation Classes 1-4	0.001	3	0.000	1.846	0.137
Error	0.341	1380	0.000		

Table A.36 One-way ANOVA Results for Fluted-Points Divided by County Area and Cultivation in the Western Region

Source	**Sum of Squares**	**df**	**Mean-Square**	**F-Ratio**	**P**
Cultivation Classes 1-4	2.260	3	0.753	0.347	0.792
Error	945.581	435	2.174		

Table A.37 One-way ANOVA Results for Fluted-Points Divided by County Area and Cultivation in Those States East of the Mississippi

Source	**Sum of Squares**	**df**	**Mean-Square**	**F-Ratio**	**P**
Cultivation Classes 1-4	0.105	3	0.035	0.563	0.640
Error	96.699	1556	0.062		

Table A.38 One-way ANOVA Results for Fluted-Points Divided by County Area and Cultivation in Those States West of the Mississippi

Source	**Sum of Squares**	**df**	**Mean-Square**	**F-Ratio**	**P**
Cultivation Classes 1-4	1.195	3	0.398	0.574	0.632
Error	1056.566	1522	0.694		

APPENDIX B

THESIS RESEARCH DATABASE

SUBMITTED AS A SUPPLEMENTAL

EXCEL FILE

BIBLIOGRAPHY

Agresti, Alan, and Barbara Finlay
1997 Statistical Methods for the Social Sciences. Third Edition. Prentice Hall, Inc. Upper saddle River, NJ.

Anderson, David G.
1990 The Paleoindian Colonization of Eastern North America: A view from the Southeastern United States. In K.B. Tnakersley and B. L. Issac (eds), Early Paleoindian economies of Eastern North America. Research in Economic Anthropology 5, pp. 163-216

Anderson, David G., and Christopher Gillam
Paloindian Colonization of the Americas: Implications from an Examination of Physiography, Demography, and Artifact Distrinution.. American Antiquity 65: 43-66.

Anderson, David G., and Michael K. Faught
1998 The Distribution of Fluted Paleoindian Projectile Points: Update 1998. Archaeology of Eastern North America 26: 163-187.

2000a A North American Paleoindian Projectile Point Database Website, Online; http//www.anthro.fsu.edu/special/paleo/paleoind.html.

2000b Paleoindian Artefact Distributions: Evidence and Implications. Antiquity 74: 507-513.

Billeck, William T
1998 Fluted Point Distribution in Loess Hills of Southwestern Iowa. Plains Anthropologist 43: 401-409.

Blackmar, Jeannete M.
2001 Regional Variability in Clovis, Folsom, and Cody Land Use. Plains Anthropologist 46: 65-94.

Buchanan, Briggs
2003 The effects of Sample Bias on Paleoindian Fluted Point Recovery in the United States. Department of Anthropology, Univeristy of Mexico, Albuquerque.

Drennan, Robert D.
1996 Statistics for Archaeologists: A Common Sense Approach. Plenum Press, New York.

Faught, Michael K.
1996 Clovis Origins and Underwater Prehistoric Archaeology in Northwestern Florida. Ph.D. dissertation, Department of Anthropology, University of Arizona, Tucson.

Largent, Floyd B., Michael R. Waters, and David L. Carlson
1991 The Spatiotemperal Distribution and Characteristics of Folsom projectile Points in Texas. Plains Anthropologist 36: 323-340.

Lepper, Bradley T.
1983 Fluted Point Distributional Patterns in the Eastern United States: A Contemporary Phenomenon. Midcontinental Journal of Archaeology 8: 269-285.

McClave, James, T. And Frank H. Dietrich II
1994 Statistics. Sixth Edition. Macmillan College Publishing. New York, NY.

Morrow, Juliet E., and Toby A. Morrow
1999 Geographic Variation in Fluted Projectile Points: A Hemispheric Perspective American Antiquity 64: 215-231.

Purdy, Barbara A.
1982 Comments on "A Compilation of Fluted Points of Eastern North America by Count and Distribution: An *AENA* Project." Archaeology of Eastern North America 11, p. 29

Rice, John A.
1995 Mathematical Statistics and Data Analysis. Second Edition. Duxbury Press. Belmont, CA.

Seeman, Mark F., and Olaf H. Prufer
1982 An Updated Distribution of Ohio Fluted Points. Midcontinental Journal of Archaeology 7: 155-169.

Shott, Michael J.
2002 Sample Bias in the Distribution and Abundance of Midwestern Fluted Bifaces. Midcontinental Journal of Archeology 27:89-123.

SPSS Science Marketing Department
1996 Systat 10 Statistics I. SPSS Inc. United States of America

Steele, James, Jonathan Adams, and Tim Sluckin
1998 Modeling Paleoindian dispersals. World Archaeology 30: 286-305.

Storck, Peter
1982 Commentary on *AENA*'s Compilation of Fluted Points in Eastern North America. Archaeology of Eastern North America 11:34-36.

Thomas, David
1986 Refiguring Anthropology: First Principles of Probability and Statistics.
Waveland Press. Prospect Heights, IL.

United States Census Bureau
2000 County Population Estimates for July 1, 1999 and population change for April,
1999 (CO-99-2). Population Estimates Program, population Division,
Washington D.C.

United States Department of Agriculture
1999 Census of Agriculture, Volume 1: United States Summary and State Data
National Agricultural Statistics Service, Washington D.C.

BIOGRAPHICAL SKETCH

CHRISTOPHER A. SCHAEFER
ARCHAEOLOGIST

Academic Background:	Terra Community College 1993-1994 B.A. The Ohio State University 1999 *M.S. Florida State University 2006*
ESI Responsibilities:	Responsibilities include underwater and terrestrial cultural resource reconnaissance and assessment surveys, site excavations at the Phase II and III levels, analysis, and report production.
Experience:	2004–Present: Environmental Services, Inc. Archaeologist. 2002–2004: Florida State University, Canoe Project, Underwater Archaeologist. 2001–2004: Archaeological Consultants, Inc. Field Archaeologist. 2001: Florida State University Teaching Assistant; Field school in Underwater Archaeology. 2002–2004: Southeastern Archaeological Research, Inc. Research Consultant.
Certifications:	Science Diver, American Academy of Underwater Science PADI open water SCUBA diver. Hazmat I and II Training, American Red Cross First Aid Training, American Red Cross
Representative Project Experience Archaeologist	Multiple underwater and terrestrial Archaeological Surveys, Mapping Projects and Excavations, Florida, Georgia and the Bahamas. Multicomponent Phase II Excavation at Peacock Slough, Suwannee County,FL. Multicomponent Phase II Excavation at Wood's Ferry, Suwannee County, FL. Dog and St. George Island's shipwreck survey, Franklin County, FL. Paleo-Aucilla Prehistory Project, Jefferson County, FL. Sub-Bottom Profiler Survey of Martin Lake, Walton County, FL. Intensive assessment survey at Summer Beach/ Amelia Island Industrial Park Tract, Nassau County, FL Survey and site excavation at Dunham Marsh Plantation and Tuscan Landing, Camden County Site excavation at Riverview, Oak Grove, Chatham County, Georgia