A NONPARAMETRIC APPROACH TO DETERMINE THE NUMBER OF OBSERVATIONS REQUIRED FOR ESTIMATING BASIN-SCALE SOIL TEST PHOSPHOROUS

By

WILLIAM RAY MARSHALL

Bachelor of Science

University of Kentucky

Lexington, Kentucky

1992

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
MASTER OF SCIENCE
July, 1998

OKLAHOMA STATE UNIVERSITY

A NONPARAMETRIC APPROACH TO DETERMINE THE NUMBER OF OBSERVATIONS REQUIRED FOR ESTIMATING BASIN-SCALE SOIL TEST PHOSPHOROUS

Thesis Approved:

Thesis Advisor

CTHaan

Wayne B. Powell
Dean of the Graduate College

ACKNOWLEDGMENTS

I would like to thank Dr. Daniel Storm for serving as my major advisor. His guidance and encouragement helped me reach this goal. I would also like to thank Dr. Tom Haan and Dr. Mike Smolen for serving on my committee. Their thoughts, comments, and suggestions are greatly appreciated. My sincere thanks is also extended to Dr. Ronald Elliott, who provided me work experiences, opportunities, and direction. To the people of the Biosystems and Agricultural Engineering Department, especially Dr. Bill Barfield, I extend my warmest wishes and gratitude for the experiences and friendships you have provided.

This research would not have been possible without the data and support supplied by the Oklahoma Conservation Commission, the Arkansas Soil and Water Conservation Commission, the Adair and Delaware County Oklahoma State University Cooperative Extension Service offices, and the City of Tulsa Public Works Department. My thanks to all of them.

I would like to thank my parents for their love, encouragement, and understanding. I give my love and admiration to my wife, Patricia, and to our new daughter, Haley Brooke. They are both treasured gifts. I thank God for my blessings and give praise to Him.

TABLE OF CONTENTS

Chapter Page

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

Figure Page

CHAPTER 1

INTRODUCTION

Background and Need

Surface runoff from agricultural related activities and other potential nonpoint pollution sources, if not properly managed, can contribute significant loadings of phosphorous and sediments to receiving surface waters. Soil phosphorous can contribute to nonpoint source pollution through runoff in soluble and sediment-bound forms. Excessive levels of phosphorous in surface waters can lead to eutrophication, an increase in the fertility status of natural waters that causes accelerated growth of algae or other water plants (Pierzynski et al., 1994).

As more emphasis is placed on nonpoint source pollution determination and prevention, the use of computer modeling and geographic information systems has come to the forefront of pollution management and evaluation technology. Computer models can be used to target critical source areas of sediment and phosphorous for priority treatment (Storm et al., 1996). Special emphasis can then be given to these critical areas to help minimize the potential for detrimental off-site water quality impacts.

Since many of these computer models are used to determine phosphorous loadings to receiving waters, an important model input parameter is soil phosphorous level. Some hydrological/water quality models require soil test phosphorous as an input parameter. Soil test phosphorous is that portion of soil phosphorous that is available for plant uptake or is in a form to readily become available during a growing season.

At present, there is no established procedure or method to predict the soil test, or plant available, phosphorous levels for various land uses, land covers, and/or soil types without employing a site-specific soil sampling program. When addressing nonpoint source pollution problems, the geographic area of interest is very often on the basin-scale of several thousand to several hundred thousand hectares. With sampling areas of this magnitude, soil sampling and analysis costs can begin to be a major part of the overall project budget.

Soil samples that are eventually collected are typically used to provide an estimate of the average, or mean, soil test phosphorous. As with any estimate, there will be some uncertainty due to the spatial variability of soil test phosphorous and soil sampling procedures. Quantifying this uncertainty will provide a measure, or degree of confidence, for the estimated mean soil test phosphorous and aid in quantifying the output uncertainty for the hydrological/water quality model employed. Thus, there is a need to predict the minimum number of soil samples required, within some specified confidence interval, to obtain an estimate of the mean soil test phosphorous level.

Classical statistical techniques are available for predicting the number of soil samples required, but are based on the assumption of a known underlying distribution, or normal distribution, of the data means. Sometimes, the assumptions associated with this approach do not apply or are not completely valid. It has been found that data from high-level soil test phosphorous basins with fields that receive poultry litter may not adhere to all the assumptions needed to use classical statistics.

The purpose of this research was to evaluate soil test phosphorous probability distributions from several watersheds and/or basins and develop a nonparametric approach for determining the minimum number of soil samples required, within a specified confidence interval, to obtain an estimate of the mean soil test phosphorous. The method was developed for basin-scale applications. Empirical distributions of the data were used so that no assumption had to be made regarding the underlying distribution of the means.

Another important component of estimating sample sizes is determining an acceptable confidence interval. An approach was also developed for identifying the confidence interval based on the expected output variance due to initial phosphorous input of a hydrological/water quality model. The Spatially Integrated Model for Phosphorous Loading and Erosion (SIMPLE) (Sabbagh et al., 1995) was the model employed. It is a continuous simulation, distributed-parameter modeling system developed to estimate watershed- and/or basin-scale sediment and phosphorous loading to surface waters. The technique used to determine the confidence interval for predicting sample size can also be used to estimate the confidence interval and model output variance associated with a predetermined sample size.

The method developed for predicting the number of soil samples required within a specified confidence interval was then used to develop a soil sampling plan for the Oklahoma portion of the Lake Eucha basin, which is located in northeastern Oklahoma.

Objectives

The overall objectives of this research were to examine the probability distributions of soil test phosphorous data and develop a nonparametric method to determine the number of observations required to estimate basin-scale soil test phosphorous. The results could also be used to apply a confidence interval to a predetermined number of samples. The empirical results were then used to develop a soil sampling plan. More specifically, the objectives were:

1. Evaluate the underlying probability distributions of soil test phosphorous data sets from four Oklahoma watersheds and the Arkansas side of the Lake Eucha basin;

2. Use regression statistics to evaluate the relationship, if any, between soil test phosphorous and soil mapping units, and soil test phosphorous and selected soil parameters, for three Oklahoma watersheds;

3. Develop a nonparametric method to determine the minimum number of soil samples required, within a specified confidence interval, to obtain an estimate of the basin-scale mean soil test phosphorous by major land use;

4. Apply the nonparametricmethod from objective 3 to develop a soil sampling plan for the Oklahoma portion of the Lake Eucha basin using the SIMPLE distributed parameter water quality computer model.

CHAPTER 2

LITERATURE REVIEW

Soil Phosphorous

Phosphorous is essential to all forms of life on earth. The lack of available phosphorous in soils can be a limiting factor in plant growth. Over the years, the addition of phosphorous to soils in the form of manures, minerals, or fertilizers has contributed to locations with elevated levels of soil phosphorous. Soils with high-level phosphorous have greater potential for phosphorous transport off-site through surface runoff in soluble and sediment-bound forms. While phosphorous is not toxic and does not represent a direct health threat to human or other organisms, it does represent a serious indirect threat to water quality (Peavy et al., 1985). Phosphorous is often the limiting nutrient in surface waters. When an increase in phosphorous occurs in a phosphorous-limited water body, the growth of algae and/or water plants can be accelerated, thus, beginning the process of eutrophication. Eutrophic conditions can negatively affect water quality by causing low dissolved oxygen levels, excessive aquatic growth, increased sedimentation, and greater turbidity. Managing our soil and water resources requires an understanding of the soil phosphorous cycle.

Soil Phosphorous Cycle

The literature is quite extensive and contains very detailed information on the topic of the soil phosphorous cycle. Rather than present such an exhaustive review on the subject, an overview of the soil phosphorous cycle is given.

Pierzynski et al. (1994) presents a discussion of the soil phosphorous cycle. Phosphorous, found in all terrestrial environments, primarily originates from the weathering of soil minerals and other more stable geological materials. As phosphorous is solubilized in soils by the chemical and physical processes of weathering, it is accumulated by plants and animals, reverts to stable forms in the landscape, or is eroded from soils and deposited as sediments in freshwaters or oceans. Soil factors that control the conversion rate of phosphorous between the inorganic and organic forms regulate the short- and long-term fate of phosphorous in the environment. The soil phosphorous cycle consists of many complex chemical and microbiological reactions.

Organic soil phosphorous includes both biologically available organic phosphorous and resistant organic phosphorous (Foth and Ellis, 1997). Common forms of organic phosphorous found in soils include inosital phosphates, phospholipids, phosphoproteins, sugar phosphates, and nucleic acids. Soil organic phosphorous transformations are primarily mineralization-immobilization reactions mediated by soil microorganisms and phosphorous uptake by plant roots. Studies have shown that as much as 50% of the phosphorous transported in runoff can be soluble organic phosphorous (Pierzynski et al., 1994). It has also been found that organic phosphorous has been somewhat correlated to extractable, soil test phosphorous (Sharpley, 1985; Sharpley et al., 1987).

Inorganic soil phosphorous can make up to 50 – 70% of the total phosphorous in soils; the major soil inorganic phosphorous transformations of importance include the fixation of phosphorous in insoluble forms by adsorption and precipitation reactions, and the solubilization of phosphorous by desorption reactions and mineralization dissolution (Pierzynski et al., 1994). In soils that are moderately weathered, the dominant minerals are apatites. In highly weathered soils, iron (Fe) and aluminum (Al) precipitates become the major mineral sources of phosphorous.

The pH is a controlling factor that determines phosphorous solubility, as described by (Johnson et al., 1997). Maximum phosphorous availability occurs in a pH range of 5.5 to 7.2, where the ions present will be either monovalent ($H_2PO_4^-$) or divalent (HPO_4^{2-}), both of which are readily available for plant uptake. At soil pH levels below 5.5, iron (Fe), aluminum (Al), and manganese (Mn) react with phosphorous to form insoluble compounds. When soil pH exceeds 7.2, phosphorous will complex with calcium (Ca) to form plant unavailable phosphorous.

The transport of phosphorous in runoff can occur in both particulate and soluble forms. Particulate phosphorous includes all solid phase forms and phosphorous sorbed by soil particles and organic matter eroded during runoff and is the major contributor (75–90%) of phosphorous transported from cultivated land (Burwell et al., 1977). Runoff from pasture and grassland tends to have higher fractions of soluble phosphorous forms. Edwards et al. (1996) reported greater than 74% of total phosphorous runoff from fescue plots treated with poultry litter was in the soluble form.

Soil Test Phosphorous

Soil test phosphorous is an availability index that is correlated to the amount of phosphorus that will be available, or become available, to a plant during a growing season. The fraction of soil phosphorous that the plant can readily use, available soil phosphorous, makes up about one percent of the total phosphorous in soils (Johnson et al., 1997). The compound solubilities present in the soil effect the availability of inorganic phosphorous. As more phosphorous from solution is extracted, it may be replaced from the precipitated or solid phase. The chance for phosphorous in soil solution increases as the amount of total soil phosphorous increases.

There are several methods used to estimate the available soil phosphorous. The overwhelmingly largest fraction of soil samples are tested for available phosphorus by extraction with dilute acid solutions (Fixen and Grove, 1990). The Mehlich III soil test method (Mehlich, 1984) is one extraction procedure that is used to measure plant available phosphorous. Cai et al. (1997) found from a comparison of four extractants (Modified Troug, Mehlich III, Olsen, and ion-exchange resin) that Mehlich III provided better detection of phosphorous-sufficient and phosphorous-deficient soils under the conditions tested. Bray P is another standard extractant method commonly used. The Soil, Water, and Forage Laboratory at Oklahoma State University uses the Mehlich III extractant method to obtain and report soil test phosphorous.

Available soil phosphorous is a vital input parameter for many hydrological/water quality models. It is used in predicting soluble phosphorous transport using a soil extraction coefficient and in sediment-bound phosphorous transport using phosphorous enrichment ratios (Sharpley et al., 1982).

Number of Soil Samples

The goal of most soil sampling programs is to obtain an average value for some soil property over the area being sampled. When presented with this task, one will inevitably need to know how many soil samples are needed and the associated level of confidence in the estimated mean. The level of variation in the parameter being sampled will impact the number of observations needed to estimate the mean. Soils, by nature, are heterogeneous and have high spatial variability.

The literature review revealed that most work to date, pertaining to estimating sample size and confidence intervals, may be generally grouped into three categories: field-scale sampling procedures, assumed underlying probability distributions, and geostatistics. Each of these categories is addressed in the following sections.

Field-scale

Classical statistical approaches to soil sampling have assumed each observation to be independent and identically distributed (Sabbe and Marx, 1987). Numerous studies have examined the spatial variability of soil properties and have described soil sampling methods for obtaining representative estimates (Cline, 1944; Rigney, 1956; Peterson and Calvin, 1983; Russo and Bresler, 1981; Sisson and Wierenga, 1981; Webster and Oliver; 1992). While different methods of sampling may have been studied, all of these cited works have focused on field-scale variations and sampling.

A study by Keogh and Maples (1967) suggested that the size of the field did not affect the coefficient of variation appreciably, especially above a minimum size of 8.1 to 12.2 ha (20 to 30 acres). It was also determined that more samples were required to determine soil test phosphorous than other soil fertility parameters. Cameron et al. (1971) also reported that the number of samples needed to estimate the field average did not increase drastically with an increase in field size. It is recommended by Zhang and Johnson (1997) to collect at least 15 random core samples per field to comprise a representative composite field soil sample.

Underlying Distributions

A known, or assumed, underlying probability distribution of the data means is the basis of classical statistical methods for estimating the required sample size, or number of observations, needed for estimating a soil parameter.

When the data tend to follow a known distribution, using classical statistical methods is a good approach. The number of observations, n, needed to achieve a desired estimation variance is given by Cline (1944),

$$n = t_\alpha^2 s^2 / (x - \mu)^2 \tag{2.1}$$

where t_α is Student's t at the chosen level of probability, α, s^2 is the estimated variance, and $x - \mu$ is the acceptable deviation from the true mean, μ. Of course, Student's t is based on the assumption of normality (Steel and Torrie, 1980). Warrick et al. (1986) states the appropriate Student t value should only be used to estimate confidence intervals; for estimating n, the two-tailed normalized deviate, z_α, should be used instead of Student's t. This method would require an estimate of the expected mean and variance for the property being sampled.

These methods assume that the means of the data follow a normal distribution. This assumption is explained by the Central Limit Theorem, and is quoted by Ott (1984) as,

> "If random samples of n measurements are repeatedly drawn from a population with a finite mean μ and a standard deviation σ, then, when n is large, the relative frequency histogram for the sample means (calculated from the sample means) will be approximately normal with mean μ and standard deviation $\sigma/\sqrt{n}$."

Haan (1977) expands the explanation of the Central Limit Theorem by stating that the population must consist of identically and independently distributed random variables and discusses some generalized conditions under which it can apply.

Many physical, chemical, and biological properties of soils display skewed distributions that are better approximated by the two parameter lognormal distribution (Parkin et al., 1988). Parkin et al. (1990) evaluated five methods for calculating confidence intervals for a lognormally distributed variable. They used four test lognormal populations, each with known means and variances. A nonparametric method was developed and was determined to be a good alternative to the others for calculating confidence intervals when sample sizes were grater than 20 and the underlying population deviated from true lognormality.

Geostatistics

Another focus in the literature pertaining to soil sampling was on the use of geostatistics, which is a form of statistics dealing with spatially referenced data. Geostatistics assumes data properties are correlated over space, so that data points close together tend to be more alike than those that are far apart. In other words, it can be used where the assumption of independent observations may not be valid.

The theory of regionalized variables, those distributed in space, was developed by G. Matheron (1963) in the 1960s. The application of this theory led to the methodology for geostatistics, which began in mining and geology for the assessment of ore bodies. The term "kriging" after D.G. Krige, described the method of producing the best estimate of the unknown value of a parameter at some location within an ore deposit. (Warrick et al., 1986; Sabbe and Marx, 1987).

McBratney and Webster (1983) presented a method for determining the required number of observations, or soil samples, needed for regional estimation of soil properties based on regionalized variables. They created semi-variograms for soil properties and showed how kriging could be used to estimate a soil property at unvisited sites. If the semi-variogram is already known, the kriging variance for any particular sampling scheme can be determined. They demonstrated the advantages of kriging from grid samples when estimating the soil properties over a region. When computing the required sample size, three-and-half-fold to nine-fold gains in efficiency over that estimated by classical theory for simple random sampling was reported.

Grid sampling was used by Gupta et al. (1997) to examine spatial variability and sampling strategies for site-specific farming at two farms. Semi-variogram models were used to describe the correlation structure of nutrients. Determination of sample size and optimum sampling interval, considering the correlation and semi-variance characteristics of the nutrients, were considered. Optimal sampling grids were determined for the nutrients based on spatial variability.

Geostatistics can be used successfully on the basin-scale where there is a gradual change spatially of the measured parameter, such as soil phosphorous. When abrupt changes occur, as along field or property boundaries, the use of geostatistics becomes limited. Basins, or watersheds, that contain poultry-related activities where the litter is spread on the fields may not be suited to geostatistics for estimating soil sampling size because there may exist significant differences in soil phosphorous levels from one field to the next. These differences are based on the history and levels of litter application that differ across field boundaries.

CHAPTER 3

WATERSHED DATA DESCRIPTIONS

Data from four Oklahoma watersheds and the Arkansas portion of the Lake Eucha basin were examined and used for this research. The watershed/basin data were chosen based on the availability and diversity of geographic location and land use. The watersheds/basins (locations) used in this study are listed below.

- Upper Little Deep Fork Creek Basin (Eastern-central Oklahoma)
- Battle Branch Watershed (Northeastern Oklahoma)
- Peacheater Creek Watershed (Northeastern Oklahoma)
- Haw Creek Watershed (Southeastern Oklahoma)
- Lake Eucha Basin (AR portion) (Northwestern Arkansas)

In general, all of the watersheds are mostly rural agricultural settings. Four of the five watersheds contain poultry production activities and have had varying time lengths of involvement in the industry. The watersheds are in close enough proximity that they experience similar climates and growing seasons. The general locations are shown in Figure 3.1.

The same types of data were not available from each watershed. For two of the locations, Lake Eucha (Arkansas portion) and Haw Creek (Oklahoma),

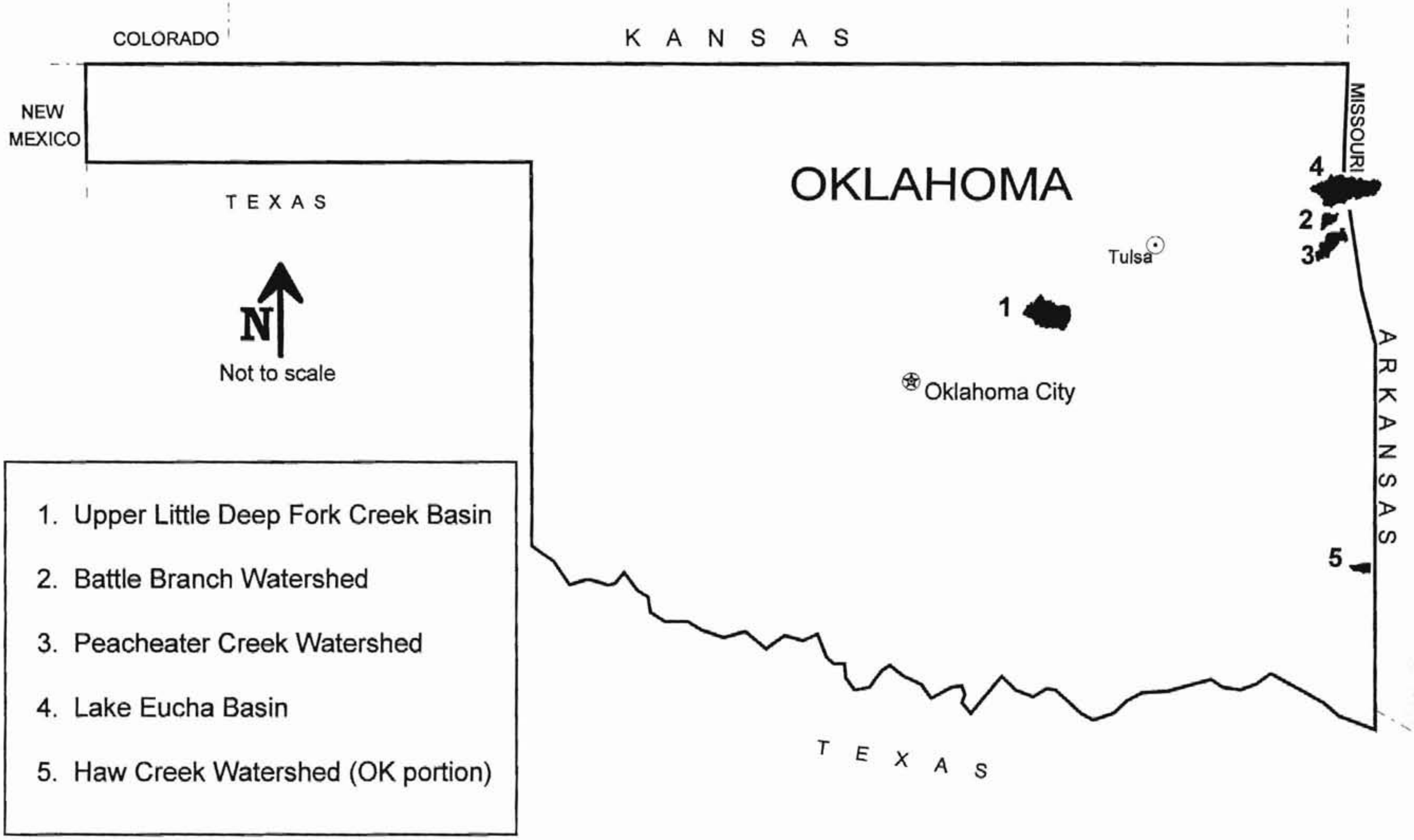


Figure 3.1. Watershed/Basin locations.

only soil test phosphorous data were available. Data layers for each of the watersheds/basins were originally obtained and developed for other projects pertaining to nonpoint source assessment of phosphorous and sediment loadings, where the computer modeling was, or is to be, performed by the Biosystems and Agricultural Engineering Department, Oklahoma State University. The watershed/basin data sets used in this study are described in the following sections. The Geographical Resource Analysis Support System (GRASS) geographic information system (GIS) developed by the U.S. Army Corps of Engineers (U.S. Army, 1991) was used to compile and organize the GIS data.

Upper Little Deep Fork Creek

The Upper Little Deep Fork Creek basin is located in the southwest corner of the northeast quadrant of Oklahoma (Figure 3.1). It covers approximately 39,500 ha (97,500 acres) and lies almost entirely in Creek County. The western 2000 ha (5,000 acres) stretch into neighboring Lincoln County. The Little Deep Fork Creek flows generally east and into the Deep Fork River, which is a tributary of the North Canadian River. A site tour of the basin revealed the major industries to include oil and gas exploration and agriculture. The agricultural activities are hay production, grazed cattle, and small grain production. The basin is approximately 40% forest and 55% grasslands, with the remaining 5% urban or other.

Soils

The soils are described by the USDA Soil Survey for Creek County. The soils are in three broad, general associations, which are sandy soils of the forested areas, dark soils of the prairies, and soils of the bottom lands. Each association is dominated by soils that developed from similar or related parent materials, have some characteristics in common, and contain many small areas that belong to one of the other two associations (SCS, 1958).

Digital soil type data boundaries for Creek and Lincoln Counties were obtained from Oklahoma soil surveys that had been digitally scanned. The attributed soil characteristic information was obtained from the U.S. Department of Agriculture, Natural Resources Conservation Service's National Map Unit Interpretation Record (MUIR) Database (USDA-NRCS, 1994) for Creek and Lincoln Counties. The MUIR data set is a collection of soil and soil-related properties, interpretations, and performance data for a soil survey area that is to be used in conjunction with county soil surveys. The data are stored in a retrievable relational database with information for most of the U.S. counties. A list of the soils within the study area, with selected attributes, can be found in Appendix A.

The majority of the soils within the basin are from the Darnell and Stephenville series. The Darnell series are very shallow acid soils developed over reddish sandstones. They are too shallow for cropping and are used mainly for woodland pasture. The Stephenville series are of medium depth over the parent materials of soft reddish sandstone and are slightly acid. Sandstone outcrops are common in both series (SCS, 1958). The distribution of the three general soil associations are depicted in Figure 3.2.

Land Use

The land use data were obtained from the Oklahoma Conservation Commission (OCC), Water Quality Division in digital format. The watershed was field inspected by OCC personnel and was divided into 29 categories. The major category ratings consisted of "poor" to "good" for grasslands and forestlands, croplands, and a few other smaller agricultural categories. Since the OCC land use data did not contain urban areas, there were some "holes" in the data set where urban activities were present. To make a complete land use data file, Soil Conservation Service (1985) Oklahoma land use digital data were used to fill in the lacking portions. This facilitated the creation of a complete land use data set. Grasslands cover approximately 55% of the basin, forestlands cover about 40%, and the remaining 5% is made up of cropland, urban, oilfield activities, and "other" land uses. Figure 3.3 shows the distribution of the major land uses.

Soil Test Phosphorous

The soil sampling was performed by OCC personnel. The soil sampling plan called for collecting a proportionate number of samples based on the percentage of land use type. The exception to this was forest land use, which was assumed to have relatively uniform soil nutrient levels. This was done for a predetermined number of composite samples over the entire basin. The basin

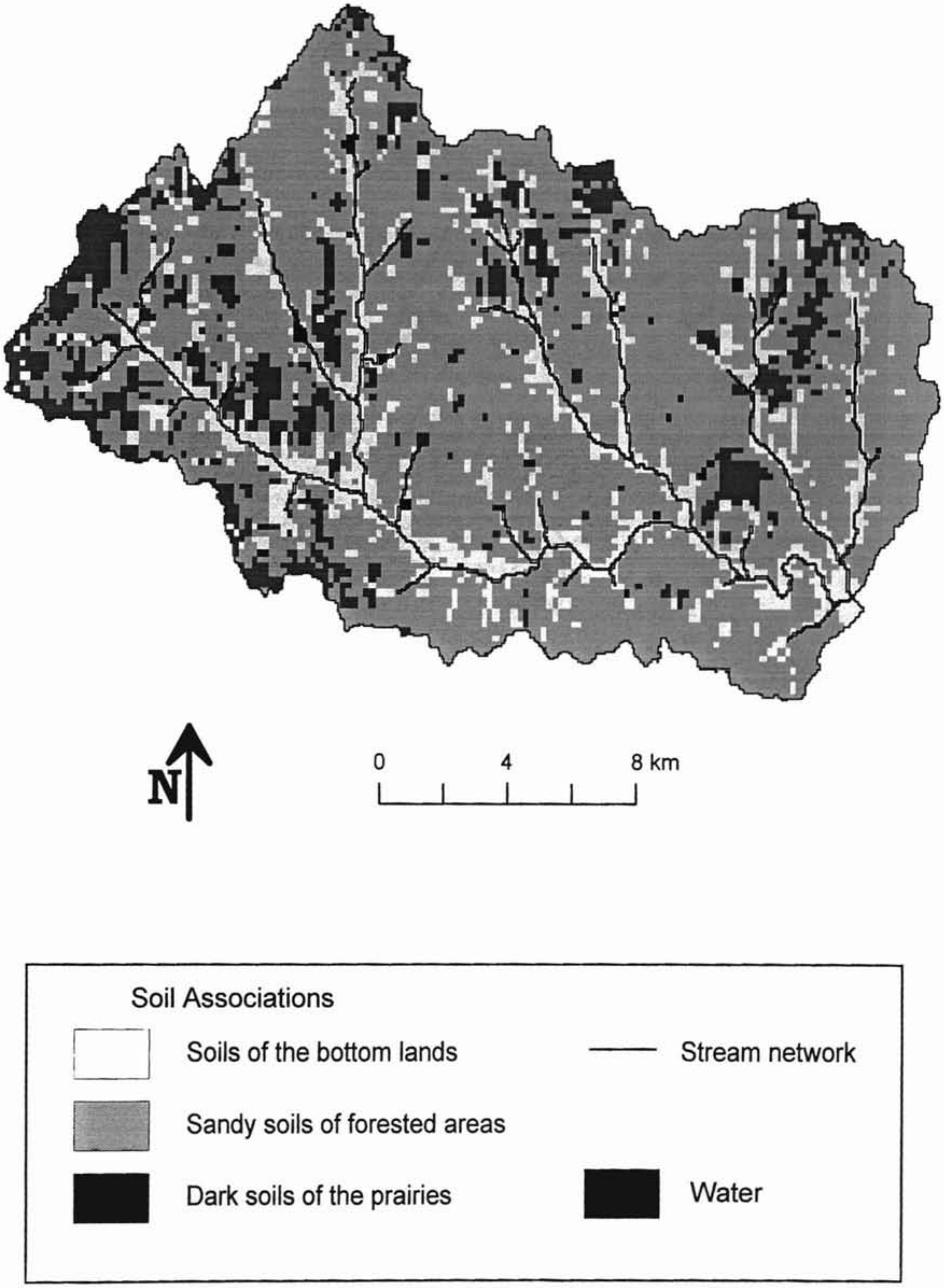


Figure 3.2. Upper Little Deep Fork Creek Basin soil associations.

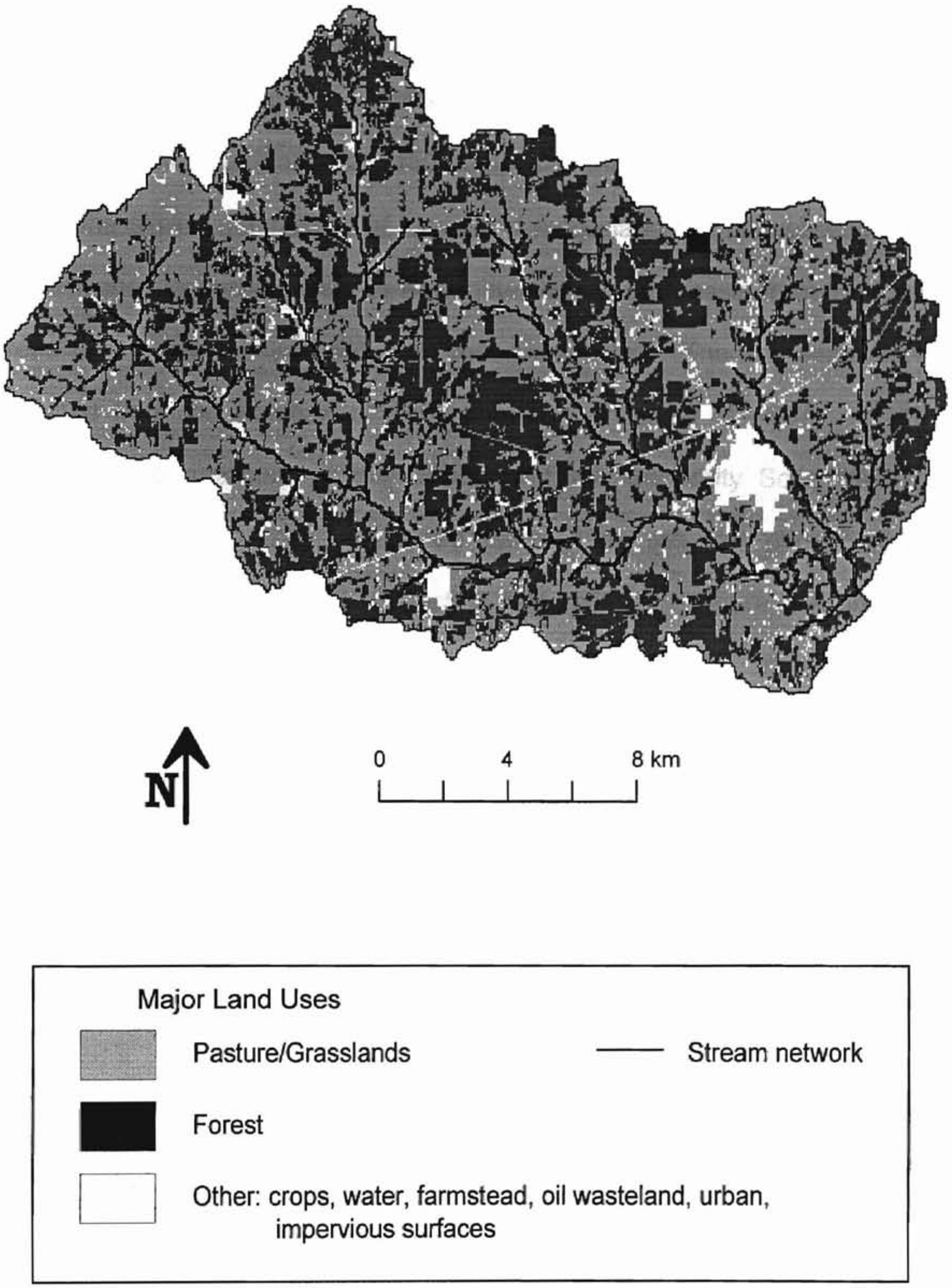


Figure 3.3. Upper Little Deep Fork Creek Basin land use.

was divided into one square mile grids and at least one composite soil sample was collected from each grid. General locations of where the samples were taken were recorded and a digital soil sample location map was developed by OCC. Originally, approximately 150 composite samples were collected. Later, due to the similar magnitudes of forest and range land soil phosphorous levels, 20 additional composite samples were obtained over the basin from the forested areas to improve the variability and mean estimates. The soil samples were collected from the Summer of 1996 through the Spring of 1997. All of the samples were analyzed by the Oklahoma State University Soil, Water, and Forage Laboratory. The results are reported as soil test phosphorous as measured by the Mehlich III extraction test method. Figure 3.4 shows the approximate locations where the composite soil samples were taken. A summary of the soil test phosphorous data is shown in Table 3.1. The complete soil test phosphorous data can be found in Appendix B.

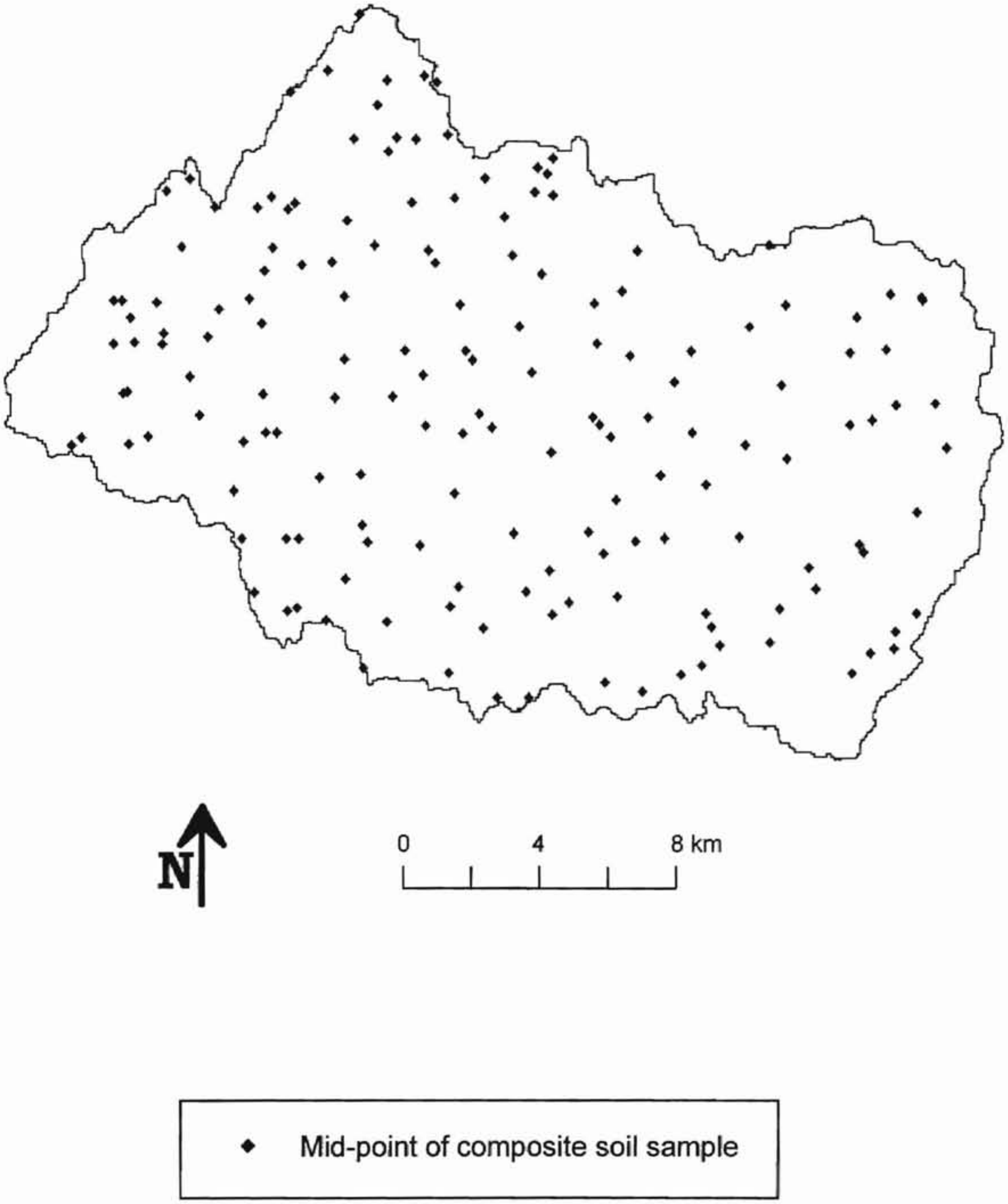


Figure 3.4. Upper Little Deep Fork Creek Basin composite soil sample locations.

Table 3.1. Upper Little Deep Fork Creek Soil Test Phosphorous* Summary

Land use	Mean (mg/kg)	Median (mg/kg)	Std. Dev. (mg/kg)	Range (mg/kg)	Count (no.)
Forestland**					
Stable	19	19	4	13 - 25	8
Moderately used	17	16	5	7 - 27	14
Heavily used	19	17	6	11 - 29	12
Grassland**					
Good condition	16	14	7	7 - 35	41
Fair condition	15	14	5	8 - 30	38
Poor condition	16	13	9	4 - 53	45
Unmanaged	17	13	11	6 - 38	6
Cropland					
Small grains	16	19	7	9 - 49	3
Salt or Oilfield Induced Erosion			9	9	3
Feedlot	275	275	304	60-490	2

* Mehlich III phosphorous

** Forestland
  Stable: undisturbed, 0 – 1% bare soil
  Moderate use: 1 –10% bare soil
  Heavy use: > 10% bare soil
 Grassland
  Good condition: < 1% bare soil
  Fair condition: 1 – 5% bare soil
  Poor condition: 5 – 20% bare soil
  Unmanaged: 20 – 100% bare soil with erosive areas

Battle Branch and Peacheater Creek

An extensive soil sampling plan for these watersheds was implemented as part a United States Department of Agriculture (USDA) hydrologic unit area project for the Illinois River Basin, which is located in northwest Arkansas and northeast Oklahoma. The project sanctioned the delineation of individual pasture fields and the soil sampling of each, as reported by Sabbagh et al. (1995). Battle Branch and Peacheater Creek watersheds are located within the Illinois River Basin. These watersheds contain extensive poultry industry activities. The data were used for this study and are described below as in the referenced report.

The Battle Branch Watershed is located in southern Delaware County in northeastern Oklahoma (Figure 3.1). The watershed area covers about 2,200 ha (5,500 acres). The topography is primarily rough steep hills with a blackjack-postoak tree cover. The major land use is agriculture. Poultry industry activities including broilers, layers, breeder hens, and pullets, are present. The survey data indicates there are approximately 25 poultry houses within the watershed.

The Peacheater Creek Watershed is located in Adair County in northeastern Oklahoma (Figure 3.1). The watershed area covers approximately 6,500 ha (16,000 acres). The watershed is in the Ozark Highland Land Resource Area. The topography is primarily rough steep hills with a blackjack-postoak tree cover and the major land use is agriculture. There are 59 poultry houses located within the Peacheater Creek watershed. These operations maintain an average of 1.1 million broilers, layers, breeder hens, and pullets per year. In addition there are nine dairies with a total of 800 dairy and about 3000 unconfined beef cattle located within the watershed.

Soils

Digital soil type data for Battle Branch and Peacheater Creek watersheds were obtained from the soil surveys for Adair County and Delaware County (SCS, 1965; SCS, 1970) that had been digitally scanned. Values for other soil characteristic, such as clay content, bulk density, slope length, erodibility factor, organic carbon content, and hydrologic group were estimated from the same soil surveys.

The Battle Branch watershed includes 19 different soil types. A complete list of the soils for Battle Branch are in Appendix A. The predominant soils in the watershed are in the Clarksville and Baxter-Locust associations. The Clarksville soils are cherty silt clay loam soils and generally have high steep slopes with high runoff potential. The Baxter and Locust soils are cherty silty clay loam soils and are found on the nearly level to gently sloping ridge tops (SCS, 1970). Figure 3.5 depicts the distribution of the soil associations within the watershed.

The Peacheater Creek watershed includes 18 different soil types and are listed in Appendix A. The predominant soils are in the Bodine-Dickson association. The Bodine soils are loamy soils and generally have steep slopes with high runoff potential (SCS, 1965). The spatial distribution of soil types are shown in Figure 3.6.

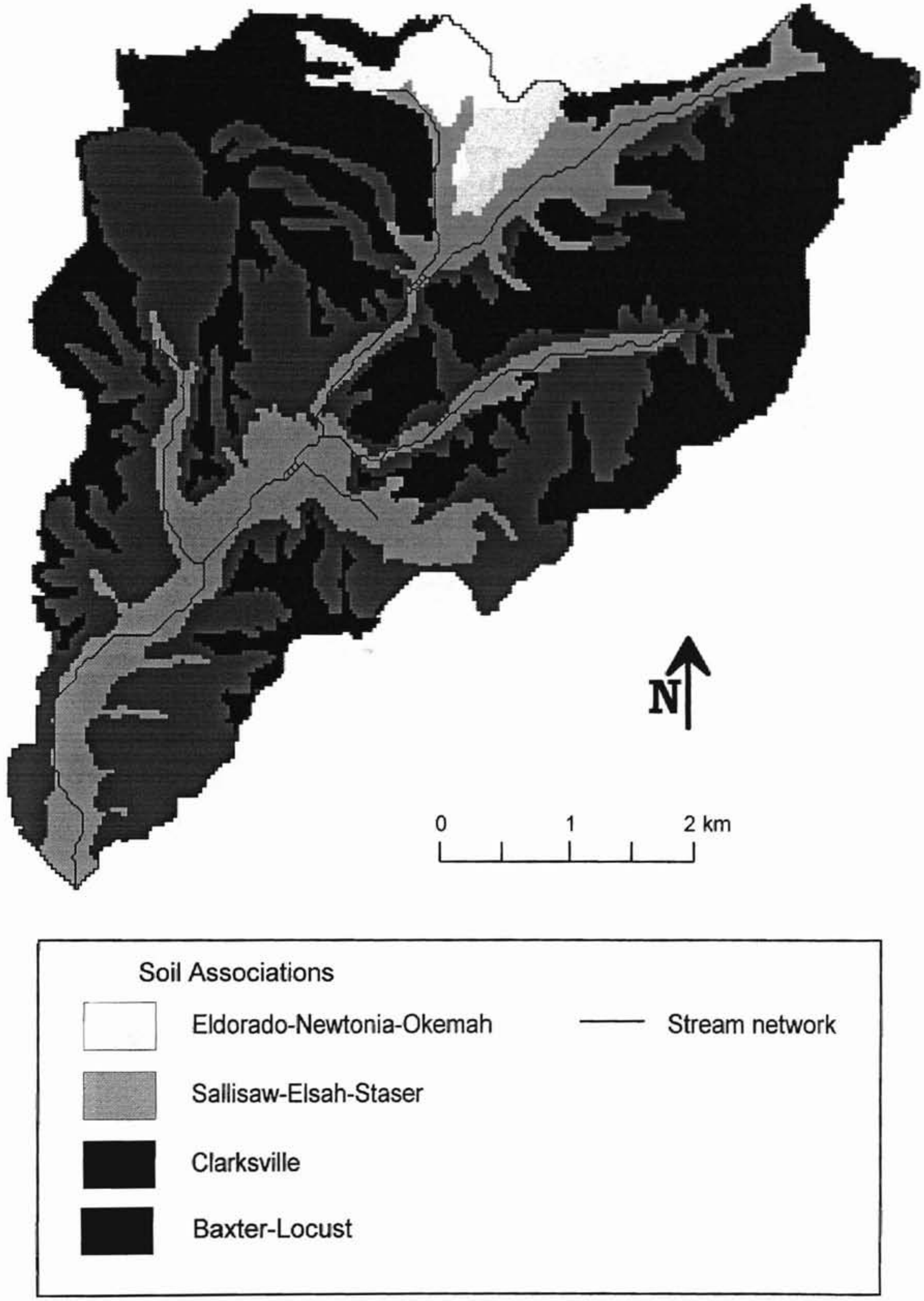


Figure 3.5. Battle Branch Watershed soil associations.

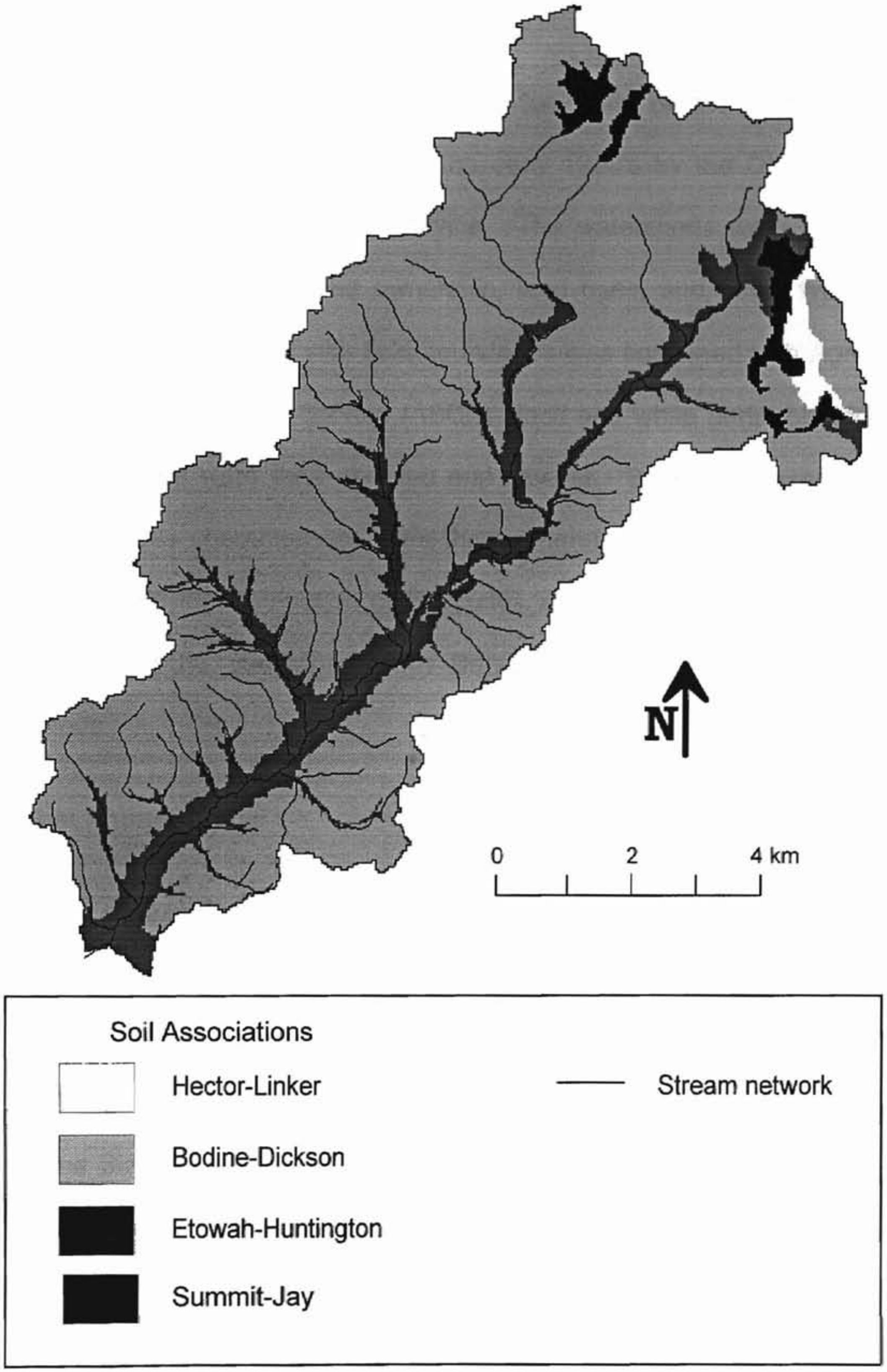


Figure 3.6. Peacheater Creek Watershed soil associations.

Land Use

A comprehensive land use inventory for Battle Branch and Peacheater Creek Watersheds was conducted in the early 1990's by the Oklahoma State University Cooperative Extension Service. The watersheds were divided into individual fields, based on land ownership, land uses, and cover types. The detailed land use inventory with field boundaries was combined with Agricultural Stability and Conservation Service (ASCS) black and white aerial photography. These boundaries were then digitized and labeled. Spatial representations of soil and land use characteristics were then generated with a GIS. Figures 3.7 and 3.8 indicate the distribution of land uses within the watersheds. There is approximately 60% pasture in both Battle Branch and Peacheater Creek watersheds.

Soil Test Phosphorous

The pasture fields in the study areas were the fields of interest, since the effects of poultry liter application to pasture were under evaluation. Soil sampling for most of the fields was done so that at least one composite sample was obtained for each of the fields sampled. Soil sampling was performed by the Oklahoma State University Cooperative Extension Service. The samples were collected over the period of 1991 through 1994. The resultant soil test phosphorus level for the composite sample was then assigned to the respective field. All of the samples were analyzed by the Oklahoma State University Soil,

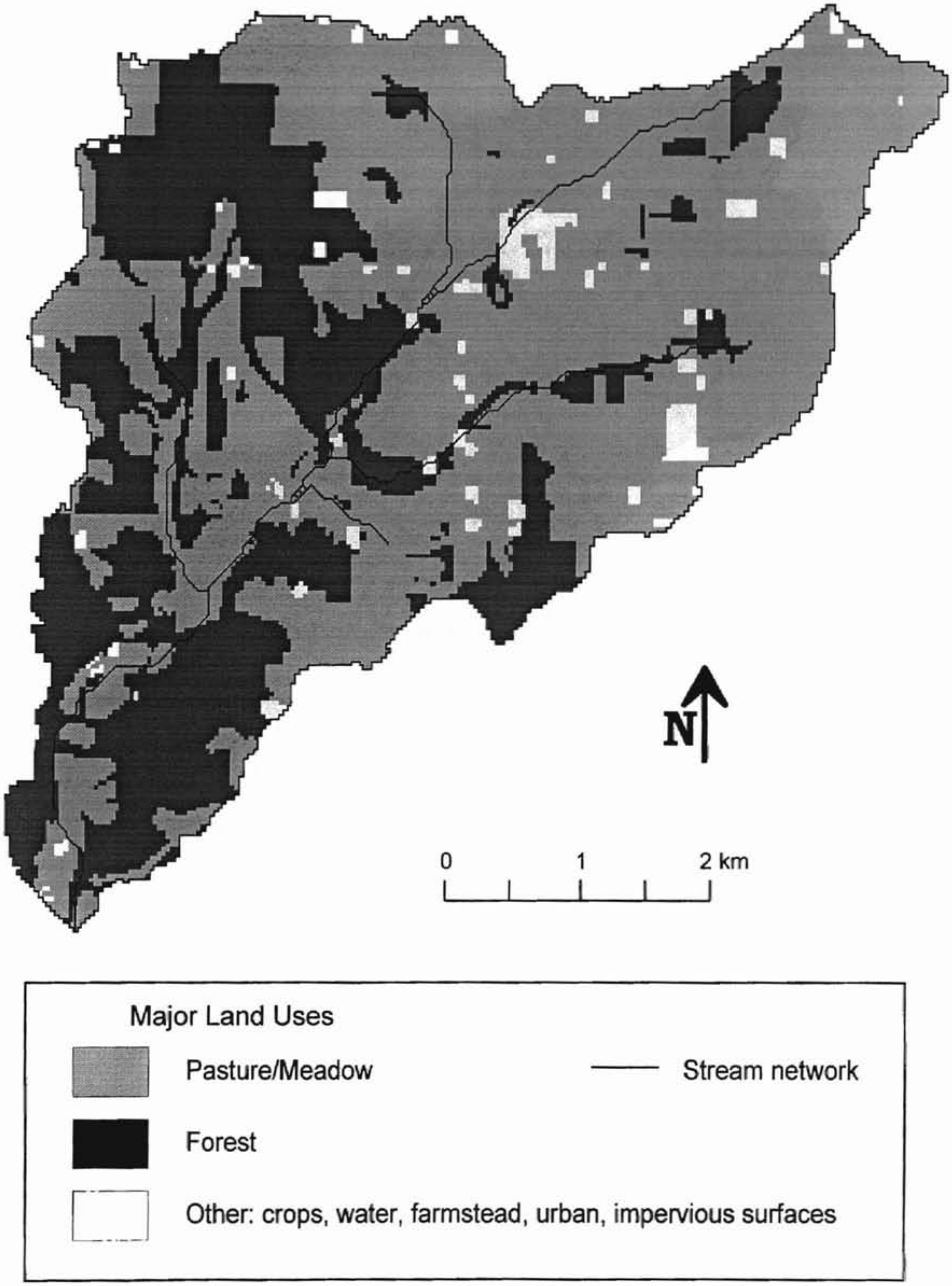


Figure 3.7. Battle Branch Watershed land use.

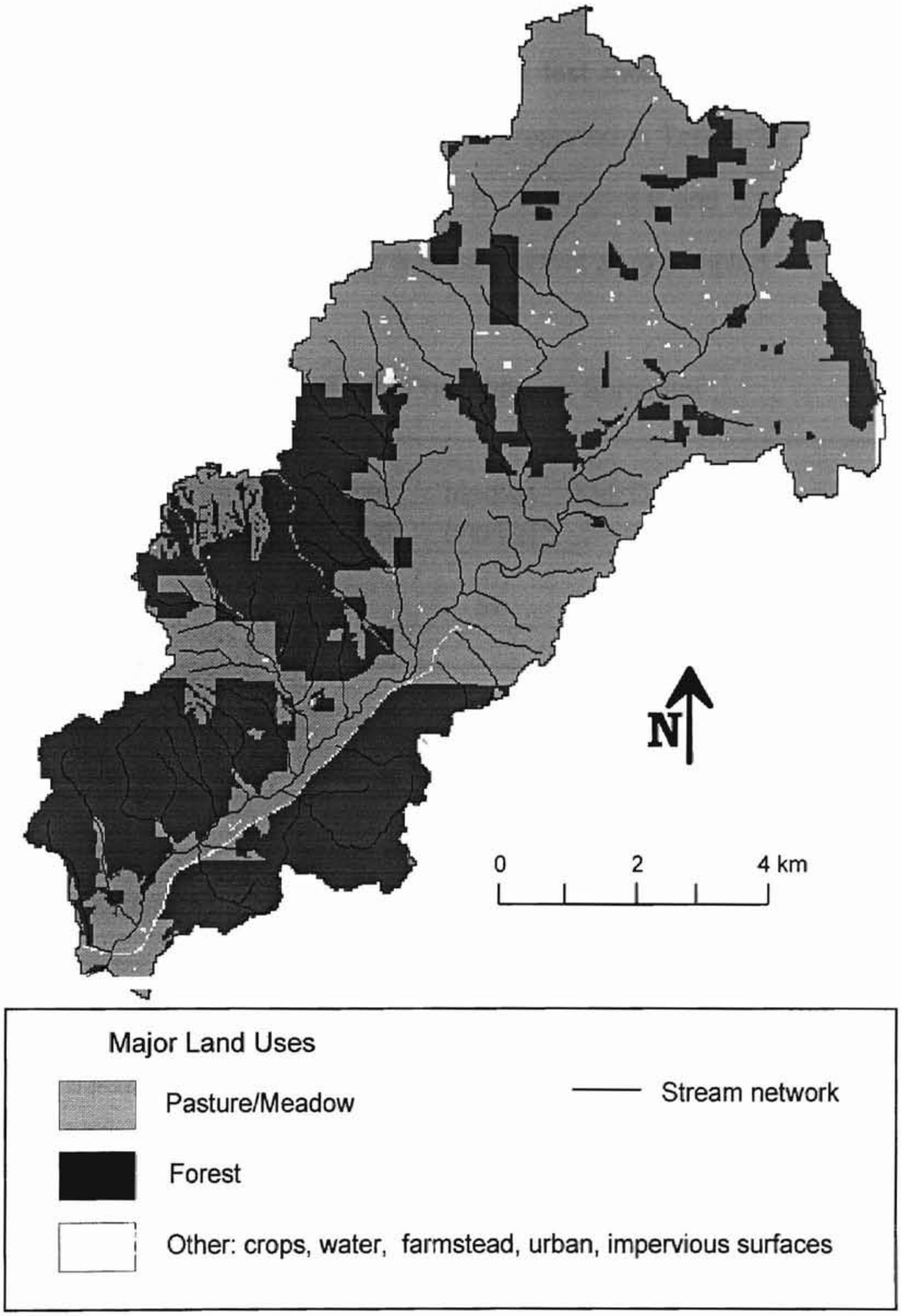


Figure 3.8. Peacheater Creek Watershed land use.

Water, and Forage Laboratory. The results are reported as soil test phosphorous as measured by the Mehlich III extraction test method. A summary of the completed soil test phosphorous results is reported in Tables 3.2 and 3.3. The detailed results are found in Appendix B. Figures 3.9 and 3.10 show the field boundaries of pastures within the watersheds that were sampled.

Table 3.2. Battle Branch Soil Test Phosphorous* Summary

Land use	Mean (mg/kg)	Median (mg/kg)	Std. Dev. (mg/kg)	Range (mg/kg)	Count (no.)
Pasture	54	52	38	7 - 164	230

* Mehlich III phosphorous

Table 3.3. Peacheater Creek Soil Test Phosphorous* Summary

Land use	Mean (mg/kg)	Median (mg/kg)	Std. Dev. (mg/kg)	Range (mg/kg)	Count (no.)
Pasture	93	74	89	9 - 490	255

* Mehlich III phosphorous

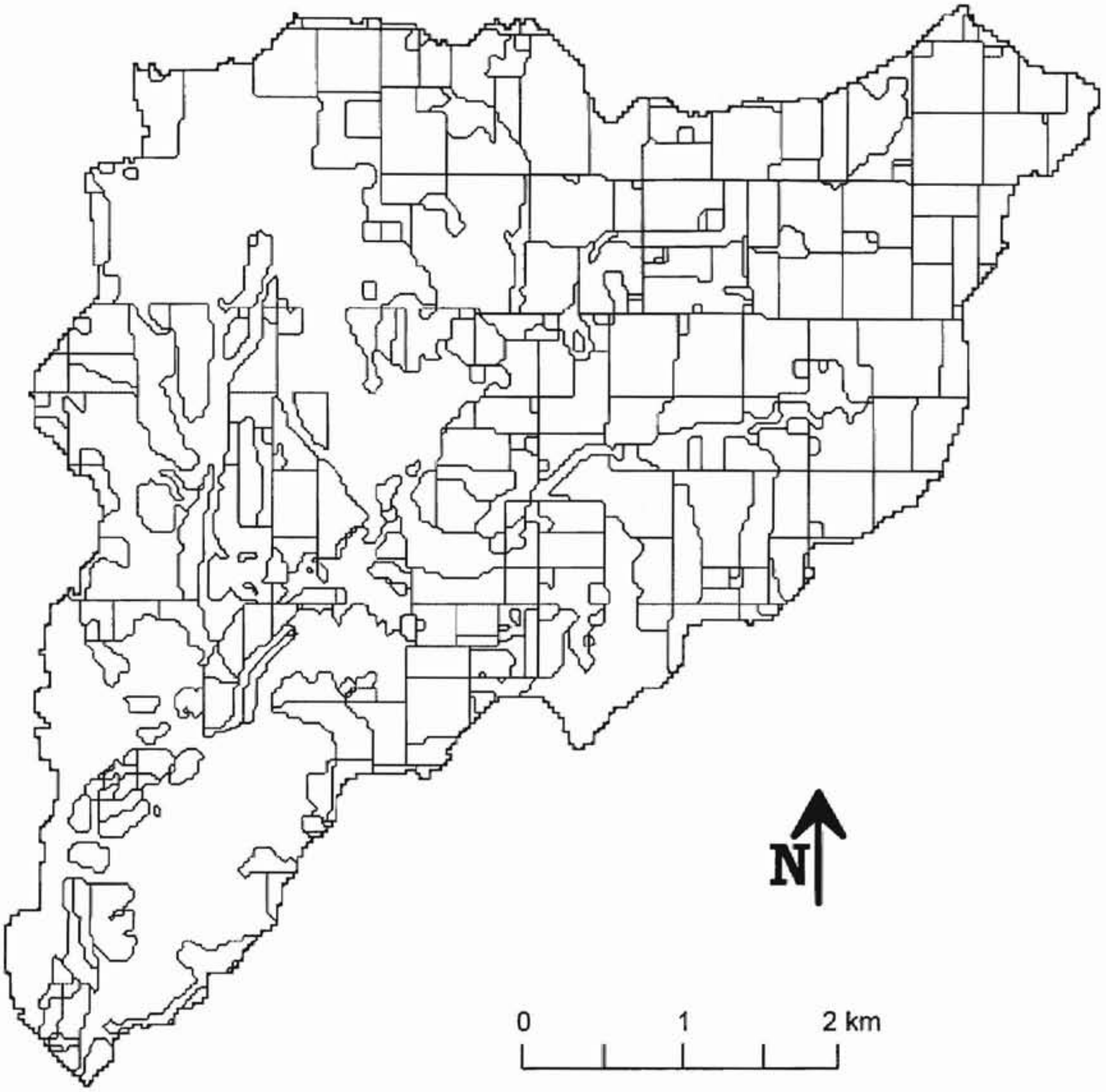


Figure 3.9. Battle Branch Watershed field boundaries.

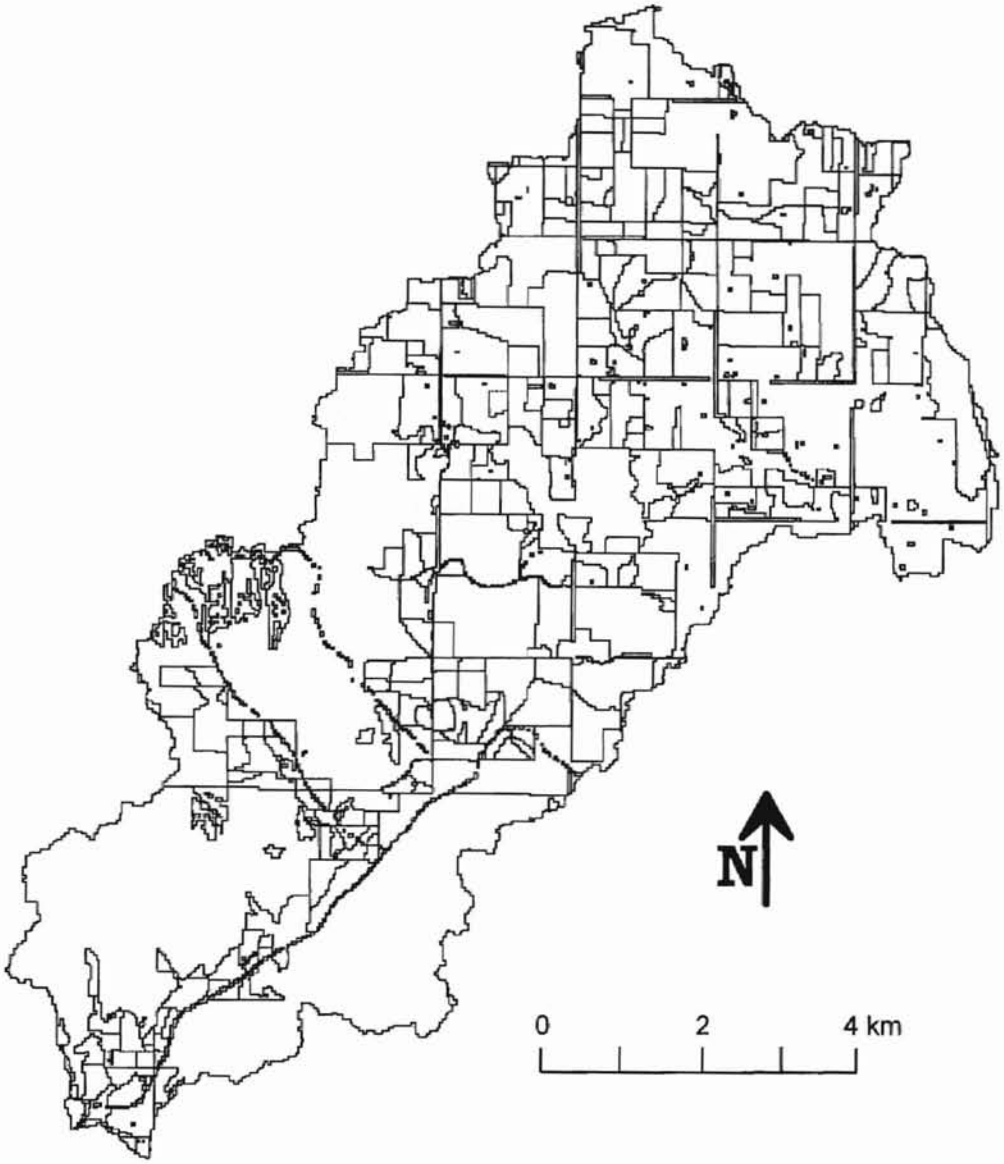


Figure 3.10. Peacheater Creek Watershed field boundaries.

Haw Creek

Haw Creek Watershed, part of the Wister Lake Basin, is located in eastern LeFlore County in southeastern Oklahoma and stretches into Arkansas (Figure 3.1). Only data from the Oklahoma portion was used for this study. It covers approximately 2,000 ha (5,000 acres) on the Oklahoma side. It is a rural watershed containing poultry industry activities. There are 20 poultry houses (approximately 20,000 birds per house) located within the Oklahoma portion of the watershed (Joe Bullard, Oklahoma State University Cooperative Extension Service, personal communication).

Soils

Soils within the watershed can be grouped into three general soil associations, and are described by the Soil Survey of LeFlore County (SCS, 1981). The Neff-Kenn-Ceda association is described as nearly level to gently sloping, moderately drained loamy soils. They are located on the flood plains and are subjected to occasional flooding. The Carnasaw-Octavia-Pirum association is deep, gently sloping to steep, well drained stony soils. They are located on ridges and mountains. The Sallisaw-Stigler association is deep, nearly level to moderately steep, well drained loamy soils. They are located on the uplands. Most of this association is in pasture.

Land Use

The major land use within the watershed is forest. Haw Creek is located within the Ouachita Mountains National Forest area, but also contains private land ownership. There is approximately 560 ha (1,400 acres) of pasture on the Oklahoma side (Joe Bullard, Oklahoma Cooperative Extension Service, personal communication), located mostly in the valleys and along the stream banks.

Soil Test Phosphorous

The pasture fields were soil sampled by the Oklahoma State University Cooperative Extension Service during the Summer and Fall of 1995. Approximately 90% of the pasture area within Haw Creek Watershed was sampled (Joe Bullard, Oklahoma State University Cooperative Extension Service, personal communication). The samples were analyzed by the Oklahoma State University Soil, Water, and Forage Laboratory. The results are reported as soil test phosphorous as measured by the Mehlich III extraction test method. Table 3.4 summarizes the soil test phosphorous results. The complete data set is found in Appendix B.

Table 3.4. Haw Creek Soil Test Phosphorous* Summary

Land use	Mean (mg/kg)	Median (mg/kg)	Std. Dev. (mg/kg)	Range (mg/kg)	Count (no.)
Pasture	104	54	110	2 - 515	82

* Mehlich III phosphorous

Lake Eucha (Arkansas)

Lake Eucha Basin is located in northeastern Oklahoma and northwestern Arkansas (Figure 3.1). The Lake Eucha basin is approximately 93,000 ha (230,000 acres), with 40% in Benton County, Arkansas and the remainder in Delaware County, Oklahoma. This basin is generally known for its extensive poultry industry activities. Soil sampling of all pasture for the Arkansas portion has been completed and the results were used in this study.

A 1996 survey by the Water Quality Division, Oklahoma Conservation Commission concluded there were 489 poultry houses on the Arkansas side of the basin (OCC, 1996). These include houses for layers and broilers.

Soils

The soil types within the Arkansas portion of the basin are reported and discussed by Wagner and Woodruff (1997). The soils are part of the Clarksville-Nixa-Captina and the Clarksville-Noark-Nixa soil mapping units. Each mapping unit contains numerous soil types, where the majority of soils within both units are cherty to very cherty silt loams. Soil thickness can range from less than one meter to several meters, but the soils are generally thin.

Land Use

Land use is primarily forest and pasture while poultry is the major agricultural commodity produced in the basin (Wagner and Woodruff, 1997). A survey performed by the Oklahoma Conservation Commission during the Spring of 1996 (OCC, 1996) found a total of 714 poultry houses in the basin, with 489 located on the Arkansas side. The survey also found 18 hog houses and 5 turkey houses in production on the Arkansas portion. Approximately 40% of the Arkansas portion is pasture/hay, 55% forest, 3% crop, 1% urban, and 1% other (NRCS, 1995). This corresponds to approximately 14,900 ha (37,000 acres) of pasture/hay land use.

Soil Test Phosphorous

The pastures within the Arkansas portion of the Lake Eucha basin have been soil sampled. Samples were collected by the Arkansas Soil and Water Conservation Commission and analyzed by the University of Arkansas Soil and Water Laboratory. The sampling occurred during the period of 1994 through 1997. The soil phosphorous results are reported as soil test phosphorous as measured by the Mehlich III extraction test method. The summarized results are shown in Table 3.5. The complete data set is found in Appendix B.

Table 3.5. Lake Eucha (AR) Soil Test Phosphorous* Summary

Land use	Mean (mg/kg)	Median (mg/kg)	Std. Dev. (mg/kg)	Range (mg/kg)	Count (no.)
Pasture	164	152	118	5 - 490	261

* Mehlich III phosphorous

CHAPTER 4

NONPARAMETRIC METHOD DEVELOPMENT

Classic statistical techniques of predicting sample size are based on a normal distribution of the data means and identical and independent distributions of the original population. The data from high-level soil test phosphorous basins with poultry-related activities, such as the case for this study, may not meet all the assumptions required for the classical approach. Soil test phosphorous levels from these basins may not be independent from field to field within a basin. In other words, information from one sample may be partially duplicated in a sample close by. In addition, the data may not be identically distributed. The application of poultry litter to fields has created a basin-scale data set of soil test phosphorus that may contain multiple, non-identical distributions.

It would appear that the use of geostatistics might be applicable in these cases. Geostatistics assume a gradual, or at least measurable, change of the interested variable over space. Basins, or watersheds, with fields that receive poultry litter applications probably will not, however, readily lend themselves to geostatistics. This is because soil test phosphorous levels can abruptly change from field to field based on field ownership, litter application rates, and litter application histories. Yet, there is no current mathematical method to account for, or model, where and when this abrupt variance from field to field will be present.

The above discussion is the basis for development of a nonparametric approach for estimating soil sample size for "poultry" watersheds or basins. Soil test phosphorous from watersheds with poultry-related activities may be elevated and exhibit bi-modal distributions, due to the fact some pastures receive poultry litter and some do not. Also, there may be a high variability of soil phosphorous among pastures that do receive litter.

Data Analysis

The initial steps in the process of developing the nonparametric method of predicting required sample sizes were to evaluate the existing data sets, and evaluate the distributions to determine if they followed typical distributions. Descriptive statistics for each data set were presented in Chapter 3. For Upper Little Deep Fork Creek, the soil test phosphorous data were statistically analyzed and it was found that there was no significant ($\alpha = 0.05$) difference among the means of various conditions (good, fair, poor, unmanaged) for the Grasslands land use. The same was found among the means of the various conditions (stable, moderate use, heavy use) for Forest land use. Since there were no significant differences between the means, the various pasture/grassland conditions were combined into one pasture land use and the various forest land use conditions were combined into one forest land use. The other data sets were used as previously presented.

The ability to predict the expected soil test phosphorous levels based on physical soil properties was also examined. Regression statistics were performed on three of the data sets in an attempt to develop a prediction equation of soil test phosphorous derived from soil mapping units and selected soil characteristics.

Soil Test Phosphorous Probability Distributions

Haan (1977) presents three ways to determine if data follows a certain probability distribution. The first is to plot the data as frequency histograms, the second is to evaluate the linearity of the data plotted on the appropriate probability paper, and the third is to use statistical tests. All three have been performed with the data for this study.

The data sets were first plotted as frequency histograms, as shown in figures 4.1 through 4.6. Since the literature revealed that soil properties tend to follow lognormal distributions, the theoretical lognormal distribution was also plotted for each data set and included on each figure. Then, the data were plotted on lognormal probability paper. If the data corresponds to the distribution as represented by the probability paper, the data will plot as a straight line. The Weibull plotting position formula was used to rank and plot the data, as presented by Haan (1977). Rather than using cumulative probability as the x-axis, the standardized normal variable, Z, was used. Figures 4.7 through 4.12 show the lognormal probability plots.

Lastly, the data sets were statistically tested to determine if they were from a lognormal distribution. The Kolmogorov-Smirnov (K-S) and Chi-square tests were used. The K-S test is used by comparing the maximum deviation between the cumulative distribution function under the null hypothesis and the sample cumulative density function based on the number of observations to a tabulated value for the chosen significance. If the maximum deviation is less than the tabulated value, the null hypothesis is accepted. In this case, the null hypothesis was that the data were from a lognormal distribution. The Chi-square test makes comparison between the actual number of observations and the expected number of observations (expected according to the distribution under test) that fall in the class intervals. The class intervals were defined so that the expected number of observations in each class interval were the same, as suggested by Haan (1977). The results of the statistical tests are summarized in Table 4.1. The tests were conducted at $\alpha = 0.05$ significance level.

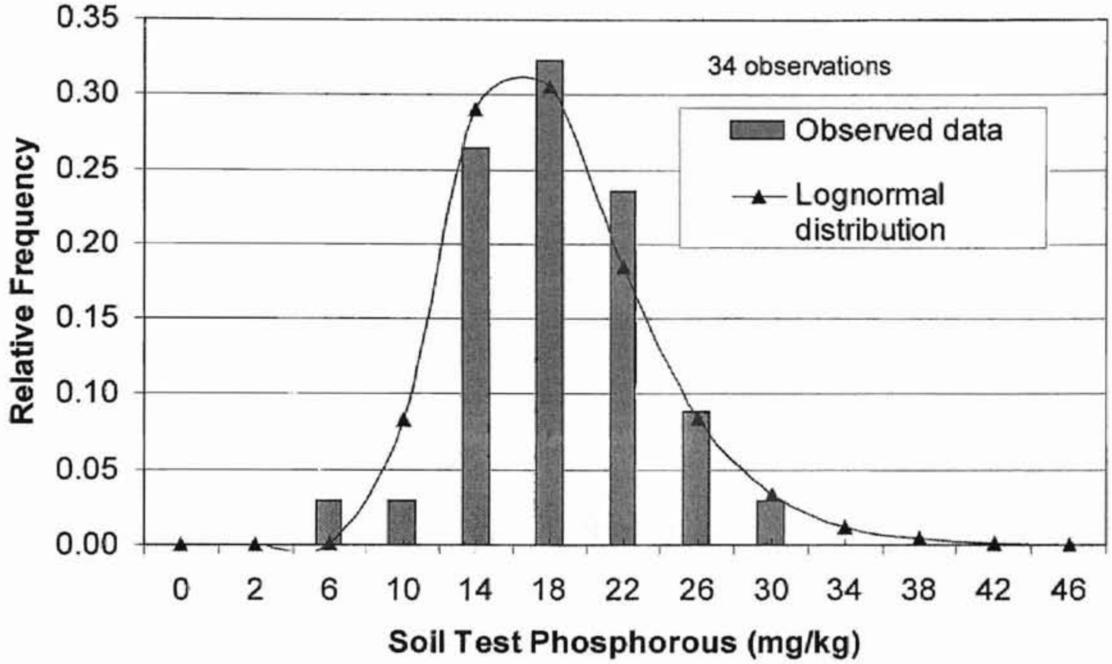


Figure 4.1. Relative frequency distribution of soil test phosphorous for forest in the Upper Little Deep Fork Creek Basin.

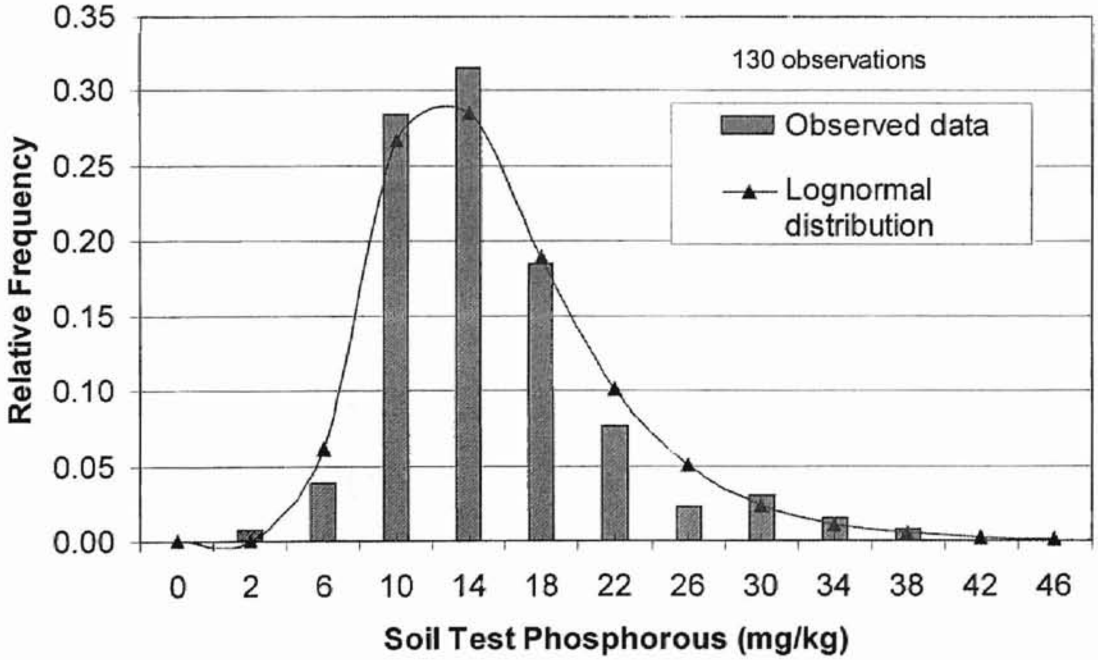


Figure 4.2. Relative frequency distribution of soil test phosphorous for pasture in the Upper Little Deep Fork Creek Basin.

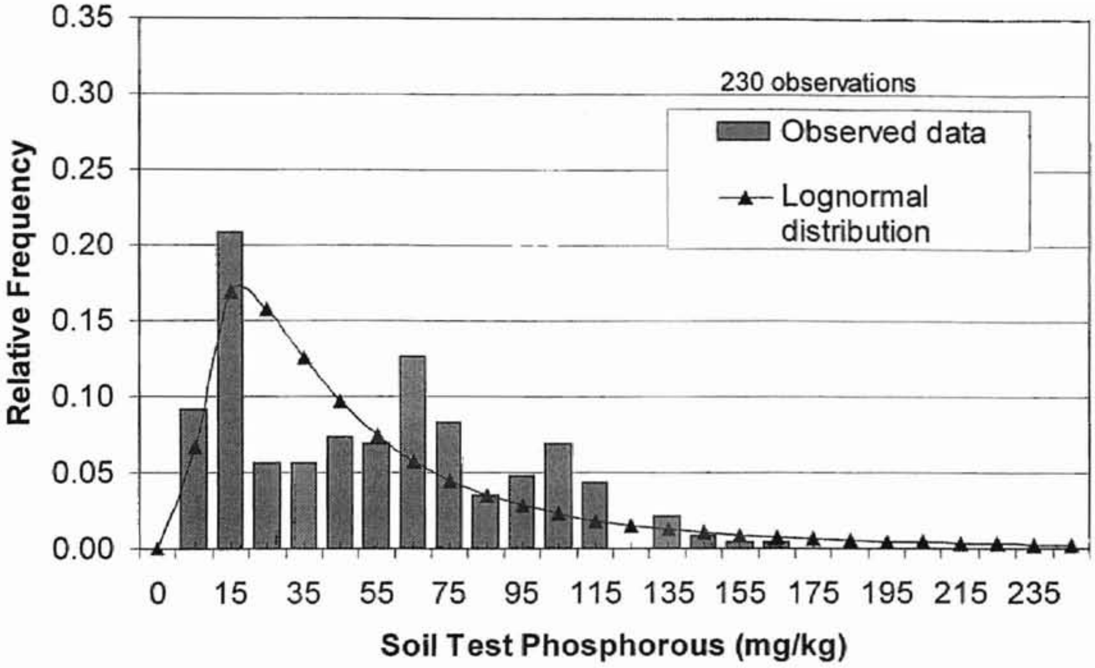


Figure 4.3. Relative frequency distribution of soil test phosphorous for pasture in the Battle Branch Watershed.

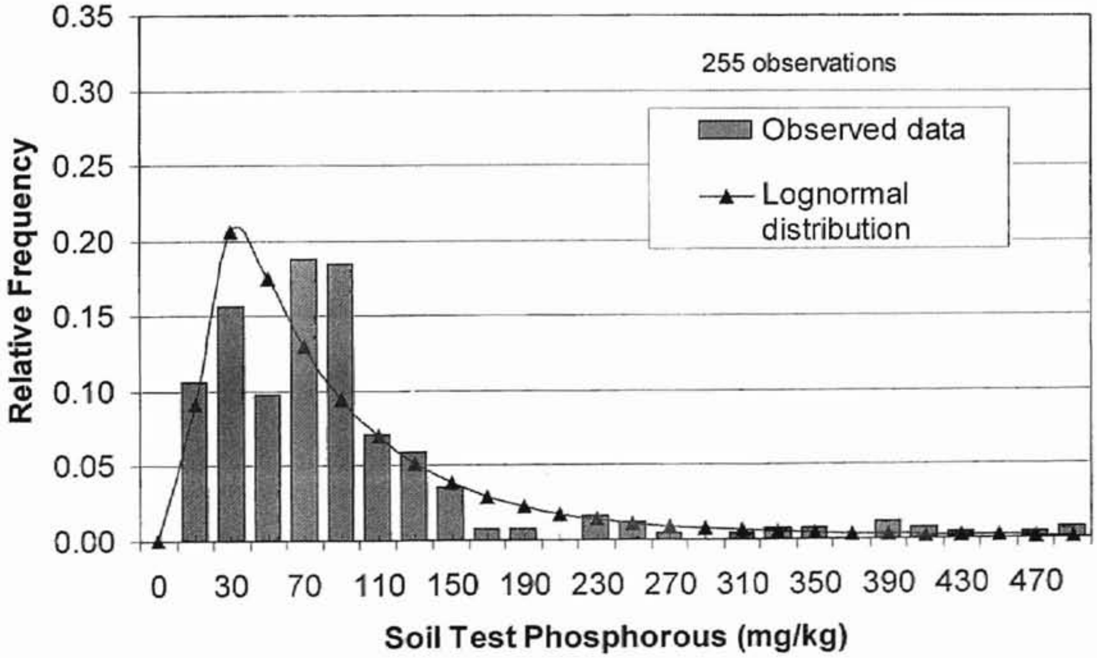


Figure 4.4. Relative frequency distribution of soil test phosphorous for pasture in the Peacheater Creek Watershed.

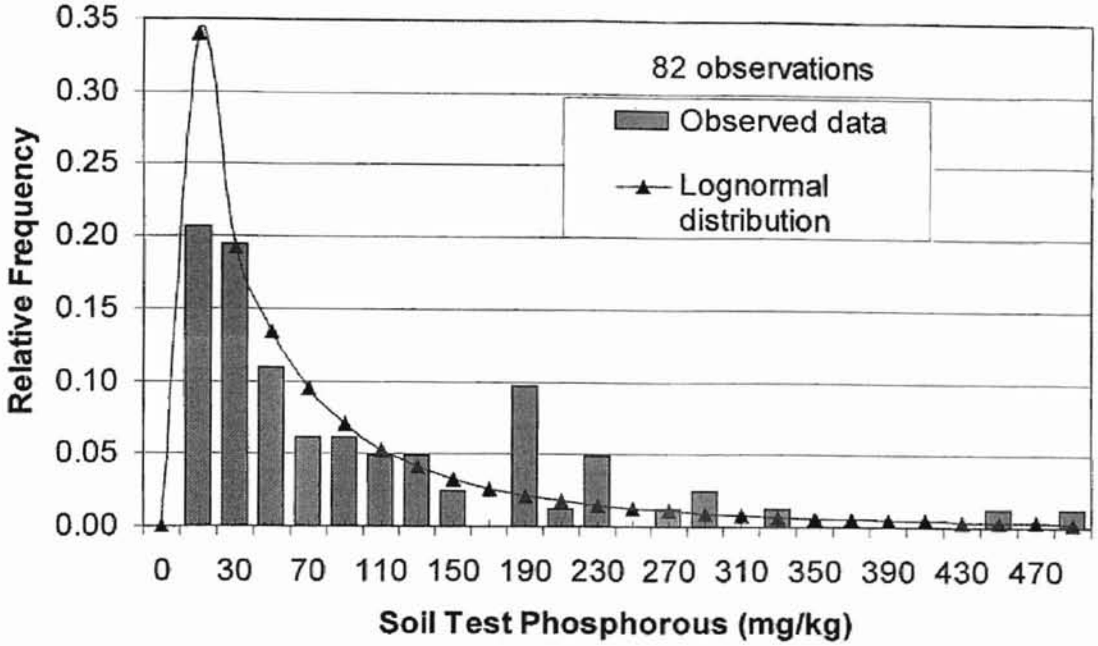


Figure 4.5. Relative frequency distribution of soil test phosphorous for pasture in the Haw Creek Watershed.

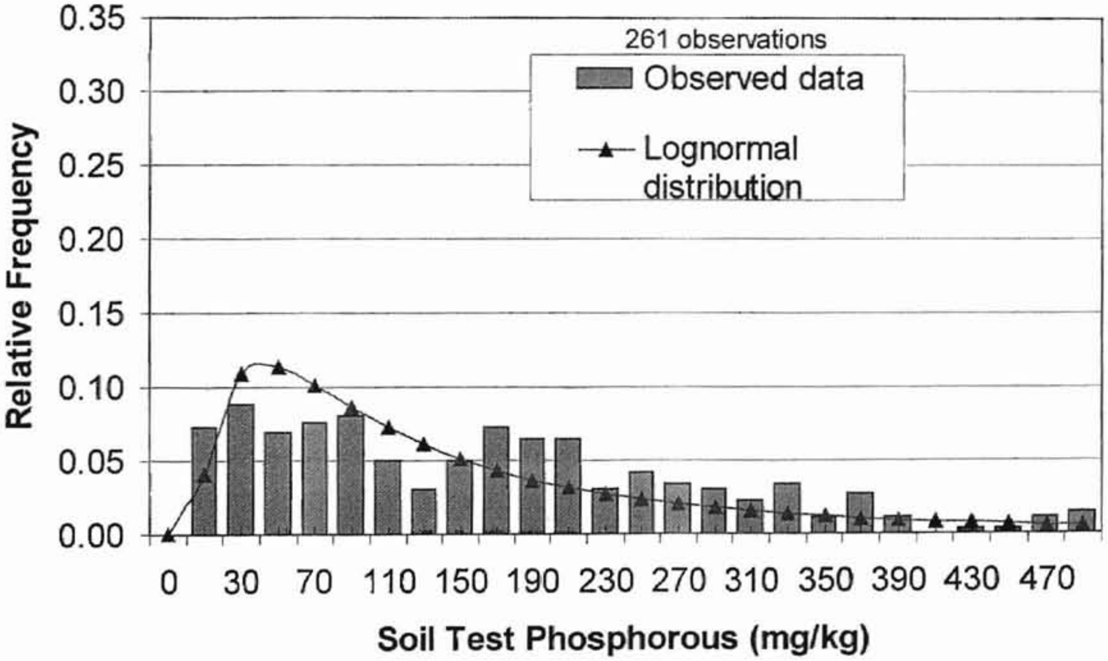


Figure 4.6. Relative frequency distribution of soil test phosphorous for pasture in the Lake Eucha Basin (AR portion).

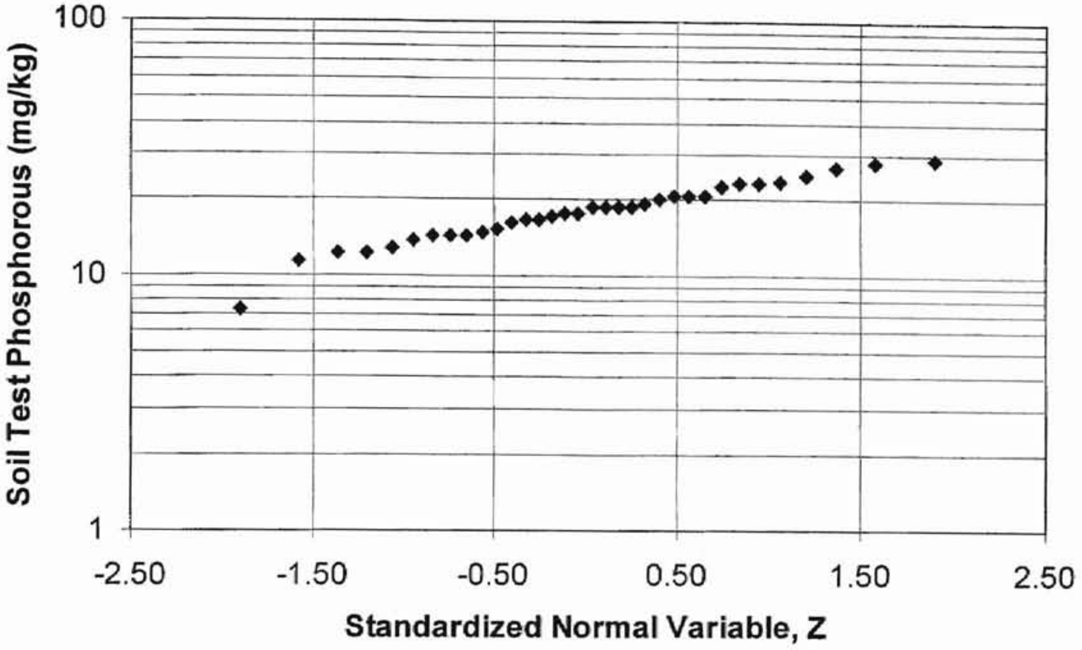


Figure 4.7. Lognormal probability plot of soil test phosphorous for forest in the Upper Little Deep Fork Creek Basin.

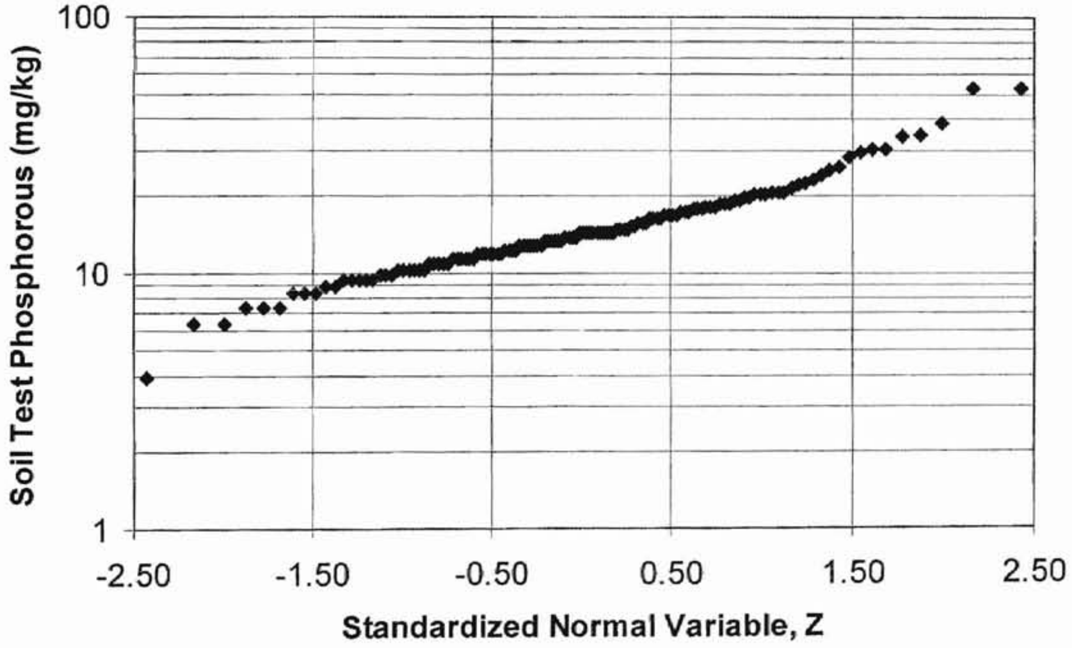


Figure 4.8. Lognormal probability plot of soil test phosphorous for pasture in the Upper Little Deep Fork Creek Basin.

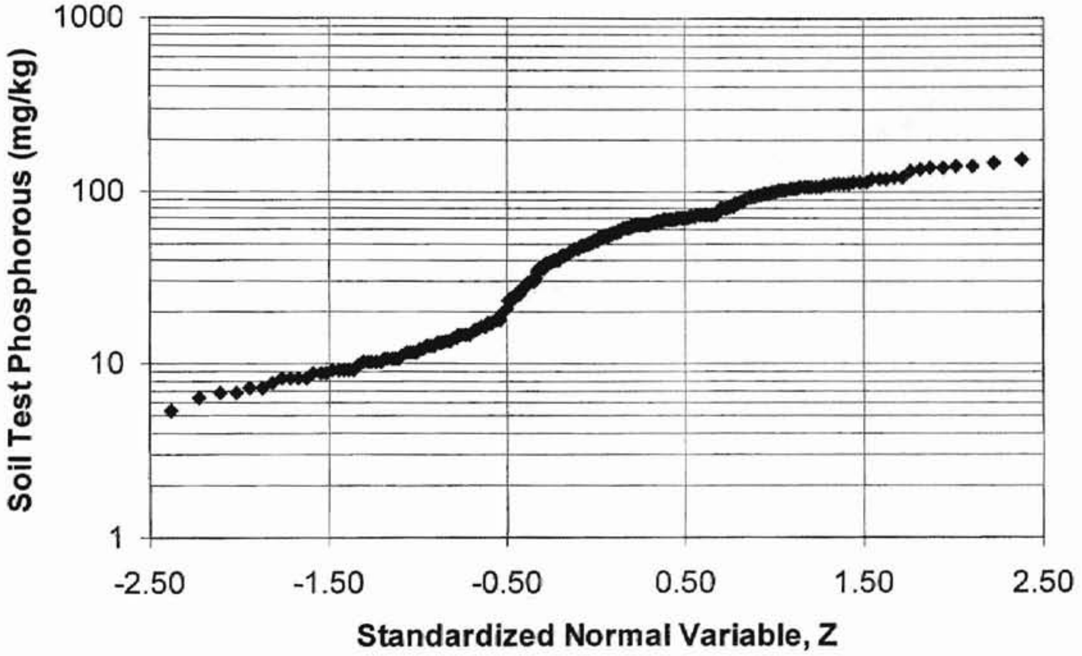


Figure 4.9. Lognormal probability plot of soil test phosphorous for pasture in the Battle Branch Watershed.

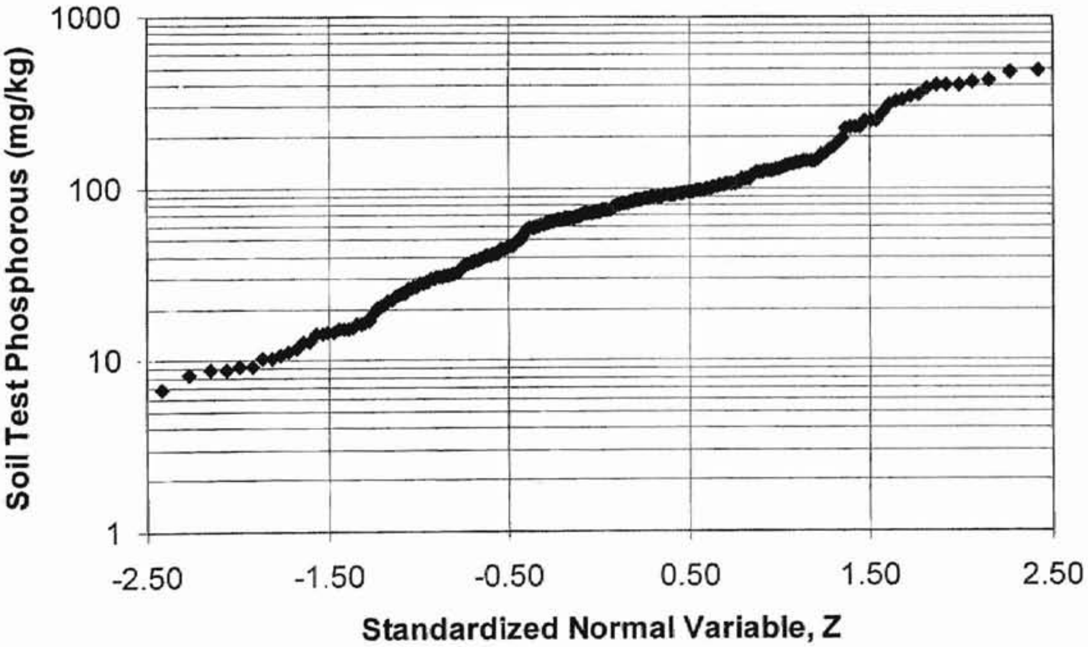


Figure 4.10. Lognormal probability plot of soil test phosphorous for pasture in the Peacheater Creek Watershed.

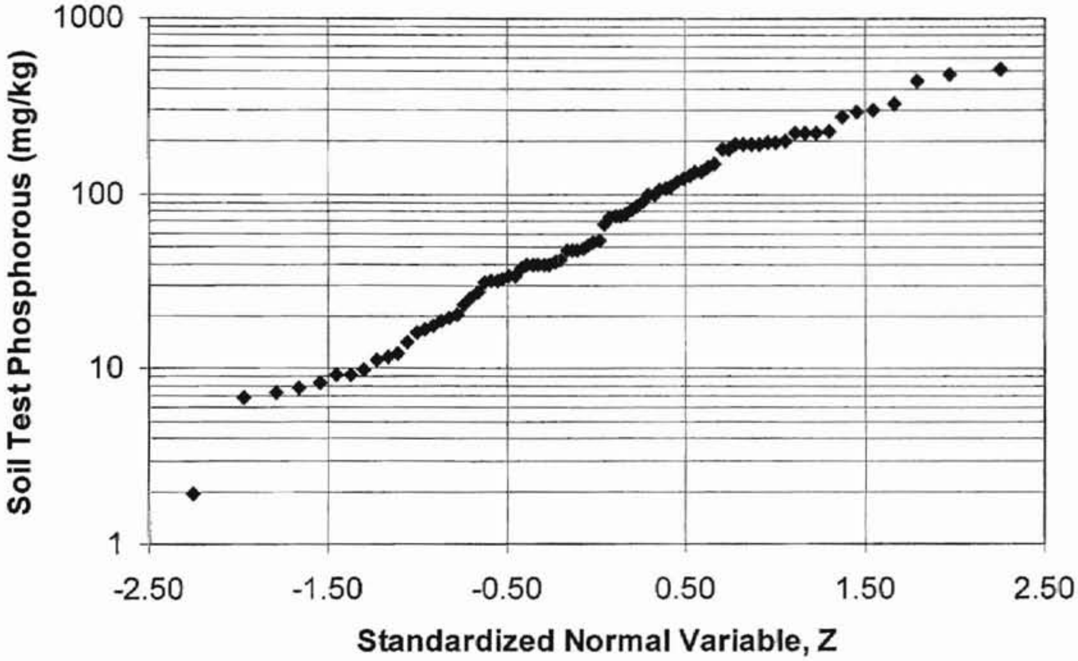


Figure 4.11. Lognormal probability plot of soil test phosphorous for pasture in the Haw Creek Watershed.

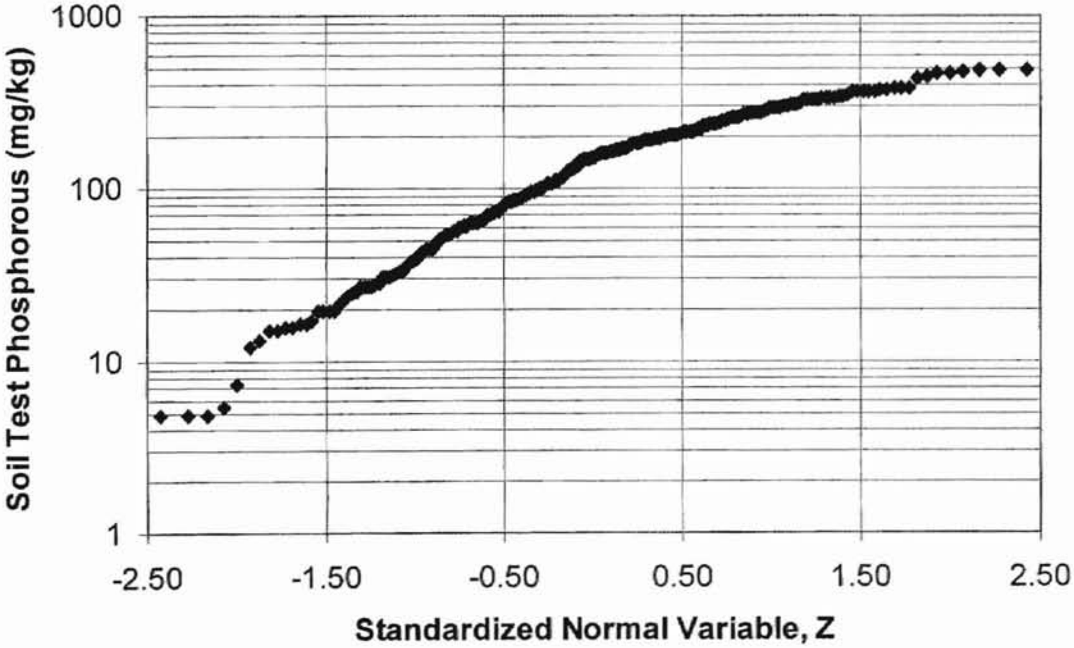


Figure 4.12. Lognormal probability plot of soil test phosphorous for pasture in the Lake Eucha Basin (AR portion).

Table 4.1. Summary of Goodness-of-Fit Tests for a Lognormal Distribution

Location/Data Set	Land Use	H_o: Data are from a lognormal distribution Kolmogorov-Smirnov Test	Chi-square Test
Upper Little Deep Fork Creek	Forest	Do Not Reject H_o	Do Not Reject H_o
Upper Little Deep Fork Creek	Pasture	Do Not Reject H_o	Do Not Reject H_o
Battle Branch	Pasture	**Reject H_o** ($p < 0.01$)	**Reject H_o** ($p < 0.005$)
Peacheater Creek	Pasture	**Reject H_o** ($p < 0.01$)	**Reject H_o** ($p < 0.05$)
Haw Creek	Pasture	Do Not Reject H_o	Do Not Reject H_o
Lake Eucha (AR portion)	Pasture	**Reject H_o** ($p < 0.01$)	**Reject H_o** ($p < 0.05$)

Notes: $\alpha = 0.05$
Ho = null hypothesis

The two visual methods (frequency histograms and probability plots) tend to agree with the statistical tests for each of the data sets. The Upper Little Deep Fork Creek data, both forest and pasture, appear to be lognormally distributed. The watersheds with poultry industry activities, Battle Branch, Peacheater Creek, and Lake Eucha (AR portion) rejected the null hypothesis of a lognormal distribution. These data sets appear to exhibit some type of bi-modal distributions. Haw Creek, which also has poultry activities, did not reject the null hypothesis of lognormally distributed soil test phosphorous. The frequency histogram and probability plot (figures 4.6 and 4.12) do appear, however, to indicate that there may be some bi-modal tendency, but apparently not enough to reject lognormality in the statistical tests.

Soil Test Phosphorous by Soil Mapping Unit and Soil Characteristics

It was desired to evaluate predicting soil test phosphorous levels based on soil mapping units or soil characteristics. If successful, soil phosphorous levels could then be estimated from soil mapping units or soil characteristic information. This would be advantageous since digital soil data sets exist for many areas. The three data sets were used where soil sample or field locations were known. For the Upper Little Deep Fork Creek, the soil data layer was overlain with the soil sample location data using GIS. This resulted in a soil mapping unit, or soil type, assignment for each soil sample. The same was performed with Battle Branch and Peacheater Creek, except the pasture field boundaries were used, since each field had been sampled separately. Where more than one soil mapping unit was present in a field, the dominant soil based on coverage area was selected.

Regression statistics were then performed. The first regression involved the use of dummy variables because the soil type was designated by a letter and not an associated numeric value, or in other words, qualitative rather than quantitative independent variables were used. The method of using dummy variables as presented by Ott (1984) was used as follows:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n + \epsilon \tag{4.1}$$

where,

y = the dependent variable;

β = the unknown parameter;

x_1 = 1 if treatment 2, x_1 = 0 otherwise;

x_2 = 1 if treatment 3, x_2 = 0 otherwise;

x_n = 1 if treatment n+1, x_n = 0 otherwise;

ϵ = the random error term.

The result is an expression for soil test phosphorous based on soil types.

The second regression performed on each data set was based on the associated soil characteristics of each soil mapping unit. This involved multiple linear regression. The results of both regressions are summarized in Tables 4.2 and 4.3. There was no apparent correlation between soil test phosphorous and soil mapping units, or between soil test phosphorous and the soil characteristics. It turned out that there was no significant difference (α = 0.05) of soil test phosphorus among soil types for data from Upper Little Deep Fork Creek, which included forest and pasture data. The regression is denoted by "N/A" for "not applicable" in Table 4.2. The coefficients of determination from the regressions are shown in the tables, but there were no significant parameters for any of the regressions.

The results of the regression analysis indicate that soil test phosphorous levels for these data sets are not related to soil type and are probably influenced primarily by the land management activities. The soil type does, however, play a role in the transport fate of phosphorous once it reaches the soil, as discussed in the literature review.

Table 4.2. Summary of Regression of Soil Test Phosphorous by Soil Mapping Unit

Location/Data Set	R^2	Adjusted R^2
Upper Little Deep Fork Creek	N/A[†]	N/A[†]
Battle Branch	0.17	0.08
Peacheater Creek	0.10	0.05

[†] N/A: no significant difference ($\alpha = 0.05$) in soil test phosphorous means among soil mapping units

Table 4.3. Summary of Regression of Soil Test Phosphorous by Selected Soil Characteristics[†]

Location/Data Set	R^2	Adjusted R^2
Upper Little Deep Fork Creek	0.01	-0.04
Battle Branch	0.10	0.07
Peacheater Creek	0.01	-0.02

[†] Soil Characteristics: K, Organic matter or Organic Carbon content, Clay content, Bulk density

Empirical Distributions

Empirical methods were used to develop a nonparametric method for determining the sample size, or number of observations, required for estimating basin-scale soil test phosphorous within a 90% confidence interval. Monte-Carlo sampling was employed to develop distributions of the soil test phosphorous means for various sample sizes from the observed data.

The data sets of interest for this study were those watersheds or basins that contained poultry industry activities, where the soil test phosphorous data did not appear to follow a lognormal or standard-type distribution. Using classical statistics, which assumes a normal distribution of the means, would be an approximation, at best, to estimate sample size for the bi-modal distributed data.

It was assumed that the data represented the parent populations for each data set since almost all pasture fields had been sampled. The data from the Upper Little Deep Fork Creek Basin for pasture, which does not contain poultry activities, was likewise used. The only data set available for forest land use was also from the Upper Little Deep Fork Creek basin.

A personal computer spreadsheet application was adapted to perform the Monte Carlo sampling for creating the empirical distributions. For each data set, the soil test phosphorous was ranked from low to high and an associated probability, from zero to one, was assigned to each data point based on its rank. A macro was written that would randomly choose a probability from a uniform distribution of zero to one and then select the corresponding soil test phosphorous value. This was performed a number of times equal to the current sample size, i.e. 25 times for a sample size of 25. Then, a mean and standard deviation were calculated for the empirical distribution of randomly chosen values. This procedure was repeatedly performed for each sample size of 5 to

250, in increments of five. The entire process was repeated 250 times for each sample size.

Once the new empirical distributions were developed, the 90% confidence intervals were calculated. They were chosen so that they were symmetrical in probability, i.e. for the 90% confidence interval, 5% of the area of the distribution is to the left and 5% is to the right. The 90% confidence interval was chosen, but any confidence interval could have been used. Figures 4.13 through 4.18 are the resultant empirical distributions for each data set for the various sample sizes.

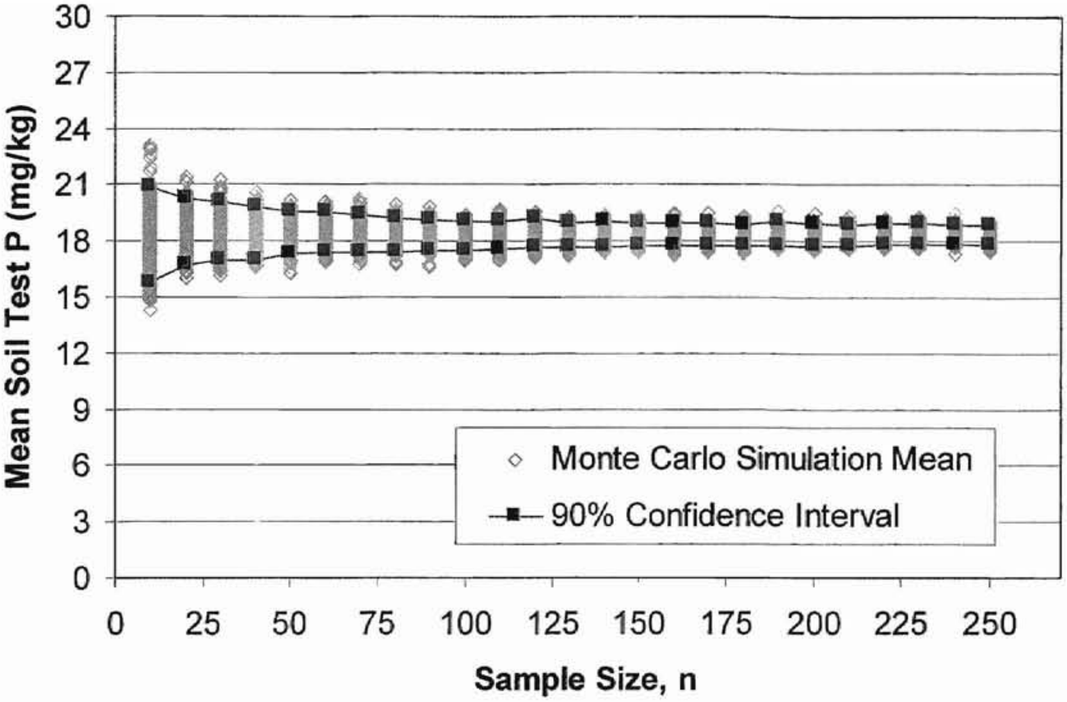


Figure 4.13. Empirical distributions of mean soil test phosphorous for various sample sizes for forest in the Upper Little Deep Fork Creek Basin.

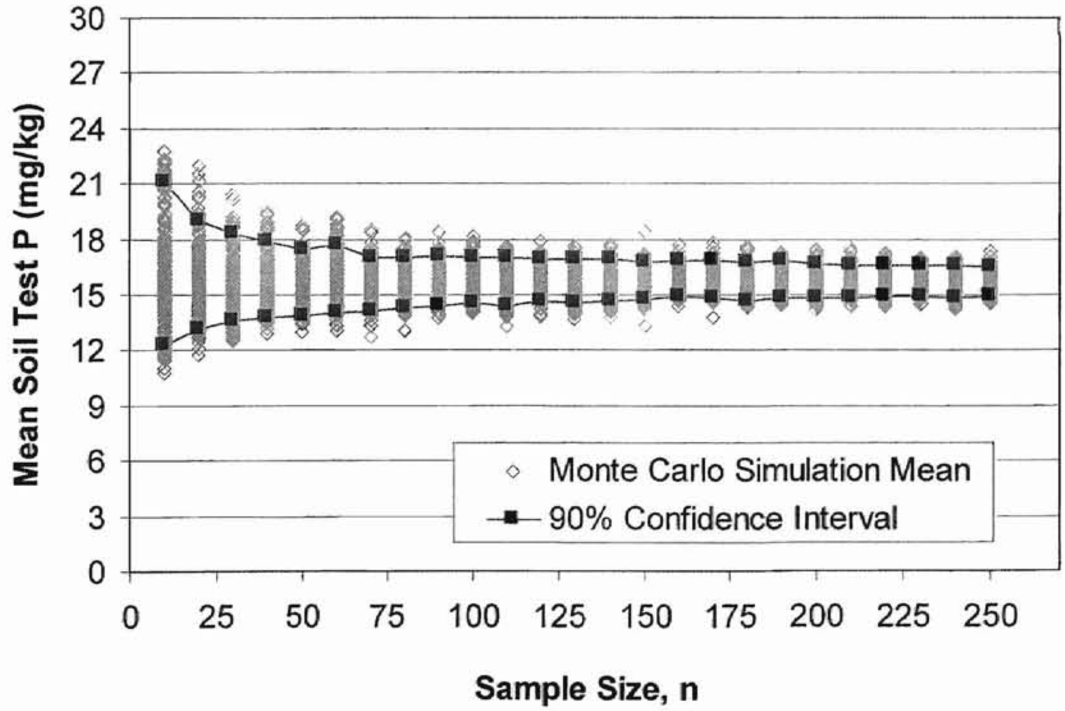


Figure 4.14. Empirical distributions of mean soil test phosphorous for various sample sizes for pasture in the Upper Little Deep Fork Creek Basin.

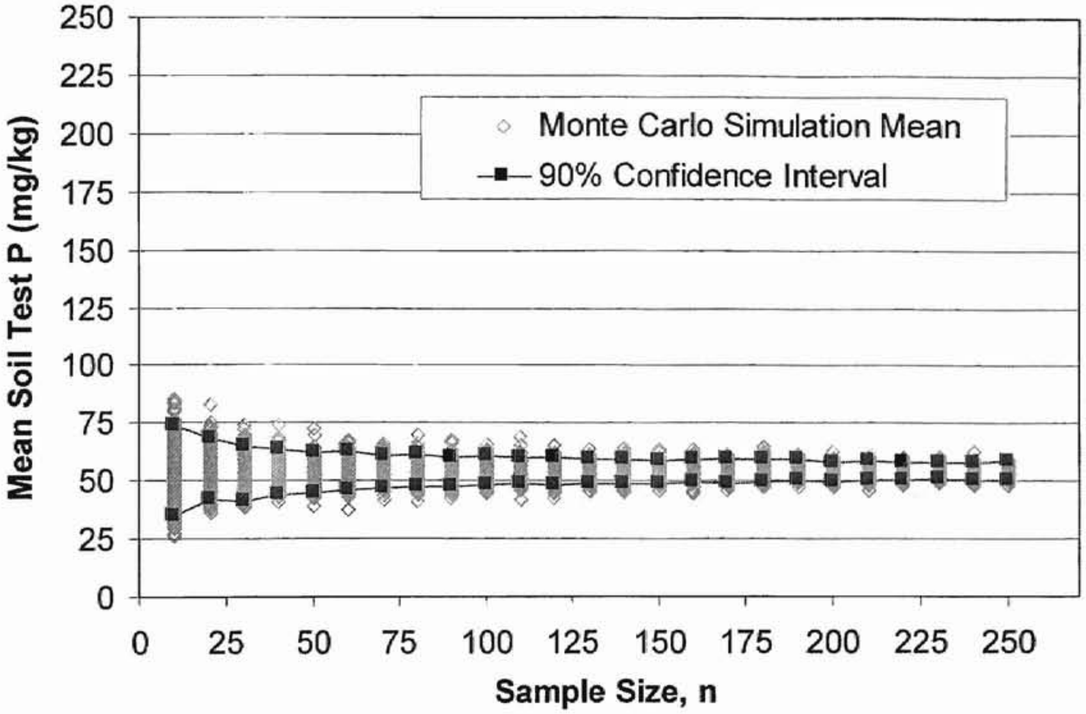


Figure 4.15. Empirical distributions of mean soil test phosphorous for various sample sizes for pasture in the Battle Branch Watershed.

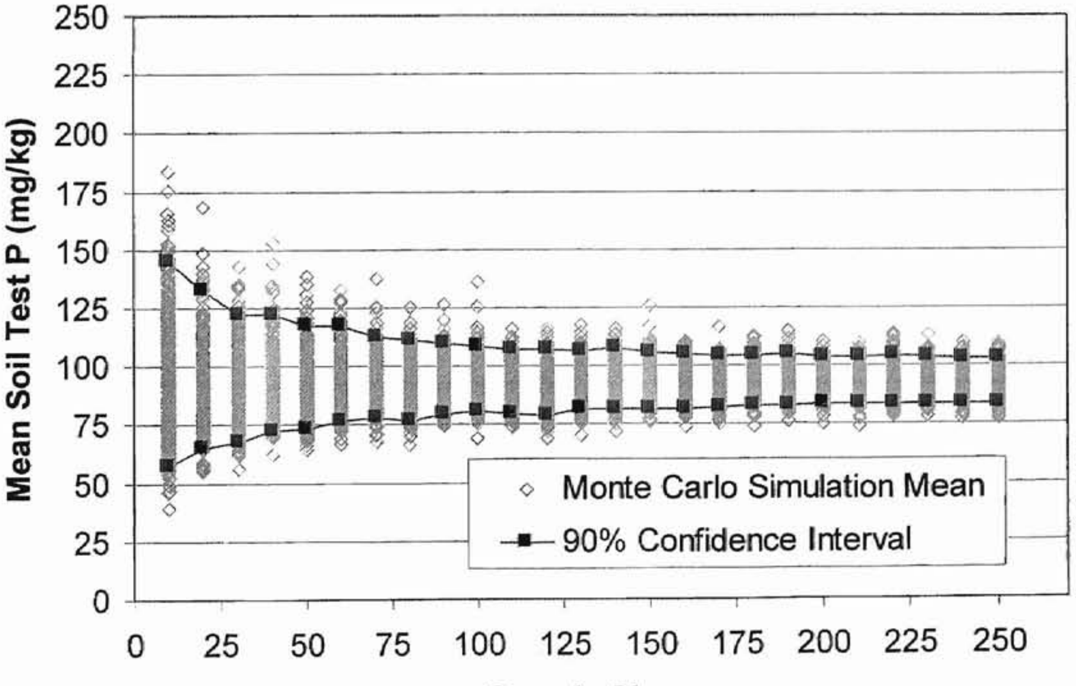


Figure 4.16. Empirical distributions of mean soil test phosphorous for various sample sizes for pasture in the Peacheater Creek Watershed.

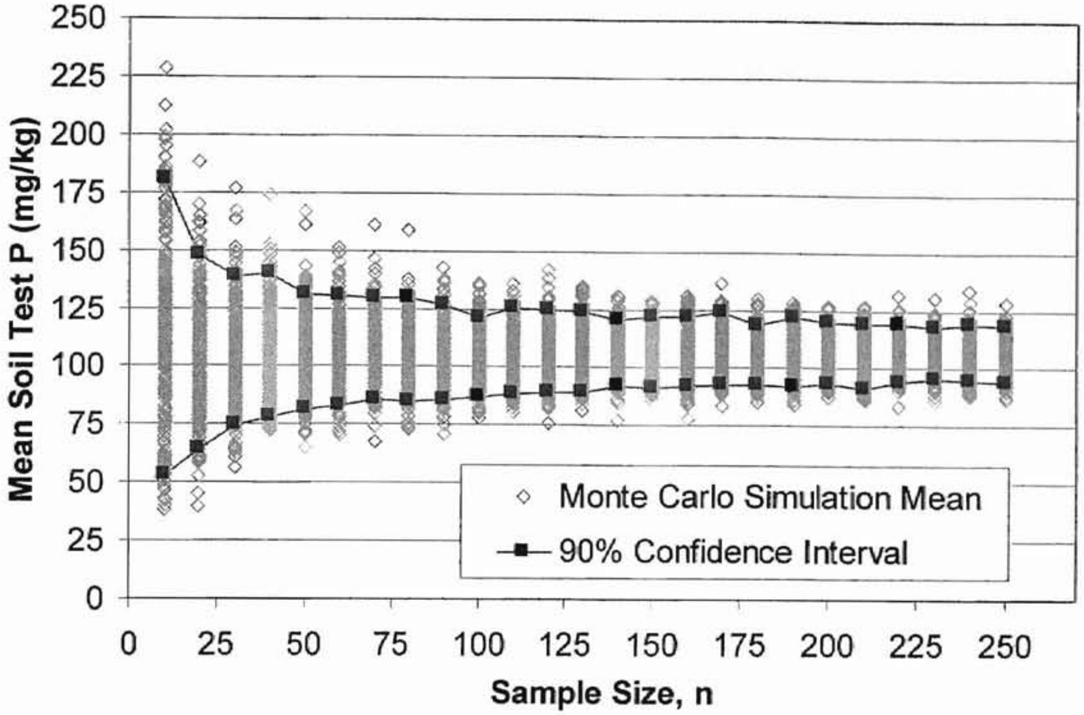


Figure 4.17. Empirical distributions of mean soil test phosphorous for various sample sizes for pasture in the Haw Creek Watershed.

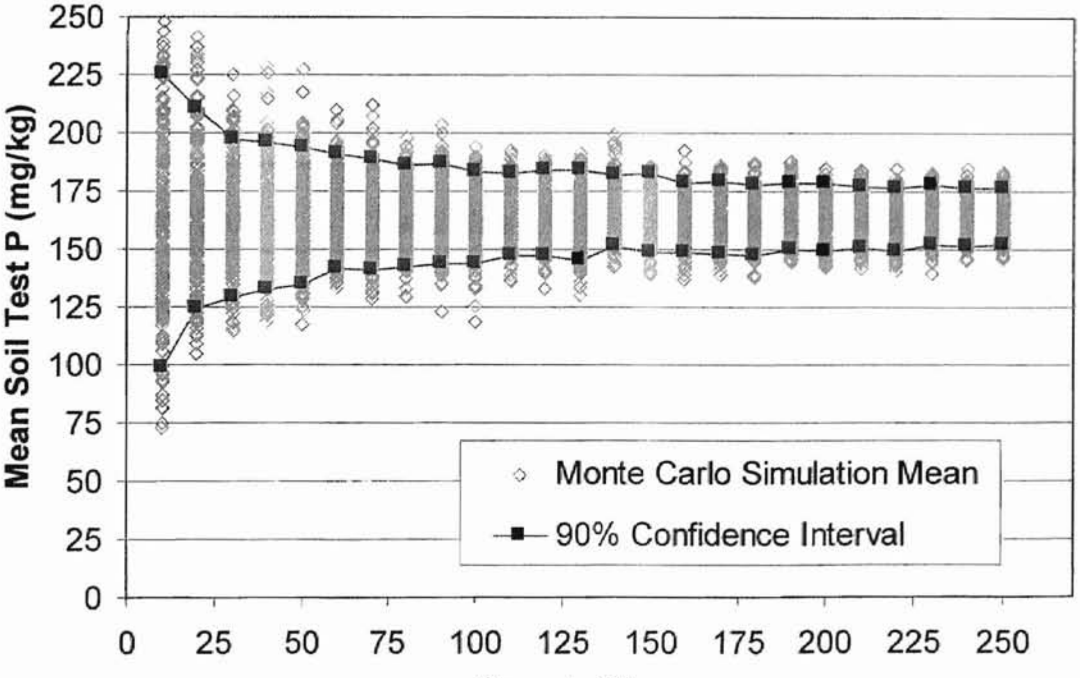


Figure 4.18. Empirical distributions of mean soil test phosphorous for various sample sizes for pasture in the Lake Eucha Basin (AR portion).

The 90% confidence interval widths were empirically determined for each sample size from each empirical distribution and plotted on a single graph. Figure 4.19 is the plot of the 90% confidence intervals for forest and Figure 4.20 is for pasture. Regression was then performed on each data set to develop an equation for each respective curve. The best-fit regression lines were of the form:

$$y = Cn^{-0.5} \tag{4.2}$$

where y represents the 90% confidence interval, C is a constant specific to each curve, and n is the sample size. Solving for n yields:

$$n = (C / y)^2 \tag{4.3}$$

A unique equation for n was developed for each of the data sets. The only difference was the constant, C. Table 4.4 lists the constants for equation 4.3 developed from each data set.

Table 4.4. Constant, C, for Equation 4.3

Location/Data Set	Land Use	Constant, C for Eqn. 4.3	R^2
Upper Little Deep Fork Creek	Forest	16.8	0.99
Upper Little Deep Fork Creek	Grasslands, Pasture	27	0.99
Battle Branch	Pasture	125	0.99
Peacheater Creek	Pasture	295	0.99
Haw Creek	Pasture	377	0.98
Lake Eucha (AR portion)	Pasture	390	0.99

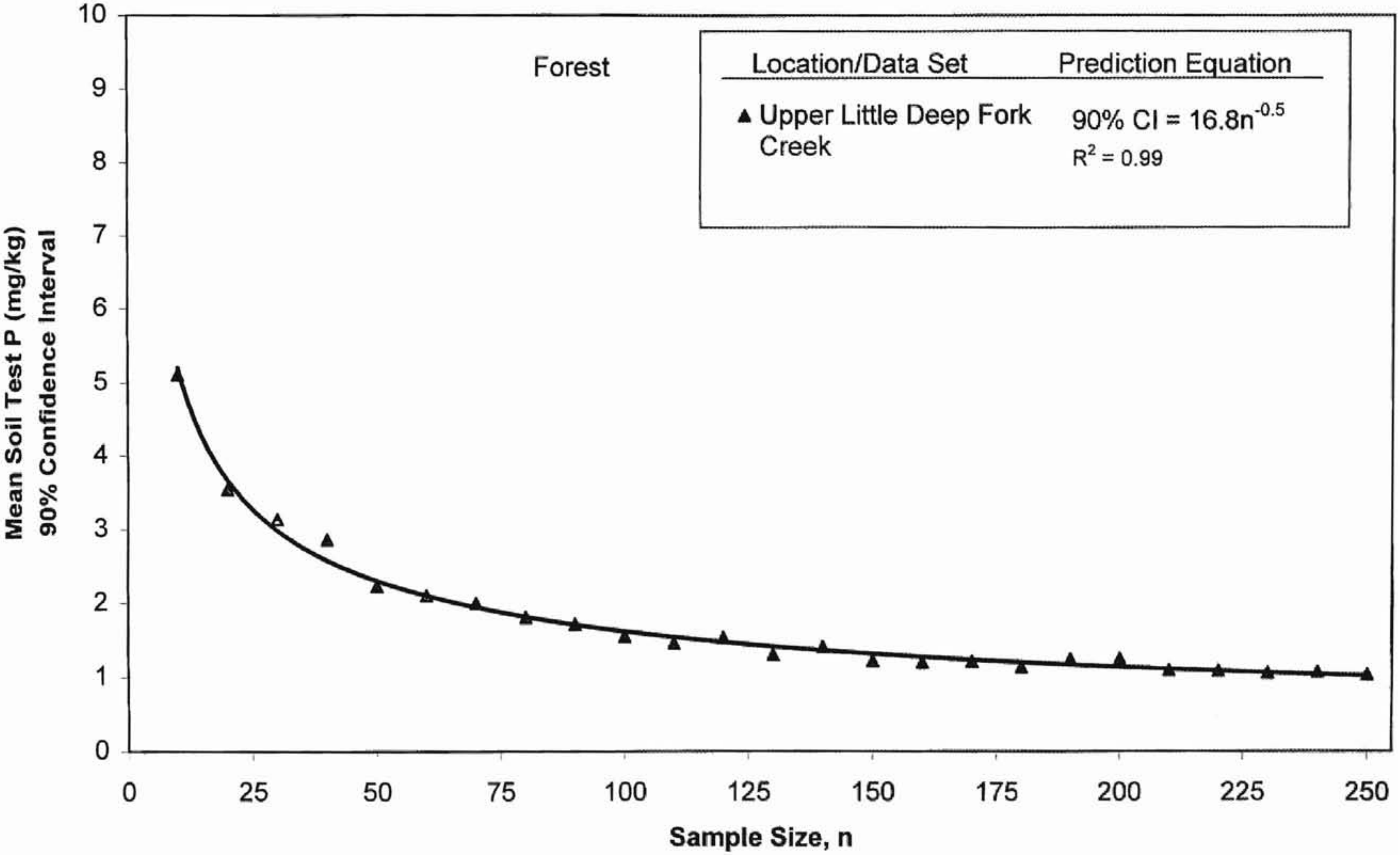


Figure 4.19. Estimated 90% confidence intervals for mean soil test phosphorous for forest for varying sample sizes.

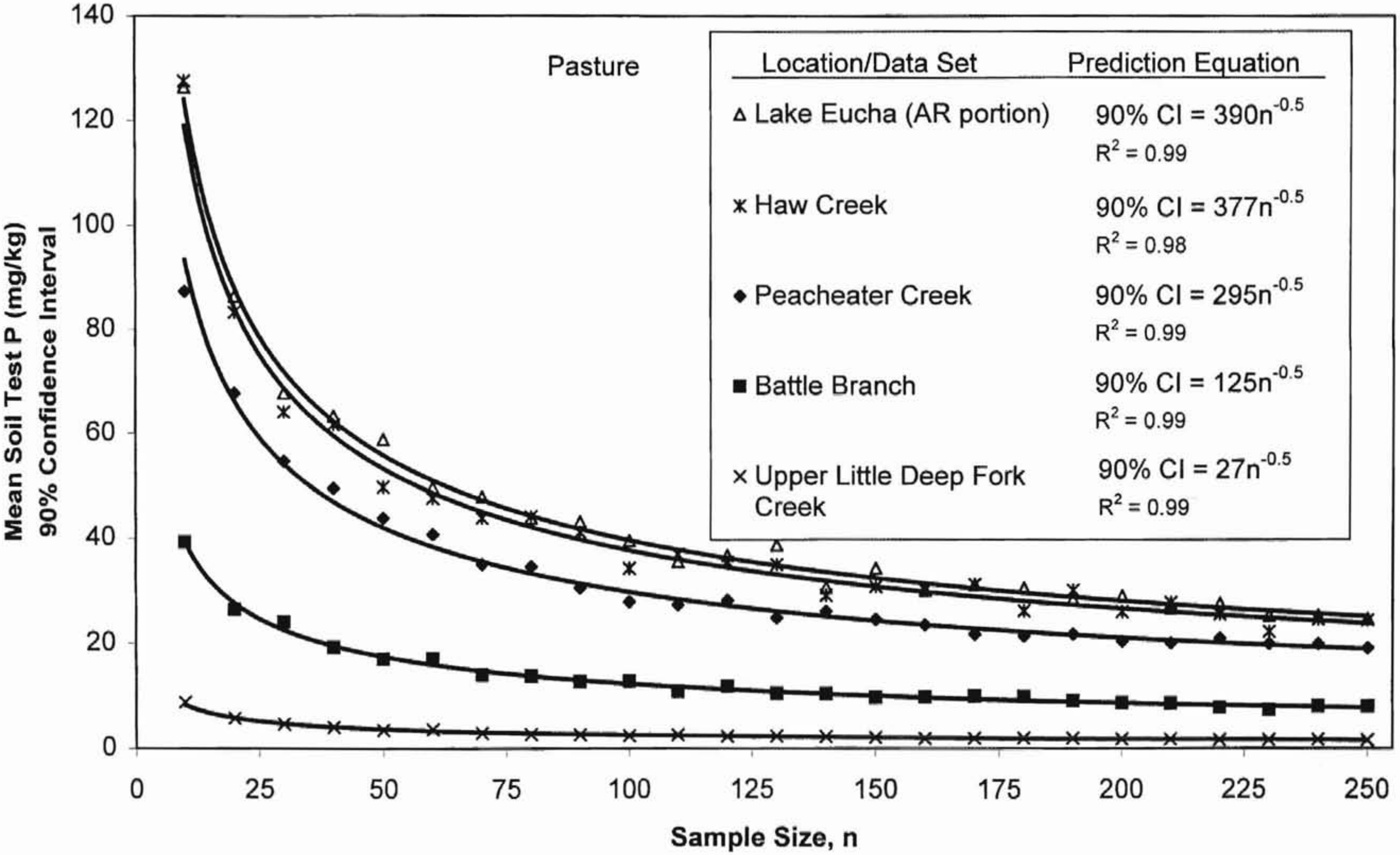


Figure 4.20. Estimated 90% confidence intervals for mean soil test phosphorous for pasture for varying sample sizes.

Comparison of Classic to Nonparametric Techniques

The form of the regression equation derived from the nonparametric approach is very similar to what would be used for an approximation of n for a normal distribution of means. For a normal distribution the lower and upper confidence limits are found from classic statistics to be:

$$L = \overline{x} - Z_{\alpha/2} \frac{\sigma_x}{\sqrt{n}} \qquad (4.4)$$

$$U = \overline{x} + Z_{\alpha/2} \frac{\sigma_x}{\sqrt{n}} \qquad (4.5)$$

where L is the lower limit, U is the upper limit, $\overline{x}$ is the sample mean, $Z_{\alpha/2}$ is the value from the standard normal distribution for the specified error level, α, and σ_x is the standard deviation. Subtracting equation 4.4 from equation 4.5 for the interval width and solving for n, yields:

$$n = \left(\frac{2Z_{\alpha/2}\sigma_x}{CI} \right)^2 \qquad (4.6)$$

where CI is the required confidence interval. Equation 4.6 is the same as that given by Steel and Torrie (1980) for estimating required sample sizes and is based on the Central Limit Theorem.

It appears that the constant, C, obtained from the nonparametric approach is comprised of a variance and probability variable. The nonparametric method does not distinguish between the two components. However, use of the nonparametric equation does not require the assumption of a known underlying distribution. As deviation from normality increases, the efficiency of parametric tests decreases, but the efficiency of nonparametric tests is not affected

(McIntyre and Tanner, 1958). Use of classical statistical techniques to calculate *n* also requires a direct estimate of the variation, or standard deviation, which is typically very difficult to estimate without sufficient data.

The nonparametric approach for determining sample size was investigated because it was not fully evident that the data sets obtained from high-level soil test phosphorous watersheds/basins would adhere to the assumptions required for using classical statistical techniques. The assumptions referred to are those related to the Central Limit Theorem, such as identically and independently distributed data. Also, application of the Central Limit Theorem is used as an approximation, assuming the means of the population are normally distributed.

To compare the two approaches, the differences in confidence intervals determined from each method were compared. For each sample size, the 90% confidence interval was computed from classical statistics for a normal distribution using equation 4.6. For σ_x, the standard deviations from the original data sets were used. The resultant interval was then compared to that obtained from the nonparametric approach. Figures 4.21 through 4.26 are plots of the percent differences of the classical 90% confidence interval to that computed by the nonparametric approach. The comparison was made for each data set. The results of the comparisons indicate there is approximately a 10% difference between the two methods for all the sample sizes. There was no definite sample size where the two converged for any data set.

As it has turned out, use of the Central Limit Theorem is probably general enough to apply to situations like those presented in this study. Due to the relative small differences, the classic approach would probably be acceptable for future estimations of sample size under the conditions studied.

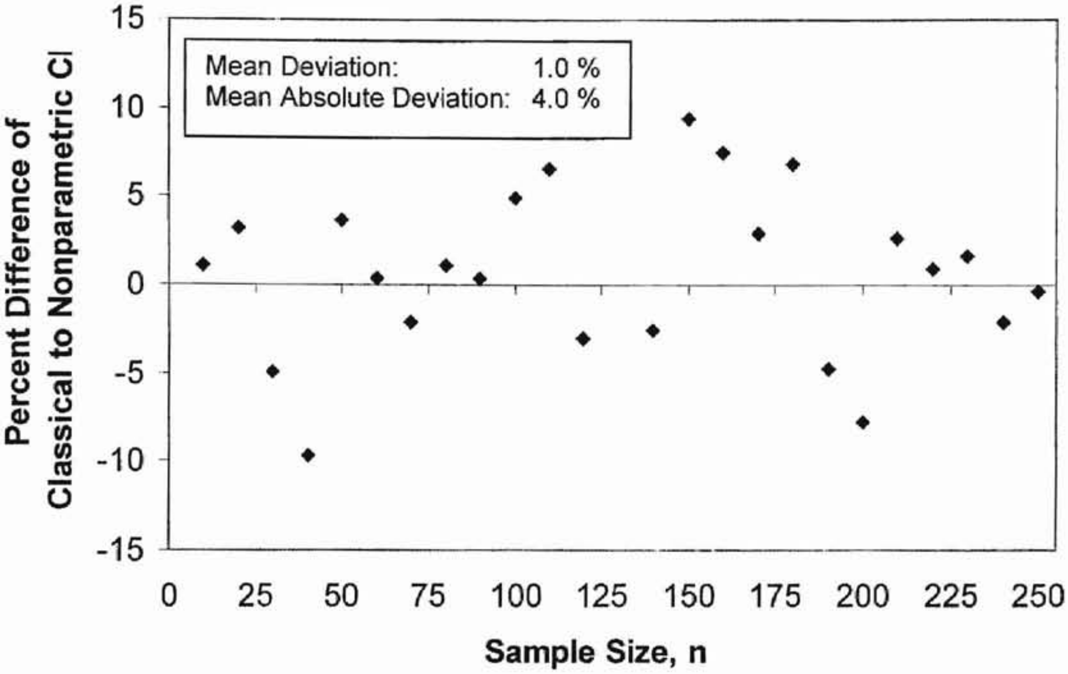


Figure 4.21. Comparison of confidence intervals computed by the classical and nonparametric methods for varying sample sizes for Upper Little Deep Fork Creek forest.

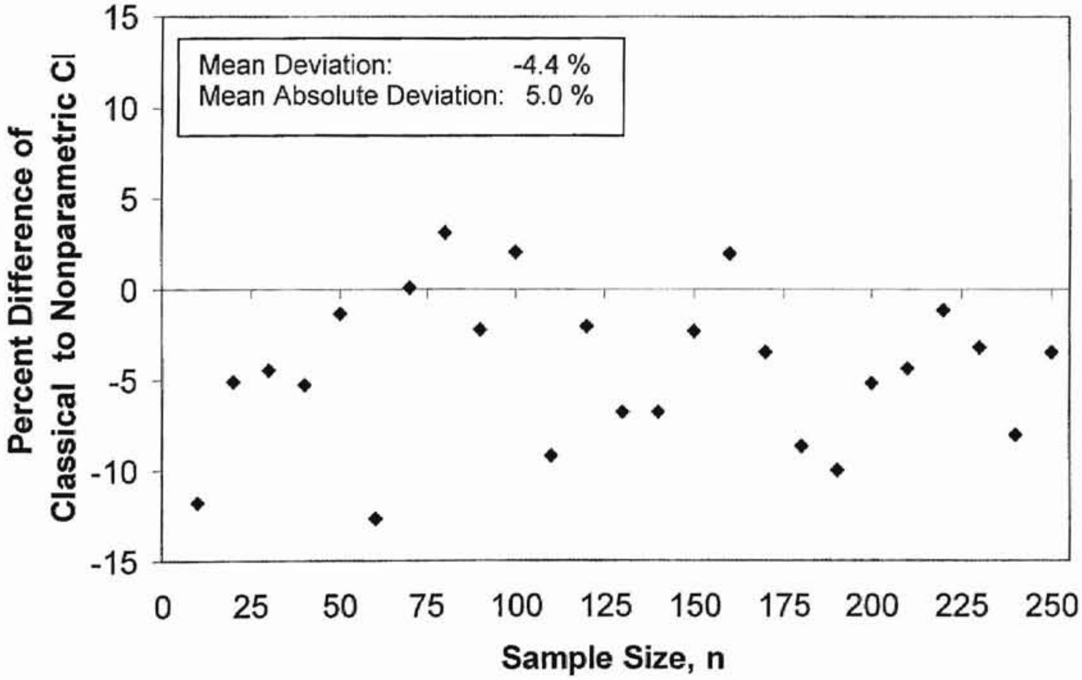


Figure 4.22. Comparison of confidence intervals computed by the classical and nonparametric methods for varying sample sizes for Upper Little Deep Fork Creek pasture.

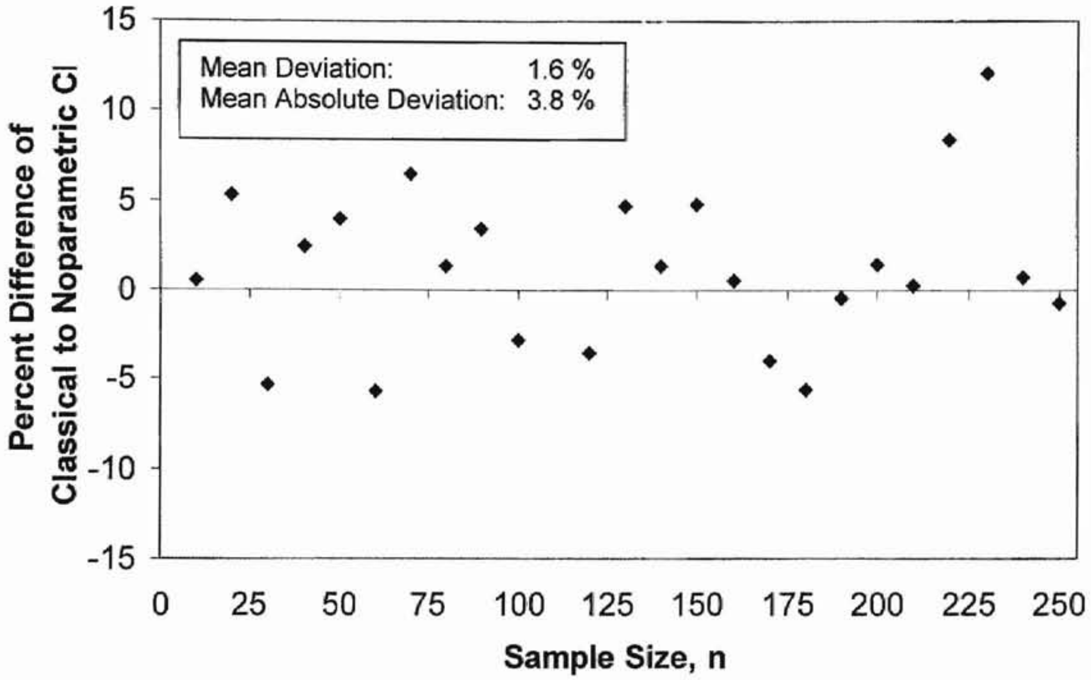


Figure 4.23. Comparison of confidence intervals computed by the classical and nonparametric methods for varying sample sizes for Battle Branch pasture.

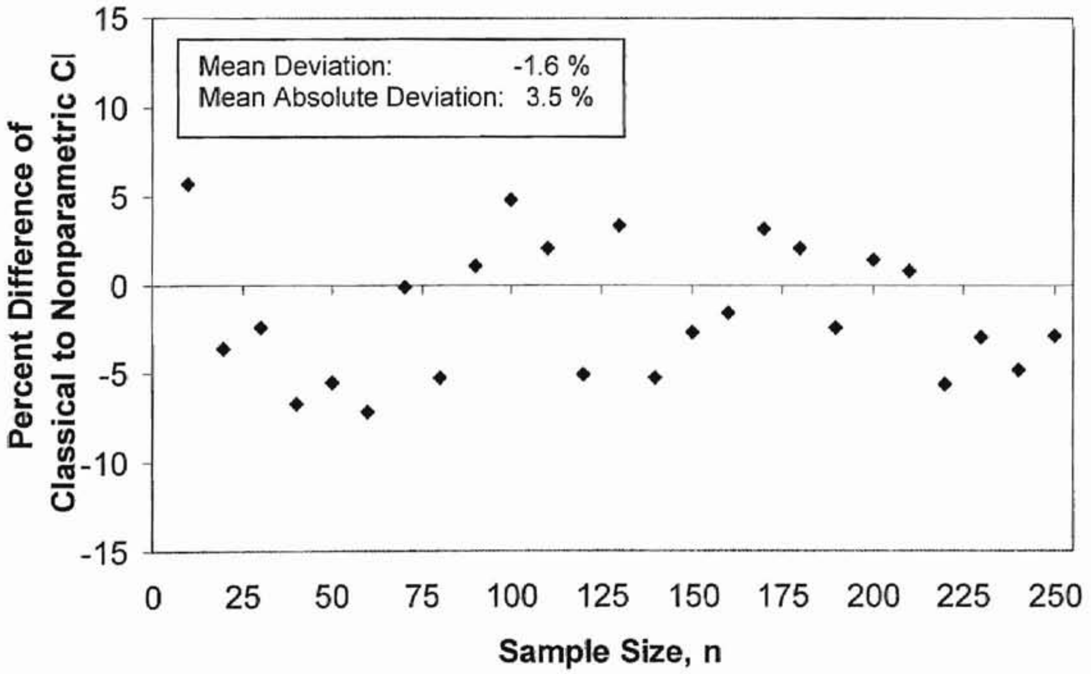


Figure 4.24. Comparison of confidence intervals computed by the classical and nonparametric methods for varying sample sizes for Peacheater Creek pasture.

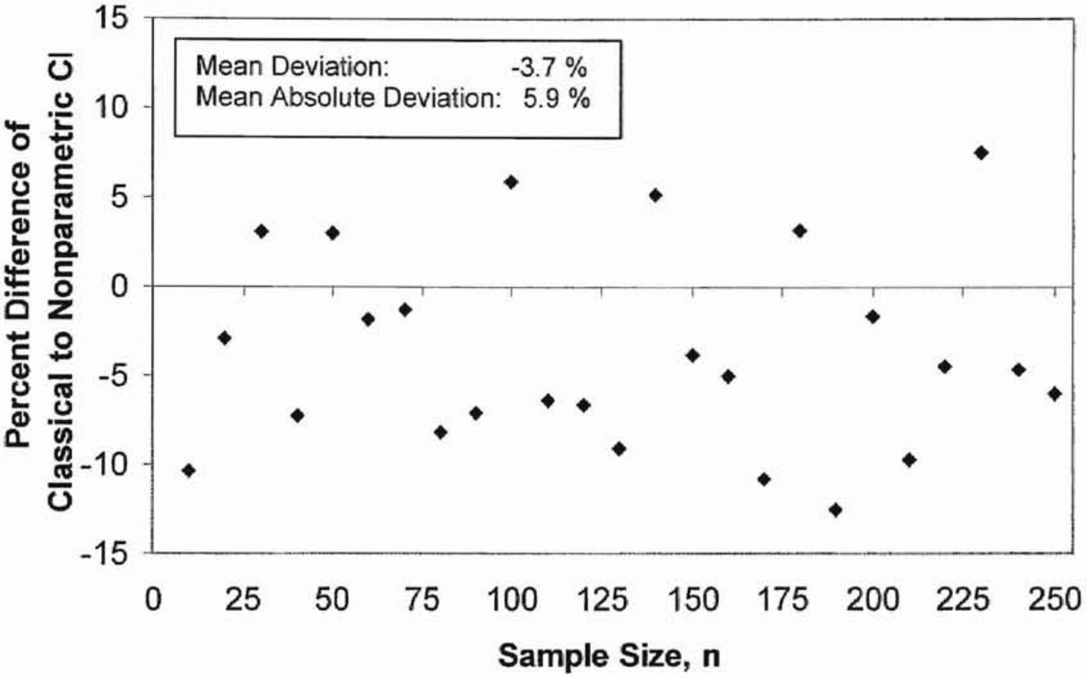


Figure 4.25. Comparison of confidence intervals computed by the classical and nonparametric methods for varying sample sizes for Haw Creek pasture.

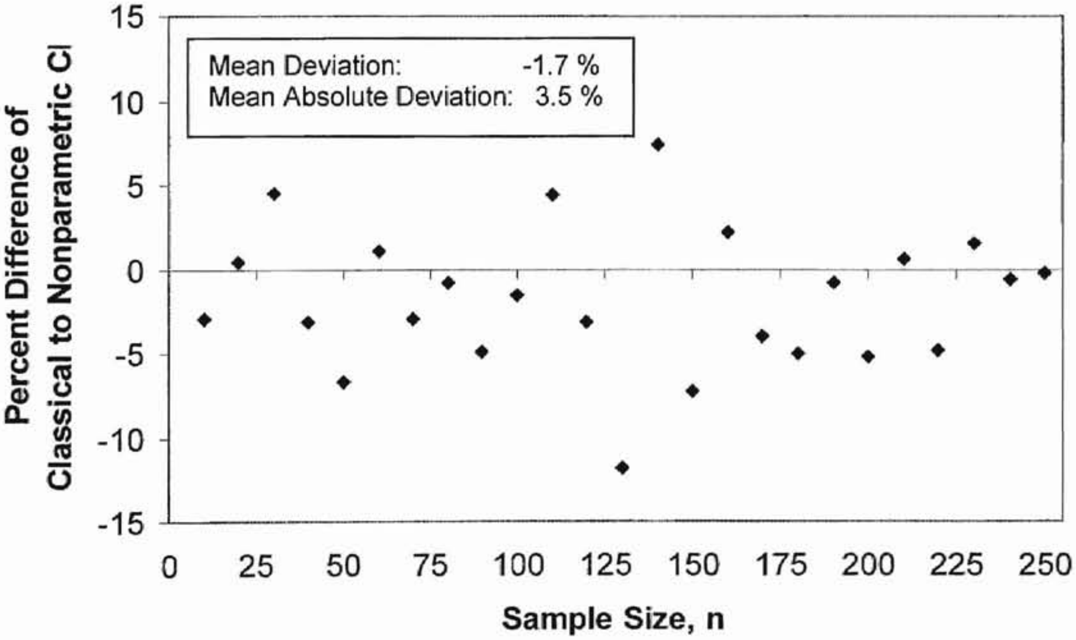


Figure 4.26. Comparison of confidence intervals computed by the classical and nonparametric methods for varying sample sizes for Lake Eucha (AR portion) pasture.

Nonparametric Method

The previous sections present the steps to develop a nonparametric approach to determine the sample size for estimating basin-scale soil test phosphorus within a 90% confidence interval. The same nonparametric approach can also be used to determine the 90% confidence interval associated with a predetermined sample size for similar watersheds or basins. These two options for using the nonparametric method are termed Option A and Option B, respectively. The method was derived assuming the soil sample locations, or fields to be sampled, are randomly selected.

Option A can be used to determine the required number of observations, or sample size. To do this, an acceptable interval of soil test phosphorous levels must first be determined. This can be thought of as an allowable plus or minus deviation from the expected mean. Since the soil phosphorous data will most likely be used as input into a hydrological/water quality model, an acceptable interval of soil test phosphorous should be determined from the chosen model. This can be obtained by running the model for varying initial soil test phosphorous levels to develop a relationship between input soil test phosphorous and output phosphorus loading. Then, after choosing an acceptable interval of phosphorous loading from the model output, an input soil test phosphorous interval can be obtained from the developed relationship. Allowing this interval width to be the acceptable 90% confidence interval, the required sample size can then be computed using the following equation and the appropriate constant, *C*, value from Table 4.5:

$$n = \left(\frac{C}{90\%CI}\right)^2 \qquad (4.7)$$

where *n* is the required sample size, *C* is the nonparametric constant from Table 4.5, and the 90% confidence interval of soil test phosphorous is in mg/kg.

Option B can be used when there is a predetermined number of samples to be collected. This may be the case when budget or time constraints limit the total number of samples available for a particular project. For this case, the appropriate constant, *C*, from Table 4.5 must be chosen and equation 4.7 solved for the 90% confidence interval. Then, based on the input versus output curve from the chosen computer model as developed and described in the previous paragraph, the expected model output phosphorous loading interval can be determined. Application of the method using Option B is demonstrated in the next chapter on a soil sampling plan for a basin-scale modeling study.

Nonparametric method for determining sample size and the 90% confidence interval for estimating basin-scale soil test phosphorous:

Option A: Determine the required sample size for the 90% confidence interval.

Option B: Determine the 90% confidence interval given a predetermined sample size.

Steps:

1. Option A: Determine an acceptable soil test phosphorous interval and let it be the 90% confidence interval.

   Option B: Use a predetermined number of soil samples.

2. Option A: Chose the appropriate constant, *C*, from Table 4.5 based on major land use, poultry house density, or other basin characteristics.

   Option B: Chose the appropriate constant, *C*, from Table 4.5 based on major land use, poultry house density, or other basin characteristics.

3. Option A: Solve equation 4.7 for the required sample size within the 90% confidence interval.

   Option B: Solve equation 4.7 for the 90% confidence interval using the predetermined sample size.

Table 4.5. Constant, *C*, for Equation 4.7 to Determine Sample Size for Estimating Basin-scale Soil Test Phosphorous within a 90% Confidence Interval

Location/Data Set	Land Use	Poultry House Density[1]	Constant, *C*, For Equation 4.7 [2]
Upper Little Deep Fork Creek	Forest	No Poultry	16.8
Upper Little Deep Fork Creek	Grasslands, Pasture	No Poultry	27
Battle Branch	Pasture	0.018 houses/ha	125
Peacheater Creek	Pasture	0.015 houses/ha	295
Haw Creek	Pasture	0.014 houses/ha	377
Lake Eucha (AR portion)	Pasture	0.03 houses/ha	390

1. Density based on pasture/grassland coverage
2. Constant, *C*, for use with equation 4.7 with 90% Confidence Interval in mg/kg

CHAPTER 5

APPLICATION OF METHOD

The Biosystems and Agricultural Engineering Department, Oklahoma State University has contracted with the City of Tulsa, Oklahoma to perform a nonpoint source computer model assessment of the Lake Eucha basin. The focus of the assessment will be on sediment and phosphorous loadings. In order to perform the modeling, initial soil test phosphorous levels are needed. To estimate basin-scale soil test phosphorous levels, the nonparametric method for determining the 90% confidence interval for sample size was used to develop a soil sampling plan for the Oklahoma portion of the Lake Eucha basin.

The basin was divided into sub-basins based on major tributaries and each sub-basin will be modeled independently. The approach for assigning initial soil phosphorus levels in the computer model is to use mean soil test phosphorous levels by major land use for each sub-basin. Thus, the number of composite soil samples, or rather the number of fields to sample, need to be determined. For this project, a predetermined number of samples (30 for forest, 170 for pasture) were to be collected over the entire basin. The nonparametric method developed in this study was used to determine the 90% confidence intervals associated with the sample sizes to be used for each sub-basin.

Lake Eucha (Oklahoma) Basin

Lake Eucha Basin is located in northeastern Oklahoma and northwestern Arkansas (Figure 3.1). The Lake Eucha basin is approximately 93,000 ha (230,000 acres), with 40% in Benton County, Arkansas and the remainder in Delaware County, Oklahoma. This corresponds to approximately 55,800 ha (138,000 acres) for the Oklahoma portion. This basin is generally known for its extensive poultry industry activities, which include layers and broilers. A 1996 survey by the Water Quality Division, Oklahoma Conservation Commission found a total of 714 poultry houses within the basin with 225 on the Oklahoma side (OCC, 1996).

The 1:250,000 scale 1985 USGS digital elevation data were used to delineate the sub-basins. The basin was divided into six sub-basins that were similar in coverage area. The 1985 NRCS digital land use data for Oklahoma were used to determine the pasture areas. The poultry house inventory data were taken from the 1996 survey by the Water Quality Division, Oklahoma Conservation Commission (OCC, 1996). Figure 5.1 presents the Oklahoma portion land use coverage and sub-basin delineations. The tabular results along with poultry house densities are shown in Table 5.1.

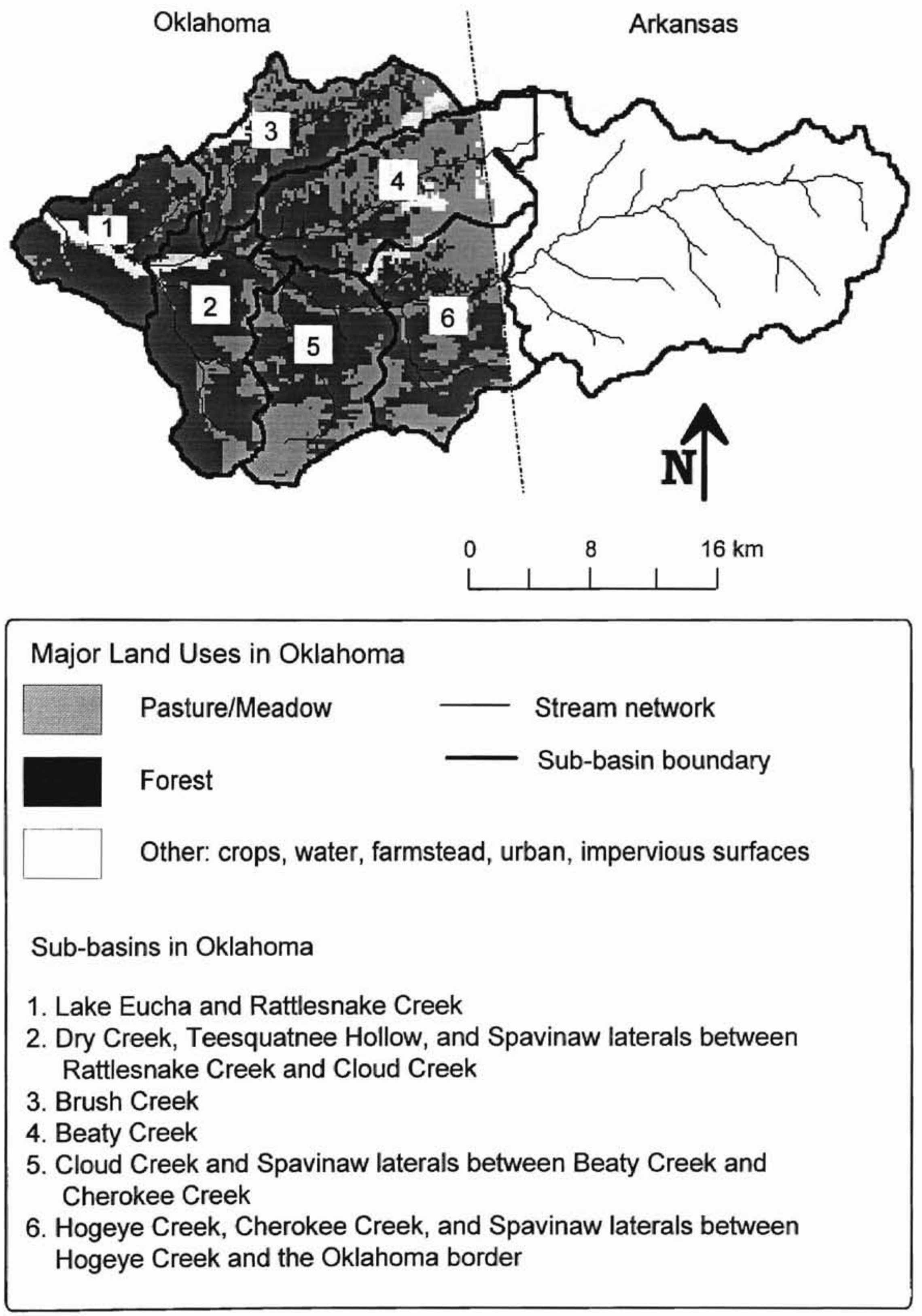


Figure 5.1. Lake Eucha Basin.

Table 5.1. Lake Eucha Basin (Oklahoma) Major Land Use and Poultry Inventory

	Sub-basin	Total Area (ha)	Forest Area (ha)	Forest Area (%)	Pasture Area (ha)	Pasture Area (%)	Poultry Houses (no.)
1.	Lake Eucha and Rattlesnake Creek	7,500	6,000	79	860	11	4
2.	Dry Creek, Teesquatnee Hollow, and Spavinaw laterals between Rattlesnake Creek and Cloud Creek	9,900	7,700	77	1,800	19	8
3.	Brush Creek	8,800	4,400	50	3,700	42	31
4.	Beaty Creek (Oklahoma portion)	10,300	4,500	44	5,200	50	80
5.	Dry Creek and Spavinaw laterals between Beaty Creek and Cherokee Creek	9,600	5,800	60	3,800	40	29
6.	Hogeye Creek, Cherokee Creek, and Spavinaw laterals between Hogeye Creek and the Oklahoma border	9,800	5,200	53	4,300	44	73

Hydrological/Water Quality Model

The hydrological/water quality model used for this project was the Spatially Integrated Model for Phosphorous Loading and Erosion (SIMPLE) (Storm et al., 1997; Sabbagh et al., 1995). SIMPLE is a continuous simulation, distributed parameter modeling system designed to predict sediment transport and phosphorous loading to surface waters from nonpoint sources on a watershed- or basin-scale. It encompasses a phosphorous transport model, a digital terrain model, a data base manager, and a menu-driven user interface. The SIMPLE modeling system can be used in conjunction with the GIS GRASS (CERL, 1988). The spatial component of SIMPLE is raster based, using either a single cell or a field consisting of multiple cells as the computational unit. SIMPLE estimates daily sediment loading, sediment-bound phosphorous, and soluble phosphorous from each cell or field. Average loading statistics are calculated on a daily, monthly, or annual basis.

Confidence Intervals

SIMPLE was used to develop relationship curves of "input soil P versus output P loading". These curves are used to determine effects on model output phosphorous loading due to variances in input soil test phosphorus. The resultant curves illustrate how varying the confidence interval width on the input initial phosphorous varies the expected output phosphorous loading interval width. Input/output curves were developed for forest and pasture separately.

SIMPLE was run on a single cell basis using typical input parameter values for forest and pasture, respectively. Ideally, the model would be run for the entire basin of interest to obtain the input/output curve. However, as with many modeling projects in the beginning stages, required data for the entire basin has not yet been obtained or developed. The sensitivity analysis for SIMPLE was used to determine the sensitive parameters. The most sensitive parameters were:

1. initial soil phosphorous,
2. curve number,
3. soil bulk density,
4. slope, and
5. USLE C factor.

The model was first run using average, or typical, values for all input parameters for a range of initial soil test phosphorous levels. Then, successive runs were made while varying a sensitive input parameter. The remaining parameters were held constant. This was done for both forest and pasture, respectively. Twenty years (1960–1980) of rainfall data from Siloam Springs, Arkansas were used for all the model runs. Independent annual results were used to compute the long-term 20 year average output phosphorus loadings. Tables 5.2 and 5.3 list the input parameters used for the model runs. The result was an "input versus output" curve for each run, as shown in Figures 5.2 and 5.3.

Table 5.2. Input Parameters for SIMPLE Single Cell Runs for Forest

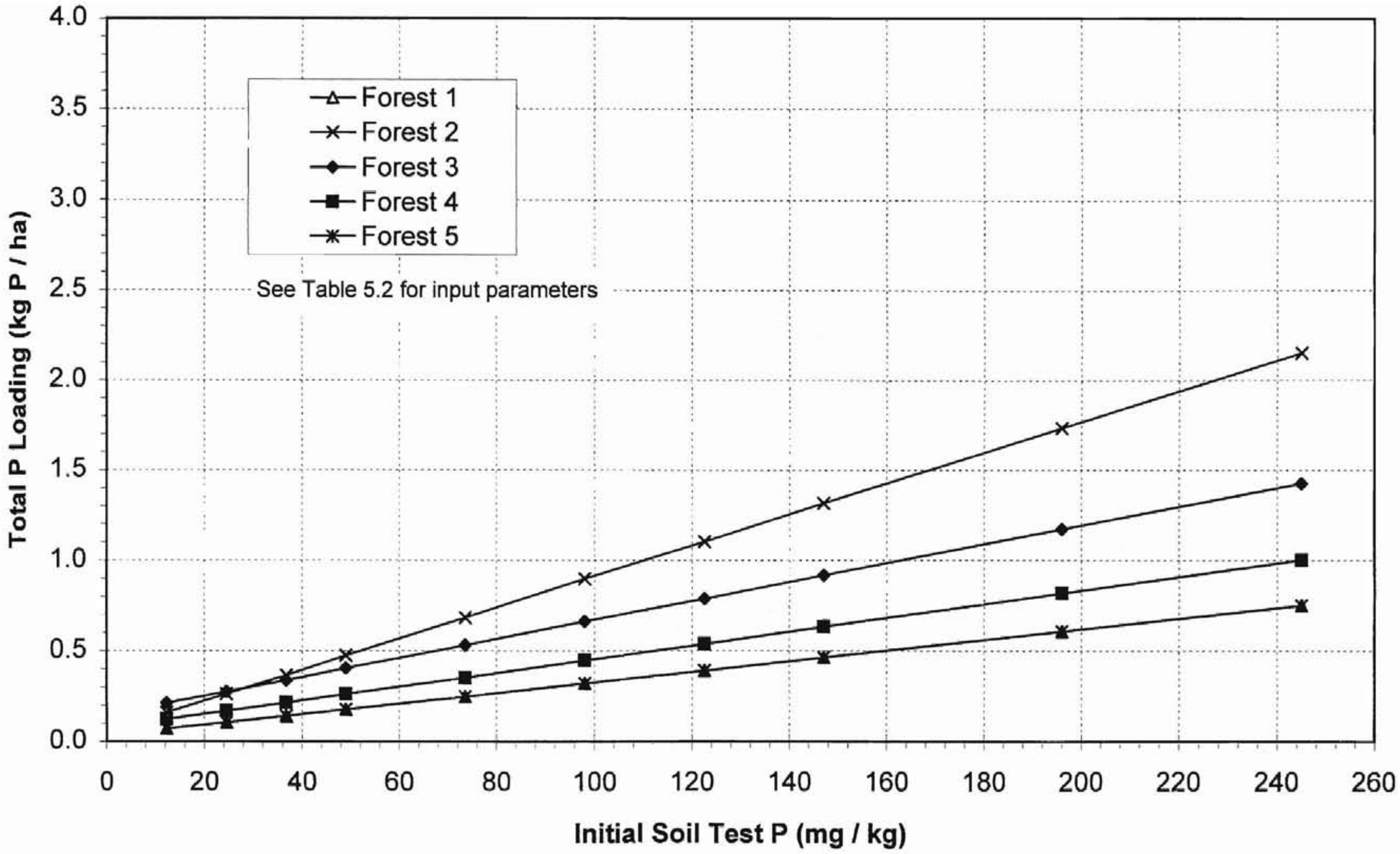
Parameter	Input Value				
	Forest 1	Forest 2	Forest 3	Forest 4	Forest 5
Initial Soil Test P* (mg/kg)	12.25 to 245	12.25 to 245	12.25 to 245	12.25 to 245	12.25 to 245
Curve Number*	55	**77**	55	55	55
USLE C factor*	0.00053	0.0005	**0.003**	0.005	0.005
Slope* (%)	5	5	5	**10**	5
Bulk Density* (g/cm^3)	1.45	1.45	1.45	1.45	**1.2**
K (English units)	0.35	0.35	0.35	0.35	0.35
Clay Content (%)	25	25	25	25	25
Organic Carbon (%)	1	1	1	1	1
pH	6.5	6.5	6.5	6.5	6.5
Hydrologic Soil Group	B	B	B	B	B
Slope to Stream (%)	10	10	10	10	10
Slope Length (m)	194	194	194	194	194

* Sensitive parameter

Table 5.3. Input Parameters for SIMPLE Single Cell Runs for Pasture

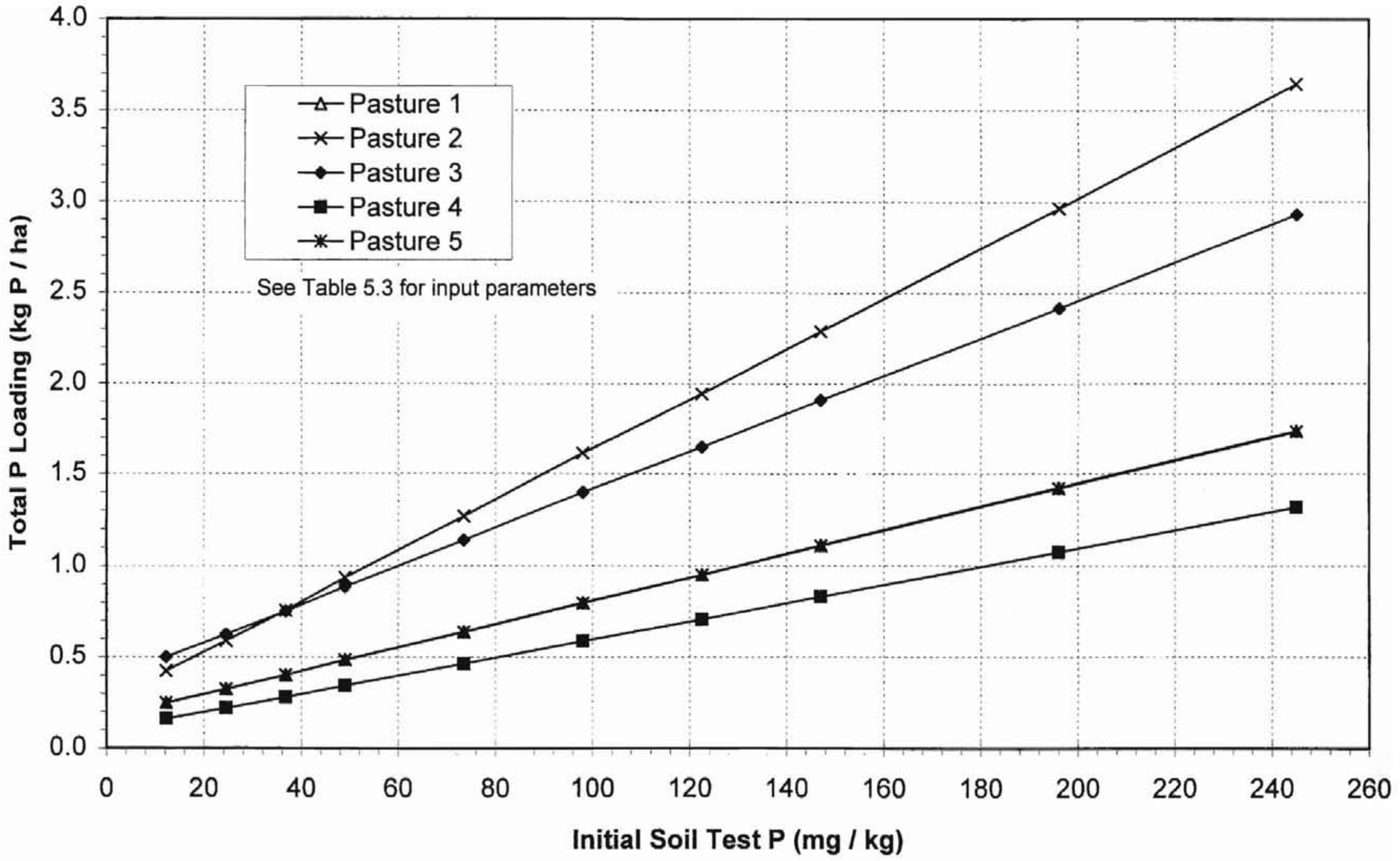
Parameter	Input Value				
	Pasture 1	Pasture 2	Pasture 3	Pasture 4	Pasture 5
Initial Soil Test P* (mg/kg)	12.25 to 245	12.25 to 245	12.25 to 245	12.25 to 245	12.25 to 245
Curve Number*	60	**80**	60	60	60
USLE C factor*	0.003	0.003	**0.008**	0.003	0.003
Slope* (%)	5	5	5	**3**	5
Bulk Density* (g/cm^3)	1.45	1.45	1.45	1.45	**1.2**
K (English units)	0.35	0.35	0.35	0.35	0.35
Clay Content (%)	25	25	25	25	25
Organic Carbon (%)	1	1	1	1	1
pH	6.5	6.5	6.5	6.5	6.5
Hydrologic Soil Group	B	B	B	B	B
Slope to Stream (%)	10	10	10	10	10
Slope Length (m)	194	194	194	194	194

* Sensitive parameter



See Table 5.2 for input parameters

Figure 5.2. SIMPLE initial P versus P loading for forest.



See Table 5.3 for input parameters

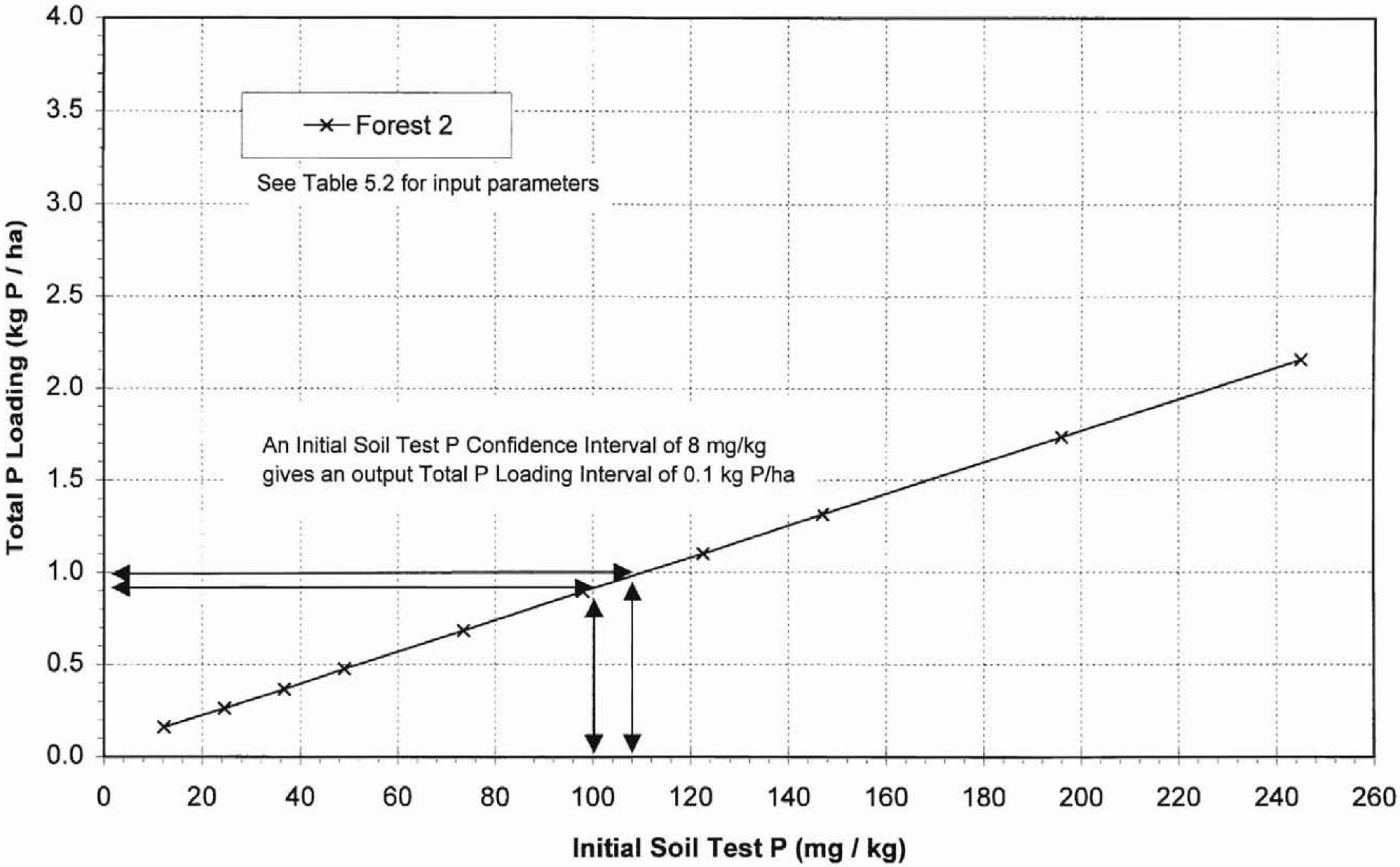
Figure 5.3. SIMPLE initial P versus P loading for pasture.

The slopes of the curves on Figures 5.2 and 5.3 vary somewhat as the input parameters change. Since all the data layers for the basin had not yet been developed, a single curve was chosen for determination of the model output confidence interval due to input soil test phosphorous variance. A conservative approach was taken to obtain the maximum expected model output variance due to input initial soil test phosphorus variance. For both forest and pasture, this corresponded to the curves for greatest expected curve numbers.

Since this soil sampling plan for the Oklahoma portion of the Lake Eucha basin dictated that a total of 200 soil samples were to be collected, Option B of the nonparametric method was used to determine the 90% confidence intervals. The 200 samples (30 for forest, 170 for pasture) were proportioned among the six sub-basins based on the percentage of forest and pasture, respectively, within each sub-basin, i.e. the sub-basin with the greatest pasture area will receive the greatest number of pasture samples.

The appropriate constant, *C*, values from Table 4.5 had to be chosen for use with the nonparametric equation developed, equation 4.7, to determine the 90% confidence intervals. Since some soil test phosphorous data (156 samples) for pasture from Delaware County were available from the Soil, Water, and Forage Laboratory, Oklahoma State University, they were used to help choose the appropriate *C* value for pasture. Only soil phosphorous and land use data were provided and sample location was not, so the data could not be directly used as fully representative of the entire Lake Eucha (Oklahoma portion) basin since no information was known of the individual locations and collection of samples. The mean of the soil test phosphorous data from Delaware County was 133 mg/kg, the median was 86 mg/kg, the standard deviation was 127 mg/kg, the minimum was 2 mg/kg, the maximum was 524 mg/kg, and the *C* value for these data was computed to be 418. The *C* value from Table 4.5 for the Peacheater Creek data was chosen as the value to use with equation 4.7 for pasture. This choice was based on Delaware County soil sample data information, proximity of the watershed to the Lake Eucha basin, watershed size similarities, and the fact that all the pasture from Peacheater Creek had been sampled. Due to the lack of any other available data, the *C* value for forest from Upper Little Deep Fork Creek was used for forest land use. These *C* values and equation 4.7 were then used to determine the 90% confidence intervals based on the number of samples to be collected.

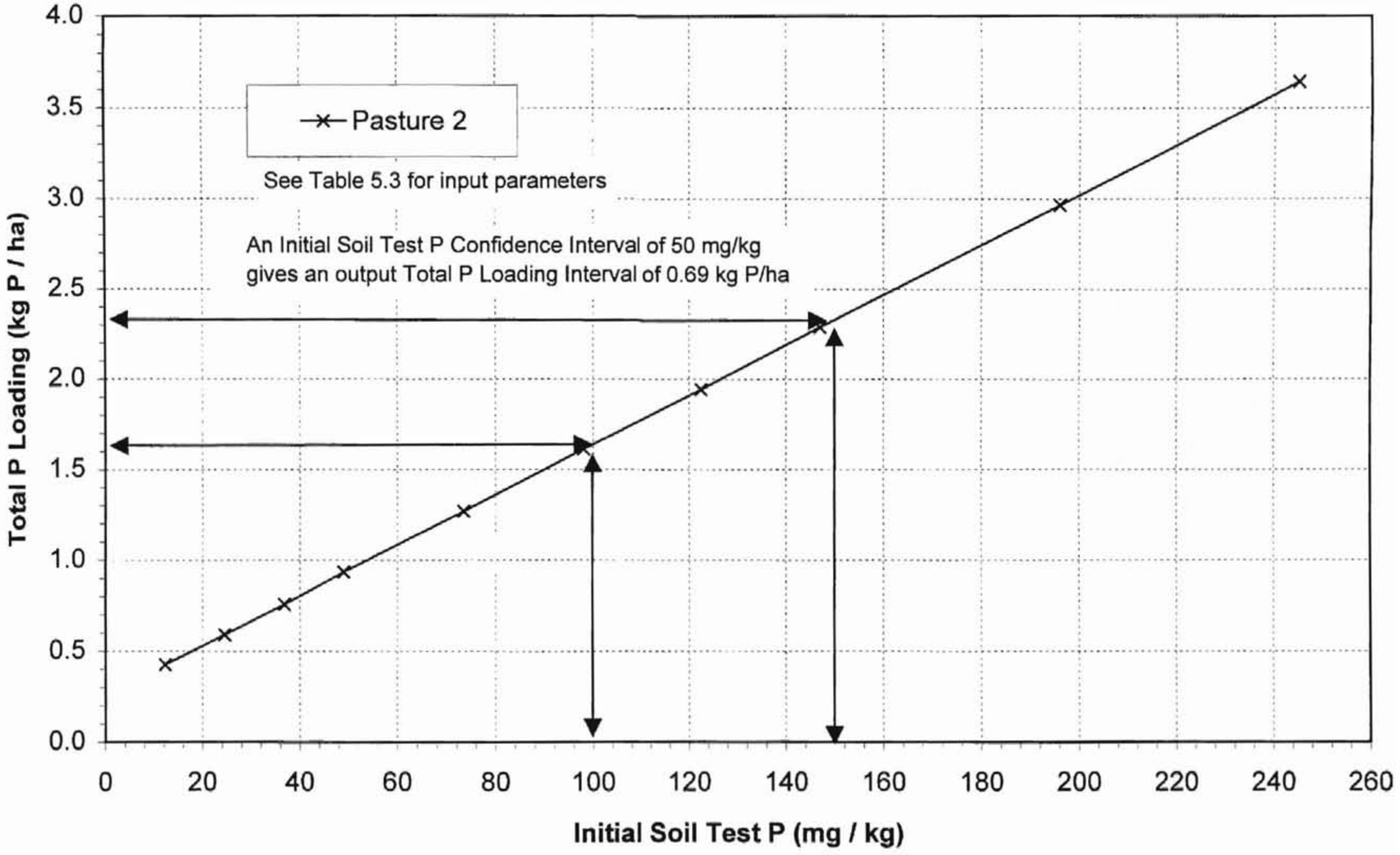
Using the curves as shown in Figures 5.4 and 5.5 for forest and pasture, respectively, the expected total phosphorous loading output interval from SIMPLE was found for each initial soil test phosphorous 90% confidence interval. The output total phosphorous loading interval from SIMPLE does not correspond to the expected 90% confidence interval from the model. The curves developed in Figures 5.2 and 5.3 were based on variation in initial phosphorus input. To obtain a true confidence interval in the model output, variance in all input parameters and inherent variance produced within the model must be taken into account. However, the curves produced do give an indication of the expected effect from initial soil phosphorous input on the model output confidence interval. The results are summarized in the next section.



See Table 5.2 for input parameters

An Initial Soil Test P Confidence Interval of 8 mg/kg gives an output Total P Loading Interval of 0.1 kg P/ha

Figure 5.4. Confidence Interval determination for forest.



See Table 5.3 for input parameters

An Initial Soil Test P Confidence Interval of 50 mg/kg gives an output Total P Loading Interval of 0.69 kg P/ha

Figure 5.5. Confidence Interval determination for pasture.

Number of Observations and Confidence Intervals

The steps of the nonparametric method developed in Chapter 4 were used to determine the 90% confidence intervals for the predetermined soil sample sizes for the soil sampling plan for the Oklahoma portion of the Lake Eucha basin. The results are shown in Table 5.4.

Option B: Determine the 90% confidence interval given a predetermined sample size.

Steps:

1. Option B: Use a predetermined number of soil samples.

   - Project requirements: 30 samples for forest, 170 for pasture. The samples were proportioned among the six sub-basins based on percentage of land use.

2. Option B: Chose the appropriate constant, *C*, from Table 4.5 based on major land use, poultry house density, or other basin characteristics.

   - The *C* value for forest from the Upper Little Deep Fork Creek data was chosen for forest. The *C* value from the Peacheater Creek data was chosen for pasture. These were chosen based on available data and watershed/basin characteristics.

3. Option B: Solve equation 4.7 for the 90% confidence interval using the predetermined sample size.

   - The samples sizes were predetermined and the corresponding 90% confidence intervals were computed. Table 5.4 shows the sub-basins, the sample sizes, and the corresponding 90% confidence intervals.

Table 5.4. Lake Eucha Basin (Oklahoma) Required Sample Sizes for the Soil Sampling Plan

	Sub-basin	Land Use	Sample Size [1]	90% Confidence Interval [2] (mg/kg)	Expected SIMPLE output interval [3] (kg P/ha)
1.	Lake Eucha and Rattlesnake Creek	Forest	5	8	0.1
		Pasture	7	111	1.53
2.	Dry Creek, Teesquatnee Hollow, and Spavinaw laterals between Rattlesnake Creek and Cloud Creek	Forest	7	6	0.05
		Pasture	15	76	1.05
3.	Brush Creek	Forest	4	8	0.1
		Pasture	32	52	0.72
4.	Beaty Creek (Oklahoma portion)	Forest	4	8	0.1
		Pasture	44	45	0.62
5.	Dry Creek and Spavinaw laterals between Beaty Creek and Cherokee Creek	Forest	5	8	0.1
		Pasture	34	50	0.69
6.	Hogeye Creek, Cherokee Creek, and Spavinaw laterals between Hogeye Creek and the Oklahoma border	Forest	5	8	0.1
		Pasture	38	48	0.66

1. Sample size based on a total of 30 for forest, 170 for pasture; distributed by percent land use coverage within each sub-basin.
2. Computed from Equation 4.7 with *C* value from Upper Little Deep Fork Creek for forest, and *C* value from Peacheater Creek for pasture.
3. Interval based only on effects of initial soil test phosphorous, from Figure 5.4 for forest and Figure 5.5 for pasture.

CHAPTER 6

SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

Summary and Conclusions

The main objectives of this research were to evaluate basin-scale soil test phosphorous probability distributions, develop and evaluate a nonparametric approach for determining required sample size for estimating basin-scale soil phosphorous, and apply the results to develop a soil sampling plan for gathering input soil phosphorous data for a hydrological/water quality computer model. The nonparametric approach was also to be used for estimating the 90% confidence interval for a predetermined number of soil samples. The nonparametric approach was investigated because data from high-level soil phosphorous basins, such as studied in this research, may not adhere to all the necessary assumptions to allow the valid use of classic parametric statistics or geostatistical techniques for determining appropriate sample size.

Soil test phosphorous probability distributions from several watersheds and basins were evaluated. The high-level soil phosphorous data were from watersheds containing poultry industry activities, such as pullets, layers, and broilers. It was found that the data tend to exhibit a bi-modal distribution in these high-level soil phosphorous watersheds. This is primarily due to varying poultry litter application rates on different pasture fields and the fact that some fields receive litter and some do not. There was no significant association found between selected soil characteristics and soil test phosphorous for the data examined.

Empirical probability distributions of the soil phosphorous data and empirical 90% confidence intervals created from the data sets were used to develop nonparametric equations for predicting soil sample sizes for estimating soil phosphorus levels for pasture and forest, respectively. It was found that the nonparametric approach did not give sample size results differing greatly from that obtained by using classic statistic techniques. The preferred use of geostatistics, however, was prohibited from use in these poultry watersheds due to abrupt changes in soil phosphorous levels across field or property boundaries. The use of geostatistics relies on the gradual change in a parameter spatially.

The nonparametric equations developed were then used to form a soil sampling plan for the Oklahoma portion of the Lake Eucha basin, which contains poultry activities. The basin was divided into sub-basins using a GIS. Since the soil sampling plan dictated a predetermined number of soil samples to be collected, the option of the nonparametric method was used for computing the 90% confidence interval based on a predetermined sample size. The appropriate nonparametric equation developed for computing the confidence intervals for pasture was selected based on limited available data from the Oklahoma portion of the Lake Eucha basin and other watershed characteristics.

An equation for determining sample size and 90% confidence intervals for forest was also developed and applied.

One difficult part was deciding which one of the developed nonparametric sample size prediction equations to use for pasture. As with most initial soil sampling plans, the site-specific variance of the parameter to be measured, soil test phosphorous in this case, is not known. Thus, judgement must be exercised in choosing an appropriate variance to apply to the sample size prediction equations.

Recommendations

It is recommended that classic statistic techniques be used for future sample size determination in similar watersheds or basins, since it would be simpler and should provide similar results. It is also recommended that further research in the area of high-level soil phosphorous determination be placed on ways of predicting soil phosphorous levels based on relatively easily obtained data. Perhaps a model could be developed that could predict soil phosphorous levels based on soil types, litter and fertilizer application history, distance of field to poultry house, etc. This could provide a means of predicting soil phosphorous levels without initiating extensive and expensive soil sampling plans.

REFERENCES

Bullard, Joe. Personal communication. May, 1998. Oklahoma State University Cooperative Extension Service.

Burwell, R.E., G.E. Schuman, H.G. Heinemann, and R.G. Spomer. 1977. Nitrogen and phosphorous movement from agricultural watersheds. J. Soil Water Conserv. 32:226-230.

Cai, T.T., T.W. Olsen, R.S. Yost, and J.A. Silva. 1997. Performance indices for tests of soil nutrient status: extractable phosphorous. Commun. In Soil Sci. Plant Anal. 28(3-5): 329-339.

Cameron, D.R., M. Nybert, J.A. Toogood, and D.H. Loverty. 1971. Accuracy of field sampling for soil tests. Can. J. Soil Sci. 51:165-175.

CERL. 1988. GRASS Reference Manual, Version 4.1. Champaign, Illinois: U.S. Army, Corps of Engineers, Construction Engineering Research Laboratory.

Cline, M.G. 1944. Principles of soil sampling. Soil Science. 58:275-288.

Edwards, D.R., P.A. Moore, Jr., T.C. Daniel, and P. Srivastava. 1996. Poultry litter-treated length effects on quality of runoff from fescue plots. Transactions of the ASAE. 39(1):105-109.

Fixen, P.E., and J.H. Grove. 1990. Testing soils for phosphorous. *In* R.L. Westerman et al. (ed.) Soil testing and plant analysis. 3rd ed. SSSA Book Series No. 3:141-180.

Foth, H.D., and B.G. Ellis. 1997. Soil fertility. 2nd ed. CRC Press. Boca Raton, FL.

Gupta, R.K., S. Mostaghimi, P.W. McClellan, M.M. Alley, D.E. Brann. 1997. Spatial variability and sampling stategies for NO_3-N, P, and K determinations for site-specific farming. Transactions of the ASAE. 40(2):337-373.

Haan, C.T. 1977. Statistical methods in hydrology. Ames: The Iowa State University Press.

Johnson, G.V., W.R. Raun, H. Zhang, and J.A. Hattey. 1997. Oklahoma soil fertility handbook. Dept. of Agronomy. Okla. State Univ., Stillwater, OK.

Matheron, G. 1963. Principles of geodstatistics. Econ. Geol. 58:1246-1266.

McBratney, A.B., and R. Webster. 1983. How many observations are needed for regional estimation of soil properties? Soil Science. 135:177-183.

McIntyre, D.S., and C.B. Tanner. 1958. Anormaly distributed soil physical measurements and nonparametric statistics. Soil Science. 88:133-137.

Mehlich, A. 1984. Mehlich 3 soil test extractant: A modification of Mehlich 2 extractant. Commun. In Soil Sci. Plant Anal. 15:1409-1416.

NRCS (Natural Resource Conservation Service), 1995. Spavinaw Creek of the Neosho River; Water Quality Incentive Proposal, Benton County, Arkansas. October 1995.

OCC (Oklahoma Conservation Commission), 1996. Confined animal inventory: Lake Eucha Watershed. OCC – Water Quality Division, Oklahoma City, OK.

Ott, L. 1984. An introduction to statistical methods and data analysis. 2nd ed. PWS Publishers. Boston, MA.

Parkin, T.B., J.J. Meisinger, S.S. Chester, J.L. Starr, and J.A. Robinson. 1988. Evaluation of statistical estimation methods for lognormally distributed variables. Soil Sci. Soc. Am. J. 52:323-329.

Parkin, T.B., S.T. Chester, and J.A. Robinson. 1990. Calculating confidence intervals for the mean of a lognormally distributed variable. Soil Sci. Soc. Am. J. 54:321-326.

Peavy, H.S., D.R. Rowe, and G. Tchobanoglous. 1985. Environmental Engineering. McGraw-Hill Pub. Co., New York.

Peterson, R.G., and L.D. Calvin. 1986. Sampling *In* A. Klute (ed.) Methods of soil analysis, Part 1. 2nd ed. Agronomy 9:33-51.

Pierzynski, G.M., J.T. Simms, and G.F. Vance. 1994. Soils and environmental quality. CRC Press. Boca Raton, FL.

Sabbagh, G.J., D.E. Storm, M.D. Smolen, C.T. Haan, and W.C. Hession. 1995. SIMPLE: sediment and phosphorous loading model. *In* Watershed management, planning for the 21st century. Watershed Management Committee of the Water Resources Engineering Division/ASCE, pp.93-102.

Sabbe, W.E., and D.B. Marx. 1987. Soil sampling: spatial and temporal variability. *In* J.R. Brown (ed.) Soil testing: sampling, correlation, calibration, and interpretation. SSSA special pub. no 21:1-14.

SAS Institute. 1995. SAS/STAT Window's NT version. Release 6.11 edition. SAS Inst., Cary, NC.

Sharpley, A.N., L.W. Reed, and D.K. Simmons. 1982. Relationships between available soil p forms and their role in water quality modeling. OK Agricultural Experiment Station Tech. Bulletin No. 157. Okla. State Univ., Stillwater, OK.

Sharpley, A.N. 1985. Phosphorous cycling in unfertilized and fertilized agricultural soils. Soil Sci. Soc. Am. J. 49:905-911.

Sharpley, A.N., H. Tiessen, and C.V. cole. 1987. Soil phosphorous forms extracted by soil tests as a function of pedogenesis. Soil Sci. soc. Am. J. 51:362-365.

SCS (Soil Conservation Service). 1958. Soil Survey of Creek County, Oklahoma. Government Printing Office. Washington, D.C.

SCS (Soil Conservation Service). 1965. Soil Survey of Adair County, Oklahoma. Government Printing Office. Washington, D.C.

SCS (Soil Conservation Service). 1970. Soil Survey of Cherokee and Delaware Counties, Oklahoma. Government Printing Office. Washington, D.C.

SCS (Soil Conservation Service). 1970. Soil Survey of Lincoln County, Oklahoma. Government Printing Office. Washington, D.C.

SCS (Soil Conservation Service). 1981. Soil Survey of LeFlore County, Oklahoma. Government Printing Office. Washington, D.C.

SCS (Soil Conservation Service). 1985. Digital Land Use data for Oklahoma. Government Printing Office. Washington, D.C.

Steel, R.G.D., and J.H. Torrie. 1980. Principles and procedures of statistics. 2nd ed. McGraw-Hill, New York.

Storm, D.E., G.J. Sabbagh, M.S. Gregory, M.D. Smolen, D.R. Gade, C.T. Haan, and T. Kornecki. 1996. Basin-wide pollution inventory for the Illinois River comprehensive basin management program. Final report, Vol. I: Estimating nonpoint and point source loading to the Upper Illinois River Basin. Submitted to the Oklahoma Conservation Commission for the U.S. Environmental Protection Agency. Departments of Biosystems and Agricultural Engineering, and Agronomy, Okla. State Univ., Stillwater, OK. August 1996.

Storm, D.E., G.J. Sabbagh, C.T. Haan, and M.D. Smolen. 1997. Phosphorous transport model for nonpoint source modeling. Submitted to the Water Resources Bulletin, not yet published.

USDA-Natural Resources Conservation Service. 1994. National Map Unit Interpretation Record (MUIR) Database. Iowa State University.

Wagner, K. and S. Woodruff. 1997. Phase I clean lakes project. Diagnostic and feasibility study of Lake Eucha. Oklahoma Conservation Commission – Water Quality Division. Oklahoma City, OK.

Warrick, A.W., D.E. Myers, and D.R. Neilsen. 1986. Geostatistical methods applied to soil science. *In* A. Klute (ed.) Methods of soil analysis, Part 1. 2nd ed. Agronomy 9:53-82.

Zhang, H., and G. Johnson. 1997. Collecting good soil samples by reducing spatial variability. OSU Production Technology 97-37, Vol. 9 no. 37. OK Coop. Ext. Service, Stillwater, OK.

APPENDICES

APPENDIX A

SOIL DATA

Table A.1. Upper Little Deep Fork Creek Soil Mapping Units

Soil Ref. No.	Soil Description	Area Coverage (ha)	(ac)	(%)
1	Bates fine sandy loam, gently sloping	356	879	0.90
2	Bates fine sandy loam, sloping	176	436	0.45
3	Bates fine sandy loam, sloping, severely eroded	13	32	0.03
4	Broken or Gullied sandy upland	122	302	0.31
6	Choteau very fine sandy loam, nearly level	9	22	0.02
7	Cleburne fine sandy loam	179	441	0.45
8	Collinsville and Bates soils, gently sloping	107	264	0.27
9	Collinsville and Talihina soils, sloping	895	2211	2.27
10	Collinsville and Talihina soils, strongly sloping	554	1368	1.40
11	Darnell and Pottsville soils, sloping	9096	22477	23.06
12	Darnell and Pottsville soils, strongly sloping	4500	11119	11.40
13	Dennis and Okemah loams, gently sloping	1167	2884	2.96
14	Dennis and Okemah loams, sloping	592	1462	1.50
15	Dennis and Okemah loams, sloping, severely eroded	121	300	0.31
16	Dougherty and Stidham fine sandy loams, gently sloping	371	916	0.94
17	Dougherty and Stidham fine sandy loams, nearly level	24	59	0.06
18	Dougherty and Stidham fine sandy loams, sloping	490	1212	1.24
23	Gullied bottom land	1224	3025	3.10
24	Mason clay loam	17	41	0.04

continued

Table A.1. (continued) Upper Little Deep Fork Creek Soil Mapping Units

Soil Ref. No.	Soil Description	Area Coverage (ha)	(ac)	(%)
25	Mason silt loam	758	1873	1.92
27	Oil-Waste land	186	460	0.47
28	Okemah and Woodson clay loams	18	44	0.04
30	Pulaski fine sandy loam	1885	4659	4.78
33	Stephenville and Darnell fine sandy loams, gently sloping	4689	11587	11.89
34	Stephenville and Darnell fine sandy loams, sloping	5360	13244	13.58
35	Stephenville & Darnell fine sandy loams,sloping,severely eroded	1891	4674	4.79
38	Teller silt loam, sloping	8	19	0.02
41	Verdigris clay loam	180	444	0.46
42	Verdigris fine sandy loam	818	2022	2.07
43	Verdigris silt loam	1365	3373	3.46
101	Bonham loam, 1 to 3 percent slopes	45	112	0.11
102	Bonham loam, 3 to 5 percent slopes	109	270	0.28
103	Bonham loam, 2 to 5 percent slopes, eroded	142	350	0.36
104	Breaks-Alluvial land complex	28	69	0.07
105	Broken alluvial land	45	110	0.11
106	Chickasha loam, 1 to 3 percent slopes	47	117	0.12
107	Chickasha loam, 3 to 5 percent slopes	17	42	0.04
108	Chickasha loam, 2 to 5 percent slopes, eroded	153	377	0.39

continued

Table A.1. (continued) Upper Little Deep Fork Creek Soil Mapping Units

Soil Ref. No.	Soil Description	Area Coverage (ha)	(ac)	(%)
109	Chickasha and Bonham soils, 2 to 6 % slopes, severely eroded	370	914	0.94
112	Darnell-Stephenville fine sandy loams, 3 to 12 % slopes	512	1265	1.30
113	Darnell-Stephenville complex, 3-12 % slopes, severely eroded	85	210	0.22
117	Konawa loamy fine sand, 0 to 3 percent slopes	15	37	0.04
120	Mason silt clay	42	103	0.11
122	Noble fine sandy loam, 3 to 8 percent slopes	6	16	0.02
129	Pulaski fine sandy loam	23	56	0.06
136	Stephenville fine sandy loam, 1 to 3 percent slopes	68	168	0.17
137	Stephenville fine sandy loam, 3 to 5 percent slopes	121	299	0.31
138	Stephenville fine sandy loam, 2 to 5 percent slopes, eroded	54	133	0.14
145	Vanoss clay loam, 3 to 5 percent slopes	8	19	0.02
146	Vernon-Collinsville complex, 5 to 20 percent slopes	131	323	0.33
	Water	264	653	0.67
	Total	39,454	97,490	100

Table A.2. Upper Little Deep Fork Creek Soil Characteristics

Soil Ref. No.	K (English units)	Organic Matter (%)	Clay (%)	Bulk Density (g/cm^3)	Hydrologic Soil Group	Slope Range (%)
1	0.2	1.50	10	1.55	B	2 - 4
2	0.2	1.50	10	1.55	B	4 - 6
3	0.2	1.50	10	1.55	B	4 - 6
4	0.24	0.75	14	1.45	B	3 - 8
6	0.43	2.00	21	1.43	C	1 - 4
7	0.24	0.75	15	1.45	B	1 - 4
8	0.32	2.00	13.5	1.43	D	2 - 4
9	0.32	2.00	13.5	1.43	D	4 - 12
10	0.32	2.00	13.5	1.43	D	12 - 20
11	0.1	0.75	15	1.48	C	4 - 12
12	0.1	0.75	15	1.48	C	12 - 20
13	0.43	2.00	18.5	1.43	C	1 - 4
14	0.43	2.00	18.5	1.43	C	4 -6
15	0.43	2.00	18.5	1.43	C	4 -6
16	0.24	0.75	14	1.45	B	2 - 5
17	0.24	0.75	14	1.45	B	0 - 2
18	0.24	0.75	14	1.45	B	5 - 8
23	0.32	2.00	23.5	1.38	B	0 - 1
24	0.32	2.00	28.5	1.45	B	0 - 1
25	0.37	2.00	19.5	1.40	B	0 - 1
27	0.24	--	15	1.45	D	0 - 4
28	0.37	2.00	31	1.45	C	0 - 1
30	0.2	0.75	14	1.45	B	0 - 1
33	0.24	0.75	15	1.45	B	2 - 4
34	0.24	0.75	15	1.45	B	4 - 7
35	0.24	0.75	15	1.45	B	4 - 7

continued

Table A.2. (continued) Upper Little Deep Fork Creek Soil Characteristics

Soil Ref. No.	K (English units)	Organic Matter (%)	Clay (%)	Bulk Density (g/cm^3)	Hydrologic Soil Group	Slope Range (%)
38	0.37	2.00	15	1.43	B	5 - 7
41	0.32	3.00	31	1.40	B	0 - 1
42	0.24	0.75	14	1.45	B	0 - 1
43	0.32	3.00	21	1.35	B	0 - 1
101	0.43	2.00	15	1.43	D	0 - 2
102	0.43	2.00	15	1.43	D	2 - 5
103	0.43	2.00	15	1.43	D	2 - 5
104	0.37	0.75	31	1.45	D	5 - 12
105	0.37	2.00	19	1.43	B	0 - 3
106	0.37	2.00	20	1.45	B	1 - 3
107	0.37	2.00	20	1.45	B	3 - 5
108	0.37	2.00	20	1.45	B	1 - 5
109	0.37	2.00	20	1.45	B	1 - 8
112	0.2	0.75	15	1.48	C	3 - 12
113	0.2	0.75	15	1.48	C	3 - 12
117	0.2	0.75	6	1.43	B	0 - 3
120	0.37	2.00	19.5	1.40	B	0 - 1
122	0.2	0.75	14	1.45	B	3 - 8
129	0.2	0.75	14	1.45	B	0 - 1
136	0.24	0.75	15	1.45	B	1 - 3
137	0.24	0.75	15	1.45	B	3 - 5
138	0.24	0.75	15	1.45	B	2 - 5
145	0.37	0.75	31	1.45	D	3 - 5
146	0.37	0.75	31	1.45	D	5 - 20

Table A.3. Battle Branch Soil Mappng Units

Soil Ref. No.	Soil Description	Area Coverage (ha)	(ac)	(%)
2	Baxter silt loam, 1 to 3 % slopes	119	294	5.32
3	Baxter cherty silt loam, 1 to 3 % slopes	274	678	12.26
4	Baxter Locust complex, 3 to 5 % slopes	286	706	12.77
5	Captina silt loam, 1 to 3 % slopes	161	398	7.19
8	Clarksville very cherty silt loam, 1 to 8 % slopes	155	382	6.91
9	Clarksville stony silt loam, 5 to 20 % slopes	274	677	12.25
10	Clarksville stony silt loam, 20 to 50 % slopes	342	845	15.28
19	Jay silt loam, 0 to 2 % slopes	18	44	0.80
21	Locust cherty silt loam, 1 to 3 % slopes	57	140	2.53
22	Newtonia silt loam, 0 to 1 % slopes	16	40	0.73
23	Newtonia silt loam, 1 to 3 % slopes	38	94	1.70
33	Sallisaw silt loam, 0 to 1 % slopes	3	6	0.12
34	Sallisaw silt loam, 1 to 3 % slopes	28	69	1.25
35	Sallisaw gravelly silt loam, 1 to 3 % slopes	40	99	1.78
36	Sallisaw gravelly silt loam, 3 to 8 % slopes	129	319	5.76
37	Staser silt loam	58	144	2.61
38	Staser gravelly loam	140	346	6.25
39	Stigler silt loam, 0 to 1 % slopes	82	203	3.67
41	Taloka silt loam, 0 to 1 % slopes	18	45	0.81
	Total	2,237	5,529	100

Table A.4. Battle Branch Soil Characteristics

Soil Ref. No.	K (English units)	Organic Carbon (%)	Clay (%)	Bulk Density (g/cm^3)	Hydrologic Soil Group	Slope Length (m)
2	0.33	1.76	19	1.37	B	152
3	0.33	1.76	19	1.37	B	152
4	0.33	1.76	19	1.37	B	121
5	0.36	1.18	12	1.43	B	152
8	0.39	0.74	12	1.46	B	15
9	0.43	0.74	25	1.43	B	60
10	0.43	0.74	25	1.43	B	30
19	0.37	1.18	18	1.51	C	167
21	0.4	0.59	12	1.48	B	152
22	0.37	1.18	18	1.41	B	182
23	0.37	1.18	18	1.41	B	152
33	0.41	0.74	33	1.46	B	15
34	0.41	0.74	33	1.46	B	15
35	0.39	0.74	12	1.46	B	15
36	0.39	0.74	12	1.46	B	15
37	0.34	1.76	25	1.35	B	15
38	0.34	1.76	25	1.35	B	15
39	0.36	1.18	12	1.43	D	182
41	0.44	0.44	25	1.45	D	182

Table A.5. Peacheater Creek Soil Mapping Units

Soil Ref. No.	Soil Description	Area Coverage (ha)	(ac)	(%)
1	Bodine very cherty silt loam, 1-8% slopes	1943	4802	30.07
2	Bodine stony silt loam, 5-15% slopes	487	1204	7.54
3	Bodine stony silt loam, steep	1653	4085	25.59
5	Dickson silt loam, 1-3% slopes	750	1852	11.60
6	Dickson cherty silt loam, 0-3% slopes	557	1377	8.62
7	Etowah silt loam, 0-1% slopes	0	1	0.01
8	Etowah silt loam, 1-3% slopes	80	198	1.24
9	Etowah gravelly silt loam, 1-3% slopes	251	620	3.88
10	Etowah and Greendale soils, 3-8% slopes	258	638	4.00
11	Gravelly alluvial land	188	464	2.91
13	Hector-Linker fine sandy loams, 1-5% slopes	23	56	0.35
15	Huntington gravelly loam	49	121	0.75
16	Jay silt loam, 0-2% slopes	139	344	2.15
17	Lawrence silt loam	3	8	0.05
20	Linker loam, 3-5% slopes	14	34	0.21
21	Linker loam, 3-5% slopes, eroded	27	68	0.42
26	Summit silty clay loam, 1-3% slopes	21	51	0.32
29	Taft silt loam	19	46	0.29
	Total	6,461	15,966	100

Table A.6. Peacheater Creek Soil Characteristics

Soil Ref. No.	K (English units)	Organic Carbon (%)	Clay (%)	Bulk Density (g/cm^3)	Hydrologic Soil Group	Slope Length (m)
1	0.28	0.44	14	1.45	B	122
2	0.28	0.44	14	1.45	B	61
3	0.28	0.44	14	1.45	B	61
5	0.43	0.74	25	1.43	B	152
6	0.43	0.74	25	1.43	B	152
7	0.37	1.18	25	1.39	B	189
8	0.37	1.18	25	1.39	B	152
9	0.37	1.18	25	1.39	B	152
10	0.37	1.18	25	1.39	B	122
11	0.21	0.01	1	1.34	B	15
13	0.19	0.85	17	1.5	C	152
15	0.28	2.65	24	1.34	B	15
16	0.43	0.01	18	1.51	C	189
17	0.43	1.47	18	1.39	C	152
20	0.28	1.03	19	1.48	B	122
21	0.28	1.03	19	1.48	B	122
26	0.37	0.1	33	1.34	C	152
29	0.43	2.06	18	1.34	D	15

APPENDIX B

SOIL TEST PHOSPHOROUS DATA

Table Page

Table B.1. Upper Little Deep Fork Creek Soil Test Phosphorous for Forest

Sample ID. No.	Land Use Classification Subclass	Soil Test P*	
		(mg/kg)	(lb/ac)
95	Stable Forest	19	38
102	(undisturbed, 0 – 1% bare soil)	21	42
109		23	47
161		18	36
165		19	38
166		13	26
169		19	38
172		25	51
	mean:	19	40
47	Moderately Used Forest	7	15
54	(1 –10% bare soil)	20	41
104		15	31
150		19	39
151		23	46
154		27	55
155		17	34
157		12	25
158		14	28
159		14	29
167		21	42
168		15	30
170		21	42
173		14	29
	mean:	17	35
14	Heavily Used Forest	23	47
32	(> 10% bare soil)	28	57
66		12	25
72		19	38
152		17	34
153		29	59
156		11	23
160		16	33
162		14	29
163		17	35
164		24	48
171		18	36
	mean:	19	39

* Mehlich III phosphorous

Table B.2. Upper Little Deep Fork Creek Soil Test Phosphorous for Grassland

Sample ID. No.	Land Use Classification Subclass	Soil Test P* (mg/kg)	(lbs/ac)
1	Good Condition Grassland	29	60
3	(< 1% bare soil)	26	53
5		13	27
7		22	44
11		11	22
16		9	19
17		13	26
19		10	20
26		9	18
29		30	62
40		14	29
41		11	23
44		13	27
48		12	24
53		18	37
55		13	26
57		15	30
58		12	24
61		19	38
62		24	49
65		14	28
68		18	37
70		10	21
78		35	71
79		14	29
80		14	28
84		8	17
91		11	23
98		15	31
103		7	15
111		14	29
113		12	24
114		20	40
125		10	21
126		11	22
129		17	35
130		12	25
135		34	70

continued

Table B.2. (continued) Upper Little Deep Fork Creek Soil Test Phosphorous for Grassland

Sample ID. No.	Land Use Classification Subclass	Soil Test P* (mg/kg)	(lbs/ac)
138	Good Condition Grassland	21	42
141	(< 1% bare soil)	11	23
144		21	42
	mean:	16	33
4	Fair Condition Grassland	18	37
6	(1 – 5% bare soil)	13	26
9		14	28
10		8	17
22		21	42
31		16	32
35		19	38
39		19	38
42		10	21
45		10	21
46		9	19
51		13	27
71		14	29
74		8	17
75		11	23
76		25	51
81		11	23
82		18	36
83		16	33
85		17	34
94		14	29
96		15	30
97		10	20
99		13	26
106		9	19
115		15	30
116		14	29
120		19	39
123		16	32
128		11	23
132		14	29
133		18	36

continued

Table B.2. (continued) Upper Little Deep Fork Creek Soil Test Phosphorous for Grassland

Sample ID. No.	Land Use Classification Subclass	Soil Test P* (mg/kg)	Soil Test P* (lbs/ac)
134	Fair Condition Grassland	15	30
139	(1 – 5% bare soil)	30	62
142		20	41
143		17	35
148		17	34
149		12	24
	mean:	15	30
2	Poor Condition Grassland	53	108
13	(5 – 20% bare soil)	23	47
18		14	29
20		7	15
23		16	33
24		20	41
25		13	26
27		12	25
28		10	21
30		12	24
33		13	26
36		14	29
37		18	36
38		17	34
43		15	31
49		13	26
50		17	34
52		11	22
59		22	45
63		13	27
64		4	8
67		16	32
69		19	39
77		11	22
86		14	29
87		9	18
88		9	19
90		12	25
92		28	58

continued

Table B.2. (continued) Upper Little Deep Fork Creek Soil Test Phosphorous for Grassland

Sample ID. No.	Land Use Classification Subclass	Soil Test P*	
		(mg/kg)	(lbs/ac)
93	Poor Condition Grassland (5 – 20% bare soil)	17	35
100		12	24
101		10	21
105		6	13
108		20	40
118		10	20
119		52	107
122		16	33
124		7	15
127		16	33
131		23	46
136		12	24
137		14	29
145		13	27
146		9	19
147		12	25
	mean:	16	32
34	Unmanaged Grassland (20 – 100% bare soil with erosive areas)	20	41
60		14	28
89		6	13
110		11	22
112		13	27
140		38	78
	mean:	17	35

* Mehlich III phosphorous

Table B.3. Upper Little Deep Fork Creek Soil Test Phosphorous for "Other" Land Uses

Sample ID. No.	Land Use Classification Subclass	Soil Test P* (mg/kg)	Soil Test P* (lbs/ac)
12	Small Grains	9	18
15		22	44
56		19	39
	mean:	17	34
8	Salt or Oilfield Induced Erosion	9	18
73		8	16
117		11	23
	mean:	9	19
107	Dairy/Feedlot	490	1000
121		60	122
	mean:	275	561

* Mehlich III phosphorous

Table B.4. Battle Branch Soil Test Phosphorous for Pasture

Rank	Soil Test P*		Rank	Soil Test P*	
	(mg/kg)	(lbs/ac)		(mg/kg)	(lbs/ac)
1	5	10	41	13	26
2	5	11	42	13	26
3	6	13	43	13	27
4	7	14	44	13	27
5	7	14	45	13	27
6	7	15	46	14	28
7	7	15	47	14	28
8	8	16	48	14	28
9	8	17	49	14	29
10	8	17	50	14	29
11	8	17	51	15	30
12	8	17	52	15	30
13	9	18	53	15	30
14	9	18	54	15	30
15	9	18	55	15	30
16	9	19	56	15	31
17	9	19	57	16	32
18	9	19	58	16	32
19	9	19	59	16	33
20	9	19	60	17	34
21	10	20	61	17	34
22	10	21	62	17	34
23	10	21	63	17	35
24	10	21	64	17	35
25	10	21	65	17	35
26	10	21	66	18	37
27	11	22	67	18	37
28	11	22	68	18	37
29	11	22	69	20	40
30	11	22	70	20	41
31	11	22	71	21	43
32	11	23	72	24	48
33	12	24	73	24	49
34	12	24	74	24	49
35	12	24	75	25	51
36	12	24	76	25	51
37	12	25	77	25	52
38	12	25	78	27	56
39	13	26	79	27	56
40	13	26	80	28	58

continued

Table B.4. (continued) Battle Branch Soil Test Phosphorous for Pasture

Rank	Soil Test P*	
	(mg/kg)	(lbs/ac)
81	29	60
82	29	60
83	30	62
84	30	62
85	32	65
86	35	71
87	36	73
88	36	74
89	36	74
90	39	79
91	39	79
92	39	80
93	39	80
94	40	81
95	40	81
96	40	82
97	42	85
98	43	88
99	43	88
100	43	88
101	44	89
102	45	92
103	46	93
104	46	93
105	47	95
106	47	96
107	48	97
108	49	100
109	49	100
110	49	100
111	49	101
112	49	101
113	51	104
114	51	104
115	51	105
116	52	107
117	54	111
118	54	111
119	54	111
120	55	112
121	55	113
122	56	114
123	56	115
124	57	116
125	58	118
126	58	118
127	60	122
128	60	122
129	60	123
130	62	126
131	62	126
132	62	126
133	62	126
134	64	130
135	64	130
136	64	131
137	64	131
138	64	131
139	64	131
140	64	131
141	65	132
142	65	132
143	65	133
144	66	134
145	66	135
146	66	135
147	67	136
148	67	137
149	68	138
150	68	139
151	68	139
152	69	140
153	69	140
154	69	140
155	69	141
156	70	142
157	70	142
158	70	143
159	70	143
160	71	144

continued

Table B.4. (continued) Battle Branch Soil Test Phosphorous for Pasture

Rank	Soil Test P*	
	(mg/kg)	(lbs/ac)
161	71	144
162	72	146
163	73	148
164	73	148
165	73	149
166	73	149
167	73	149
168	74	150
169	74	150
170	74	150
171	74	150
172	74	150
173	74	150
174	77	157
175	79	162
176	80	163
177	81	165
178	81	165
179	82	167
180	82	168
181	85	174
182	86	175
183	87	177
184	87	178
185	92	187
186	93	189
187	93	189
188	93	190
189	95	194
190	96	196
191	97	197
192	97	197
193	98	200
194	99	203
195	100	204
196	101	206
197	102	208
198	102	209
199	103	210
200	103	210
201	104	212
202	105	214
203	105	214
204	105	215
205	106	216
206	106	216
207	106	216
208	109	222
209	109	222
210	110	224
211	110	224
212	111	226
213	112	228
214	114	232
215	114	233
216	114	233
217	117	239
218	119	242
219	119	243
220	120	244
221	120	244
222	131	268
223	135	276
224	137	280
225	138	281
226	140	285
227	140	286
228	147	300
229	154	314
230	164	335
mean:	54	110

* Mehlich III phosphorous

Table B.5. Peacheater Creek Soil Test Phosphorous for Pasture

Rank	Soil Test P*	
	(mg/kg)	(lbs/ac)
1	6	13
2	7	14
3	8	17
4	9	18
5	9	18
6	9	19
7	9	19
8	10	21
9	10	21
10	11	22
11	11	23
12	12	24
13	13	26
14	13	26
15	14	29
16	14	29
17	15	30
18	15	30
19	15	31
20	15	31
21	15	31
22	16	32
23	16	33
24	16	33
25	17	34
26	17	35
27	19	38
28	20	41
29	21	42
30	21	43
31	22	45
32	23	46
33	24	48
34	24	49
35	25	50
36	25	50
37	26	54
38	26	54
39	27	55
40	27	55
41	28	57
42	28	57
43	28	58
44	29	59
45	30	61
46	30	61
47	30	62
48	31	63
49	31	63
50	31	64
51	31	64
52	31	64
53	32	65
54	32	65
55	33	67
56	33	67
57	34	70
58	36	73
59	36	74
60	36	74
61	37	76
62	37	76
63	38	77
64	38	77
65	39	79
66	39	80
67	40	81
68	41	83
69	41	83
70	41	83
71	41	84
72	41	84
73	42	85
74	42	86
75	44	90
76	45	91
77	45	91
78	45	92
79	46	94
80	46	94

continued

Table B.5. (continued) Peacheater Creek Soil Test Phosphorous for Pasture

Rank	Soil Test P*		Rank	Soil Test P*	
	(mg/kg)	(lbs/ac)		(mg/kg)	(lbs/ac)
81	47	95	121	72	146
82	47	96	122	72	147
83	48	98	123	72	147
84	51	104	124	72	147
85	51	104	125	73	148
86	52	107	126	73	149
87	55	113	127	74	150
88	58	118	128	74	150
89	58	118	129	74	150
90	59	120	130	74	151
91	59	120	131	74	151
92	59	120	132	74	151
93	60	123	133	74	151
94	61	124	134	75	154
95	61	124	135	77	157
96	61	125	136	77	158
97	62	126	137	78	160
98	62	127	138	79	162
99	62	127	139	80	163
100	63	129	140	80	163
101	64	130	141	81	166
102	64	130	142	82	167
103	65	132	143	82	168
104	65	133	144	82	168
105	65	133	145	83	170
106	66	134	146	84	171
107	66	135	147	84	172
108	66	135	148	85	174
109	67	136	149	85	174
110	67	137	150	85	174
111	67	137	151	86	175
112	68	138	152	86	175
113	68	138	153	86	176
114	68	138	154	87	177
115	68	139	155	88	179
116	68	139	156	88	179
117	69	140	157	89	181
118	70	142	158	89	181
119	71	144	159	89	182
120	71	145	160	90	183

continued

Table B.5. (continued) Peacheater Creek Soil Test Phosphorous for Pasture

Rank	Soil Test P*	
	(mg/kg)	(lbs/ac)
161	90	183
162	90	183
163	90	183
164	90	184
165	90	184
166	91	185
167	91	185
168	91	185
169	91	186
170	92	187
171	92	188
172	93	189
173	93	190
174	93	190
175	94	192
176	95	194
177	95	194
178	96	195
179	96	196
180	97	197
181	97	197
182	97	198
183	98	200
184	98	200
185	99	202
186	99	203
187	99	203
188	102	208
189	102	209
190	102	209
191	103	210
192	103	211
193	104	213
194	105	215
195	106	216
196	106	216
197	106	217
198	107	218
199	107	219
200	109	222
201	112	229
202	113	231
203	114	232
204	114	233
205	119	243
206	122	248
207	123	252
208	124	254
209	124	254
210	125	256
211	125	256
212	126	258
213	126	258
214	129	263
215	129	264
216	131	267
217	135	275
218	135	276
219	136	278
220	137	280
221	141	287
222	142	290
223	143	292
224	144	293
225	145	295
226	145	295
227	146	298
228	157	321
229	159	325
230	168	343
231	175	358
232	184	376
233	196	400
234	221	450
235	225	460
236	227	463
237	229	467
238	245	501
239	246	503
240	247	505

continued

Table B.5. (continued) Peacheater Creek Soil Test Phosphorous for Pasture

Rank	Soil Test P* (mg/kg)	Soil Test P* (lbs/ac)
241	278	567
242	306	625
243	323	660
244	330	674
245	341	696
246	350	715
247	382	780
248	396	809
249	400	816
250	401	819
251	419	855
252	427	872
253	478	976
254	490	999
255	490	999
mean:	93	190

* Mehlich III phosphorous

Table B.6. Haw Creek Soil Test Phosphorous for Pasture

Rank	Soil Test P*	
	(mg/kg)	(lbs/ac)
1	48	98
2	274	560
3	191	390
4	108	220
5	328	670
6	294	600
7	485	990
8	299	610
9	150	307
10	92	188
11	201	410
12	40	81
13	117	239
14	99	203
15	31	64
16	20	40
17	225	460
18	34	69
19	41	84
20	18	36
21	40	81
22	191	390
23	515	1050
24	86	175
25	225	460
26	196	400
27	446	910
28	76	156
29	111	226
30	105	214
31	8	17
32	225	460
33	228	465
34	21	42
35	14	29
36	181	370
37	191	390
38	42	86
39	48	98
40	191	390
41	123	250
42	134	273
43	19	38
44	48	98
45	12	24
46	7	14
47	8	16
48	128	262
49	49	99
50	134	274
51	75	154
52	16	33
53	12	25
54	74	152
55	9	19
56	7	15
57	9	19
58	2	4
59	32	65
60	25	52
61	32	66
62	67	137
63	53	108
64	40	81
65	23	47
66	17	34
67	40	81
68	34	70
69	54	111
70	196	400
71	27	56
72	51	104
73	32	65
74	11	23
75	10	20
76	142	290
77	99	202
78	74	150
79	38	78
80	181	370

continued

Table B.6. (continued) Haw Creek Soil Test Phosphorous for Pasture

Rank	Soil Test P* (mg/kg)	Soil Test P* (lbs/ac)	
81	82	168	
82	40	81	
mean:	104	212	

* Mehlich III phosphorous

Table B.7. Lake Eucha (Arkansas portion) Soil Test Phosphorous for Pasture

Rank	Soil Test P*	
	(mg/kg)	(lbs/ac)
1	5	10
2	5	10
3	5	10
4	5	10
5	5	11
6	7	15
7	12	25
8	13	27
9	15	31
10	15	31
11	16	32
12	16	32
13	17	34
14	17	34
15	17	35
16	20	40
17	20	40
18	20	40
19	20	40
20	22	44
21	23	47
22	24	49
23	25	51
24	25	52
25	27	55
26	27	55
27	27	55
28	27	56
29	28	57
30	28	58
31	31	63
32	31	63
33	31	63
34	32	65
35	32	66
36	33	67
37	33	68
38	35	72
39	36	74
40	38	77
41	38	78
42	40	81
43	41	83
44	43	88
45	44	90
46	45	91
47	45	92
48	45	92
49	47	95
50	49	101
51	51	105
52	52	106
53	53	109
54	54	110
55	54	111
56	56	114
57	56	115
58	57	117
59	59	121
60	60	122
61	61	124
62	61	124
63	61	125
64	63	128
65	63	129
66	63	129
67	64	130
68	64	130
69	65	133
70	66	134
71	66	135
72	70	142
73	70	143
74	71	145
75	72	146
76	72	147
77	73	149
78	74	150
79	75	153
80	79	161

continued

Table B.7. (continued) Lake Eucha (Arkansas portion) Soil Test Phosphorous for Pasture

Rank	Soil Test P*	
	(mg/kg)	(lbs/ac)
81	81	165
82	83	169
83	83	170
84	83	170
85	84	171
86	85	174
87	86	175
88	86	176
89	88	180
90	89	181
91	90	183
92	93	189
93	93	189
94	94	191
95	94	191
96	97	197
97	97	197
98	97	198
99	98	199
100	99	202
101	100	204
102	100	205
103	103	210
104	105	215
105	106	217
106	106	217
107	106	217
108	109	223
109	111	227
110	111	227
111	112	229
112	113	230
113	116	236
114	118	240
115	120	245
116	121	246
117	126	258
118	126	258
119	128	262
120	129	263
121	133	272
122	135	276
123	144	293
124	144	294
125	147	301
126	148	302
127	148	303
128	149	304
129	149	305
130	150	307
131	151	308
132	152	311
133	155	316
134	158	322
135	159	324
136	160	327
137	160	327
138	161	328
139	161	329
140	162	330
141	165	336
142	165	337
143	166	338
144	167	341
145	169	345
146	170	346
147	170	347
148	171	349
149	172	350
150	172	351
151	173	353
152	177	362
153	178	364
154	179	365
155	183	374
156	184	375
157	184	375
158	185	377
159	186	380
160	188	384

continued

Table B.7. (continued) Lake Eucha (Arkansas portion) Soil Test Phosphorous for Pasture

Rank	Soil Test P*	
	(mg/kg)	(lbs/ac)
161	190	388
162	191	390
163	193	394
164	193	394
165	194	395
166	195	398
167	196	400
168	196	401
169	196	401
170	198	404
171	199	406
172	200	409
173	202	412
174	204	416
175	204	417
176	205	418
177	206	420
178	206	421
179	206	421
180	208	425
181	210	428
182	212	432
183	213	434
184	216	440
185	216	440
186	216	440
187	216	441
188	220	448
189	223	456
190	224	457
191	225	459
192	232	473
193	233	475
194	235	480
195	237	483
196	237	484
197	240	490
198	241	491
199	243	496
200	245	499
201	246	503
202	250	510
203	253	516
204	255	520
205	258	527
206	259	528
207	260	530
208	262	534
209	265	540
210	271	553
211	271	554
212	274	559
213	277	566
214	278	567
215	278	568
216	279	570
217	280	572
218	282	576
219	292	595
220	294	599
221	294	600
222	295	603
223	297	606
224	299	610
225	300	613
226	303	619
227	308	628
228	308	629
229	309	630
230	319	650
231	329	672
232	330	673
233	330	674
234	332	677
235	334	681
236	335	683
237	337	688
238	338	690
239	339	691
240	342	697

continued

Table B.7. (continued) Lake Eucha (Arkansas portion) Soil Test Phosphorous for Pasture

Rank	Soil Test P* (mg/kg)	Soil Test P* (lbs/ac)
241	343	700
242	349	713
243	365	745
244	366	747
245	366	747
246	369	753
247	370	755
248	373	761
249	378	771
250	385	785
251	387	789
252	389	793
253	435	888
254	448	915
255	463	944
256	468	955
257	476	971
258	488	996
259	490	999
260	490	999
261	490	999
mean:	164	334

* Mehlich III phosphorous

VITA

William Ray Marshall

Candidate for the Degree of

Master of Science

Thesis: A NONPARAMETRIC APPROACH TO DETERMINE THE NUMBER OF OBSERVATIONS REQUIRED FOR ESTIMATING BASIN-SCALE SOIL TEST PHOSPHOROUS

Major Field: Biosystems Engineering

Biographical:

Personal Data: Born in Nebo, Kentucky, On February 23, 1969, the son of Bob and Norma Marshall.

Education: Graduated from West Hopkins High School, Rabbit Ridge, Kentucky in May 1987; received Bachelor of Science degree in Agricultural Engineering from the University of Kentucky, Lexington, Kentucky in December 1992. Completed the requirements for the Master of Science degree with a major in Biosystems Engineering at Oklahoma State University in July, 1998.

Experience: Worked on a farm during high school; employed as a carpenter during high school and undergraduate summers; CAD Draftsman, Agricultural Engineering Department, University of Kentucky, Summer 1991; Environmental Scholarship Student Intern, Kentucky Department for Environmental Protection, Summer 1992; Environmental Engineer, Kentucky Division of Waste Management, January,1993 to November, 1994; Research Engineer, Biosystems and Agricultural Engineering Department, Oklahoma State University, December, 1994 to August, 1998.

Professional Memberships: American Society of Agricultural Engineers; National Society of Professional Engineers; Alpha Epsilon; Gamma Sigma Delta; Registered Engineer Intern.