KWAME NKRUMAH UNIVERSITIY OF SCIENCE AND TECHNOLOGY

KUMASI, GHANA

COLLEGE OF SCIENCE

FACULTY OF PHYSICAL SCIENCES

DEPARTMENT OF MATHEMATICS

FACTOR ANALYSIS OF TREE DISTRIBUTION PATTERNS OF SIX FOREST RESERVES IN ASHANTI REGION

BY

NANCY BOAKYE (BSC. MATHEMATICS)

A thesis submitted to the Department of Mathematics,

Kwame Nkrumah University of Science and Technology,

In partial fulfilment of the requirements for the degree

Of

Master of Philosophy in Mathematics

July, 2010

DECLARATION

I hereby declare that this submission is my own work towards MPhil. and that, to the best of my knowledge, it contains no materials previously published by another person nor material, which has been accepted for the award of any other degree of the University except where due acknowledge has been made in the text.

Miss Nancy Boakye (PG2888108)   ....

Student name and ID   Signature   Date

Certified by: ....

Dr S.K. Amponsah   Signature   Date

Supervisor

Certified by:

Dr. S.K. Amponsah   ....

Head of Department   Signature   Date.

ABSTRACT

This study demonstrates aspects of three recognized principles, which are important in phytosociological interpretations using multivariate analysis techniques. First, when a restricted geographic area is analyzed (Ashanti region, in this case), species with different total ranges may be correlated. Secondly, when the utilized data are on a relatively course scale, microhabitat preferences are not reflected in the plant association found. Third, species may have different ecological characteristics in different portion of their range. In view of this, factor analysis was used to elucidate general distribution patterns for hundred tree species and some environmental variables for six of the selected forest reserves in the ecological zone of Ashanti Region. The three factors combined various tree species and environmental variables into groups and reflect general patterns that were meaningful in relation to the ecological characteristics of the various plants included. The general factors group together species from six forest reserves in Ashanti Region namely Onyimsu forest reserve, Bobiri forest reserve, Mirasa Hills, Prakaw forest reserve, Pra Anum forest reserve and Onuem Bepo forest reserve. Each of the three factors indicates a group of species and/or environmental variables which is superficially explained first by distribution and then by environmental variability, past history, competitive ability and habitat preference. There are a number of tree species that do not fit into these general patterns, but instead show independent patterns of distribution. The species positively associated with factor I are all characteristics of Onyimsu and Bobiri forest. Factor I loadings were influenced by considering distributions only for a relatively small area in contrast to conducting the analysis on all the reserves in Ghana. Factor III is the most clearly defined of all the

factors. The environmental variables with the highest correlations are useful descriptors in differentiating the Onuem Bepo forest from the others. There are other variables with higher loadings than rainfall variables in factor II and factor III and also a temperature variable shows a higher correlation than rainfall variable.

This does not mean that precipitation (rainfall) is not important, but suggests that it alone may have less effect than other variables. Moisture directly available to tree species is of course, governed by other environmental variables such as topographic relief and the water retaining characteristics of soils in a particular area.

CONTENTS

LIST OF FIGURES

LIST OF PLATES

ACKNOWLEDGEMENT

I would like to thank the Almighty God for his immense protection, guidance, blessing and grace He steadfastly demonstrated to me throughout this research work.

One cannot engage in a research of this magnitude without the help, support and encouragement of many other individuals, so I also wish to acknowledge the following people whose numerous efforts and helpful criticism have sustained the preparation of this research work.

I am greatly indebted to my supervisor, Dr. S. K. Amponsah, currently the Head of Mathematics Department KNUST, who through his supervisory work I have been able to complete this research successfully.

Again I thank Mr. Nana Kena Frempong, a Lecturer of Mathematics Department KNUST for his help.

Also, I thank Mr. Kofi Affum Baffoe of Forestry Commission for his help in this research work. I am thankful to my friends and love ones, the ones who helped me to type this project to a successful completion and the ones who read it.

Finally, my appreciation also goes to my parents for the support they gave me in diverse ways to complete this project.

CHAPTER 1

INTRODUCTION

Ghana is located on West Africa's Gulf of Guinea only a few degrees north of the Equator. Half of the country lies less than one hundred and fifty two (152) metres above sea level, and the highest point is eight hundred and eighty three (883) metres. The five hundred and thirty seven (537)-kilometre coastline is mostly a low, sandy shore backed by plains and scrub and intersected by several rivers and streams, most of which are navigable only by canoes. A tropical rain forest belt, broken by heavily forested hills and many streams and rivers, extends northward from the shore, near the Cote d'Ivoire frontier. This area, known as the "Ashanti," produces most of the country's cocoa, minerals, and timber. North of this belt, the country varies from ninety one (91) to three hundred ninety six (396) metres above sea level and also lies an undulating savanna country that is drained by the Black and White Volta rivers, which join to form the Volta, which then flows south to the sea through a narrow gap in the hills. Southern Ghana contains evergreen and semi deciduous forests, consisting of tall silk cottons, kolas, and valuable West African hardwoods such as mahogany, odum, ebony etc. The northern two-thirds of the country are covered by savanna (tropical grassland with a scattering of shrubs and trees), featuring Shea trees, acacias, and baobabs. The oil palm is found throughout the south and the Ashanti uplands, and the lagoons of the coast contain mangroves (Ghana Homepage, 1994-2009).

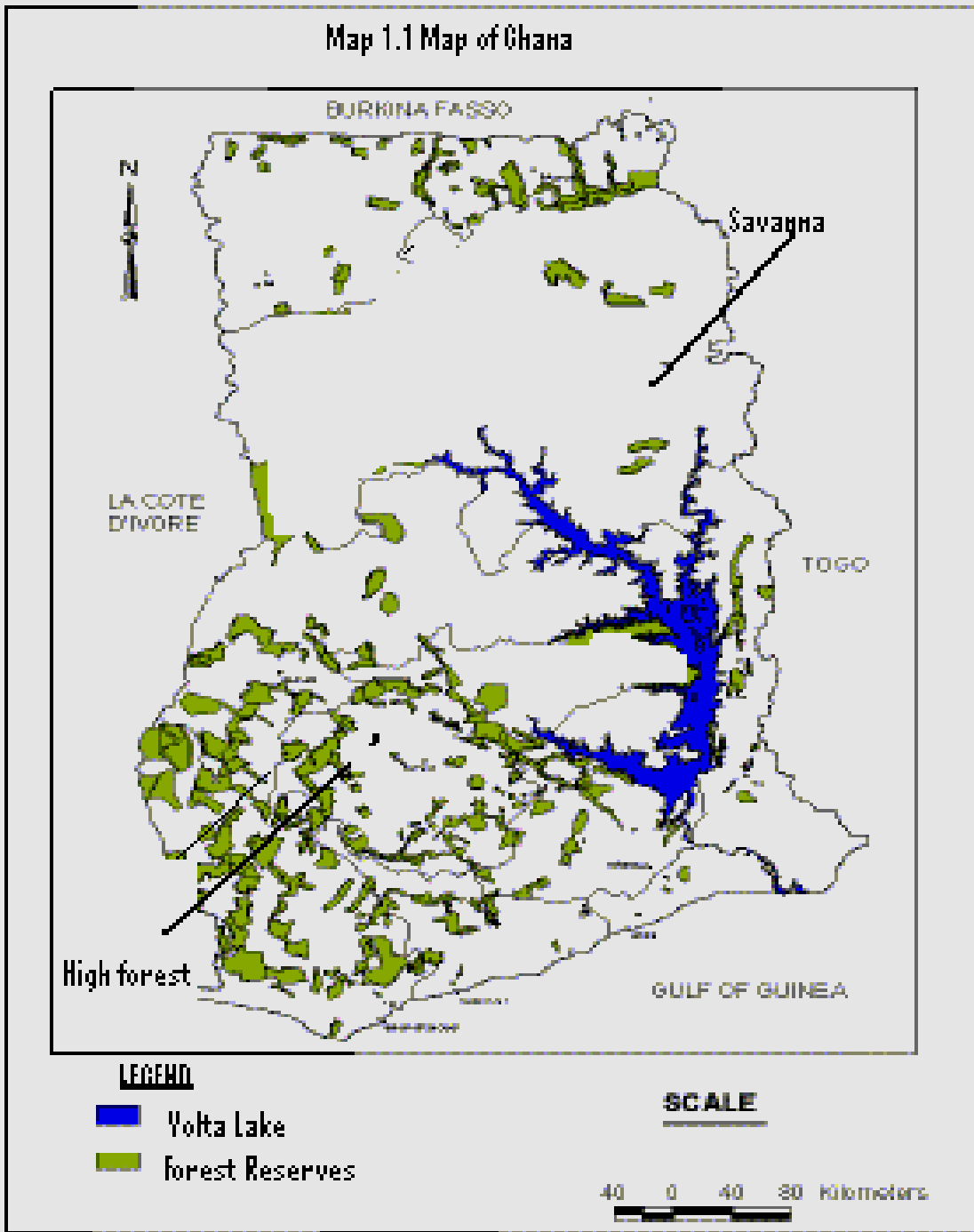**Map 1.1 Map of Ghana**

BURKINA FASSO

N

Savanna

LA COTE D'IVORE

TOGO

High forest

GULF OF GUINEA

LEGEND
- Volta Lake
- Forest Reserves

SCALE

40 0 40 80 Kilometers

Figure 1.1: The Map of Forest Reserve in High Forest Zone of Ghana

1.1 CLIMATE

1.2 TEMPERATURE

The climate of Ghana is tropical, but temperatures vary with season and elevation. The harmattan, a dry desert wind, blows from the northeast from December to March, lowering the humidity and creating hot days and cool nights in the north. In the south the effects of the harmattan are felt in January. However, in the forest zone, the average daily range is eight to nine degree Celsius (8-9) ^{0}C and the seasonal range of daily mean temperature is (3-4.5) ^{0}C. The mean monthly maximum in the hottest months is 31-34 ^{0}C, and the mean monthly minimum in the coldest months is 19-21 ^{0}C (Wills, 1962).

1.2.1 RAINFALL

The principal feature of rainfall in Ghana is its seasonal character and variability from year to year. Very considerable variations exist between successive rainy seasons in time of onset, duration and amounts received. In some seasons, individual rainfalls are numerous and well distributed and in some others, scattered and infrequent. There are two distinct rainfall patterns that divide the country into two zones; the northern (dominated by savanna) and the southern (dominated by high forest) zones.

In the north two rainy seasons occur, from April to July and from September to November. The rainy season begins in April and lasts until September. Annual rainfall ranges from about one thousand and one hundred (1,100) mm in the north to about two thousand and one hundred (2,100) mm in the southeast. The southern zone is characterized by a bi-modal rainfall pattern with peaks in May-June and September-

October followed by a short dry spell. Generally the months of December, January and February are dry as a result of a north-easterly wind that blows from the Sahara desert. Another short and mild dry period is experienced in August. The mean annual rainfall for the forest zone ranges from one thousand to two thousand one hundred and fifty (1000-2150) mm (Wills, 1962).

1.2.2 VEGETATION

The two main vegetation types that cover the surface of Ghana are the high forest which has an area of approximately 8.25 million ha and savanna with an estimated area of 15.6 million ha. (Wills, 1962). The changing seasons are reflected by the leafy and leafless phases in the life of trees. In the spring, as the weather gets warmer, new sets of leaves grow covering the tree with bright green foliage. During the summer the leaves absorb and use the energy of sunlight to produce organic food materials. Each leaf contains the substance chlorophyll, with the light energy, makes chemical substances, which are retained and can be released as energy for growth or to build up the tree. Carbon dioxide, which is present in the atmosphere, is combined with water absorbed by the roots to form energy rich sugars. Together with mineral salts, which the roots absorb, an intricate chemical process takes place and forms the trees substance. In species growing in areas with seasonal climate changes, wood growth produced at different times of the year may be visible as alternating light and dark, or soft and hard, rings of wood. In temperate climates, and tropical climates with a single wet-dry season alternation, the growth rings are annual, each pair of light and dark rings being one year of growth; these are known as annual rings. In areas with two wet and dry seasons each year, there may

be two pairs of light and dark rings each year; and in some (mainly semi-desert regions with irregular rainfall), there may be a new growth ring with each rainfall. In tropical rainforest regions, with constant year-round climate, growth is continuous and the growth rings are not visible nor is there a change in the wood texture. In species with annual rings, these rings can be counted to determine the age of the tree, and used to date cores or even wood taken from trees in the past, a practice is known as the science of dendrochronology. Very few tropical trees can be accurately aged in this manner. A tree which is a perennial woody plant is most often defined as a woody plant that has many secondary branches supported clear of the ground on a single main stem or trunk with clear apical dominance. A minimum height specification at maturity is cited by some authors, varying from three (3) m to six (6) m; some authors set a minimum of ten (10) cm trunk diameter thirty (30) cm girth. Woody plants that do not meet these definitions by having multiple stems and/or small size are called shrubs. Compared with most other plants, trees are long-lived, some reaching several thousand years old and growing to up to one hundred and fifteen (115) m (379 ft) high. The parts of a tree are the roots, trunk(s), branches, twigs and leaves. A tree is a plant form that occurs in many different orders and families of plants. Trees show a variety of growth forms, leaf type and shape, bark characteristics, and organs. The tree form has evolved separately in unrelated classes of plants, in response to similar environmental challenges, making it a classic example of parallel evolution. With an estimate of one hundred thousand (100,000) tree species, the number of tree species worldwide might total twenty five (25) percent of all living plant species. The majority of tree species grow in tropical regions of the world and many of these areas have not been surveyed yet by scientist, making

species diversity and ranges poorly understood. The earliest trees were tree ferns, horsetails and lycophytes, which grew in forests in the Carboniferous Period; tree ferns still survive, but the only surviving horsetails and lycophytes are not of tree form. Later, in the Triassic Period, conifers, ginkgos, cycads and other gymnosperms appeared, and subsequently flowering plants in the Cretaceous Period. Most species of trees today are flowering plants (Angiosperms) and conifers. A small group of trees growing together is called a grove or copse, and a landscape covered by a dense growth of trees is called a forest. Several biotopes are defined largely by the trees that inhabit them; examples are rainforest and taiga. A landscape of trees scattered or spaced across grassland (usually grazed or burned over periodically) is called a savanna. A forest of great age is called old growth forest or ancient woodland (in the UK). A young tree is called a sapling (Wikipedia Foundation, Inc., 2009).

Trees being a long-lived dominant organism are not as sensitive indicators of current environmental conditions as some other organisms. Indeed, the occurrence of a species may be more the result of a historical event than current favorable environmental conditions. The existence of an individual tree indicates a compatibility with the present integrated environment and the capability of coping with the longer-term variation in that environment. Simply realizing that tree, by their presence, indicates an integrated environment is not particularly informative in the quest for more explicit relationships between the species and various environmental factors. As Waring and Major (1964) discussed, two approaches to the analysis of factors influencing plant growth and occurrence are Operational and Correlative.

- **Operational approach**

  This investigation looks at those recognizable and measurable habitat factors, which can be shown to influence the behaviour of the plant.

- **Correlative approach**

  This investigation simply assumes a correlation between habitat factors and their direct influence on the plant.

The later approach is not helpful in determining mechanisms, but rather offers the possibility of indicating probable important environmental factors which can subsequently be tested in a rigorous manner. There have been numerous attempts to examine the correlation between environmental factors with the occurrence or behavior of individual species (example Loucks 1962, Waring and Major 1964). In most instances, the strategy is to consider a series of environmental variables and determine which of these most closely corresponds with some indicator of species behavior. While the present study includes some environmental variables, it is not intended to look specifically at the relationship between various habitat factors and individual species. Rather it is an investigation using factor analysis to examine, groups of species and to address hypotheses concerning separate groups of species and environmental variables. The specific hypothesis addressed include

(i)   In Ashanti Region, where the generalized moisture (rainfall) variables used here will show common variation with selected groups of species other than their temperature or soil variables.

(ii)   In this region where there is a relatively uniform and strong climatic gradient, each group of species represented by a factor will have contiguous geographical distribution.

Examination of the hypothesis and the ecological characteristics of these resulting species group should lead to specific experiments aimed at providing operational explanations for the geographic distributions of the species.

1.3 FOREST TYPES IN GHANA

A closer observation of the high forest of Ghana reveals distinctive characteristic features that group them into types. These groupings are a reflection of local climatic and altitudinal differences prevailing in the area which stretch beyond the boundaries of Ghana. Hall and Swaine (1981), classified the high forest according to the gradual change in forest composition, from the south west, where the rainfall is highest and the forest are evergreen, towards the savanna boundaries in the east and north, where the forest is dry and deciduous. The types are as follows:

Wet Evergreen (WE), Moist Evergreen (ME), Moist Semi-deciduous made up of South East (MSSE) and North West (MSNW) subtypes and the Dry Semi-deciduous (DS) which is also sub-divided to inner and outer (fire zone) subtypes. All these ecological zones have distinctive characteristics such as species diversity and vertical structure.

However, this study which is in Ashanti Region falls under the Moist Semi-deciduous made up of South East (MSSE).The Table1.1 below describes the general characteristics of this major forest zone.

Table 1.1: General characteristics of the Moist Semi-deciduous made up of South East (MSSE).

Vegetation Type	Forest Reserve Area(km^2)	Mean annual Rainfall (mm)	General Characteristics
Moist Semi-deciduous South East (MSSE)	2289.89	1250-1750	It is the wet part of the semi-deciduous forest. A mixture of evergreen and deciduous species. It is the most productive among the forest zones. Few tree species (about 100) but majority of commonest species in Ghana achieves their greatest frequency here. It has the most tallest trees, heights often exceeding 50-60 m. Less depletion of soil nutrients compared to WE and ME, pH=5-6. Less fire damage to the forest reserve than MSNW and DS. Mean basal area is estimated at $23.2 \text{ m}^2\text{ha}^{-1}$.

1.3.1 ECONOMIC IMPORTANCE OF TREES IN GHANA

Trees have been of great importance to mankind since prehistoric days. Sixty (60) percent of the earth was once covered with trees (forest). With the development of civilization, large areas have been cleared to make way for farms, mines, towns and roads. Today about (30) percent of earth is still forested.

The following are some of the role trees play in building the economy of Ghana:

(i)   Trees help to preserve agricultural land from the danger of erosion. Thus they help to conserve soil by preventing rapid runoff of water after heavy rain and minimizing flooding (keeping top soil from being washed away).

(ii)   Trees are the main source of timber exportation, for building (homes and shelters for people and animals, rafts, canoes and other small boats) and furniture as well as firewood and cooking charcoal.

(iii)   Pulp and paper are made from the cellulose of trees. Processed wood products include cellophane, plastics.

(iv)   Trees also provide raw material for a number of other industries like sports goods, matches, and toothpaste etc.

(v)   Many kinds of pharmaceuticals, as well as tobacco, coffee, alcohol, and other drugs, synthetic fibers like rayon, nylon, and other useful materials like gums, resin, turpentine oil, are made from the raw material that can be found in trees.

(vi) Trees also anchor the soil and prevent the winds from blowing the surface soil away.

(vii) Rivers and streams receive water that flows down from very steep mountain rocks; trees on these slopes restrict the speed which would otherwise produce floods which would have adverse effect on the economy.

(viii) Latex from trees such as the rubber tree goes to make tyres, tunes and a wide range of rubber goods.

(ix) Real estate values increase when trees beautify a property or neighborhood. Trees can increase the property value of your home by fifteen (15) percent or more.

Individual trees and shrubs have value, but the variability of species, size, condition, and function makes determining their economic value difficult. The economic benefits of trees can be both direct and indirect. Direct economic benefits are usually associated with energy costs. Air-conditioning costs are lower in a tree-shaded home. Heating costs are reduced when a home has a windbreak. Trees increase in value from the time they are planted until they mature. Trees are a wise investment of funds because landscaped homes are more valuable than nonlandscaped homes. The savings in energy costs and the increase in property value directly benefit each home owner.

The indirect economic benefits of trees are even greater. These benefits are available to the community or region. Lowered electricity bills are paid by customers when power companies are able to use less water in their cooling towers, build fewer new facilities to meet peak demands, use reduced amounts of fossil fuel in their furnaces, and use fewer measures to control air pollution. Communities also can save money if fewer facilities must be built to control storm water in the region. To the individual, these savings are small, but to the community, reductions in these expenses are often in the thousands of cedis (International Society of Arboriculture, 2005).

1.4 OBJECTIVES OF STUDY

The principal objectives are as follow:

(i)   To investigate using factor analysis to examine groups of tree species.

(ii) To address hypotheses concerning separate groups of tree species and environmental variables.

(iii) To address environmental parameters that may be limiting tree distribution.

(iv) To obtain a number of subsequent analyses (both correlative and operational) that could be done to characterize better the ecological relationships of forested and other areas in Ashanti Region.

1.5 JUSTIFICATION OF STUDY

Studies such as this one are helpful in indicating environmental parameters that may be limiting on tree (plant) distributions (or in suggesting that a group of species are not limited by particular parameters). The results of this study point out a number of

subsequent analyses that could be done to characterize better the ecological relationships of forested and other areas in Ashanti Region. Also this investigation uses factor analysis to examine groups of tree species. Examination of the hypothesis and the ecological characteristics of these resulting species group should lead to specific experiments aimed at providing operational explanations for the geographic distributions of the species.

1.7 LIMITATIONS OF THE STUDY

The limitations of the study could be attributed to factors including time, cost and current statistical data. Lack of funding for this project makes it difficult to consider all the ten regions in the country and even with one region chosen to represent the country as a whole, it was restricted to a limited area, only 6 were selected. Again for this data which is secondary, its reliability or authenticity cannot be 100% guaranteed.

1.8 STRUCTURE OF THE THESIS

This study is made up of five chapters. Chapter one contains the introduction. Literature review is in chapter two. Also, we have in chapter three, Basic concepts in Factor analysis and Factor theorems. The Data analysis of the study can be found in chapter four and last but not the least chapter five which contains the discussion, conclusion and recommendation of the study. Finally, we have the list of references, tables, figures and appendix.

CHAPTER 2

LITERATURE REVIEW

2.1: BRIEF INTRODUCTION OF THE CHAPTER

The existence of natural groupings of species is a fundamental premise for most phytosociological analyses. However, there is no one uniformly recognised method for determining or delimiting these natural associations or groups. Academically, a natural group would be one in which each member or entity showed a greater affinity for all other members than for any other non-members, and where a marked discontinuity was recognisable between groups (McIntosh, 1973). Tree distribution patterns are shaped by a variety of biotic and abiotic factors. If different plant species respond to environmental variation in a similar way, the result would be a series of coincident geographic distributions reflecting changes in the environmental variables. Also if groups of plant species respond in different ways to the same environmental variables, or if are limited by different set of variables, a series of overlapping distributional patterns should be observable in a given geographic area. Of course, finding that patterns of tree distributions parallel those of environmental parameters is not determinative proof that the species involved are limited by these parameters; however, such a demonstration would be strongly suggestive as to the factors that have marked effects on plant distributions.

Several authors' have used different techniques to elucidate relationships between distribution patterns of trees. One important research study that demonstrates this was conducted by He Fangliang, Pierre Legendre, and James V. LaFrankie. Distribution

patterns of tree species were studied in a fifty (50)-ha tropical rain forest plot in the Pasoh forest, Malaysia. This forest is characterized by a high diversity and very high number of rare species. The forest tract under study is a plot of 50 ha forming a rectangle 1 km long and 0.5 km wide. The plot lies in a mostly level plain of relatively uniform terrain between two meandering streams. About half of the plot lies within a range of two (2) m of topographic change; a hill rises in the centre of the plot to ca. 24 m above the lowest point. The survey consisted of enumerating all freestanding trees and shrubs at least 1 cm in diameter at breast height, positioning each one by geographic coordinates on a reference map and identifying it to species. Several statistics have been proposed and patterns detected have been interpreted in terms of species interactions. (Haase, 1995; Haase et al., 1996). The nearest-neighbor distance statistic is applied to detect whether a point pattern departs from an assumed random Poisson point pattern (Clark & Evans 1954). The method is biased if there is a strong edge effect. Donnelly (1978) modified the statistic after simulation experiments to account for edge effects. The value of the Donnelly statistic is transformed into a standard normal deviate z which is amenable to a z-test. In the Pasoh forest, the Donnelly statistic and the clumping index were used to determine the spatial patterns of the tree species. Seven hundred and forty-five (745) species were retained for the analysis. Any population in a community, at a given scale of observation, presents one of three distributions: aggregated, random, or regular, depending on the underlying processes. It has been found that few species in nature are distributed in a regular way; on the contrary, most of them are clumped, or appear to be randomly distributed at some given observation scale (Greig-Smith, 1983). Species in the Pasoh forest follow this general pattern. Only Galearia maingayi is

regularly distributed; aggregated patterns are dominant at the scale of the whole plot. The spatial patterns of rare species are more variable than for common species. Different factors may be at work in different types of species distributions. An aggregated pattern may occur if seeds are randomly dispersed over a heterogeneous environment, which in turn makes germination vary from site to site; or in a homogeneous environment, if individual trees grow in family groups due to vegetative reproduction, or by seeds with small dispersal distance (Feller, 1943). Hubbell (1979) showed that in a tropical dry forest, different reproductive systems do not necessarily lead to different spatial patterns, while the methods of seed dispersal result in different clumping intensities of populations. Spatially heterogeneous environmental conditions, such as topography, can also lead to patchy distributions for some species (Hubbell& Foster 1983). Of course, the factors that are responsible for the pattern of one species may not influence another one.

A much different technique that yielded similar results was conducted by William J. Wolfe. The study area straddles the divide between the Duck and Elk River drainage basins and typifies the low-relief upland topography of The Barrens (Burchett, 1977; Smalley, 1983).

Tree distribution patterns were determined at three of the sinkhole-wetland sites and across near by intermittent drainage ways at Arnold Engineering Development Center (AEDC), near Manchester, Tennessee. The data were collected to assess tree distribution patterns, flooding regime, and ground-water interaction. The geomorphically controlled water regimes of the sinkhole wetlands at Arnold Engineering Development Center (AEDC) have ecological significance. Seasonal patterns of flooding and soil saturation

16

are important controls of wetland tree distribution (Cowardin, 1979; Carter, 1986; Gill, 1970). At Arnold Engineering Development Center (AEDC) and similar settings, karst wetlands support a wide variety of trees (Benham Group, 1989; Ellis and Chester, 1989; Patterson, 1989). The water regimes of karst wetlands enable wetland trees to survive in isolated pockets far from their normal ranges. This investigation examines the relation of water regime to the distribution of local wetland tree species in karst. Wetland trees at Arnold Engineering Development Center (AEDC) can be grouped according to their normal geographic range and their inherent affinity to wetland settings. Two of these species, water tupelo and overcup oak occur nearly exclusively in wetland sites under natural conditions. The third species, willow oak, is common in The Barrens where it usually occupies wet depressions and stream bottoms. The three remaining species, sweetgum, black tupelo, and red maple, are among the most widely distributed and adaptable trees in North America. All are common in The Barrens where they occupy a wide range of wetland and upland sites.

Other local tree species that are able to exploit some wetland sites, but are not restricted to wetlands include American hornbeam, eastern hornbeam, and pawpaw. A Wilcoxon rank test (Wannacott and Wannacott, 1985) is the technique used. This study has documented relatively strong associations between tree distribution and flooding patterns. However, many of the rare or threatened species at Arnold Engineering Development Center (AEDC) are understory plants which may respond strongly to environmental factors other than surfaces flooding. Finally, the direction and environmental controls on ecological succession associated with trees distribution is unclear. The peripheral (shallow) transects through Sinking Pond and Tupelo Swamp

indicates that the tree distributions are interspersed with local species near the edges of relatively pure stands.

Also, Patterns of regional distribution of tree species were analyzed on three spatial Scales at Brazil by Samuel Jorge Leite and Frederico Santos Lopes. Data were collected in seven forest fragments dispersed over approximately seven thousand and two hundred (7200) kilometers squared; remnants of a former continuous forest in the south of Mato Grosso do Sul state, southwestern Brazil in the sub-basins of the Dourados and Guiraí rivers, within the basin of the Ivinhema, a tributary of the Paraná River. The semi-deciduous seasonal forests of Mato Grosso do Sul state are usually associated with the presence of fertile soils. Its floristic structure is very similar to that of the Brazilian Atlantic Forest formation and presents three very well differentiated strata: two of them arboreal and one herbaceous shrubby. This physiognomy is related to the seasonal climate of the region, to the rainy, hot season occurring from October to March, and to the dry and cold season extending from May to September. This seasonality is responsible for the leaf fall in the arboreal stratum, which is particularly evident for individuals in the upper storey, whose heights range from twenty (20) to thirty (30) metres. Brown's 1984 niche based model, Hanski's 1982 original core-satellite model and Tokeshi's 1992 model was the technique used. In static models, distribution and abundance of species do not vary over time, except as a consequence of major climatic or habitat changes. In contrast, dynamic models posit that both the geographic range and the local abundance of a species vary over time. Some static models, as that of Brown (1984), consider that the relationships between a species local abundance and its geographic distribution are based on its niche amplitude. Brown's model further predicts

that both the abundance and the frequency of a species decrease from the center of its geographic distribution toward all its boundaries. In addition, Brown emphasized that the distribution of species by sites is usually unimodal, and suggested that a bimodal pattern, such as that proposed by Hanski (1982), is a sample artifact derived from a small number of samples on a very local scale. Scheiner and Rey-Benayas (1997), analyzing data for terrestrial vascular plants from 74 landscapes throughout the world, concluded that niche-based models are robust on scales greater than 1 km2, while current metapopulation models are not. Levins's (1969) metapopulation model, which is dynamic, is not based on any consideration about niche, but on colonization and extinction rates instead. The metapopulation approach is directly related to the life history of organisms, their life span, fecundity, dispersal, and recruitment (Eriksson and Jakobsson 1998). Hanski's (1982) core-satellite metapopulation model, derived from Levins's model, applies to a suite of species where dispersal ability between sub-populations is high enough to allow all individuals to reach all sites in a region (Collins and Glenn 1991, 1997; Scheiner and Rey-Benayas 1997). Another assumption of Hanski's (1982) model is that the environment should be nearly homogeneous for the set of species analyzed, which is a corollary of the assumption that only colonization and extinction rates determine the spatial distribution and abundance patterns of the species. Results obtained by Scheiner and Rey- Benayas and by Collins and Glenn suggested that scale might be a key factor in contrasting Brown's (1984) and Hanski's (1982) models. Most of the studies on this issue were done with small and short life cycles organisms, and very few tests were performed using long-lived and large-sized organisms. The reasons for this paucity are obviously the great effort and time needed to obtain data for

organisms such as trees, for example. One of the few studies of this kind was that of Scheiner and Rey-Benayas (1997), on forests in Italy, South Africa, and Ghana. The Tokeshi's model is a variant of the island-biogeography model, in which the extinction rate is a decreasing function of the species frequency. The more widespread species tend to have a higher local abundance, with a positive correlation between frequency and abundance. The abundant species should a have a lower extinction rate than rare species. The result is a negative correlation of species frequency and extinction rate. This model assumes a positive correlation between frequency and abundance and predicts that species frequency distributions will either be flat or have a mode at or near one (1), depending on immigration and extinction probabilities. The Tokeshi's model predicts as much the unimodality as the bimodality (Tokeshi 1992, Colins and Glenn 1997, Scheiner and Rey- Benayas 1997). For each species, the following parameters were employed: regional frequency (expressed as the ratio between the number of fragments where the species was found and the total number of fragments); sub-regional frequency (the ratio between the number of plots possessing that species and the total number of plots of a sub-region); and local frequency (the ratio between the number of plots possessing that species and the number of plots in a fragment). The parameter abundance was considered as the mean abundance of the plots where the species occurred. Values thus obtained were tested for modality. Although no general tests for bimodality exist (Ray Benayas 1997), Tokeshi (1992) proposed a particular test that allows assessing the exact probability of some classes having more occurrences that those expected by chance. Collins and Glenn's (1997) data suggest that, for different kinds of organisms, different patterns of distribution and abundance should be expected

on different scales, and that such differences are caused by differences in the dispersal ability of the organisms. The larger the region being considered, the greater the dispersal ability required of organisms to conform to the core-satellite predictions. Plant dispersal ability, on average, is not as great as that of free-living animals. If Collins and Glenn (1997) are right, one should expect the predictions of the core-satellite model to be verified for plants on very local scales. However, their data presented unimodality on local scale for the prairie plants studied by them, as did our data for trees. Furthermore, these authors suggested that longer-living organisms are more likely to conform to the core-satellite predictions than short-living ones with the same dispersal abilities, and that this occurs because chance dispersal events in a long-living organism can be compounded over time. Trees, nonetheless, are very long-living organisms, but our data did not show any evidence that could reinforce the predicted bimodality. We notice that Tokeshi's (1992) metapopulation model is more robust than both Hanski's (1982) and Brown's (1984) models in describing patterns of frequency and abundance for tree species.

In this study, similar investigation is performed using factor analysis to elucidate general distribution patterns in Ashanti Region. Ashanti Region is ideal for this analysis since it encompasses a considerable degree of environmental variation. Because of the range of variability, our results should assist in interpreting the species distribution patterns found in surrounding regions

2.2 THE TREE AND THE ENVIRONMENT

Trees are important, valuable and necessary to our very existence. It's not too hard to believe that, without trees we humans would not exist on this beautiful planet. They are essential to life as we know it and are the ground troops on an environmental frontline. Our existing forest and the trees we plant work in tandem to make a better world. There are many ways in which trees can be valuable to us and to the environment as a whole.

(i)   First of all, trees are important because they keep the environment pleasant by absorbing moisture from the earth through their roots and spreading it in the air through their leaves thereby cleaning the air we breathe. Many people don't realize that the substance that trees are mostly made of (the carbon) comes not from the ground but from the air. Trees convert carbon dioxide CO_2 into oxygen which is essential to people, animals and plant life on the globe. Increasing green house effect is a growing menace for all sorts of life. Only trees (forests) can fight this danger thereby stabilizing the environment.

(ii)   Secondly, wild animals and other beings get enough protection from the solar heat, temperature and the leaf cover formed on the earth gives cooling effect to the earth. Since trees absorb heat, we shouldn't cut them down. Natural wildlife

is important because it is part of the natural circle of life. If bears start to die off, as it would be too hot without trees (forests) then the trout are going to become overpopulated and other scavengers that rely on bears will begin to starve. Without eagles and other birds of prey, rodents' population will increase which will cause more rodents in the cities and towns.

(iii)   Trees offer privacy that is they are used to make fences. Some tree species are planted long fence lines to support and form part of the fence as living fence posts.

(iv)   Also they offer a sound barrier. Trees muffle urban noise almost as effectively as stone walls. Trees, planted at strategic points in a neighborhood or around your house, can abate major noises from freeways and airports.

(v)   They reduce light reflection that is; trees are an important source of shade. Some species provide an almost total blockage of the sunlight while others provides a more pleasant shade.

(vi)   Trees guide wind direction and speed. During windy and cold seasons, trees located on the windward side act as windbreaks. A reduction in wind can also reduce the drying effect on soil and vegetation behind the windbreak and help keep precious topsoil in place.

(vii)   They help in keeping the environment healthy and beautiful. Some species of trees are used as ornamental trees to beautify gardens, streets, and parks. Also they provide background, soften, complement, and enhance architecture. By using trees in the cities, we are able to moderate the heat-island effect caused by pavement and buildings in commercial areas.

(viii)   Trees are historical landmarks. Thus they are important national, state and local symbols.

(ix)   By planting trees and shrubs, we return to a more natural, less artificial environment. Birds and other wildlife are attracted to the area. The natural cycles of plant growth, reproduction, and decomposition are again present, both above and below ground. Natural harmony is restored to the environment (Blurtit, 2007-2009).

2.3 USES OF TREES

Trees as a source of sustenance.

Trees are one of the major sources of sustenance: food; sugars; starches; spices and condiments;

Beverages; fumitories, masticatories and narcotics; medicines; essential oils; fatty oils and vegetable fats; waxes; soap substitutes; vegetable ivory; fodder; fuel, bioenergy or biofuel;

Fertilizers; fiber; Tannins; dyes; and cork. These are described separately under the following headings:

Food

Trees as a source of food include edible fruits, vegetables and tree legumes. Botanically, a fruit is a matured or ripened ovary, along with its contents and adhering accessory structures, if any. The seeds inside the fruits are the fertilized ovules. Sometimes seeds are formed without fertilization. This phenomenon is called "agamospermy". A fruit that

matures without seed formation is called "parthenocarpic fruit". Fruits are eaten raw. Vegetables are edible plants that store reserve food (mainly carbohydrates) in roots, stems, leaves or fruits and that are eaten either cooked or raw. Legumes or pods are the proteinaceous fruits of family Leguminosae. Some of these are edible.

Sugars

Sugar is a plant product surpassed in importance only by cereals and potatoes. It is one of the most important reserve food supplies, not only for the plant in which it is found but also because it serves as the most necessary food source of energy for humans. Sugar in plants may occur in the form of sucrose (a disaccharide of glucose and fructose that, to humans, tastes sweeter than either of its constituent monosaccharides) or cane sugar, glucose or grape sugar and fructose or fruit sugar. It occurs in roots (beets, carrots, parsnips, etc.), stems (sugarcane, maize, sorghum, and sugar maple), flowers (palms), bulbs (onion) and many fruits.

Starches

Starch, a complex carbohydrate, is a polymer of glucose units linked by alpha bonds. It exists in two forms in plants: unbranched or linear polymers called "amyloses", in which hundreds of glucose molecules form coiled molecules of starch; and branched polymers called "amylopectins", in which only forty (40)–sixty (60) glucose molecules that form branched chains do not coil. Soluble starch (starch grains soaked in hot water until they burst and form a thin, clear solution or paste) is used in the textile industry for strengthening fibers and cementing loose ends together, making the thread smoother and

easier to weave and thus giving a finish to the goods. It is used as a mordant in calico printing and a thickener or vehicle for colors. It is also used in laundry work, in toilet powders, in medicine, as a sizing agent in the paper industry, as binding material for china clay and many derivatives or products like dextrin, glucose, industrial alcohol and nitrostarch. Starch is one of the main reserve foods for green plants, which store it in thin-walled cells in the form of grains of different sizes, shapes and microscopic and physical characteristics. The chief sources of commercial starch are maize, potato, wheat, rice, sago, cassava and arrowroot, of which the last two are obtained from shrubs and sago is obtained from trees. Arrowroot starch is obtained from the tubers of many tropical plants. Sago starch is obtained from the starchy pith of the stems of Metroxylon sagu. Starchy pith is removed after the trees are cut, and, after washing, the starch is freed by sedimentation. Dried, it is known as "sago flour", it is made into a flour and then dried in the sun or in ovens to obtain shiny, granular starch, called "pearl sago." Both are used almost entirely for food purposes, like macaroni and spaghetti.

Spices and Condiments

Spices and condiments are flavoring agents obtained from plants. They are difficult to distinguish, so the terms are used interchangeably. Because they have little nutritive value, they are not classified as foods. They contain essential oils, which impart flavor and aroma to food and add greatly to the pleasure of eating. They stimulate the appetite and increase the flow of gastric juices. For these reasons they are often referred to as "food accessories" or "adjuncts."

Nonalcoholic Beverages

Beverage plants are those plants which yield beverages or drinks nonalcoholic or alcoholic that are palatable and refreshing. Nonalcoholic beverages usually contain caffeine, an alkaloid, which has stimulating and refreshing qualities. Alcoholic beverages are those that contain one or more hydroxyl $-OH$ groups; example, ethanol $CH_3 - CH_2 - OH$. They may be fermented or distilled. Fruit juices and other beverages that contain neither caffeine nor alcohol are called "soft drinks". They have high sugar content and thus are a good source of energy.

Fumitories, Masticatories and Narcotics

Some narcotic substances are smoked or chewed by humans for pleasure or to seek a "world full of new sensation or some flight from reality." Narcotic substances that are used for smoking purposes are called "fumitories," and those that are used for chewing purposes are called "masticatories." They have a distinct stimulating or even narcotic effect due to the presence of various alkaloids. They are also used in religious ceremonies.

Essential Oils

Like all other necessities of humans, oils are one of the main necessities of daily life. Oils are of two types: essential, volatile or distilled oils; and fatty, nonvolatile, expressed or fixed oils. Essential oils are by-products of carbohydrate and fat metabolism. They occur in small concentrations, from minute traces to as much as (1–2) percent, or even more, in specialized cells, glands or ducts, either in one particular organ of the plant or

distributed over many parts. They may be present in flowers (e.g., roses), fruits (e.g., oranges), leaves (e.g., eucalyptus), bark (e.g., cinnamomum), roots (e.g., ginger), woods (e.g., cedar) or seeds (e.g., cardamon) and many resinous exudations. The utility of essential oils to the plant itself is obscure. The characteristic aroma and flavor they impart to flowers, fruits and seeds probably attract insects and other animals, which play an important role in pollination and/or in the dispersal of fruits and seeds. When essential oils are present in high concentrations, the unpleasant odor may serve to repel enemies like parasites, animals and insects. The essential oils may have antiseptic and bactericidal properties, and may thus act as a wound fluid. They affect transpiration and other physiological processes by minimizing the effect of heat on transpiration. They play a vital role as hydrogen donors in oxide-reduction reactions as potential sources of energy. Because of their odor and high volatility, essential oils are also put to a variety of uses by humans. They are extensively used in the manufacture of perfumes, sachets, soaps and other toilet preparations. The perfumes are stored in closed, compactly filled containers since they deteriorate due to oxidation and polymerization when they come into contact with air. In confectionary and aerated waters they are used as flavoring materials or essences for ice creams, candies, cordials, liqueurs, nonalcoholic beverages, tobacco, etc. They are very valuable in medicine, dentistry and pharmaceuticals because of their therapeutic, antiseptic and bactericidal properties. They are used as insecticides and deodorants, as solvents in paint and varnish industries and in the manufacture of several synthetic odors and flavors, such as attars and scents. Some of the essential oils (e.g., clove oil) are used as clearing or cleaning agents in histological work. They are

also used in such diversified products as chewing gum, toothpaste, incense, shoe polish, library paste and fish glue.

Medicines

Several trees are a source of important drugs. These are obtained from the bark of Bauhinia variegata (kachnar), Barringtonia acutangula (hijjal), Cinnamomum zeylanicum (dalchini),C. calisaya, C. ledgerina, C. officinalis, C. robusta, C. succirubra (all yielding quinine), Mimusops elengi (maulsari), Myrica nagi (kaiphal), Symplocos racemosa (lodh), Saraca indica (ashok), Terminalia arjuna (arjun) and Toddalia asiatica (kanj). The stems and wood of Acacia catechu (katha), Pinus roxburghii (chir) and Santalum album (safed chandan) yield drugs. Drugs are also obtained from the fruit of Aegle marmelos (bael), Cassia fistula (amaltas), Emblica officinalis (amla), Terminalia bellerica (bahera) and T. chebula (harar) . The seeds of Croton tiglium (jamalgota), Pongamia pinnata (karanja), Ricinus communis (arand) and Strychnos nux-vomica (kuchla) are also used for obtaining drugs.

Fatty Oils and Vegetable Fats

Vegetable fatty oils are called "fixed oils" or "nonvolatile oils" because they do not evaporate or become volatile like the essential oils. They are also called "nondistilled oils" because they cannot be distilled without being decomposed. Chemically, fatty oils consist of glycerin in combination with a fatty acid. The so-called fats or tallows are solids at ordinary temperatures and contain stearic or palmitic acid. Their iodine number (the number of grams of iodine absorbed by one hundred (100) g of the fats in a medium

in which it is soluble) is below seventy (70). On the other hand, oils are liquids at ordinary temperatures and contain oleic acid. Oils are of three types: drying, semidrying and nondrying. The drying oils are able to absorb oxygen and, on exposure, dry into thin elastic film. They are used mainly in the paint and varnish industry. Their iodine number is higher than one hundred and fifty (150). The semidrying oils absorb oxygen slowly and only in limited amounts. They form a soft film only after long exposure. Their iodine number is between one hundred (100) and one hundred and fifty (150). The nondrying oils remain liquid at ordinary temperatures and do not form a film. Their iodine number is between seventy (70) and one hundred (100). The fatty oils are insoluble in water but soluble in various organic solvents. When a fat is boiled with an alkali, it decomposes, and the fatty acid unites with the alkali to form soap. If soda is used, a hard soap is obtained; and if potash or lye is used, a soft soap is obtained. When fats break down, they yield fatty acids and glycerin, of which they are composed, and usually develop a rancid odor and taste. The fatty oils are bland (balmy) and lack the strong taste, odor and antiseptic qualities of essential oils. Thus they are available as indispensable articles in human food.

Waxes

Waxes are quite similar to fats but are esters of monohydric alcohols rather than glycerides.

They are harder than fats and have a high melting point. They are less easily hydrolyzed and do not become rancid. Waxes are usually found on the epidermis of leaves and fruits. They serve to prevent excessive loss of water through transpiration, because of

their impervious character. Wax is also obtained from the leaves of the raffia and licuri palms, sugarcane and esparto.

Soap Substitutes

Saponins are a group of water-soluble glucosides that yield soap froth in water, form emulsions with oils and fats, and are capable of absorbing large amounts of gases such as carbon dioxide. Because of these properties they are used for cleansing and other purposes, both at home and in industry. It may be added here that leaves of a familiar garden plant, bouncing bet or soapwort

(Saponaria officinalis, family Caryophyllaceae), when placed in water, produce a lather that is utilized for washing and imparting luster to silk and woolen fabrics. Similarly, bulbs of the

Californian soaproot (Chlorogalum pomeridianum, family Liliaceae) yield a good lather, which is utilized for washing fabrics.

Vegetable Ivory

The seeds of Phytelephas macrocarpa, in the family Arecaceae, commonly called "ivory nut" or "tagua palm tree," is the chief source of vegetable ivory. It is extensively used as a substitute for true ivory. It can be carved and used in the manufacture of buttons, chess pieces, poker chips, dice, knobs, inlays, billiard balls, toys, etc. Metroxylon amicarum, in the Arecaceae family, can likewise be used for these purposes.

Fodder

The leaves of trees and shrubs are rich in calcium and phosphorus. Although considered inferior to grasses, trees in different parts of India are lopped for fodder, especially when grasses are scarce. The important fodder-yielding trees are Acacia nilotica (A. arabica), A. catechu, Acer spp., Aegle marmelos, Bauhinia variegata, Celtis australis, Dendrocalamus strictus, Ficus glomerata, F. religiosa, Grewia spp., Helicteres isora, Kydia calycina, Leucaena leucocephala, Melia azedarach, Millettia auriculata, Morus australis, M. serrata, Ougeinia oojeinsis, Populus

ciliata, Quercus glauca, Q. incana, Zizyphus mauritiana and Z. nummularia (Singh, 1982; Anonymous,1983).

Fuel, Bioenergy or Biofuel

Bioenergy is the energy available from biological sources, both living and immediate remains.

Fuel is any material that burns readily in air. Biofuel are materials of biological origin that are used for producing heat and other forms of energy. Fuel is a great necessity of modern life. Wood, peat and coal, which represent three stages in the carbonization of the original woody plant tissue, are important fuel substances. Because their moisture content is lower than that of green wood, seasoned or oven-dried wood makes excellent fuel: ninety (99) percent of it is combustible, so it leaves only a small amount of ash. Hardwoods, such as ash, beech, hickory, maple and oak, which burn for a longer time and provide more uniform heat than does gymnospermic wood, are excellent fuel woods. The mean calorific value of oven-dried Indian hardwoods is about nine thousand (9000) btu. The qualities needed for fuel wood are physical properties of the wood as

well as environmental and silvicultural properties of the species. Small-diameter, thorn less shrubs and trees, which are easy to cut with primitive tools and easy to transport, are generally preferred. Likewise, fuel wood that is easy to split and either has low moisture content or dries rapidly is preferred over other wood, because considerable heat is lost in burning moist wood. Such wood is also nontoxic and produces less smoke.

Fertilizers

Several species of nitrogen-fixing bacteria of Rhizobium, including R. leguminosarum, R. lupini, R. meliloti and R. phaseoli, live inside the root nodules of leguminous trees. Similarly, Frankia, a nitrogen-fixing mycelial bacterium, is associated symbiotically with the root nodules of several nonlegume plants, including Alnus, Casuarina, Coriaria, Myrica and Rubus. Both Rhizobium and Frankia are capable of fixing atmospheric nitrogen. When the roots of these plants decay, they enrich the soil with nitrogen salts.

Fibers

Botanically, a fiber is a special type of cell (sclerenchymatous) that has thick walls, a narrow lumen and tapering ends. Chemically, it is made up of cellulose and lignin. Commercially, a plant fiber is a strand consisting of one or hundreds of cells that varies in length from a fraction of a millimeter to two (2) meters or more. Depending on how fibers are used, they can be classed as textile fibers (for fabrics, netting, and cordage), brush fibers, plaiting and rough weaving fibers (for hats, sandals, baskets, chairs, etc.), filling fibers, natural fabrics and papermaking fibers. The important fiber-yielding woody plants, including trees and shrubs, are Abroma angusta, Abutilon spp., Acacia

leucocephala, Ananas comosus, Antiaris toxicaria, Boehmeria nivea, Borassus flabellifer, Butea monosperma, Caryota urens (leaves), Cordia dichotoma, C. rothii, Ficus bengalensis, F. cunia, F. religiosa, Grewia glabra, G. elastica, G. optiva, G. tiliaefolia, G. vestita, Hardwickia binata, Hibiscus spp., Malachra capitata, Marsdenia volubilis, Pandanus spp. (leaves), Sterculia foetida, S. urens, S. villosa, Trema orientalis and Urena lobata. Most fibers are obtained from the bark of these plants. Silky flosses produced in the fruits of Bombax ceiba, Ceiba pentandra and Cochlospermum religiosum are also used as fibers for filling purposes. The well-known coir fiber is obtained from the fibrous mesocarp of the coconut palm, Cocos nucifera. It is coarse, stiff, buoyant and elastic and is therefore used for ship ropes, mats, brushes, ropes, etc (Watt, 1889–1893; Anonymous, 1983; Maithani et al., 1991).

Tannins

Tannins are soluble, astringent, bitter and complex phenolic substances of plant origin. These are glycosidal in nature and acidic in reactions. They may be hydrolysable or condensed in nature. Whereas hydrolysable tannins are easily split into alcohols and acids by water, condensed tannins are not, for they are made up of polymers of cyclic compounds. Tannins may be present in individual cells or in special containers known as "tannin sacs." In individual cells, tannins are found in the cell sap or are impregnated in the cell's walls, often accumulating in large quantities in dead tissues such as cork or present in bark, wood, leaves, roots, fruits and galls. The biological functions of tannins are not very clear. It is thought that tannins protect the protoplast against desiccation, decay and injury by animals. It may be concerned with the formation of cork or with

protection of the plant. They have the ability to unite with certain types of proteins, such as those in animal skins (hides), to form a strong, flexible, resistant and insoluble substance known as "leather." The process and art of converting raw hides and skins of animals into leather, usually through the use of certain chemicals, is called "tanning." Tannins react with salts of iron to form dark blue, blue-black or greenish black compounds, which are the basis of tannin or writing inks. Tannins are also useful in medicine, because of their astringent nature. Tanning materials are often utilized in oil drilling to reduce the viscosity of the drill without reducing the specific gravity

Dyes

Dyes are colored compounds (pigments) that are capable of being fixed to fabrics permanently; i.e., they neither fade on exposure to light nor wash out with soap. Therefore, a colored organic substance is not necessarily a dye. For example, trinitrotoluene, which is yellow in color, cannot fix to a cloth and therefore is not a dye. On the other hand, picric acid, which is also yellow in color, can fix to a cloth and therefore is a dye. A large number of plants secrete or contain pigments, but only about one hundred and fifty (150) are commercially important. In addition, synthetic or aniline dyes are now obtained from coal-tar products. These are cheaper, brighter, more permanent and easier to use, and they offer a wider range of colors. Among the chief uses of dyes is in coloring fabrics in the textile industry, where they are used with weak salt solutions of various metals like iron, chromium, aluminum or tin. A fine layer of metallic oxide, which forms an insoluble compound with the dye, is deposited on the

cloth. Such salts of metals that increase the adherence of various dyes to the fabrics are called "mordants." These actually form

a chemical bridge between the fiber molecules and the dye. Dyes are also used for coloring paints, varnishes, leather, ink, paper, wood, furs, food, cosmetics and medicines.

Cork

Commercial cork is obtained from the outer bark (phellem) of cork oak, Quercus suber, an evergreen tree of the family Fagaceae. It is native to the western Mediterranean region: about seventy (70) percent of the world's commercial cork comes from Portugal alone. Cork is nothing more than thin-walled but strong cellulosic cell walls, which are heavily coated with suberin, a substance that is impervious to water. Cell lumens, which represent nearly fifty three (53) percent of the total cork volume, are filled with air, thus making cork very light. Cork is buoyant, light and highly compressible, but it is resilient, chemically inert to moisture and common liquids, resistant to deterioration, an excellent insulator, a nonconductor of electricity, and a low thermal conductor and impervious to water and other liquids. It imparts no flavor or odor to substances, is slow to catch fire, absorbs sound and vibrations and has a high coefficient of friction. All of these properties render commercial cork invaluable in the world market, and it is used either as natural cork or as composition cork, the latter as linoleum, linotiles, binder-coated cork and cork (insulation) boards. Cork is used in the preparation of stoppers, hats and helmets, tips for cigarettes, carburetor floats, fishing-net floats, golf-club handles, penholders, fishing rods, life preservers, floats and life jackets, surf balls, seals for jars,

sealing liners, shoe insoles, sporting goods, picture frames, small cork balls in referees'

whistles, etc (Seth, 2004).

CHAPTER 3

METHODOLOGY

The study involved the use of factor analysis which is most familiar to researchers as an exploratory tool for unearthing the basic empirical concepts in a field of investigation. Representing patterns of relationship between phenomena, these basic concepts may corroborate the reality of prevailing concepts or may be as new and strange as to defy immediate labeling. Factor analysis is often used to discover such concepts reflecting unsuspected influences at work in a domain. The delineation of these interrelated phenomena enables generalizations to be made and hypotheses posed about the underlying influences bringing about the relationships. In other words, you may want to break a large set of variables down into smaller sets of related data. Factor analysis is the statistic used to determine if any of the independent variables comprise common underlying dimensions called "factors." Through factor analysis, you can find variables that are correlated with each other, but relatively independent of other sets of data. Therefore, factor analysis is one of the most popular data reduction tools. Now, coming to its technicalities, this analysis uses correlations among the items and items that show the highest correlation among each other is generally grouped together. Note that a single original variable can have more than one factor explaining that variable and it is called communality. The portions of the different original variables that cannot be explained by any factor at all are the independent terms and have to be considered as they are. These are demonstrated in the subsequent chapters. The statistical Product for Social Scientists (SPSS) was used for the computation of the analysis.

3.1 DEFINITION OF FACTOR ANALYSIS

Factor analysis is a statistical technique used to reduce a set of variables to a smaller number of variables or factors. Factor analysis examines the pattern of intercorrelations between the variables, and determines whether there are subsets of variables (or factors) that correlate highly with each other but that show low correlations with other subsets (or factors). The statistical approach involves finding a way of condensing the information contained in a number of original variables into a smaller set of dimensions (factors) with a minimum loss of information.

Cureton and D'Agostino (1983) described factor analysis as "a collection of procedures for analyzing the relations among a set of random variables observed or counted or measured for each individual of a group". The purpose, they said, "is to account for the intercorrelations among n variables, by postulating a set of common factors, considerably fewer in number than the number, n, of these variables". Bryman and Cramer (1990) broadly defined factor analysis as "a number of related statistical techniques which help us to determine them (the characteristics which go together)".

Gorsuch (1983) reminded the reader that "all scientists are united in a common goal: they seek to summarize data so that the empirical relationships can be grasped by the human mind". The purpose of factor analysis, he said, "is to summarize the interrelationships among the variables in a concise but accurate manner as an aid in conceptualization". These definitions most likely make a great deal of sense to those "left-brained" individuals who understand complex things fairly easily. Kerlinger (1979) gave both a left-brained and a right-brained definition of factor analysis. For the left-

brainers: "Factor analysis is an analytic method for determining the number and nature of the variables that underlie larger numbers of variables or measures". And for the right-brainers he noted: "It (factor analysis) tells the researcher, in effect, what tests or measures belong together, which ones virtually measure the same thing, in other words, and how much they do so". He further commented on factor analysis in terms of curiosity and parsimony. He noted, "Scientists are curious. They want to know what's there and why. They want to know what is behind things. And they want to do this in as parsimonious a fashion as possible. They do not want an elaborate explanation when it is not needed." He sounds like a very right-brained individual!

Each definition of factor analysis has common elements. Each refers in some way to the correlations among variables as reflected by the use of the words interrelationships, intercorrelations and relations. Further, each definition makes clear the notion of reducing the number of variables into a smaller set of factors. In short, factor analysis helps to explain things by reducing large amounts of information into a manageable form and size.

3.2 A BRIEF HISTORY OF FACTOR ANALYSIS

Factor analysis was invented nearly hundred (100) years ago by Charles Spearman, who hypothesized that the enormous variety of tests of mental ability-measures of mathematical skill, vocabulary, other verbal skills, artistic skills, logical reasoning ability, etc., could all be explained by one underlying "factor" of general intelligence that he called g. He hypothesized that if g could be measured and you could select a subpopulation of people with the same score on g, in that subpopulation you would find

no correlations among any tests of mental ability. In other words, he hypothesized that *g* was the only factor common to all those measures. It was an interesting idea, but it turned out to be wrong. Today the College Board testing service operates a system based on the idea that there are at least three important factors of mental ability; verbal, mathematical, and logical abilities and most psychologists agree that many other factors could be identified as well. If a statistical method can have an embarrassing history, factor analysis is that method. Around (1950) the reputation of factor analysis suffered from over promotion by a few over enthusiastic partisans. In retrospect there were three things wrong with the way some people were thinking about factor analysis at that time.

First, some people seemed to see factor analysis as the statistical method rather than a statistical method. Second, they were thinking in absolute terms about problems for which a heuristic approach would have been more appropriate.

Third, they were thinking of overly broad sets of variables "we want to understand all of human personality" rather than "we want to understand the nature of curiosity". Thus in three different ways, they were attempting to stretch factor analysis farther than it was capable of going. In recent decades factor analysis seems to have found its rightful place as a family of methods which is useful for certain purposes.

3.3 FUNDAMENTAL FACTOR THEOREM

The simplest mathematical model for describing a variable in terms of several others is a linear one. Several types of factors may be distinguished as common factors and unique factors. The common factors are necessary to account for the inter-correlations among the variables, while each unique factor represents that portion of a variable not ascribable to its correlations with other variables in the set. Factor analysis can be considered as an approximation to the observed covariance matrix **S** and whether the data are consistent with the prescribed structure.

We employ the notation $F_1, F_2, ..., F_m$ for m common factors and $U_1, U_2, ..., U_n$ for n unique factors, for convenience, all factors are in standardized form i.e. each factor has mean zero and unit variance. Let $x_1, x_2, ..., x_n$ represent the observed variables and z_{ji} represent the standardized value of variable j for species i, i.e. $z_j = x_{ji} / \sigma_j$. The set of all values z_{ji} (i = 1, 2,...,N) is the statistical variable z_{ji} in standard form. The complete linear expression for any variable z_{ji} (j =1, 2,…,N) may be presented in the form :

$$z_{ji} = a_{j1}F_{1i} + a_{j2}F_{2i} + \cdots + a_{jm}F_{mi} + a_j U_{ji} \qquad . \quad . \quad . \ E_1$$

In equation E_1, it is assumed that there is a value for each of the N species of the sample. The basic problem is to determine the coefficients $a_{j1}, a_{j2}, ..., a_{jm}$. The factor model can be expressed in matrix notation. In this chapter some of the fundamental properties of factor patterns and structures are represented in matrix form.

N values for each of the n variables will be represented by matrices as

$$Z_{(n \times N)}, \; F_{((m+n) \times N)}, \; M_{(n \times (m+n))} \text{ and } \; S_{(n \times (m+n))}$$

demonstrated below:

$$Z = \begin{pmatrix} z_{11} & z_{12} & z_{13} & \cdots & z_{1N} \\ z_{21} & z_{22} & z_{23} & \cdots & z_{2N} \\ . & . & . & \cdots & \cdots \\ z_{n1} & z_{n2} & z_{n3} & \cdots & z_{nN} \end{pmatrix} \quad ; \; F = \begin{pmatrix} F_{11} & F_{12} & \cdots & F_{1N} \\ . & . & \cdots & \cdots \\ F_{m1} & F_{m\hat{2}} & \cdots & F_{mN} \\ U_{11} & U_{12} & \cdots & U_{1N} \\ . & . & \cdots & \cdots \\ U_{n1} & U_{n2} & \cdots & U_{nN} \end{pmatrix} = \{ f \mid U \} \qquad \cdots \quad E_2$$

Where Z represent the complete matrix of values and F represent the complete set of N

values for each the $(m+n)$ factors. The coefficients of the factors in a factor pattern are

represented by the matrix M below.

$$M = \{ A \mid a \} \quad \cdots \quad E_3$$

Where the total pattern matrix M is made up of a matrix $A_{(n \times m)}$ of common factor

coefficients and the diagonal matrix **a** of unique-factor coefficients on the diagonals.

Matrix A is referred to as the pattern matrix. In addition to a pattern, a factor analysis

also yields a structure i.e. the correlations of the variables with the factors. The complete

factor structure may be represented by:

$$S = \{ s \mid a \} \quad \cdots \quad E_4$$

If the correlations with the common factors are defined by:

$$s_{jp} = r_{z_j F_p} \quad , \quad \left(j = 1, 2, \ldots, m \right)$$

Usually, the n x m matrix **s** of correlations of variables with the common factor is referred to as the factor structure. We now formulate certain relationship between the quantities defined above. The factor pattern equations for the complete and common factor portion is given by **Z=MF** and **Z=Af** respectively. Now,

$$Z\dot{F} / N = M \left(F\dot{F} / N \right) = S \quad \cdots \quad E_5$$

The expression in E_8 is also a matrix of correlation coefficient among the factors since all factors are in standard form. This matrix may be represented by:

$$\varphi = F\dot{F} / N = \begin{bmatrix} \phi & 0 \\ 0 & 1 \end{bmatrix} \quad \cdots \quad E_6$$

In E_6 where the correlations among the unique factors constitute an identity matrix of order n and the $m \times m$ matrix of correlations among the common factors is defined by

$$\Phi = \begin{bmatrix} 1 & r_{F_1 F_2} & \cdots & r_{F_1 F_m} \\ r_{F_2 F_1} & 1 & & \\ . & . & \cdots & . \\ r_{F_m F_1} & r_{F_m F_2} & \cdots & 1 \end{bmatrix} \quad \cdots \quad E_7$$

Where Φ is symmetric, upon substituting E_7 into E_6, the expression in E_5 reduces to:

$$S = M\Phi \quad \cdots \quad E_8$$

Equation E_8 is the fundamental relationship between a factor pattern **M** and a factor

structure **S**, the structure matrix is equal to the pattern matrix post-multiplied by the matrix of correlations among factors. The explicit expression for the pattern matrix can be obtained from E_8 by post-multiplying both sides by $\Phi^{\hat{I}^1}$, i.e. $M=S\Phi^{\hat{I}1}$

By definition, the matrix of observed correlations is given by:

$$R=Z\overline{Z} \, / \, N \quad \cdots \quad E_9$$

When dealing with the common-factors only, the matrix **M** is replaced by **A** then the matrix $R^+$of reproduced correlations with communalities in the diagonal becomes:

$$R^+ = A\Phi\overline{A} \quad \cdots \quad E_{10}$$

Equation E_{10} involves the factor pattern and the correlations among the common factors. If one assumes that the factors are uncorrelated, the matrix Φ reduces to an identity matrix and expression E_{10} simplifies to the following expression for the reproduced correlations:

$$R^+ = A\overline{A} \quad \cdots \quad E_{11}$$

This equation is known as the fundamental factor theorem. The factor problem is concern with fitting a set of data (the observed correlations) to model the factor pattern and structure. Under the assumption of such pattern, the correlations are reproduced by means of common factor coefficients alone. For the reproduced correlation matrix R^+ to be an appropriate fit to the observed correlation matrix **R,** the diagonal elements must also be reproduced from the common factor portion of the pattern.

3.4 CLASSIFICATION OF FACTOR ANALYSIS

There are basically two types of factor analysis namely, exploratory and confirmatory. Next to exploratory factor analysis, confirmatory factor analysis exists. This study is about exploratory factor analysis, and will henceforth simply be named factor analysis.

3.4.1 EXPLORATORY FACTOR ANALYSIS (EFA)

For this study, exploratory factor analysis is used, as this analysis technique allows for the extraction of as many significant factors as possible from the data set. Thus as an initial analysis, to identify the number of constructs and the underlying factor structure, exploratory factor analysis is considered. By performing exploratory factor analysis (EFA), the number of constructs and the underlying factor structure are identified.

- **The Orthogonal Factor Model With m common factors**

The observable random vector **X** of interval measurements, with n components, has mean μ and covariance matrix Σ. From the previous section, the factor model postulates that **X** is linearly dependant upon a few unobservable random variables $F_1, F_2, ..., F_m$ called common factors and n additional sources of variations $\varepsilon_1, \varepsilon_2, ..., \varepsilon_n$. The orthogonal model is presented below:

$$X = \mu + AF + \varepsilon \quad \cdots \quad E_{12}$$

Where the dimensions are $X_{(nx1)}, \mu_{(nx1)}, A_{(nxm)}, F_{(mx1)}$ and $\varepsilon_{(nx1)}$. μ is the mean vector of the n variables, ε is the specific vector of the n variables, a_{ij} is the loading of the *ith* variable on the *jth* factor. Additional assumptions about the random vectors $\mathbf{F}$ and ε are made. $\mathbf{F}$ and ε are independent, so

$$\operatorname{cov}(\varepsilon, F) = E(\varepsilon F^{'}) = 0_{(nxm)}, E(F) = 0, \operatorname{cov}(F) = 1, E(\varepsilon) = 0, \operatorname{cov}(\varepsilon) = \Psi$$

Where Ψ is a diagonal matrix. The orthogonal factor model implies a covariance structure for X which is derived from E_{12} as:

$$(X-\mu)(X-\mu)^{'} = (AF+\varepsilon)(AF+\varepsilon)^{'} = AF(AF)^{'} + \varepsilon(AF)^{'} + AF\varepsilon^{'} + \varepsilon\varepsilon^{'} \quad \cdots \quad E_{13}$$

From E_{13} it can be deduced that:

$$\Sigma = \operatorname{cov}(x) = E(X-\mu)(X-\mu)^{'} = AE(FF^{'})A^{'} + E(\varepsilon F^{'})A^{'} + AE(F\varepsilon^{'}) + E(\varepsilon\varepsilon^{'})$$

$$= AA^{'} + \Psi$$

$$Var(X_i) = a_{i1}^2 + \; \ldots \; + a_{im}^2 + \Psi_i$$

Or

$$Cov(X_i, X_k) = a_{i1}a_{k1} + \; \ldots \; + a_{im}a_{km} \quad \cdots \quad E_{14}$$

The portion of the variance in E_{14} of the *ith* variable contributed by the *m* common factors is called the *ith communality*. The other portion Ψ_i due to specific factor is called the uniqueness. The covariance structure can be derived as;

$$Cov(X, F) = E(X-\mu)F^{'} = AE(FF^{'}) + E(F\varepsilon^{'}) = A.$$

The factor model assumes that the n(n+1)/2 variances and covariances for **X** can be reproduced from the *nm* factor a_{ij} loadings and the *n* specific variances ψ_i.

The sample covariance matrix $\mathbf{S}$ is an estimator of the unknown population variance matrix Σ. If the off-diagonal elements of $\mathbf{S}$ are small or those of the sample correlation matrix $\mathbf{R}$ are essentially zero, a factor analysis will not prove useful. Two of the most popular methods of parameter estimation in factor analysis are the *principal component* (and related *principal factor)* method and the *maximum likelihood method.* Both methods were considered in the statistical analysis.

- **Principal component (and related principal factor)**

The covariance matrix Σ can be decomposed spectrally to obtain the square root matrix. This spectral decomposition allows us to express the inverse of a square matrix in terms of its eigenvalues and eigenvectors. Let Σ have eigenvalue-eigenvector pairs (λ_i, e_i) with $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_n \geq 0$ and satisfying a positive definite, then

$$\Sigma = \lambda_1 e_1 e_1' + \lambda_2 e_2 e_2' + ... + \lambda_n e_n e_n'$$

$$= \left[\sqrt{\lambda_1} e_1 \mid \sqrt{\lambda_2} e_2 \mid ... \sqrt{\lambda_n} e_n \right] \begin{bmatrix} \sqrt{\lambda_1} e_1' \\ \sqrt{\lambda_2} e_2' \\ \vdots \\ \sqrt{\lambda_n} e_n' \end{bmatrix} \quad \cdots \quad E_{15}$$

Expression E_{15} fits the prescribed covariance structure for the factor analysis model having as many factors as variables *(m=n)* and specific variance $\psi_i = 0$ for all *i*. The loading matrix has a *jth* column given by $\sqrt{\lambda_j} e_j$, i.e. we can write

$$\Sigma = \mathbf{A}\mathbf{A}' + 0 = \mathbf{A}\mathbf{A}' \quad \cdots \quad E_{16}$$

The factor analysis representation of Σ in E_{16} is exact, not particular useful because it employs as many common factors as there are variables. Models that explain the

covariance structure in terms of just a few common factors are preferred. An approach in achieving is when the last $n\text{-}m$ eigenvalues are small, is to neglect the contribution of $\lambda_{m+1}e_{m+1}e'_{m+1} + \lambda_{m+2}e_{m+2}e'_{m+2} + \ldots + \lambda_n e_n e'$ to Σ in E_{15}. Neglecting this contribution, we obtain the following approximation.

Let $m<n$ be the number of common factors. Then the matrix of estimated factor loadings $\{\tilde{a}_{ij}\}$

$$\tilde{A} = \left[\sqrt{\hat{\lambda}_1}\hat{e}_1 \mid \sqrt{\hat{\lambda}_2}\hat{e}_2 \mid \cdots \mid \sqrt{\hat{\lambda}_m}\hat{e}_m \right]$$

The estimated specific variances are provided by the diagonal elements of the matrix

$S - \tilde{A}\tilde{A}$, so $\tilde{\psi}_i = s_{ii} - \sum_{j=1}^{m} \tilde{a}_{ij}^2$ for $i=1, 2,\ldots, n$. The communalities are estimated as

$$\tilde{h}_i^2 = \tilde{a}_{i1}^2 + \tilde{a}_{i2}^2 + \ldots + \tilde{a}_{im}^2 \quad \cdots \quad E_{17}$$

For the principal component solution, the estimated factor loadings for a giving factor do not change as the number of factors is increased. The contribution to the sample variance s_{ii} from the first common factor is $\tilde{a}_{i1}^2$. The contribution to the total sample variance $s_{11} + s_{22} + \ldots + s_{nn} = \text{tr}(S)$, from the first common factor is then

$$\tilde{a}_{11}^2 + \tilde{a}_{21}^2 + \ldots + \tilde{a}_{n1}^2 = \left(\sqrt{\tilde{\lambda}_1}\tilde{e}_1 \right)' \left(\sqrt{\tilde{\lambda}}\tilde{e} \right) = \tilde{\lambda}_1 \quad$$ since the eigenvector e_1 has unit length. In general, Proportion of total variance due to the *jth* factor $= \tilde{\lambda}_j / s_{11} + s_{22} + \ldots + s_{nn}$ for analysis of **S**. likewise, Proportion of total variance due to the *jth* factor $= \tilde{\lambda}_j / n$ for a factor analysis of **R**.

Methods described above are purely based on standard mathematical principles.

- **Maximum Likelihood Method**

If the m common factors are assumed to be multivariate normally distributed random variables with mean vector $\mathbf{0}$ and covariance matrix $\boldsymbol{\Phi}$ and the specific factors ε are assumed to be normally distributed random variables with mean zero and diagonal covariance matrix $\boldsymbol{\Psi}_{\delta.}$ The likelihood function depends on the factor loadings $\mathbf{A}$ and specific variances Ψ through equation E_{14} . The maximum likelihood estimates $\hat{A}$ and $\hat{\psi}$ are obtained by numerical maximization.

3.5 CRITERIA FOR EXTRACTING FACTORS

To minimize subjectivity, three factor selection criterions MINEIGEN, SCREE TEST, and INTERPRETABILITY were used to come out with a sufficient number of factors required.

(i) Kaiser's criterion, suggested by Guttman and adapted by Kaiser, considers factors with an eigenvalue greater than one as common factors.

(ii) Cattelll's (1966) scree test. If there is a point below which factors explain relatively little variance and above which they explain substantially more, this usually appears as an "elbow'' in the plot. This amounts to keeping the factors that contribute most to the variance.

(iii) Interpretability criteria where there are at least three (3) items with significant loadings, the variables that load on a factor share some conceptual meaning, the variables that load on different factors seem to measure different constructs and whether the rotated factor pattern demonstrate a simple structure.

3.6 THE SOLUTION SPACE

The analysis of the factor model proceeds by imposing conditions that allows one to uniquely estimate **A** and Ψ**.** However, the solution may not always be consistent with the statistical interpretation of the coefficients. For instance, when a coefficient estimates for a specific factor ψ_i for variable i is non-positive. Also there are instances where some of the factor loadings a_{ij} are greater than unity. In such situation, the loading matrix **A** is then rotated, where the rotation is determined by some "ease-of interpretation'' criterion.

Let **T** any m x m orthogonal matrix, so that $TT'=T'T=1$, then the expression in E_{12} can be written as:

$$X\text{-}\mu=AF+\varepsilon=ATT'F+\varepsilon=A^*F^*+\varepsilon \qquad \cdots \qquad E_{18}$$

Where $A^*=AT$ and $F^*=T'F$ i.e. the factors F and F^* have the same statistical properties even though the loadings A^* are different from the loadings A , they both generate the same covariance matrix Σ i.e. from E_{16} , and we can get this expression:

$$\Sigma = AA'+\Psi = AT'TA'+\Psi = \left(A^*\right)\left(A^*\right)' + \Psi$$

3.7 TEST FOR THE NUMBER OF COMMON FACTORS AND DIAGNOSTICS

The assumption of a normal population leads directly to test of the adequacy of the model. Suppose m factors holds, testing the adequacy of the m common factor model is equivalently to testing: $\mathbf{H}_0: \Sigma = AA'+\Psi$; $\mathbf{H}_1: \Sigma \qquad \cdots \qquad E_{20}$

any other positive definite matrix. Under H_0 in E_{20}, the likelihood ratio statistic is used with an approximation to the Chi-square distribution with $1/2\left((n-m)^2 - n - m\right)$ degrees of freedom. In implementing the test, the adequacy of the m common factor model by comparing the generalized variances $\left|\hat{A}\hat{A}' + \hat{\psi}\right|$ and $\left|S_n\right|$. The estimated values of the common factors called factor scores and plots of the principal factor loadings against the maximum likelihood factor loadings were used as tool to detect possible outliers. Model diagnostics proceeded to assess the normality assumption for the maximum likelihood method using the gamma plot which is the plot of the squared generalized distances with quantiles of a chi-squared distribution. The squared generalized distances behave as a chi-squared distribution when the parent population from which the sample is selected is multivariate normal and the sample size is relatively larger as compared to the number of variables or random variables.

3.8 A STEPWISE TREATMENT OF FACTOR ANALYSIS

As the goal of this study is to show and explain using SPSS which is a computer programme to solve factor analysis of tree distribution patterns. The theoretical aspects of factor analysis will now be discussed from a practical, applied perspective. Below is a flow diagram from (Rietveld & Van Hout, 1993) that offers an overview of the steps in factor analysis. Treating the theory in such a stepwise and practical fashion should ultimately result in a systematic and practical background that makes it possible to successfully carry out a factor analysis. The flow diagram (Fig1) presents the steps in factor analysis.

```
                              ┌─────────────┐
                        ╱     Reliable    ╲
NO ────────────────────◇                  ◇
                        ╲                 ╱
                              └─────────────┘
                                    │
                                   YES
                                    │
                        ┌───────────────────────┐
                        │   Correlation matrix   │
                        └───────────────────────┘
                                    ↕
                              ╱           ╲
                NO ─────────◇    Factor     ◇
                   │         ╲           ╱
                   │              └───┘
                   │                │
                   ↓               YES
     ┌──────────────────┐         │
     │ Principle        │    ┌──────────────────┐
     │ component        │    │  How to estimate  │
     │ analysis: unities│    │  Communalities    │
     │ in diagonal of   │    └──────────────────┘
     │ correlation matrix│            │
     └──────────────────┘            │
              │                       ↓
              └──────────────→┌──────────────────────┐
                              │  How many factors to be │
                              │  retained?              │
                              └──────────────────────┘
                                       │
                                       ↓
                              ┌──────────────────────┐
                              │ Factor rotation?       │
                              │ Orthogonal/oblique?    │
                              └──────────────────────┘
```
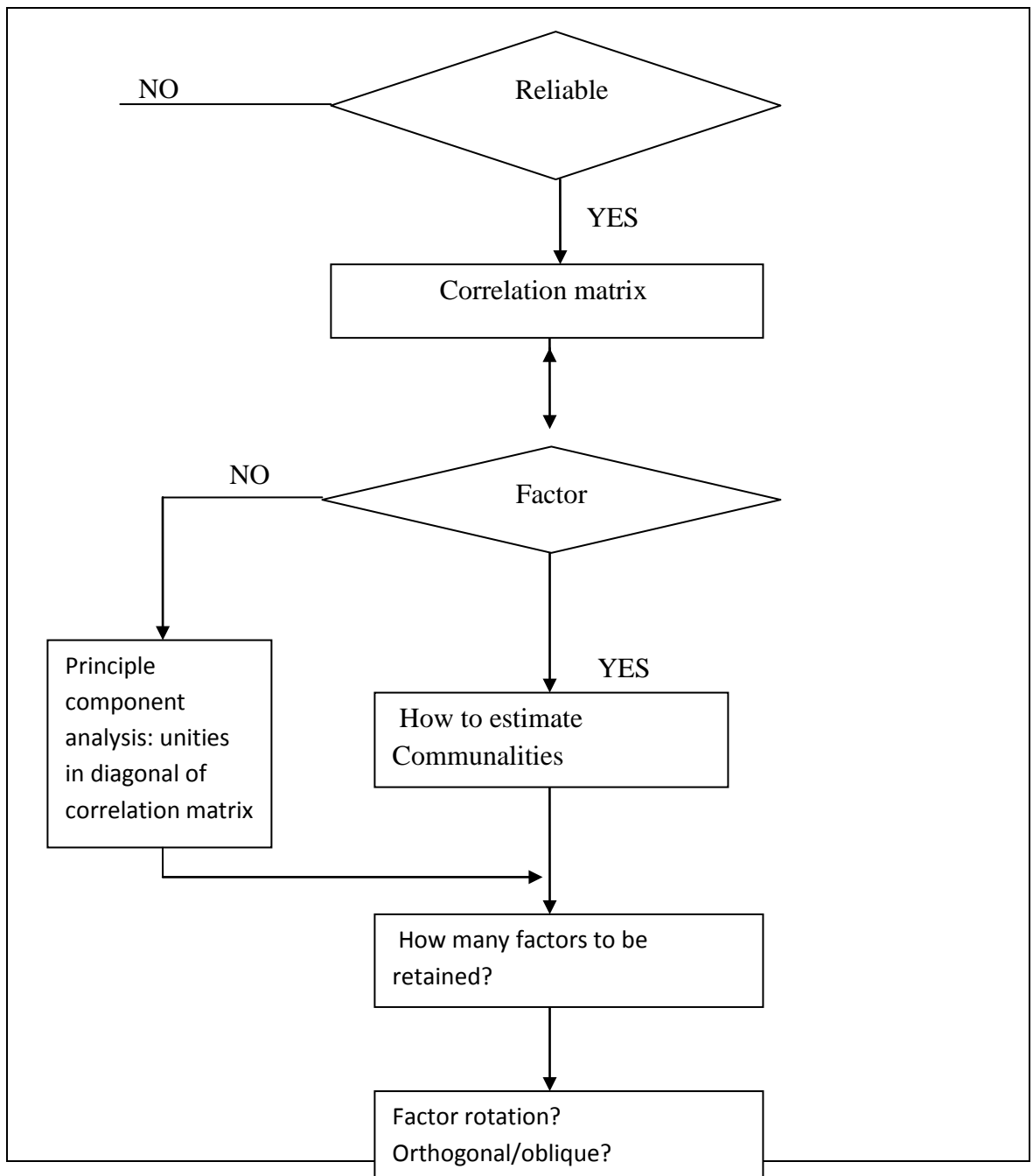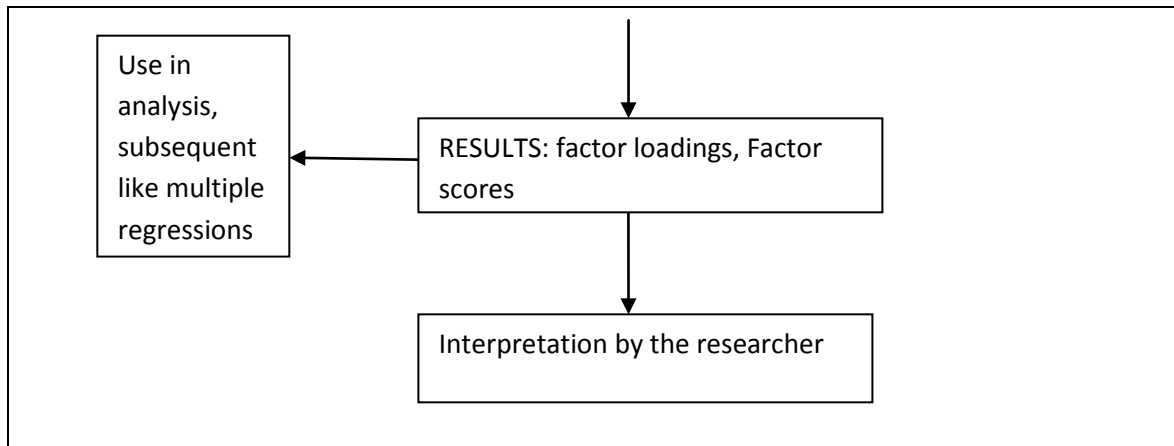
Figure 1: overview of the steps in a factor analysis

As can be seen, it consists of seven main steps: reliable measurements, correlation matrix, factor analysis versus principal component analysis, the number of factors to be retained, factor rotation, and use and interpretation of the results. Below, these steps will be discussed one at a time.

Measurements

Since factor analysis departures from a correlation matrix, the used variables should first of all be measured at (at least) an interval level. Secondly, the variables should roughly be normally distributed; this makes it possible to "generalize the results of your analysis beyond the sample collected". Thirdly, the sample size should be taken into consideration, as correlations are not resistant, and can hence seriously influence the reliability of the analysis. According to Field, "much has been written about the necessary sample size for factor analysis resulting in many 'rules-of-thumb'". Field himself, for example, states that a researcher should have "at least ten to fifteen (10-15) subjects per variable". Habing, however, states that "you should have at least 50 observations and at least 5 times as many observations as variables". Fortunately, Monte

Carlo studies have resulted in more specific statements concerning sample size (Field, 2000; Habing, 2003). The general conclusion of these studies was that "the most important factors in determining reliable factor solutions were the absolute sample size and the absolute magnitude of factor loadings". The more frequent and higher the loadings are on a factor, the smaller the sample can be. Field communality of a variable represents the proportion of the variance in that variable that can be accounted for by all 'common' extracted factors. Thus if the communality of a variable is high, the extracted factors account for a big proportion of the variable's variance. This thus means that this particular variable is reflected well via the extracted factors, and hence that the factor analysis is reliable. When the communalities are not very high though, the sample size has to compensate for this.

Correlation matrix

When the data are appropriate, it is possible to create a correlation matrix by calculating the correlations between each pair of variables. The following (hypothetical) matrix offers an example of this:

Table 1: a hypothetical correlation matrix.

$$
\begin{pmatrix}
1.00 & & & & & \\
0.77 & 1.00 & & & & \\
0.66 & 0.87 & 1.00 & & & \\
0.09 & 0.04 & 0.11 & 1.00 & & \\
0.12 & 0.06 & 0.10 & \mathbf{0.51} & 1.00 & \\
0.08 & 0.14 & 0.08 & \mathbf{0.61} & \mathbf{0.49} & 1.00
\end{pmatrix}
$$

In this matrix two clusters of variables with high intercorrelations are represented. As has already been said before, these clusters of variables could well be "manifestations of the same underlying variable". The data of this matrix could then be reduced down into these two underlying variables or factors. With respect to the correlation matrix, two things are important: the variables have to be intercorrelated, but they should not correlate too highly (extreme multicollinearity and singularity) as this would cause difficulties in determining the unique contribution of the variables to a factor. When the correlation matrix is an identity matrix, there would be no correlations between the variables. Multicollinearity, then, can be detected via the determinant of the correlation matrix, which can also be calculated in SPSS: if the determinant is greater than 0.00001, then there is no multicollinearity.

Factor analysis versus principal component analysis

After having obtained the correlation matrix, it is time to decide which type of analysis to use, factor analysis or principal component analysis. The main difference between these types of analysis lies in the way the communalities are used. In principal component analysis it is assumed that the communalities are initially one. In other words, principal component analysis assumes that the total variance of the variables can be accounted for by means of its components (or factors), and hence that there is no error variance. On the other hand, factor analysis does assume error variance. This is reflected in the fact that in factor analysis the communalities have to be estimated, which makes factor analysis more complicated than principal component analysis, but also more conservative. Thus theoretically, factor analysis is more correct. However, the

solutions generated from principal component analysis differ little from those derived from factor analysis techniques. The difference between factor analysis and principal component analysis decreased when the number of variables and the magnitudes of the factor loadings increased. The choice between factor analyses thus depends on the number of variables and the magnitude of the factor loadings (Field, 2000). After having made this choice, the question arises how many factors there are to be retained.

Number of factors to be retained

As can be inferred from the last section, the number of factors to be retained is similar to the number of positive eigenvalues of the correlation matrix. This may, however, not always lead to the right solutions, as it is possible to obtain eigenvalue that are positive but very close to zero. Therefore, some rules of thumb have been suggested for determining how many factors should be retained.

(i) Retain only those factors with an eigenvalue larger than one (Guttman-Kaiser rule).

(ii) Keep the factors which, in total, account for about seventy to eighty (70-80) percent of the variance;

(iii) Make a scree-plot, keep all factors before the breaking point or elbow.

It is furthermore always important to check the communalities after factor extraction. If the communalities are low, the extracted factors account for only a little part of the variance, and more factors might be retained in order to provide a better account of the variance.

Factor rotation

After factor extraction it might be difficult to interpret and name the factors / components on the basis of their factor loadings. Remember that the criterion of principal component analysis that the first factor accounts for the maximum part of the variance; this will often ensure that "most variables have high loadings on the most important factor, and small loadings on all other factors". Thus, interpretation of the factors can be very difficult. A solution for this difficulty is factor rotation. Factor rotation alters the pattern of the factor loadings, and hence can improve interpretation. Rotation can best be explained by imagining factors as axes in a graph, on which the original variables load. By rotating these axes, then, it is possible to make clusters of variables load optimally. There are two types of rotation: orthogonal and oblique rotation. In orthogonal rotation there is no correlation between the extracted factors, while in oblique rotation there is. It is not always easy to decide which type of rotation to take; as Field states, "the choice of rotation depends on whether there is a good theoretical reason to suppose that the factors should be related or independent, and also how the variables cluster on the factors before rotation". A fairly straightforward way to decide which rotation to take is to carry out the analysis using both types of rotation; "if the oblique rotation demonstrates a negligible correlation between the extracted factors then it is reasonable to use the orthogonally rotated solution". There are several methods to carry out rotations. SPSS offers five: varimax, quartimax, equamax, direct oblimin and promax. The first three options are orthogonal rotation; the last two oblique. It depends on the situation, but mostly varimax is used in orthogonal rotation and direct oblimin in oblique rotation. Orthogonal rotation results in a rotated component / factor matrix that presents the 'post-rotation' loadings of the original variables on the extracted

factors, and a transformation matrix that gives information about the angle of rotation. In oblique rotation the results are a pattern matrix, structure matrix, and a component correlation matrix. The pattern matrix presents the 'pattern loadings' "regression coefficients of the variable on each of the factors" while the structure matrix presents 'structure loadings' "correlations between the variables and the factors"; most of the time the pattern matrix is used to interpret the factors. The component correlation matrix presents the correlation between the extracted factors / components, and is thus important for choosing between orthogonal and oblique rotation. The choice of oblique rotation in favour of orthogonal rotation is because of the fact that the factor scores in oblique rotation correlate rather highly, as the component score covariance matrix makes clear.

Results: factor loadings and factor scores

In the last paragraph it was already reported that the factor loadings are represented in the rotated component matrix. As may be known by now, these factor loadings are important for the interpretation of the factors, especially the high ones. One can wonder, however, how high a loading has to be in order to determine the interpretation of the factor in a significant way.

This is dependent of the sample size: the bigger the sample the smaller the loadings can be to be significant. Stevens (1992) made a critical values table to determine this significance. Field (2000) states, on the other hand, that "the significance of a loading gives little indication of the substantive importance of a variable to a factor". For this to determine, the loadings have to be squared. Stevens and Field then "recommends interpreting only factor loadings with an absolute value greater than 0.4 (which explain

around 16% of variance)". This is only possible in principal component analysis, though. Thus, Stevens's recommendation should be approached with care!

The second result of factor analysis is the factor scores. These factor scores can be useful in several ways.

CHAPTER 4

DATA COLLECTION AND ANALYSIS

For this study, data was collected on six forest reserves in Ashanti Region namely Onyimsu forest reserve, Bobiri forest reserve, Mirasa Hills, Prakaw forest reserve, Pra Anum forest reserve and Onuem Bepo forest reserve. The data is secondary, based mainly on specimens obtained from Forestry commission, at the resource management support centre (RMSC) in Kumasi, Ashanti Region.

Assumption:

These three important assumptions are made when using factor analysis.

(i)   Individual variables and factors are linearly interrelated.

(ii)   Factors act additively in respect to any given variable, and

(iii)   There are no interaction effects among variables.

The plates below shows the pictorial view of some of the forest reserves used for this study.

Plate 1: Onyimsu forest reserve

Plate 2: Bobiri forest reserve

Plate 3: Mirasa Hills

Plate 4: Prakaw forest reserve

4.1 DATA ANALYSIS

Starting with initial data of hundred (100) tree species and some environmental variables for each of the six selected forest reserves in the ecological zone of Ashanti Region, the statistical Product for Social Scientists (SPSS) was used as a tool to perform the analysis. We computed correlations between each pair of coefficients as estimates of communalities. By using standardized data we computed the factor scores.

The outcome and analysis of the research are as follows:

4.2 DESCRIPTIVE STATISTICS

Table 4.1:   Descriptive Statistics

	Mean	Std. Deviation	Analysis N
Diameter of tree	73.10	20.772	100
Soil pH level [ph]	5.48	.401	100
Mean annual rainfall	1625.97	70.663	100
March temperature (hot months)	32.37	1.178	100
December temperature (cold months)	19.89	.803	100

Mean - These are the means of the variables used in the factor analysis.

Std. Deviation - These are the standard deviations of the variables used in the factor

analysis.

Analysis N - This is the number of cases used in the factor analysis.

4.3 COMMUNALITIES

Table 4.2:   Communalities

	Initial	Extraction
Diameter of tree	1.000	.884
Soil pH level [ph]	1.000	.897
Mean annual rainfall	1.000	.257
March temperature	1.000	.764
December temperature	1.000	.610

This is the squared multiple correlation for the variable as dependent using the factors as

predictors. The communality measures the percent of variance in a given variable

explained by all the factors jointly and may be interpreted as the reliability of the indicator. In table 4.2 above, the extracted factors explain over 89% of Soil ph level but only 25% for Mean annual rainfall. In general, communalities show for which measured variables the factor analysis is working best and least well.

4.4 TOTAL VARIANCE EXPLAINED

Table 4.3: Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total
1	1.326	26.527	26.527	1.326	26.527	26.527	1.312
2	1.058	21.156	47.683	1.058	21.156	47.683	1.085
3	1.028	20.557	68.240	1.028	20.557	68.240	1.035
4	.957	19.132	87.372				
5	.631	12.628	100.000				

Eigenvalues measure the amount of variation in the total sample accounted for by each factor. The eigenvalue is not the percent of variance explained but rather a measure of

the amount of variance in relation to total variance, since variables are standardized to have means of 0 and variances of 1; total variance is equal to the number of variables. When components are correlated, sums of squared loadings cannot be added to obtain a total variance. From table 4.3 above, 5 components (factors) would be needed to explain 100% of the variance in the data. In using the conventional criterion of stopping when the initial eigenvalue drops below 1.0, only 3 of the 5 factors were actually extracted in this analysis. These three account for 68% of the variance in the data.

4.5 CORRELATION MATRIX

Table 4.4: Correlation Matrix

		Diameter	Soil pH level [ph]	Mean annual rainfall	March temperature	Dec temperature
Correlation	Diameter	1.000	.056	.000	-.098	.044
	Soil pH level [ph]	.056	1.000	.089	.061	-.147
	Mean annual rainfall	.000	.089	1.000	-.154	-.042
	March temperature	-.098	.061	-.154	1.000	.225
	Dec temperature	.044	-.147	-.042	.225	1.000

Sig. (1-tailed)	Diameter		.290	.498	.166	.331
	Soil pH level [ph]	.290		.188	.273	.072
	Mean annual rainfall	.498	.188		.063	.339
	March temperature	.166	.273	.063		.012
	Dec temperature	.331	.072	.339	.012	

Determinant = 0.871

4.6 COMPONENT SCORE

Table 4.5:   Component Score Coefficient Matrix

	Component		
	1	2	3
Diameter	-.058	.037	.907
Soil pH level [ph]	-.069	.865	.028
Rainfall	-.366	.131	.081
March	.608	.307	-.156
December	.502	-.256	.335

Computing factor scores allows one to look for factor outliers. Outliers can impact correlations heavily and thus distort factor analysis. One may use Mahalanobis distance to identify cases which are multivariate outliers, then remove them from the analysis prior to factor analysis. One can also create a dummy variable set to 1 for cases with high Mahalanobis distance, and then regress this dummy on all other variables. If this regression is non-significant then the outliers are judged to be at random and there is less danger in retaining them. The ratio of the beta weights in this regression indicates which variables are most associated with the outlier cases.
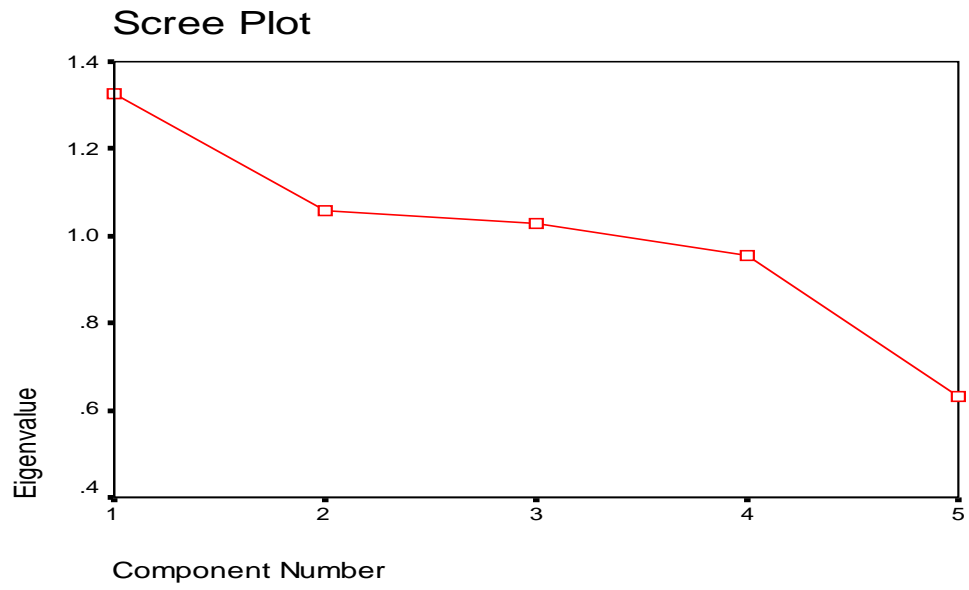
Scree Plot

Fig.4.1 Scree plot of distribution pattern of trees.

CHAPTER 5

CONCLUSION AND RECOMMENDATION

5.1 BRIEF CHAPTER INTRODUCTION

The three factors combined various tree species and environmental variables into groups and reflect general patterns that were meaningful in relation to the ecological characteristics of the various plants included. These factors were named using the common names of the species showing the highest correlation. Correlations (Loadings) for the variables on the three factors are given in the Tables 5.1 to 5.3. Positive indicates a positive correlation and negative also indicates a negative correlation. Variables that exhibited correlations of less than .40 are only loosely associated with any of the three factor groups, and these marginal associates can be thought of as having essentially independent distribution patterns.

A total of 19 species are positively correlated with factor I, the Akasa group. There are 5 marginally associated (Table 5.1) that is species with essentially independent distributions are listed with this group. This factor statistically explains 10.9% of the total and 26% of the common variance. Factor scores are high in factor 1, and species with high correlations as .80 and above are restricted to these reserves.

TABLE 5.1: Variables associated with tree factor 1

Factor Loadings (Correlations)	Variables	
0.88	*Chrysophyllum spp.*	AKASAA
0.88	*Nauclea diderrichii*	KUSIA
0.87	*Afzelia bella*	PAPAO
0.77	*Piptadeniastrum afri*	DAHOMA
0.77	*Cylicodiscus gabunen*	DENYAO
0.74	*Albizia ferruginea*	AWIEMFOSAMINA
0.72	*Mallotus oppositifol*	ANYANYANFOROWA
0.72	*Mareya micrantha*	DUBRAFO
0.63	*Macaranga hurifolia*	OPAMFUFUO
0.60	*Alstonia boonei*	SINURO
0.59	*Vernonia amygdalina*	AWONWENE
0.49	*Antrocaryon micraste*	APROKUMA
0.49	*Bombax brevicuspe*	ONYINAKOBEN
0.47	*Copaifera salikounda*	ENTEDUA
0.46	*Entandophragma angol*	EDINAM
0.37	*Hexalobus crispiflor*	DUABAHA
-0.35	*Zanthoxyllum leprieu*	OYAA
0.32	*Erythrophleum suaveo*	ODOM
0.27	*Guarea thompsonii*	KWADWUMA
0.22	*Millettia rhodantha*	TETETOA

TABLE 5.2: Variables associated with tree factor 2

Factor Loadings	Variables	
0.94	*Sterculia tragacanth*	SOFO
0.92	*Amphimas pterocarpoi*	YAYA
0.92	*Daniellia spp.*	HYEDUA
0.85	*Triplochiton sclerox*	WAWA
0.83	*Guarea cedrata*	KWABOHORO
0.80	*Terminalia superba*	OFRAM
0.80	*Aningeria altissima*	SAMFENA
-0.77	*Zanthoxylum gilletii*	OKUO
0.73	*Khaya ivorensis*	DUBINI
-0.73	Soil ph	
0.72	*Mallotus oppositifol*	ANYANYANFOROWA
-0.71	*Trichilia prieuriana*	KAKADIKURO
0.66	*Stereospermum acumin*	ESONO-TOKWAKOFU
-0.62	*Celtis zenkeri*	ESAKOKO
0.60	*Antiaris toxicaria*	KYEN-KYEN
-0.59	*Trichilia monadelpha*	TANDRO
-0.58	*Petersianthus macroc*	ESIA
-0.52	*Ceiba pentandra*	ONYINA
0.51	*Daniellia ogea*	EHYEDUA
0.49	*Guibortia ehie*	ANOKYE-HYEDUA
-0.48	*Dialium aubrevillei*	DUABANKYE
0.47	*Majidea fosteri*	ANKYEWA
0.47	*Cynometra ananta*	ANANTA
0.44	*Klainedoxa gabonensi*	KROMA
-0.43	*Sterculia oblonga*	OHAA
0.42	*Uapaca guineensis*	KONTAN
0.38	*Anopyxis klaineana*	KOKOTE
0.32	*Heretiera utilis*	NYANKOM
0.31	*Diospyros sanza-mini*	SANZA-MULIKA
0.31	*Tabernaemontana spp.*	OBANAWA
0.28	*Aningeria robusta*	SAMFENANINI

TABLE 5.3: Variables associated with tree factor 3

Factor Loadings	Variables	
0.99	*Daniellia thurifera*	SOPI
0.95	*Albizia zygia*	OKORO
0.93	*Sterculia rhinopetal*	WAWABIMA
0.90	*Vitex spp.*	OTWENTOROWA
0.87	*Morinda lucida*	KONKROMA
0.87	*Xylopia quintasii*	OBAA
0.86	*Scottellia klaineana*	TIABUTUO
0.86	*Trichilia tessmannii*	TANUNINI
0.85	*Tieghemella heckelii*	BAKU
0.81	*Millettia zechiana*	FAFRAHA
0.80	*Aningeria altissima*	SAMFENA
-0.69	*Mammea africana*	BOMPADYA
0.65	March temperature	
0.63	Mean annual rainfall	
0.57	Soil ph level	
0.52	*Xylia evansii*	ABOBABEMA
0.48	*Erythrophleum ivoren*	POTRODOM
0.44	*Zanthoxylum spp.*	OYAA/OKUO
0.43	*Khaya anthotheca*	DUBINIBIRI
0.42	*Zanthoxylum lemairei*	OKUONINI
0.39	*Distemonanthus benth*	BONSAMDUA
0.39	December temperature	
0.36	*Pterygota macrocarpa*	KYEREYE
0.35	*Monodora tenuifolia*	MOTOKURADUA
0.35	*Mansonia altissima*	OPRONO
0.33	*Xylopia staudtii*	OBAAKOKO
0.32	*Entandophragma cylin*	PENKWA
0.31	*Entandophragma cando*	PENKWA-AKOA
0.29	*Parkia bicolor*	ASOMA
0.29	*Entandophragma utile*	EFOOBRODEDWO
0.28	*Bombax buonopozense*	AKATA
0.25	*Lovoa trichilioides*	DUBINIBIRI
0.22	*Vitex doniana*	ABISAA
-0.21	*Macaranga barteri*	OPAM

Factor II, the Sofo group, includes 21 species with positive correlations and 9 with negative loadings (Table 5.2). Five species show weak association. This factor explains 7.8% of the total and 21% of the common variance. Species with high correlations such as .80 and above are distributed throughout the main portion of the region and are typical of species with high correlation on factor II. One soil ph has negative correlation with this factor. Factor III, entitled the Sopi group (Table 5.3), includes 33 species largely restricted to the Onuem Bepo forest.

Two species have negative correlations with this factor, one of which has a marginal negative association. The factor explains 17.04% of the total and 20% of the common variance. Three environmental variables are part of this factor.

The species positively associated with factor I are all characteristics of Onyimsu and Bobiri forest. Factor I loadings were influenced by considering distributions only for a relatively small area in contrast to conducting the analysis on all the reserves in Ghana. Papao (*Afzelia bella*) is an early successional species which primary occur in fairly dense forests. The fact that none of the temperature variables have high loadings on factor I suggests that the species associated in this group are not limited by temperature alone.

The species which show correlation greater than .60 on factor II are widespread species and are likely to be found in all the forest reserves. The other 3 species with high scores on this factor, sofo, yaya, and hyedua are widespread, but there is some distinction in the apparent habitat preference. Sofo is commonly found in Mirasa Hills, but occurs in Prakaw forest. Yaya has a particularly wide ecological tolerance in that it frequently

73

occurs as relatively small trees, but reaches its maximum development in the lowland forest. The remainder of the species is widespread, but more restricted to the Pra Anum forest. The negative loading on soil ph on factor II is an indication that, the plants associated positively with this factor tend to occur in areas with the lowest rainfall. They may be important in terms of limiting distributions of the species showing this distribution pattern.

Factor III is the most clearly defined of all the factors. The environmental variables with the highest correlations are useful descriptors in differentiating the Onuem Bepo forest from the others. Similarly, the top 11 species are also essentially restricted to this area, although there are variations within this forest. The fact the March temperature has a positive association with factor III indicates that the tree species with positive loadings on this factor can survive in the areas of stressful summers.

5.2 CONCLUSION

Each of the three factors indicates a group of species and/or environmental variables which is superficially explained first by distribution and then by environmental variability, past history, competitive ability and habitat preference. There are a number of tree species that do not fit into these general patterns, but instead show independent patterns of distribution.

Reexamination of the original two hypotheses reveals that the first, concerning the predominant influence of precipitation (rainfall) is disproved. None of the generalized

moisture variable is found in factor I. There are other variables with higher loadings than rainfall variables in factor II and factor III and also a temperature variable shows a higher correlation than rainfall variable.

This does not mean that precipitation (rainfall) is not important, but suggests that it alone may have less effect than other variables. Moisture directly available to tree species is of course, governed by other environmental variables such as topographic relief and the water retaining characteristics of soils in a particular area. The factors are compatible with the second hypothesis which has to do with the relatively uniform and strong climatic gradient and contiguous geographical distribution since the factors include many which are found throughout most of the six forest reserves under study. Also they demonstrated a wide tolerance to the represented environmental conditions.

5.3 RECOMMENDATIONS

Trees are important, valuable and necessary to our very existence. It's not too hard to believe that, without trees we humans would not exist on this planet. They are essential to life as we know it and are the ground troops on an environmental frontline. Our existing forest and the trees we plant work in tandem to make a better world. In view of this crucial role trees play in society, every effort must be put into fulfilling the objectives of protecting and building our forest reserves.

It is in this regard that the following recommendations are made for the necessary action to be taken by the stakeholders.

(i)   Studies such as this one are helpful in indicating environmental parameters that are limiting tree distributions. So there is the need for the government and other concerned bodies to promote and encourage such investigation in future to give much reliable and current information about the distribution patterns of trees.

(ii)   The study points out a number of subsequent analyses both correlative and operational, which could be done to characterize better the ecological relationships of forested and other areas in Ashanti Region and Ghana as a whole. For example, the forestry commission of Ghana should undertake detailed field and laboratory studies of the relationship between individual tree species and environmental variables associated on a given factor to give more precise statements concerning possible cause and effect relationships.

(iii)   It would be of interest for stakeholders to analyze in more detail the factors underlying the irregular tree distribution patterns in Ghana as a whole.

(iv)   In addition, investigations similar to this one, but involving other types woody and/or herbaceous plant species may result in further clarification of the influence of environmental variation on tree distributions in Ashanti Region and other areas in Ghana.

REFERENCES

1. Blurtit, www.samtrees.com © (2007-2009).

2. Brown, J. H. (1984). On the relationship between abundance and distribution of species. Am. Nat. 124: 255-279.

3. Bryman, A. and Cramer, D. (1990). Quantitative data analysis for social scientists. New York: Routledge.

4. Burchett, CR. (1977), Water resources of the upper Duck River basin, central Tennessee: Nashville, Tennessee Department of Conservation, Division of Water Resources, Water Resources Series, no. 12, 103p.

5. Clark, P. J. and Evans, F. C. (1954). Distance to nearest neighbor as a measure of spatial relationships in populations. Ecology 35: 445-453.

6. Collins, S. L. and S. M. Glenn. (1997). Effects of organismal and distance scaling on analysis of species distribution and abundance. Ecol. App. 7: 543-551.

7 Cowardin, L. M., Carter, Virginia, Golet, EC., and LaRoe, E. T. (1979), Classification of wetlands and deepwater habitats in the United States: U.S. Fish and Wildlife Service, Biological Services Program, FWS/OBS70/31, 131p.

8 Cureton, E. E., and D'Agostino, R. B. (1983). Factor analysis: An applied approach. Hillsdale, NJ: Erlbaum.

9 Donnelly, K. P. (1978). Simulations to determine the variance and edge effect of total nearest-neighbor distance, pp. 91- 95.

10 Eriksson, O. and Jakobsson A. (1998). Abundance, distribution and life histories of grassland plants: a comparative study of 81 species. J. Ecol. 86: 922-933.

11 Feller, W. (1943). On a general class of contagious distributions. Ann. Math. Stat. 14: 389-400.

12 Field, A. (2000). Discovering Statistics using SPSS for Windows. London – Thousand Oaks –New Delhi: Sage publications. Habing, B. (2003). Exploratory Factor Analysis. Website: http://www.stat.sc.edu/~habing/courses/530EFA.pdf (accessed 10 May 2004).

13 Ghana Home Page , (1994-2009), Ghanaweb.com

14 Gorsuch, R. L. (1983). Factor analysis (2nd Ed.). Hillsdale, NJ: Erlbaum.

15 Greig-Smith, P. (1983). Quantitative plant ecology, 3rd ed. Blackwell Scientific, Oxford.

16 Haase, P. (1995). Spatial pattern analysis in ecology based on Ripley's K-function: Introduction and methods of edge correction. J. Veg. Sci. 6: 575-582.

17 Haase, P., Pugnaire, F. L., Clark, S. C. and Incoll, L. D. (1996). Spatial patterns in a two-tiered semi-arid shrub land in southeastern Spain. J. Veg. Sci. 7: 527-534.

18 Hanski, I. (1982). Dynamics of regional distribution: the core and satellite species hypothesis. Oikos 38: 210-221.

19 Hubbell, S. P. and Foster, R. B. (1983). Diversity of canopy trees in neotropical forest and implication for conservation. In: Sutton, S. L., Whitmore, T. C. and Chadwick, A. C. (eds.) Tropical rainforest: ecology and management, pp. 24-41. Blackwell Scientific, Oxford.

20 Hubbell, S. P. (1979). Tree dispersion, abundance, and diversity in a tropical dry forest. Science 203: 1299-1309. Longman, London.

21 International Society of Arboriculture, (2005). **Bureau Insurance Services Charter House, 43 St Leonards Road, Bexhill on Sea, East Sussex TN40 1JA.**

22 Kerlinger, F. N. (1979). Behavioral research: A conceptual approach. Dallas: Holt, Rinehart and Winston.

23 McIntosh, R. P. (1973). Matrix and plexus techniques, p 157-191. Ecological Society o America.

24 Rietveld, T. and Van Hout, R. (1993). Statistical Techniques for the Study of Language and Language Behaviour. Berlin – New York: Mouton de Gruyter.

25 Scheiner, S. M. and J. M. Rey-Benayas (1997). Placing empirical limits on metapopulation models for terrestrial plants. Evol. Ecol. 11: 275-288.

26 Smalley, G. W. (1983), Classification and evaluation of forest sites on the eastern Highland Rim and Pennyroyal: New Orleans, USDA Forest Service, Southern Forest Experiment Station, General Technical Report SO-43, 123p.

27 Stevens, J. P. (1992). Applied Multivariate Statistics for the Social Sciences (2nd edition). Hillsdale, NJ: Erlbaum.

28 Tokeshi, M. (1992). Dynamics of distribution in animal communities: Theory and analysis. Res. Popul. Ecol. 34: 249-27.

29 Wannacott, R. J. and Wannacott, T. H. (1985), Introductory statistics, 4th ed.: New York, Wiley, 649p.

30 Watt, (1889-1893); Anonymous, (1983); Maithani et al., (1991). Seth, (2004).

31 Wikipedia Foundation, Inc., (2009).

APPENDIX

Figure 1.1: Map of Forest Reserve in High Forest Zone of Ghana. Source: Hawthorne and Abu Juam, 1995.
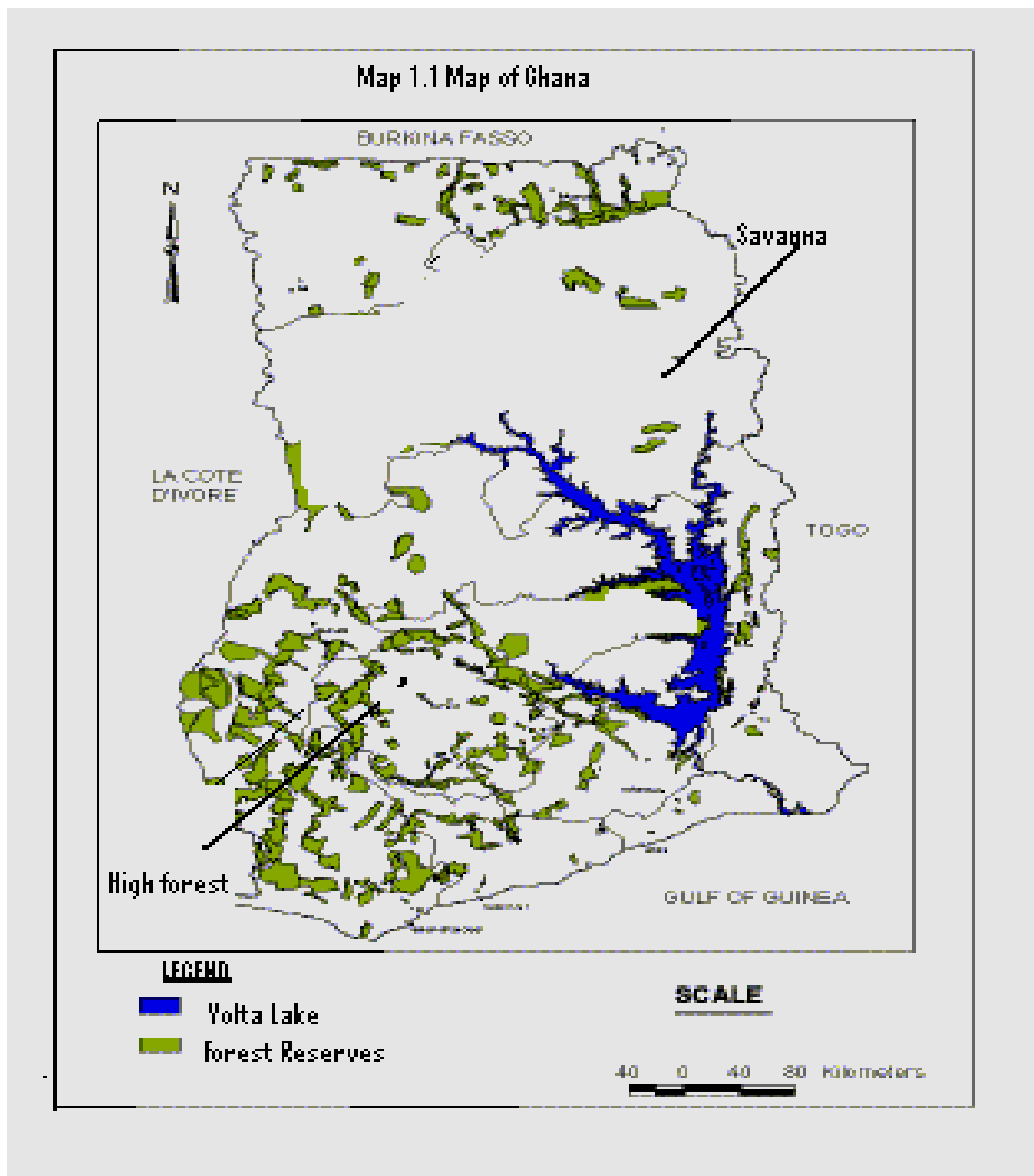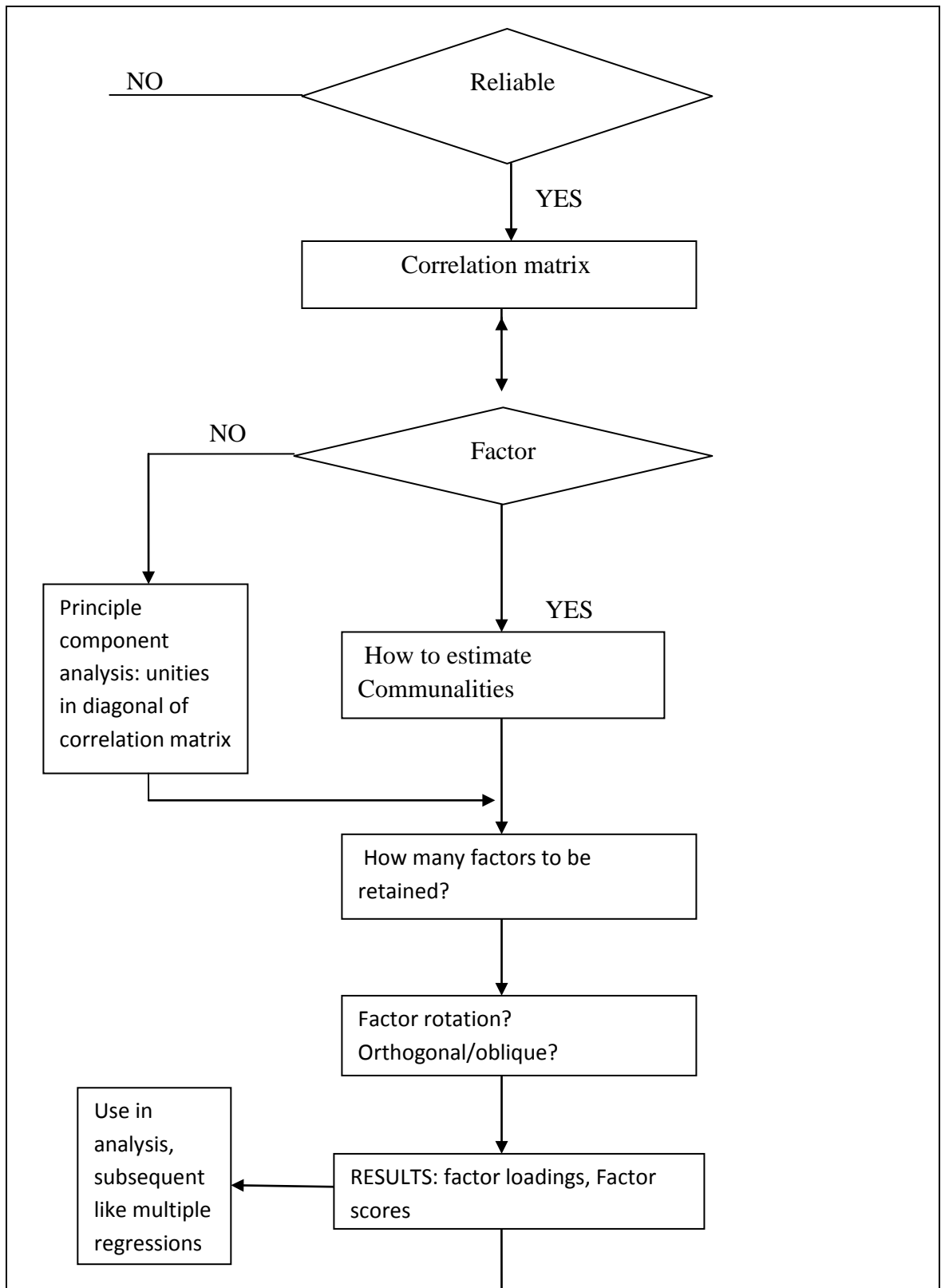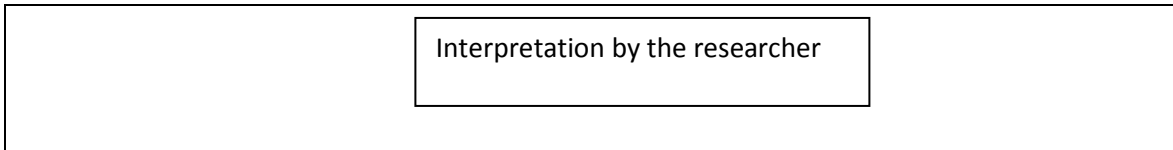
Figure 3.1: Overview of the steps in a factor analysis.

Table 4.1:   Descriptive Statistics

	Mean	Std. Deviation	Analysis N
Diameter of tree	73.10	20.772	100
Soil pH level [ph]	5.48	.401	100
Mean annual rainfall	1625.97	70.663	100
March temperature (hot months)	32.37	1.178	100
December temperature (cold months)	19.89	.803	100

Table 4.2:   Communalities

	Initial	Extraction
Diameter of tree	1.000	.884
Soil pH level [ph]	1.000	.897
Mean annual rainfall	1.000	.257

March temperature	1.000	.764
December temperature	1.000	.610

Table 4.3: Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total
1	1.326	26.527	26.527	1.326	26.527	26.527	1.312
2	1.058	21.156	47.683	1.058	21.156	47.683	1.085
3	1.028	20.557	68.240	1.028	20.557	68.240	1.035
4	.957	19.132	87.372				
5	.631	12.628	100.000				

Table 4.4: Correlation Matrix

		Diameter	Soil pH level [ph]	Mean annual rainfall	March temperature	Dec temperature
Correlation	Diameter	1.000	.056	.000	-.098	.044
	Soil pH level [ph]	.056	1.000	.089	.061	-.147
	Mean annual rainfall	.000	.089	1.000	-.154	-.042
	March temperature	-.098	.061	-.154	1.000	.225
	Dec temperature	.044	-.147	-.042	.225	1.000
Sig. (1-tailed)	Diameter		.290	.498	.166	.331
	Soil pH level [ph]	.290		.188	.273	.072
	Mean annual rainfall	.498	.188		.063	.339
	March temperature	.166	.273	.063		.012
	Dec temperature	.331	.072	.339	.012	

Table 4.5:   **Component Score Coefficient Matrix**

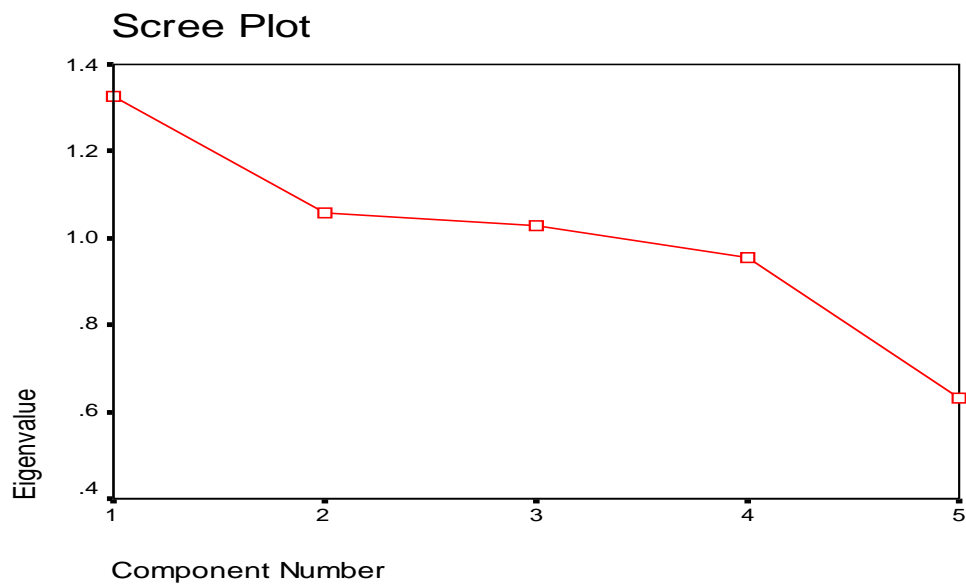|	Component		
	1	2	3
Diameter	-.058	.037	.907
Soil pH level [ph]	-.069	.865	.028
Rainfall	-.366	.131	.081
March	.608	.307	-.156
December	.502	-.256	.335

Fig.4.1 Scree plot of distribution pattern of trees.