UNDERSTANDING AND MODELING PLANT BIODIVERSITY OF NALLIHAN (A3-ANKARA) FOREST ECOSYSTEM BY MEANS OF GEOGRAPHIC INFORMATION SYSTEMS AND REMOTE SENSING

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
THE MIDDLE EAST TECHNICAL UNIVERSITY

BY

HAKAN METE DOĞAN

IN PARTIAL FULLFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
IN
THE DEPARTMENT OF BIOLOGICAL SCIENCES

AUGUST 2003

Approval of the Graduate School of Department of Biological Sciences

Prof. Dr. Canan Özgen
Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Doctor of Philosophy.

Prof. Dr. Mesude İşcan
Head of Department

This is to certify that we have read this thesis and that in our opinion it is fully adequate, in scope and quality, as a thesis for the degree of Doctor of Philosophy.

Assoc. Prof. Dr. Vedat TOPRAK Co-Supervisor	Prof. Dr. Musa DOĞAN Supervisor

Examining Committee Members

Prof. Dr. Zeki Kaya

Prof. Dr. Musa Doğan

Prof. Dr. İnci Togan

Prof. Dr. Osman KETENOĞLU

Assoc. Prof. Dr. Vedat TOPRAK

ABSTRACT

UNDERSTANDING AND MODELING PLANT BIODIVERSITY OF NALLIHAN (A3-ANKARA) FOREST ECOSYSTEM BY MEANS OF GEOGRAPHIC INFORMATION SYSTEMS AND REMOTE SENSING

DOĞAN, Hakan Mete
Phd. Department of Biology
Supervisor: Prof. Dr. Musa DOĞAN
Co-Supervisor: Assoc. Prof. Dr. Vedat TOPRAK

August 2003, 164 Pages

In this study, geographic information systems (GIS) and remote sensing (RS) tools were integrated and used to investigate the plant species diversity of the Nallıhan forest ecosystem. Two distinct indices, Shannon Wiener and Simpson, were employed in order to express species diversity. The relationships between the indices and pertinent independent variables (topography, geology, soil, climate, supervised classes, and Normalized Difference Vegetation Index (NDVI) classes) were investigated to develop two distinct models for each index. After detecting important components with *factor analysis*, two models were developed by using *multiple regression* statistics. Running the models, two plant species diversity maps in grid format were produced. The validity of the models were tested by (1) mapping residuals to predict the locations where the models work perfectly, and (2) logical interpretations in ecological point of view. Elevation and climatic factors formed the most important component that are effective on plant species diversity. Geological formations, soil, land cover and land-use characteristics were also found influential for both models. Considering the disturbance and potential evapotranspiration (PET), the model developed for Shannon Wiener index was found out more suitable comparing the model for Simpson index.

Keywords: Geographic Information Systems, Remote Sensing, Spatial Analysis, Mapping, Modeling, Plant Biodiversity, Plant Ecology, Plant Community, Plant Taxonomy, Flora, Phytogeography, Species Diversity

ÖZ

NALLIHAN (A3-ANKARA) ORMAN EKOSİSTEMİNİN BİTKİ BİYOLOJİK ÇEŞİTLİLİĞİNİN COĞRAFİ BİLGİ SİSTEMLERİ VE UZAKTAN ALGILAMA İLE KAVRANMASI VE MODELLENMESİ

DOĞAN, Hakan Mete
Doktora, Biyoloji Bölümü
Tez Yöneticisi: Prof. Dr. Musa DOĞAN
Ortak Tez Yöneticisi: Doç. Dr. Vedat TOPRAK
Ağustos 2003, 164 Sayfa

Bu çalışmada coğrafi bilgi sistemleri (CBS) ve uzaktan algılama (UA) teknikleri bütünleştirilerek, Nallıhan orman ekosistemindeki bitki tür çeşitliliği incelenmiştir. Tür çeşitliliğini açıklamak için Shannon Wiener ve Simpson indeks değerleri kullanılmıştır. Bağımlı (Shannon Wiener ve Simpson indeksleri) ve ilgili bağımsız değişkenler (topografya, jeoloji, toprak, iklim, uzaktan algılama sınıfları ve Normalleştirilmiş Vejetasyon Farklılık İndeks (NDVI) sınıfları) arasındaki ilişki her bir indeks değeri için iki ayrı model geliştirilmek için araştırılmıştır. *Faktör analizi* ile önemli bileşenler belirlendikten sonra, *çoklu regresyon* istatistiği kullanılarak iki ayrı model geliştirilmiştir. Modellerin çalıştırılması ile iki ayrı bitki tür çeşitliliği haritası üretilmiştir. Modellerin doğruluğu (1) kalan değerlerin (modelden hesaplanan değerle gerçek değer arasındaki farkların) haritalanarak modellerin doğru çalıştığı alanların tahmini ve (2) ekolojik bakış açısından mantıksal yorumlama kullanılarak kontrol edilmiştir. Yükselti ve iklimsel faktörler her iki modelde de bitki tür çeşitliliğine etki eden en önemli bileşeni oluşturmuştur. Jeolojik formasyonlar, toprak, arazi örtüsü ve kullanımı özellikleri ise diğer önemli bileşenler olarak bulunmuştur. Potansiyel evepotranspirasyon (PET) ve bitki örtüsündeki bozulma göz önüne alındığında, Shannon Wiener için geliştirilen modelin Simpson`a göre daha uygun olduğu görülmüştür.

Anahtar Kelimeler: Coğrafi Bilgi Sistemleri, Uzaktan Algılama, Alansal Analiz, Haritalama, Modelleme, Bitki Biyoçeşitliliği, Bitki Ekolojisi, Bitki Toplumu, Bitki Taksonomisi, Flora, Fitocoğrafya

To my family

ACKNOWLEDGMENTS

I would like to thank my supervisor Prof. Dr. Musa Doğan for introducing me to the field of biological sciences. His patience and guidance have helped me throughout the study as well as professionally. I also would like to thank my co-supervisor Assoc. Prof. Dr. Vedat Toprak who guided me during my research and gave me better insight in the field of geology and remote sensing as well as statistics. I have to mention Prof. Dr. Osman Ketenoğlu who taught me a lot about plant species during the identification process of plant samples in the Ankara Herbarium. In addition, special thanks to Prof. Dr. Michael DeMers, who gave me the priceless insight about the geography and spatial thinking during my master study in the U.S.A., should be stated here.

I also would like to thank Prof. Dr. Ünal Şorman, Assoc. Prof. Dr. Zuhal Akyürek, and Dr. Lütfi Süzen. They helped me about geographic information systems and modeling issues throughout the office works of this study. I want to mention my special emphasis to Sargun Tont who taught me a lot about ecology, and guided me as a teacher and as well as a friend. I also would like to thank Prof. Dr. Zeki Kaya, Prof. Dr. İnci Togan, Assoc. Prof. Dr. Can Bilgin and Assoc. Prof. Dr. Meryem Beklioğlu for their help during the courses.

I am also grateful to my colleagues Dr. Ali Mermer, Ediz Ünal, Tuncay Porsuk, Öztekin Urla, Nihal Ceylan, and Hakan Yıldız in GIS & RS Department of Central Research Institute for Field Crops. I would like to give my special thanks to the team of Nallıhan Forest Management District who helped me a lot in the field studies. I especially would like to mention the two extraordinary men who are forest engineer Murat Çetiner and forest officer İrfan Artuç in this district. In addition, appreciation is also expressed to the faculty and staff of the Biology Department of the Middle East Technical University for their cooperation and friendship during my Phd. study.

Finally, I would like to thank my wife and my daughter for their patience and endless help during my studies.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

AET	:	Annual actual evapotranspiration
Arc/Info	:	A GIS software package from ESRI Inc. for establishing and editing spatial database in vector and raster format. Map coverages and grids created in ARC/INFO may be used in Arc/View.
Arc/View	:	A GIS software package from ESRI Inc. for managing and analyzing spatial database including views, tables, charts, layouts, and scripts.
CRIFC	:	Central Research Institute for Field Crops
dBASE file	:	A file format native to dBASE database management software. Arc/View can read, create, and export tables in dBASE format.
DEM	:	Digital elevation model
ERDAS-Imagine	:	A software package from ERDAS Inc. for image processing.
ESRI	:	Environmental Systems Research Institute
FAO	:	Food and Agriculture Organization of the United Nations
GIS	:	Acronym for geographic information systems
GPS	:	Global positioning system that is based on a satellite system in order to determine locations on the earth.
HGK	:	General Commandership of Mapping
IDW	:	Inverse Distance Weighted method that is used interpolating surfaces by using point data
IUCN	:	The World Conservation Union
KHGM	:	General Directorate of Rural Affairs
LANDSAT-TM	:	Land satellite system-thematic mapper
LOCCLIM	:	A powerful local climate estimator software developed by FAO
MTA	:	General Directorate of Mineral Research and Exploration
NATO	:	North Atlantic Treaty Organization; a military alliance of western nations for the purpose of collective defense
NDVI	:	Normalized difference vegetative index that is sensitive to vegetation cover
NOAA	:	National Oceanic and Atmospheric Administration satellite system

PCA	:	Acronym for principle component analysis
PET	:	Potential evapotranspiration
RS	:	Acronym for remote sensing
SPSS	:	A statistical software developed by SPSS Inc.
SYSTAT	:	A statistical software developed by SPSS Inc.
TIFF	:	A generic image recording format
TIN	:	Triangulated network used for mapping purpose
TUBIVES	:	Turkey's Plant Database
UNEP	:	United Nations Environment Program
UNESCO	:	United Nations Educational, Scientific, and Cultural Organization
UNIX	:	A computer operating system different from PC and NT.
USGS-NPS	:	United State Geological Survey-National Park Survey
UTM	:	Universal Transverse Mercator map projection system

CHAPTER 1

INTRODUCTION

1.1. Motivation

To understand the priceless value of biodiversity at a glance, McNeely`s (1994) words may be a good starting point. McNeely (1994) addressed that “The Earth's genes, species and ecosystems are the basis for the survival of our own species, and they are the result of over 3000 million years of evolution. Biological diversity or biodiversity, the measure of the variation in genes, species and ecosystems, is valuable because future practical values are unpredictable, because variety is inherently interesting and more attractive, and because our understanding of ecosystems is insufficient to be certain of the impact of removing any component.”

Not only the unknown facts but also apparent evidences are enough to emphasize the importance of biodiversity. For instance, the earth's biodiversity plays a considerable role as a vast source of livestock, crops and pollinators of crops, the biological agents, agricultural pesticides, pharmaceuticals, and numerous ecosystem services essential to agriculture, including the creation of soils and the renewal of their fertility (Anon., 1999). The world has at least 5 to 7 million different species of plants, animals, and microorganisms (May, 1999), many of which have contributed to one of the most dramatic changes that has occurred on earth since the emergence of humans as the dominant species (Diamond, 1997).

The rapid expansion of human activities is having unprecedented impacts on a global scale (Vitousek *et al.*, 1997). Humans not only control, and use for their benefit, almost half of the world's land surface but also dominate the global cycles of nitrogen, carbon, and water, and are changing global climate. Such activities of modern societies, including the destruction of native habitats and their fragmentation into ever-smaller areas, have initiated an episode of extinctions that may prove to be the most extreme extinction event ever. Indeed, the current rate of species extinctions is at least a thousand times faster than at any time within the last 10.000 years (Pimm *et al.*, 1995).

Today, 80 plant crops provide about 90% of the world's food from plants. Fifty animal species account for most domestic animal production of food and fiber. Thousands of other plant species are actively farmed, and tens of thousands of plant species are known to have edible parts. Hundreds of animal species are regularly harvested for food, and additional species are being domesticated. Hundreds of thousands of animal species, mainly insects, are essential for pollinating crops and protecting them from pests. Tens of thousands of microbial species, most of them living in soil and on plants, provide for nutrient cycling, crop residue decomposition, and enhanced crop growth. Humans always have been, presently are, and always will be dependent on the diversity of organisms to provide food for the growing human population. Humankind's agricultural successes have stemmed from its ability to use biological diversity to its advantage. However, expanding human activities are threatening this diversity, which threatens the stability and sustainability of society (Anon., 1996).

Depending on its importance, there is a strong concern to preserve biological diversity, and this issue has been depicted in many international conferences such as; Rio Declaration in Brazil (Anon., 1992), and Antalya Declaration in Turkey (Anon., 1997). Furthermore, the issue has been repeatedly emphasized in many of the reports, decisions and recommendations of Food and Agriculture Organization of the United Nations (FAO), United Nations Environment Program (UNEP), The World Conservation Union (IUCN) and other international or national organizations. A broad capacity-building effort is urgently needed so that countries can monitor their forests (an essential requirement of Rio Declaration). Governments and institutions should establish and/or strengthen national assessment and observation systems for forests, forest resources and forest programs. This will require new data systems and statistical modeling, ground surveys and other technological innovations such as geographical information systems (GIS) and remote sensing (RS).

1.2. Biodiversity

Biological diversity, or biodiversity, refers to all forms of life, including all species and genetic variants within species and all ecosystems that contain and sustain those diverse forms of life (Anon. 1999). Cole (1994) defined very well and summarized the inter-relationships between country-specific biological diversity and various biological resources used for sustainable economic and social development (Figure 1.01).

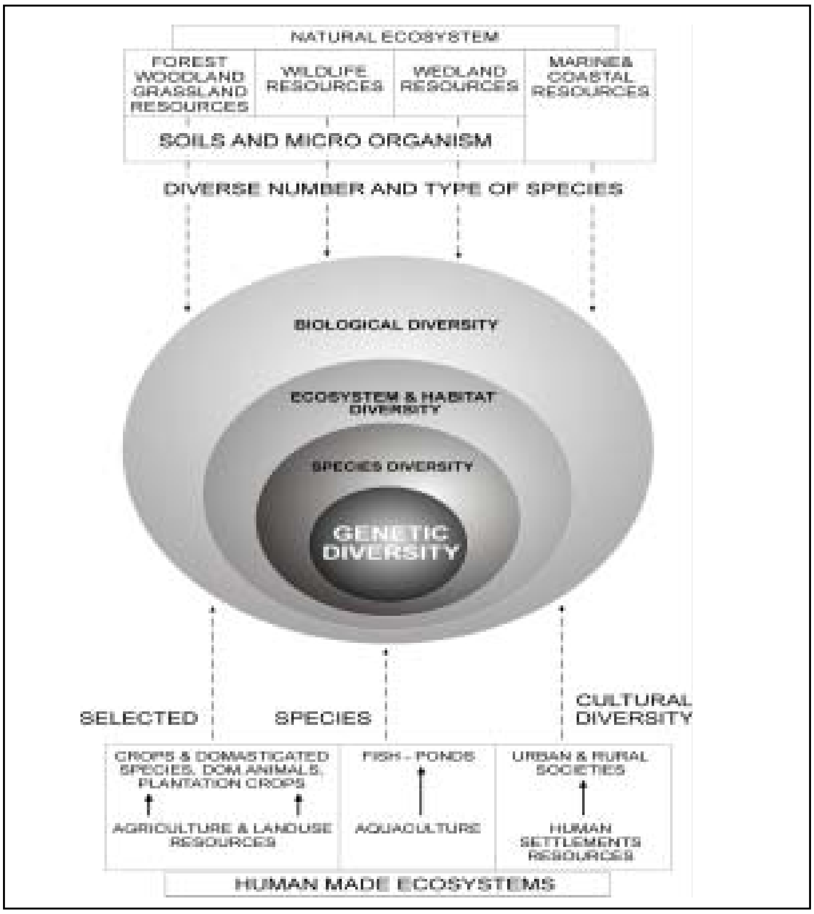


Figure 1.01. Inter-relationships between Country-Specific Biological Diversity and Various Biological Resources Used for Sustainable Economic and Social Development (Cole, 1994).

“Nature is an essential partnership; essential, because each species has its space and role, and performs a function essential to the whole; a partnership, because the living components of nature - the species - can only thrive and survive together, because together they create a dynamic equilibrium. Nature is a dynamic entity that is never the same, that changes, that adapts, that evolves; an equilibrium that remains, in essence, unchanged, because it always accommodates evolution and diversity. We all live in the same home and cannot tear it down to rebuild it. We can only improve it by respecting other entities in it, and around our own immediate living space” (Krattiger, 1994).

Although biological diversity has a vitally important role in human life including all living organisms, the available evidence indicates that human activities have been eroding biological resources and greatly reducing the planet's biodiversity. Given the projected growth in population and economic activity, the rate of loss of biodiversity is far more likely to increase than stabilize. Almost 40% of the Earth's net primary terrestrial photosynthetic productivity is now directly consumed, converted or wasted as a result of human activities. A very considerable body of work in the field of conservation biology over the past several decades has shown that reducing the area of habitat reduces not only the population of each species (and hence its genetic diversity), but also the number of species the habitat can hold. As a broad general rule, reducing the size of the habitat by 90% will reduce the number of species that can be supported in the long run by about 50%. It might be concluded that major habitat changes and associated losses of biodiversity are the inevitable price that people are willing to pay for progress as humans become an ever more dominant species. On the other hand, society has cause for concern when habitats are degraded to lower productivity, especially when accompanied by species losses, which can have world-wide ramifications. (Jeffrey and McNeely, 1994).

Productive and efficient agriculture, which is the foundation of modern successful societies, has depended on biological diversity, and will be even more dependent on it in the decades and centuries ahead. The earth's biodiversity is the source of all livestock, of all crops and pollinators of crops, of the biological agents that control crop pests, of many agricultural pesticides and pharmaceuticals, and of numerous ecosystem services essential to agriculture, including the creation of soils and the renewal of their fertility (Anon. 1999).

Measuring human pressure on the Earth, and how that pressure is distributed among countries and regions were described in the Living Planet Report-2002. Living planet index and ecological footprint were defined to explain this pressure (Figure 1.02). Basically, the living planet index is a measure of the state of natural ecosystems while the ecological footprint compares countries` consumption of natural resources with the Earth's biological capacity to regenerate them. It is the average of three sub-indices measuring changes in forest, freshwater, and marine ecosystems. It fell by 37 per cent between 1970 and 2000 (Figure 1.02-a). The ecological footprint is a measure of humanity's use of renewable natural resources. It grew by 80 per cent between 1961 and 1999, to a level 20 per cent above the Earth's biological capacity (Figure 1.02-b). These two measures do not take into account all of the conditions necessary to achieve sustainable development. But unless we recognize the ecological limits of the biosphere, we cannot claim to be sustainable (Anon. 2002).

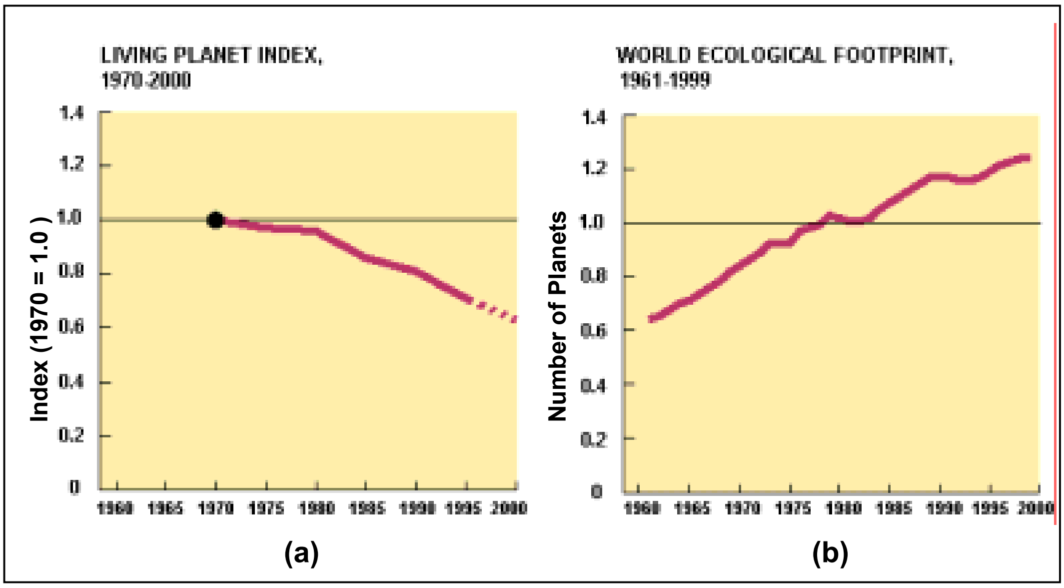


Figure 1.02. The planet index and ecological footprint; a) The Living Planet Index (the dotted line indicates the most recent period, for which fewer data are available) b) The ecological footprint (it is expressed as number of planets, where one planet equals the total biologically productive capacity of the Earth in any one year) Natural resource consumption can exceed the planet's productive capacity by depleting the Earth's natural capital, but this cannot be sustained indefinitely (Anon. 2002).

1.3. Community structure and measurement of biodiversity

A community is an association of interacting species inhabiting some defined area. This might be the plant community on a mountainside, the insect community associated with a particular species of tree, or the fish community on a coral reef. Understanting how various abiotic and biotic aspects of the environment influence the structure of communities is important. The number of species, the relative abundance of species, and the kinds of species constitute the attributes of community structure (Molles, 1999).

One of the most fundamental aspects of community structure is the relative abudance of species (Molles, 1999). According to Sugihara (1980); this fundamental property can be defined as "minimal community structure". There are regularities in the relative abundance of species in communities such as "most species are moderately abundant; few are very abundant or extremely rare" (Molles, 1999). This pattern that is one of the best documented pattern in natural communities was first quantified by Preston (1948, 1962a, 1962b).

Ecologists define species diversity on the basis of two factors: species richness, and species evenness. The number of species in the community is called species richness, while the relative abundance of species is described as species evenness. How environmental structure affect species diversity is one of the most fundamental subject of investigation about communities. Mainly, species diversity increases with environmental complexity or heterogeneity. Therefore, the ecological requirements of species (niches) should be known to predict how environmental structure affects their diversity. The niche summarizes the environmental factors that influence the growth, survival, and reproduction of a species. In another words, a species niche consist of all the factors necessary for its existence (Molles, 1999). According to Hutchinson (1957); the fundamental niche defines the physical conditions under which a species might live, in the absence of interactions with other species. Interactions such as; competitions may restrict the environments in which a species may live and referred to these more restricted conditions as the realized riche.

Understanding the mechanisms relating species richness variation to habitat heterogeneity is an important issue in ecology and conservation biology. Continuous areas of habitat have been progressively transformed into a patchy mosaic of isolated "islands" of available habitat as a result of human alterations (Holt *et al.*, 1995; Hanski 1999). Despite the ubiquity of these highly fragmented habitats and their implications for biodiversity, a lack of knowledge still exists on how community diversity varies from sites within large, contiguous habitat areas to those within smaller, fragmented areas.

Recent theoretical work has shown that the decline of species richness with habitat loss is a non-linear process, with species extinctions becoming more and more frequent as habitat continues to disappear (Tilman *et al.*, 1994; Stone, 1995). However, these studies did not use spatially explicit models, making it difficult to infer relationships between spatial patterns of habitat arrangement and species richness.

Maximizing species richness is often an explicit or implicit goal of conservation studies (May 1988), and current and background rates of species extinction are calibrated against patterns of species richness (Simberloff, 1986). Therefore, it is important to examine how ecologists have quantified this fundamental measure of biodiversity and to highlight some recurrent pitfalls (Gotelli and Colwell, 2001).

Although species richness is a natural measure of biodiversity, it is an elusive quantity to measure properly (May, 1988). The problem is that, for diverse taxa, as more individuals are sampled, more species will be recorded (Bunge & Fitzpatrick, 1993). The same, of course, is true for higher taxa, such as genera or families. This sampling curve rises relatively rapidly at first, then much more slowly in later samples as increasingly rare taxa are added. In principle, for a survey of some well-defined spatial scope, an asymptote will eventually be reached and no further taxa will be added.

Gotelli and Colwell (2001) distinguished four kinds of taxon sampling curves, based on two dichotomies (Figure 1.03). The first dichotomy concerns the sampling protocol used to assess species richness. For instance; if someone wishes to compare the number of tree species in two contrasting 10 ha forest plots, one approach is to examine some number of individual trees at random within each plot, recording sequentially the species identity of one tree after another.

Gotelli and Colwell (2001) referred to such an assessment protocol as individual-based (Figure 1.03). Alternatively, one could establish a series of quadrats in each plot, record the number and identity of all the trees within each, and accumulate the total number of species as additional quadrats are censused. This is an example of a sample-based assessment. The second dichotomy distinguishes accumulation curves from rarefaction curves. A species (or higher taxon) accumulation curve records the total number of species revealed, during the process of data collection, as additional individuals or sample units are added to the pool of all previously observed or collected individuals or samples (Figure 1.03).

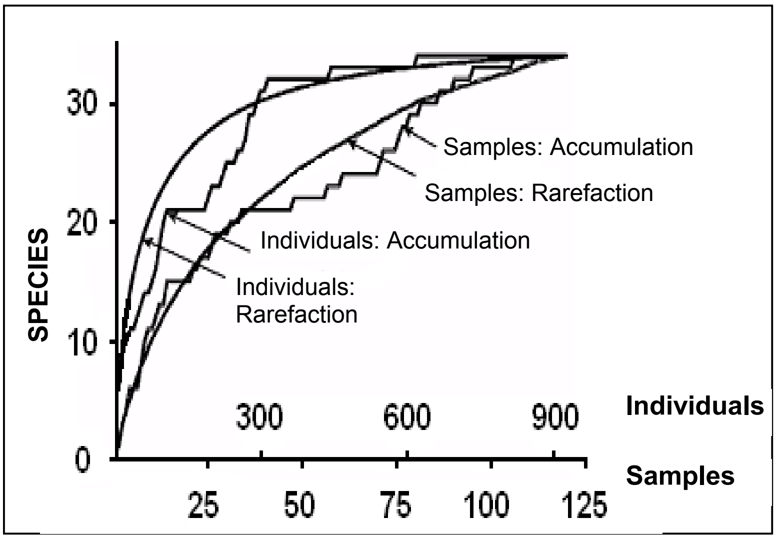


Figure 1.03. Sample and individual based rarefaction and accumulation curves (Gotelli and Colwell, 2001)

Accumulation curves may be either individual-based (Clench 1979, Robbins *et al.*, 1989) or sample-based (Novotny and Basset, 2000). In Figure 1.03, accumulation curves represent a single ordering of individuals or samples, as they are successively pooled. Rarefaction curves represent the means of repeated re-sampling of all pooled individuals or all pooled samples. The rarefaction curves thus represent the statistical expectation for the corresponding accumulation curves. The sample-based curves lie below the individual-based curves because of the spatial aggregation of species.

Quantification of diversity is essential for hypothesis generation and testing in ecology. Simply, quantification can be reached by counting the number of species. This is possible only to a limited extent. Not all members of the community can be counted, because they are unobtainable. Counting results may be corrected by rarefaction or jackknifing methods compensating sampling error. Furthermore, estimations of species number by means of regression or other statistical methods are possible. These estimations are based on partial sets of the total species number (correlates). Basically, this means that only a limited number of groups are sampled and total species number is extrapolated. The number of species is highly correlated to number of higher taxa such as families and orders (Wiegleb, 2002).

Instead of counting species number, several diversity indices have been proposed which represent diversity more exactly. Besides species number the frequency distribution of individuals is recognized. As all indices are calculated using species number there is a high correlation between species number and each index. There are two approaches to quantify species richness either empasis is laid on α-diversity or on evenness (Table 1.01). Both reflect different aspects of diversity (difference or similarity). Measurement is usually based on information theoretical indices, such as the Shannon-Wiener index, the Simpson index etc. (Wiegleb, 2002).

The problem of all indices of the α-diversity family is that they do not contain semantic information. Ecological interpretation remains difficult. Furthermore, different states of the system may result in the same numerical value of the index. Thus by condensing two levels of observation in one index, not only additional information is gained but also information is lost (Wiegleb, 2002).

Table 1.01. Measurements of biodiversity depending on two different approaches (Wiegleb, 2002).

MEASURE	**Subunit**	**Object (unit)**	**Pattern of interest**	**Pattern model for observation and analysis**	**Example**
α-diversity	Species, individuals	Community	Number and frequency distribution of species	Index based on information theory	Any Shannon-Wiener, Simpson, etc. index
Evenness	Species, individuals	Community	Number and frequency distribution of species	Index based on information theory	Any evenness index

1.4. GIS and RS applications

The demand for maps of specific themes of the earth`s surface, such as natural resources, has accelerated recently. Remotely sensed imagery has given the earth resource scientists -the geologist, the soil scientist, the ecologist, the land-use specialist- enormous advantages for reconnaissance and semi-detailed mapping. The resulting thematic maps have been a source of useful information for resource exploitation and management (Burrough, 1986).

Effective management of wildlife populations also depends upon understanding and predicting their habitat needs. For this reason, the use of multivariate statistics to asses habitat suitability has increased in recent years because the multidimensional nature of habitat limits use of simple univariate statistical techniques (Rexstad *et al.*, 1988).

Moreover, spatial modeling of ecosystems is essential if one`s modeling goals include developing a relatively realistic description of past behavior and predictions of the impacts of alternative management policies on future ecosystem behavior (Constanza *et al.*, 1990).

Development of these thematic maps and models has been limited in the past by the large amount of input data required and the difficulty of hardware. These two limitations have begun to erode with the increasing availability of remote sensing data and geographic information systems (GIS) to manipulate it, and the development of hardware systems that allow computation of large complex spatial arrays (Constanza and Maxwell, 1991).

Visualization is an important element of human intellectual capability because so much information can be gained by seeing rather than any other way. Graphic transformations appeal to our need for visualization and they provide a method to observe the environment in a new fresh way. In this way, our potential for creative and productive thought is enhanced (Muehrcke and Muehrcke, 1992). Therefore, maps are useful tools and support our understanding about our environment in a visual way.

Visual analysis and communication are historically essential elements of cartography that help to make visible facts and concepts. This is especially important for the information that might remain hidden otherwise (Muehrcke and Muehrcke, 1992; Taylor, 1994). Maps are a handsome blend of science and art, because of their genius of omissions and ability to present the essentials (Muehrcke and Muehrcke, 1992). According to Campbell (1993), maps are useful tools for exploring, organizing, and analyzing information about patterns on the Earth’s surface. They also give suggestions for a better understanding of the origins of those patterns. Moreover, they provide historical information that can be compared with recent changes and used for planning and assessment purposes (Campbell, 1993). Seeing whole scenes both in natural and spatial form is possible by maps. Therefore, maps capture the spatial character of our surroundings and help us to understand geographical relations more than any other communication device. (Muehrcke and Muehrcke, 1992).

Geographic data that contain the measurements of three-dimensional space and time have gained great importance for the studies of spatial and temporal relationships in landscapes recently. Scientific visualization supports the analysis and communication of these kinds of data (Brown *et al.*, 1995). Recent technological changes involving computer cartography and computer graphics has resulted in modern cartographic visualization that is different in both qualitative and quantitative ways. In quantitative terms a wide range of different cartographic products can be produced much faster and much more cheaply. In qualitative terms interaction with visual displays greatly increase comprehension in a wide variety of subject areas (Taylor, 1994).

Both digital mapping and quantitative analysis of spatial phenomena require that spatial objects (or geographic units) be quantified and evaluated based on a clearly defined measurement. The measurement of attribute data is dependent on the nature of the variable under consideration and data availability. Variables are classified into four measurement levels: nominal, ordinal, interval, and ratio (Chou, 1997). Nominal level of measurements are qualitative and categorical. Therefore, a nominal value serves merely as a label or name, and no assumption of ordering or distances between categories is made. Ordinal level measurements are meaningful in terms of rank order in that each category has unique position relative to the other categories. Although ordinal level measurements can be coded alphabetically or numerically, the degree of difference (distance) between one category and another is not known. Interval level measurements are meaningful in terms of ordering and distance between categories, and they permit examination of the differences between phenomena, but not their proportionate magnitude such as; temperature (Fahrenheit or Centigrade). In interval level measurements, there is no meaningful starting point. For instance, 0^0 C is arbitrarily defined by the freezing point of water under specific conditions, and does not imply the absence of heat. The ratio level of measurement has all the properties of the interval level, with the additional property that the zero point is inherently defined by the measurement scheme (Chou, 1997).

Elichirigoity (1999) reported the views of United Nation's Global Resource Information Database, and stated that “Traditional access to environmental data—in shelves of reports and proceedings as well as in fast-aging maps and charts—no longer meet the demands of planners faced with a world in which the nature of environmental change is infinitely complex. With the development of computers that can handle large quantities of data, a global database is now possible.

”Geographic Information Systems (GIS), the application of computer graphic visualization in 3-dimensions, is a tool that expands our understanding on a variety of spatial relationships by graphically visualizing all those spatial data (Watson 1992; Habb 1995). GIS needs for visualization contain technological, conceptual, and evaluatory solutions that can be seen in three broad domains namely analysis, illustration, and decision-making (Buttenfield and Ganter, 1990). Research tools such as maps, 3-D surface plots and other visual presentations expand our perspective by selectively displaying an array of spatial information for analysis (Watson, 1992) and promote awareness through symbolization (Mayoraz *et al.*, 1992).

On the other hand there are some difficulties in mapping stage. Things which do not have easy boundaries, may set up some classification problems. Soils, for example, enter into some biodiversity databases as related to particular kinds of communities of flora and fauna. Different agencies within a nation often adopt different soil classifications, rendering the pooling of data complex. Gray (1980) discusses the problems of classifying something that does not break up into natural units. Researchers in different disciplines may have very different perspectives on soil.

Similarly, there are difficulties with the concept of communities in ecology. It is argued that there is no simple metadata solution to the problems of integrating information from multiple scientific disciplines in biodiversity and thus to problems of adopting GIS for producing maps of biodiversity. To the contrary, it is maintained that within this field, GIS needs to take account of and represent an irreducible ontological diversity in the many biodiversity databases being produced (Bowker, 2000).

There are also a set of intriguing problems that have arisen as scientists from many different disciplines have begun working together to attempt to build up a unified picture of global biodiversity useful for both basic scientific research and for science policy. In order to produce such a map, there must be some basic agreement about data standards, classification systems and spatiotemporal frameworks. It has been argued that such agreements are hard won and frequently involve necessary but difficult trade-offs between the needs of different scientific fields and governmental agencies (Bowker, 2000).

Bowker (2000) said that without using multiple classification systems, data storage and map production is impossible. Thus a representation of tropical rainforest cover might involve not only a classification of vegetation covers, but also the various surrogate measures (indirect means of sensing that can be used to infer cover). A major problem for mapping biodiversity globally is that while classification systems are about kinds of things (flora, fauna, communities etc), the world of biodiversity data is radically singular. Just as species can be endemic to very small areas, so too can data about species.

There is also the question of what kinds of frames to wrap the data in so as ultimately to produce useful maps. For example, in order to produce generalizations about biodiversity, it would be useful to be able to attach plant or animal communities to particular landscapes. However, landscape topography has proven very difficult to classify (Bowker, 2000).

Remote Sensing is the acquisition of data about an object or scene by a sensor that is far from the object (Colwell, 1983). According to Lillesand and Kiefer (1994), “Remote sensing is the science and art of obtaining information about an object, area, or phenomenon through the analysis of data acquired by a device that is not in contact with the object, area or phenomenon under investigation." Aerial photography, satellite imagery, and radar are all forms of remotely sensed data. Usually, remotely sensed data refer to data of earth collected from sensors on satellites or aircraft.

Remote sensing is a new technical discipline that expanded rapidly within the past decade, and adding greatly to our ability to perceive and analyze the physical, chemical, biological, and cultural character of the earth's surface (Strahler and Strahler, 1996). Satellite-based remote sensing has produced detailed images of essentially every square meter of the earth's surface. These images provide very useful information to ecologists, especially for landscape and geographic ecology (Molles, 1999). Many research studies showed that satellite based remote sensing can gather large amounts of data over large areas.

Comprehensive information on the distribution of species as well as information about spatio-temporal changes is required in order to design meaningful conservation strategies. Remotely sensed data provide a systematic, synoptic view of earth cover at regular time intervals, and have been successfully used for these kinds of purposes (Lubchenco *et al.*, 1991, Roughgarden *et al.*, 1991, Stoms and Estes, 1993, Debinski and Humphrey, 1997, Innes and Koch, 1998). Coupled with Geographical Information Systems (GIS), they can provide information about landscape history, topography, soil, rainfall, temperature and other climatic conditions, as well as about present day habitat and soil coverage—factors on which the distribution of species depend (Noss, 1996). Relationships between species distribution patterns and remotely sensed/GIS data, if known, can be used to predict the distribution of single species or sets of species over large areas (Debinski and Humphrey, 1997).

When we encountered remote sensing, the notion of resolution should be explained first. In remote sensing, there are four aspects of resolution. These four aspects can be summarized as spatial, radiometric, spectral, and temporal resolutions. Spatial and spectral resolution may be the most important for many scientific research in different disciplines. Spatial resolution points out the size of the smallest object that can be resolved on the ground. In a digital image, the resolution is limited by the pixel size. Spectral resolution, on the other hand, can be defined as width of a band on electromagnetic spectrum (Lillesand and Kiefer, 1994).

The amount of information in a remotely sensed image of a given area is determined by spatial resolution. If spatial resolution is too low, discrimination of object classes becomes difficult. If too high, intra-class variability may increase and classification accuracy decrease correspondingly (Meyer *et al.*, 1996). The ratio of spatial resolution to the size of the objects being classified (tree crowns, single plant species or patches of a species), plays a crucial role in deciding whether species separation during classification is adequate or not (Nagendra, 2001).

Information on the structural characteristics of a plant association can be gained with high spatial resolution. On the other hand, there is also loss of information on species type and abundance at this high resolution level. Consequently, the problems associated with pixel-wise classification of these high-variance images could be expected hereafter (St-Onge and Cavavas, 1997). McGraw *et al.* (1998) and Lobo et al (1996) reported that image segmentation into training fields comprising the objects mapped, is a convenient way to avoid these kinds of problems. An ideal spatial resolution would minimize within-object variance, while maximizing between-object variance (Meyer *et al.*, 1996). O'Neill *et al.* (1996) suggested that spatial resolution should be two to five times smaller than the objects of interest.

Trees are the individuals which could be directly mapped in species level by using remotely sensed data. Even with these fairly large plants, the spatial resolution required for identification is fairly high (Nagendra, 2001). Biging *et al.* (1995) conducted a study based on visual discrimination of RECON 1 (video imagery emulating LANDSAT-TM) to discriminate tree species. They concluded that pixel sizes of 0.5 m (at a 1:12 000 scale) are not capable of assigning individual tree crowns to species. Distinguishing conifer species from hardwood also proved difficult. Biging *et al.* (1995) reported that a spatial scale higher than 1:12 000 could be useful for this process.

According to Hyppanen (1996); spectral bands also affect the optimal resolution. In a boreal forest, for instance, local maximum variance was found 3 meters in the infrared and green bands, and 2 meters in the red band. Consequently, the optimal spatial resolution varies depending upon what one is classifying. In the electromagnetic spectrum, different species of plants respond differently to light (Verbyla, 1995). Theoretically, remotely sensed data of adequate spectral resolution can be used to distinguish different plant species, but identification of the appropriate spectral bands is a challenging task (Nagendra, 2001). Near infrared (Verbyla, 1995), middle infrared (Everitt *et al.*, 1987) and thermal infrared bands (Salisbury and Milton, 1987) have been strongly suggested for species discrimination. Near infrared data responds to green biomass, and is considered useful for differentiating species according to their foliage content (Nixon *et al.*, 1985; Taylor 1993). The middle infrared band responds to leaf water content, and has been demonstrated useful for the separation of succulent plants from non-succulents (Everitt *et al.*, 1987).

Inter-specific as well as intra-specific variability also affects the capability of high spectral resolution sensors to discriminate between plant species in spectral reflectance. Although a family of spectra can represent a particular species, all species are not easily separated using remote sensing Sunflower, alfalfa and corn, for instance, have very similar reflectance values in visible–infrared range of 0.5–2.5 mm (Price, 1994).

According to Allen and Starr (1982); all observations depend upon the scale of study which could be expressed by extent and grain. Extent refers to the size of the study area investigated, while grain is the resolution of the remote sensor. As extent increases, the level of detail (grain) can be maintained, given constraints on time, effort and money, will decrease, and vice versa. The amount of information that can be retrieved, on numbers of species or numbers of habitat types, critically depends on these factors.

A chief use of remote sensing data is in classifying the myriad of features in a scene (usually presented as an image) into meaningful categories or classes. The image then becomes a thematic map (the theme is selectable, e.g., land use; geology; vegetation types; rainfall). This is done by creating an unsupervised classification when features are separated solely on their spectral properties and a supervised classification when we use some prior or acquired knowledge of the classes in a scene in setting up training sites to estimate and identify the spectral characteristics of each class (Lillesand and Kiefer, 1994).

The relationship among the spatial resolution of the image, the size of the objects being classified, and the size of the individual components at a smaller scale is the most critical issues through classification process (Simmons *et al.*, 1992). Varying habitat patches in different regions are another important aspect through the classification process. Breininger *et al.* (1995) reported that habitat patches in Florida scrub range from 20m in width upwards. On the contrary, habitat heterogeneity is higher and patch size can be as small as 0.1 ha in Indian tropical landscapes (Nagendra and Gadgil, 1999). For an appropriate classification, Cohen *et al.* (1990) recommended the pixel sizes lower than 1 m for conifer canopies in the Pacific Northwest region of the United States, while Weishampel *et al.* (1994) suggested the pixel sizes lower than 24 m for a mixed conifer/hardwood forest located in Maine, USA.

Achard and Mayaux (2001) reported that global to regional datasets from coarse spatial resolution satellite sensors are increasingly becoming available. Land cover and vegetation extent maps are among the terrestrial products derived from such data. Whilst they provide a previously unavailable synoptic view of global situations, accuracy assessment methodological issues are often ignored when creating these digital databases or producing statistics from such sources.

Sader *et al.* (1990) reported that although primary forest (relatively undisturbed forest) could be identified with high accuracy using LANDSAT Thematic Mapper (TM) data, old secondary forest and disturbed primary forest could not be distinguished. According to Mausel *et al.* (1993); the different succession stages could be mapped using TM data. For instance, in the work of where three classes of secondary succession forests were mapped, 82–88% accuracy was obtained. TM data were also found to be useful to detect the presence or absence of under story vegetation for varying canopy closures. The spectral information provided by band 5 appeared as the key ingredient for classification (Stenback and Congalton 1990). However, detailed mapping at the level of species and forest type could not be reliable by using TM data (Apan, 1997). Wolter *et al.* (1995) used TM data and found bands 3, 4 and 5 particularly useful for separating forest from non-forest and stratification of forested regions into conifer, hardwood and mixed conifer–hardwood classes.

Helmer *et al.* (2000) conducted a research and used multi-date, LANDSAT Thematic Mapper (TM) imagery to map secondary forests, agricultural lands and old-growth forests in the Talamanca Mountain Range in southern Costa Rica. They intended to determine the feasibility of using Landsat imagery from multiple dates to map land use and forest successional stage at a single date in a mountainous tropical region for which limited reference data were available. Accordingly, they used several approaches. First, they examined various spectral indices for distinguishing land-cover classes, including spectral differences between image dates. Secondly, they sought to determine whether adding the dimension of temporal changes in spectral response would improve classification accuracy. They also tested the usefulness of merging spectral data from multiple decades for detecting land use and land cover, including successional forest. Finally, they applied these techniques to create a land-use/land-cover (LULC) map. In this study, Helmer *et al.* (2000) obtained 87% overall accuracy with a Kappa coefficient of 0.83.

Baugh *et al.* (2001) conducted a research to measure oak tree abundance near Oak Ridge in Tennessee, and used LANDSAT-TM data. By doing this, they aimed to estimate fluxes of isoprene in their study area. Accordingly, they performed a classification process by using multi-date, supervised classification techniques, and an iterative approach. They also collected field data from transects (size:10x10 m). Training sites were selected based on transect data, and ten vegetation classes were mapped. Baugh *et al.* (2001) used empirical vegetation emission data to estimate the isoprene flux from each of the vegetation classes. The resultant isoprene flux maps were compared with concentrations measured in the field, and a good correspondence was observed. Results from these landcover classifications are used as input for models that predict tropospheric ozone production and are used to investigate ozone control strategies.

Trisurat *et al.* (2000) used LANDSAT Thematic Mapper (TM) false colour composite (BGR-2, 3, 4) to produce forest cover maps in Thailand. They used supervised and unsupervised classification methods with different band combinations to discriminate vegetation types in the Khao Yai National Park. They reported that thematic classes derived from supervised classification produced higher overall accuracy than unsupervised classification. In addition, the combination of ratio bands R4/3, R5/2, R5/4 and R5/7 ranked the highest in terms of accuracy (65% for unsupervised and 79% for supervised) and the combination of bands 2, 3 and 4 gave the lowest (56% for both methods). They said that even within the limit of spectral information available in the image, the digital classification can improve the result of visual interpretation.

According to Nagendra (2001); most remote sensing studies of species diversity focus on land area, because animal species cannot be normally observed using remote sensors without high spatial resolution. Relatively few studies of animal species distribution have therefore been carried out using remotely sensed data. Most direct observations of individual plants or animals require them to be fairly macroscopic and slow moving. The vast majority of macroscopic sessile organisms on land are colonies of lichens or plants. Consequently, most remote sensing studies of vegetation concentrate on larger plants on land areas. Smaller organisms like mosses or fungi are usually too small to consider mapping. However, there are few exceptions such as moss and lichen cover on tundra areas which are usually present in the lower strata of vegetation.

Salisbury and Milton (1987) investigated leaves of 13 evergreen and deciduous species in the visible, middle infrared and thermal infrared range in laboratory conditions. Although, they did not observe species-specific differences in the visible and middle infrared ranges, they recognized peak intensities of leaves from the same species differed to the order of 0.1 mm in the thermal infrared bank. Differences between leaves of different species were much greater, of the order of at least 0.5 mm, thus enabling the correct identification of species. The applicability of these conclusions to in situ remote sensing of species is not clear because of the differences between the real world and laboratory conditions.

Except microwave imagery, most remotely sensed imagery is incapable of penetrating through the top canopy of vegetation to receive information about lower strata. For this reason, remote sensing provides little information on the lower strata, such as herbs or shrubs (Verbyla 1995). As a result, a majority of studies on plant species diversity using remote sensing have been carried out on tree species in the uppermost canopy (McGraw *et al.* 1998), or crops and weeds in fields where they form larger associations in the topmost strata of vegetation (Dietz and Steinlein, 1996; Atkinson 1997; Atkinson and Curran, 1997).

According to Price (1994); trees with broad leaf canopies may reflect most of the non-absorbed radiation back. Needle-leaf canopy trees, especially those with highly random needle orientation, scatter most of the received radiation in random and different directions. For this reason, the information reaching the remote observer is minimum, and the chance of discrimination of different species is low. Moreover, there are some uncertain variables such as; sunlight, dust, humidity as well as contributions from branches, bark and the underlying soil (Atkinson *et al.*, 1997). In practice, even data of sufficiently high spectral resolution, that collected at high spatial resolution, might not permit the accurate identification of virtually any plant species (Price, 1994).

Nagendra (2001) evaluated the potential of remote sensing for assessing species diversity. According to Nagendra (2001); studies of species distribution patterns using remote sensing could be categorized into three types. Those are (1) direct mapping of individuals and associations (mapping individual plants or associations of single species existing in relatively large, spatially contiguous units that can be distinguished using the remote sensor), (2) mapping of habitats using remotely sensed data, and predicting species distribution based on habitat requirements, (3) Modeling the relationship between species distribution patterns and remotely sensed data (Investigations of direct relationships between spectral radiance values recorded from remote sensors, and species distribution patterns recorded from field observations).

According to Nagendra (2001); nature of the land under consideration highly affects the specific suitability of these various techniques. Consequently, methods applied in the heterogeneous and species rich mountains of India may be very different from the relatively species poor semi-arid Kalahari. Direct mapping is applicable over smaller extents, for detailed information on the distribution of certain canopy tree species or associations. Estimations of relationships between spectral values and species distributions may be useful for the limited purpose of indicating areas with higher levels of species diversity, and can be applied over spatial extents of hundreds of square kilometers. Therefore, direct mapping of individuals and associations may be useful to map small numbers of dominant tree canopy species. Habitat maps appear most capable of providing information on the distributions of large numbers of species in a wider variety of habitat types. However, the studies mostly focused on plant species, and it is not clear whether species–habitat relationships will prove equally strong for mobile taxa like birds or butterflies. This technique is strongly limited by beta diversity, and it is applicable at the landscape scale of tens of square kilometers generally. The third major technique appears capable of differentiating between areas with different levels of species diversity, but it has not been demonstrated capable of delineating the distributions of large numbers of species. While it provides indicators for further data collection on the ground, relationships between spectral values and species diversity may have to be calculated afresh for each image, thereby reducing its generality. Furthermore, this technique has not been demonstrated in areas where vegetation is relatively heterogeneous, and species diversity large.

Landform parameters like elevation, slope or aspect are important input parameters for spatial analysis and modeling of vegetation distribution in mountain landscapes. In a complex system of site factors topography is the major (indirect) factor for vegetation distribution (Barrio *et al.*, 1997). Thus, topography creates a patchwork-like pattern of small scale habitats and realized niches within the ecological space. Besides natural environmental factors, natural disturbance and the history of human impact play a major role for the distribution of vegetation types (Tappeiner *et al.*, 1998).

In most cases, spatially referenced data on historic (and sometimes even recent) land use practice and disturbance frequency is lacking, corresponding information is difficult to achieve and data handling is often time consuming. Thus, the actual vegetation distribution is a result of the complex interaction of historic and recent environmental, human and disturbance factors. Even in a landscape which is marked by human impact the overall influence of topography on the distribution of vegetation types is indisputable. Digital vegetation information as well as landform and topography-dependent microclimatic conditions of small relief patches are commonly assessed and represented using GIS (Hoersch *et al.*, 2002).

Hoersch *et al.* (2002) indicated three main approaches of performing habitat analyses in current research literature. First approach relies on analyzing the relation between vegetation and direct influence factors which means a reliable spatial database based on measurement/mapping of climate and edaphic data, geomorphological processes and human influence factors. Reliable point data could be inter-/extrapolated by the use of digital elevation data using regression procedures for instance. This approach is often subject to incalculable error due to insufficient data quantity and also quality. Additionally, reliable spatial data on climatic variables covering whole areas of research are not likely to be gained in the near future. Second approach depends on analyzing the relation between vegetation and the entire set of site factors including direct and also indirect (i.e. topographic) environmental variables. Using statistical techniques often requires non-redundant data; considering the entire set of site factors in statistical analyses can therefore cause extremely unstable results. Furthermore, computing time is increased considerably and results are often hardly interpretable because of interacting variables. Third approach related to analyzing vegetation habitats, and it assumes that any direct influence factor can be indicated or parameterized by landform parameters. The spatially heterogeneous pattern of landform derivatives many variables such as slope, aspect, and curvature exert. Those variables not only have strong influence on the spatial distribution of irradiation, precipitation, air and soil temperature, soil water and nutrients, snow accumulation and winds, but also control geomorphic processes and human interference. Microclimatic conditions have to be indicated by the analysis of topodiversity which is one of the key factors increasing the habitat diversity of high mountain landscapes. The major advantage of this approach is that spatially referenced data on topography are available for large areas at different spatial resolutions, thus offering a much more reliable predictor database compared with direct climatic or edaphic site factors.

For the integration of diversely scaled data of in homogeneous sources a GIS based approach seems most appropriate (Blaszczynski, 1997), offering the possibility to handle and homogenize spatial data on vegetation, landform and natural environment in terms of orthorectification, georeferencing and spatially analyzing common structures and patterns. Remotely sensed data have been widely used for assisting in vegetation mapping in the last few years and have proved an effective tool. They offer the possibility of extrapolating mapping results, especially in large and hardly accessible remote areas (Kalliola and Syrjänen, 1991).

“The ecological space of a vegetation type corresponds to its fundamental niche; in contrast to the fundamental the realised niche is defined through interaction with other vegetation types; finally the geographic space of vegetation types or species is equal to its spatial distribution, caused by natural factors and human impact (Hoersch *et al.*, 2002)." The physical environment is often regarded as one of the most important factors controlling the spatial heterogeneity of the landscape in mountain areas (Bolstad *et al.*, 1998; Tappeiner *et al.*, 1998).

According to Hoersch *et al.* (2002); the analysis of habitat factors for the distribution of vegetation based on the analysis of landform characteristics is an important aspect to understand high mountain ecology. They conducted a study in the Western Alps (Switzerland), and followed a GIS and remote sensing based approach to produce different scale vegetation maps for the study area. Hoersch *et al.* (2002) reported that "As spatial information on site factors is commonly lacking in mountain areas, the use of a Digital Elevation Model (DEM) is a potential substitute for use in vegetation analyses, as it highly correlates with temperature, moisture, geomorphological processes and disturbance factors." Consequently, it is essential to analyse the capabilities of a DEM for indicating habitat conditions in a landscape characterised by high topodiversity and a patchwork of microclimatic habitats. Then, appropriate landform parameters could be derived, indicating temperature and moisture distribution, exposure towards wind, snow etc. Hoersch *et al.* (2002) analysed the overall influence of topography and landform on vegetation distribution by using contingency tables and principal components analysis. The lack of information on the human dimension remained some uncertainties in the interpretation of spatial patterns of vegetation in their research. Moreover, landform classification schemes separating the landscape into basic landform-elements only proved useful for characterising azonal, non-altitudinal vegetation classes.

Cohen *et al.* (2001) modelled forest vegetation attributes as continuous variables across western Oregon using a multi-image mosaic of Thematic Mapper (TM) data. They modelled four specific attributes by using regression analysis. Specific attributes were (1) percent green vegetation cover, (2) percent conifer cover, (3) conifer crown diameter, and (4) conifer stand age. Airphotos were used to derive reference data for the cover and diameter attributes, and existing agency polygon databases were used for stand age. Cohen *et al.* (2001) developed and applied a new method for regional mapping called applied radiometric normalization. The method involved development of a set of models for a centrally located 'source' scene which were then extended to 'destination' scenes (neighboring scenes in the TM mosaic). According to Cohen *et al.* (2001); use of airphotos and existing digital databases in combination with applied radiometric normalization translates to a cost-effective procedure for regional mapping with TM data. Cohen *et al.* (2001) concluded that "Modelling forest attributes as continuous variables enables creation of a flexible forest cover information base, containing important fundamental building blocks for a variety of related classification schemes."

Advanced Very High Resolution Radiometer (AVHRR) composite data from National Oceanic and Atmospheric Administration's (NOAA) satellite data have been used to monitor agricultural lands of Turkey since 1998 (Dogan, 2000a; Doğan *et al.*, 2000; Doğan, 2001). The Geographic Information Systems and Remote Sensing Department of the Central Research Institute for Field Crops, that was established and funded by the World Bank and Global Environmental Facilities in 1997, has ground stations to receive AVHRR and METEOSAT satellite data. Vegetative indices, especially the Normalized Difference Vegetative Index (NDVI), have been used to analyze AVHRR data. Processing and analyzes of these data have been conducted by using Land Analysis System (LAS) software. The research areas of this department can be summarized as monitoring crop vigour, meteorological data overlay, the difference image with no crop mask, satellite images as raster-based grids, and area sampling frame overlay. NDVI image masks have supplied the opportunity to focus on the vegetation indices of specific crops. Consequently, focusing crop areas have been distinguished from other different areas that were excluded. NDVI composite images that contains 15 days period supplied very valuable information to detect the unusual drought periods that can be important for field crops.

Doğan (1998, 2002b) carried out a study to investigate the desertification process in northern Chihuahuan desert in New Mexico, Las Cruces. He used archival plant data collected from the transects in the field, and employed statistical analyses and GIS visualization procedures to display the changing conditions of the Chihuahuan Desert Rangeland Research Center from 1982 to 1993. According to Doğan (1998 and 2002b); results denoted a proof of the notion that visualization is a substantial tool for monitoring range condition. Results also indicated the dramatic mesquite (*Prosopis glandulosa*) change between the years (1982 and 1993) and some statistically significant relationships between the dependent and independent natural variables.

Doğan *et al.* (2000) conducted a study to determine and classify the rangeland areas in Kargalı and Gölköy pilot areas in Ankara-Polatlı, Turkey. They investigated the relationship between the topography and rangeland distribution by using geographic information systems (GIS) and global positioning system (GPS). The results of this study pointed out the importance of coordinated field data and digital elevation model. Basically, important relationships between the topography and land-use were detected to develop a model.

1.5. Plant biodiversity in Turkey

Turkey is a center of origin and still a source of important genetic diversity for numerous globally important agricultural, horticultural, medicinal and ornamental plants. Diverse geological and climatic conditions have given rise to a number of unique species represented nowhere else in the world. According to Doğan (1998, 2002); the first researches about Turkey`s flora were started by the visits of botanists from western countries at the beginning of 18th century. During that period some studies were conducted in different parts of the country, and collected plant material were stored in several institutes in developed countries (mostly in Europe). These visits delineated the importance of Turkey`s flora, and attracted the attention of many researchers.

Davis and Edmonson (1974) gave detailed information about previous botanical studies in Turkey starting from 18th century and ending by the publication of "Flora of Turkey Vol. 1" in 1965. According to Davis and Edmonson (1974); all research studies within this period could be investigated in four main parts. The first part of studies were started by the visits of foreign scientists at the beginning of 18th century, and ended by the visit of Swiss botanist E. Boissier in 1842. The flora of Central and West Anatolia were investigated mostly during this period. The researchers who visited Anatolia for research purposes in this period can be summarized as; Tournefort, Sibthorp, Clarke, Fleishcher, Aucher-Eloy and Monbred, Grisebach, Jaubert, and Thirke. Well-known Swiss botanist C. Liune also examined the plant specimens collected within this period.

The second part of studies were started by the visit of Swiss botanist E. Boissier in 1842, and ended by the publication of "Flora Orientalis" in 1888. Within this period, many field works were conducted in Anatolia, and Boissier prepared the "Flora Orientalist" at the end of this period. Not only Turkey but also the broad area that contains the Balkans in the west, Pakistan in the east, Crimea in the north, Egypt and Arabia in the south was investigated within this period. Plant species in these regions were investigated in the modern scientific point of view for the first time. In this period, the researchers who contributed to botanical studies in Anatolia can be summarized as; Kotschy, Pınard, Koch, Nöe, Heldreich, Tchihatcheff, Clementi, Pavillon, Bourgeau, Hawsknecht and Stapf (Davis and Edmonson, 1974).

The period between "Flora Orientalis" (1888) and the Second World War constitutes the third part of studies. During this period, East and Southeast Anatolia were investigated mostly. Scientists who visited Turkey in this period can be summarized as Aznavour, Rechinger, Sintenis, Post, Handel-Mazetti, Nabelek, Schischkin and famous

Swiss botanist Bornmüller. The Flora of Aegean was published at the end of this period. (Davis and Edmonson, 1974).

After the Second World War, the fourth period studies took place. In this period, the detailed studies that were started by the English Botanist P.H. Davis in 1938 were important. Every regions of Turkey were investigated and documented in a very detailed manner in a 48-year period. Depending on these studies, the first volume of the “Flora of Turkey” was published in 1965. The study was completed by the publication of 10^{th} volume in 1986. For this reason, this period is called the “Flora of Turkey” epoch. Approximately 100 botanists who were recognized internationally took part of this work. Many Turkish researchers also contributed to this study, and their names can be summarized as Hikmet Birand, A. Rıza Çetik, K. Karamanoğlu, Hasan Peşmen, Haydar Bağda, Baki Kasaplıgil. Demiriz (1993) compiled the literature dealing with the flora and the vegetation of Turkey, and an additional supplement of the flora, Vol. 11, was published in 2000 (Güner et. al., 2000).

To understand the species richness of Turkey, someone should realize that the size and geographical position of the country. Davis (1965-1986) used a grid system that is based on two degrees of latitude and longitude as the primary division for the citation of specimens. According to the grid system, Turkey was divided into twenty-nine squares as it was shown in Figure 1.04. The country is also the meeting ground of three phytogeographical regions: Euro-Siberian, Mediterranean, and Irano-Turanian (Figure 1.04).

In Flora of Turkey (Davis, 1965-1988), 163 families 1225 genera and 8745 species were recognized and revised. The majority of them (8617) are angiosperm followed by ferns (86) and gymnosperms (23). By adding bryophytes, it is clear that there are approximately 9000 species in Turkey. Endemic species, confined to Turkey particularly, constitute 30% of the flora. Consequently, it can be said that approximately 3000 species are endemic in the country (Davis, 1965-1988; Davis, 1971; Doğan, 1998).

Figure 1.04. Grid system used in “Flora of Turkey”, and Phytogeographical regions (Davis, 1965-1986 and Davis, 1971) E.S.(CE/B): Euro-Siberian Region probably central European/Balkan province, E.S.(EUX): Euro-Siberian Region Euxine province, E.S.(Col.): Colchic sector of Euxine province, MED(T): Mediterranean Taurus district, MED (W.A.): Mediterranean west Anatolian district, IR.-TUR.(C.A.): Irano-Turanian central Anatolia, IR.-TUR.(E.A.): Irano-Turanian east Anatolia, IR.-TUR.(Mes.): Irano-Turanian Mesopotamia (NOTE: This map was produced in digital format by Hakan Mete Doğan in 2003.)

Turkey’s Plant Database (TUBIVES, 2003) contains very detailed information about the plant species of Turkey. According to the records of TUBIVES; 119 family, 553 genera, and 1350 species were recorded in the A3 region (Figure 1.04) where the study area is located.

The reason of high endemism in Turkey is hidden in its geographic position. Turkey is a country that is isolated by surrounding seas (Blacksea in the north, Marmara in the northwest, Aegean in the west, and Mediterranenan in the south) in three sides, and not affected by the ice ages for a period over 40 000 years.

Turkey contains a rich array of plant communities and is especially rich in a diversity of forest ecosystems. Approximately 26% of the land base of the country is covered by forests (Konukçu, 1998). These forests have been valued primarily as sources of traditional wood products such as lumber, wood fiber and fuel wood.

Often overlooked is the fact that these forests, both closed canopy and steppe types, are the home of numerous relatives of woody crop species including walnut, chestnut, almond, pistachio, pear, apple, olive, hazelnut and many other fruit and nut crops. They are also rich source of other products including food, valued medicinal products, and special products such as resin. On the other hand, intense human activities have seriously threatened this high biodiversity as in the many countries in the world.

1.6. Justification of study site

According to the Flora of Turkey (Davis, 1965-1988); the location of the study area is found in the A3 grid square of Central Anatolia (Figure 1.04). The area is also assessed in the Irano-Turanian phytogeographical region with some Mediterranean penetrations (Davis, 1971). Irano-Turanian territory is the most important considering the species richness of the flora. On the other hand, its vegetation has been heavily used, and recently exterminated through extension of dry farming (Zohary, 1971).

The following criteria have been considered for the selection of this site. First of all, this place is a convenient sampling area because of its identical plant cover and diverse species characteristics. The study area also shows diverse geographical characteristics between 144 and 1740 meters. Moreover, proximity to Ankara (capital) supplies some advantages such as; easy and fast access to the study area, saving money from the transportation expenses, and gaining more time for the detailed field studies. Availability and easy access of the complementary data (maps and remotely sensed images) of the area constitutes another advantage, and reduces the costs.

1.7. Previous works in the study area

Akman (1974) conducted a study in the area, and determined 616 different plant species belonging to 72 families in Beypazarı Nallıhan region. The proportions of those species according to families were given in Figure 1.05-a. Leguminosae is the family that constituted the highest proportion. Gramineae, Compositae, Cruciferae, Labiatae, Caryophyllaceae, and Rosaceae are the families that follow Leguminosae. The proportions of the species corresponding to their phytogeographical regions were also determined and given in Figure 1.05-b. Results showed that 20.3% of the species were the Mediterranean element, while 17% of the species were Irano-Turanian (Akman, 1974).

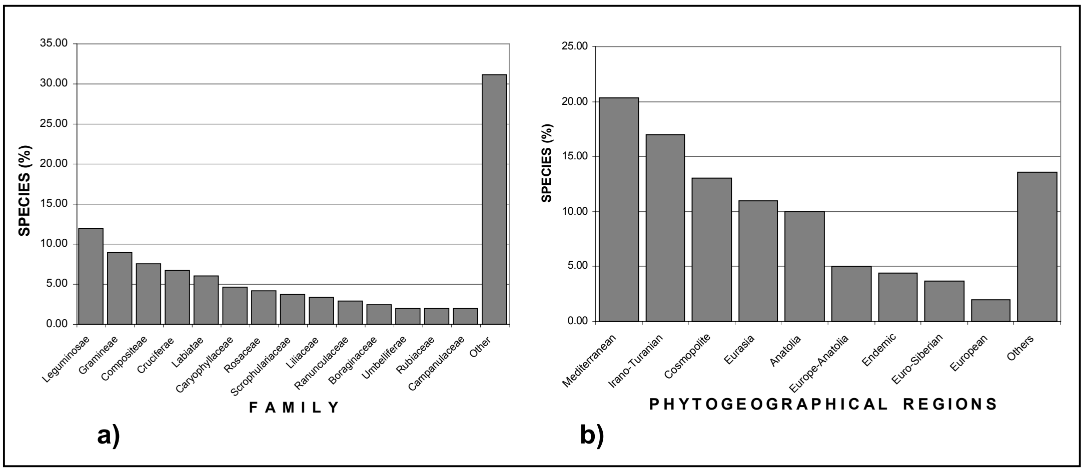


Figure 1.05. General view of the plant species in the study area; a) the proportions of species according to families, b) the proportions of the species corresponding to their phytogeographical regions (Akman, 1974).

Türker (1990) carried out a study in the area that is surrounded by Ayaş, Güdül, Beypazarı and Polatlı counties between the years 1989 and 1990. He recorded 313 taxa that belong to 41 families in the study area that is covered by steppe vegetation.

Güner (2000) studied Doğandede Hill that is located between Beypazarı and Çayırhan provinces of Ankara. He collected 598 plant specimens from the area, and identified 59 families, 184 genera and 302 taxa which are specific and infraspecific rank. He reported that 86 species and infraspecific taxa are new records for A3 region, and 64 taxa are endemic (21,19%). According to Güner (2000); phytogeographical elements of the area as follows: Irano- Turanian elements 93 (30,79 %), Mediterranean elements 25 (8,27 %), Euro-Siberian elements 9 (2,98 %) and the others 175 (57,95 %) which are widely distributed, and their phytogeographical origins are unidentified.

Yılmaz (1996) conducted a study on Sarıçal Mountain (Nallıhan). After evaluating 464 specimens, she determined 58 families, 179 genera, and 321 species. Yılmaz (1996) reported that 53 species are new records for A3 square, and endemism rate was 9.9%. Phyto-geographical elements of the species can be summarized as Irano-Turanian 12.7%, Euro-Siberian 10.9%, and Mediterranean 9.3%.

Pazarlıkçı (1998) investigated the vegetation around Sarıyar Dam Lake, and identified 380 species and subspecies taxa from 700 plant samples. She reported that 69 of them are endemic to Turkey. According to their phytogeographical regions the numbers and rates of the species can be stated as follows; Irano-Turanian 109 (% 28.6), Mediterranean 33 (% 8.6), Euro-Siberian 10 (% 2.6), and unknown or cosmopolitans 228 (% 60.2). Pazarlıkçı (1998) concluded that 11 of the total 380 taxa are new for A3 square.

Aksoy (2001) studied the flora of Karakiriş Mountain (Seben-Nallıhan), and collected 1227 plant specimens. He identified 511 taxa that belong to 72 families and 291genera. Total 150 (29.35%) are new records for A3 region. Phyto-geographical elements of the species can be summarized as Irano-Turanian 134 (26.22%), Euro-Siberian 69 (13.50%), Mediterranean 40 (7.83%), and not decided 268 (52.44%). Recently, Doğan (2002) published a new species of *Acantholimon* from Nallıhan.

As far as possible, all previous works that were conducted in the nearby areas were summarized in Table 1.02. Except some important ecological and phytosociological studies conducted by Akman (1974) and Akman and Ketenoğlu (1976) in the areas which are so close to Nallıhan, majority of the studies have floristic characteristics. In another words, ecological and synecological studies are so few comparing to the other sub categories. Consequently, all previous works have indicated the high plant biodiversity of the region and necessity of geographical approach to plant biodiversity.

By the development of technology, especially in computer hardware and software, geographic information systems (GIS) and remote sensing became substantial tools in many disciplines. One example of these kind of studies is the National Biological Survey / National Park Service (NBS/NPS) Vegetation Mapping Program that is used successfully in many biological and ecological research in the United States (Grossman *et al.*, 2003). Therefore, it is supposed that geographical dimensions of vegetation studies will be increased very soon by the widespread usage of GIS and RS in Turkey.

Table 1.02. Previous works completed in the nearby areas of Nallıhan

Major Category	**Sub Categories**	**Study**	**Interest Area**
Plant Biology	Plant Ecology	Akman, 1974	Beypazarı-Karaşar (Nallıhan)
Plant Biology	Plant Taxonomy (flora)	Akman, 1974	Beypazarı-Karaşar (Nallıhan)
Plant Biology	Plant Ecology Plant Sociology	Erik, 1976	Karagöl
Plant Biology	Plant Ecology Plant Sociology	Akman and Ketenoğlu, 1976	Ayaş Mountain
Plant Biology	Plant Sociology	Akman and Ketenoğlu, 1978	Köroğlu Mountain
Plant Biology	Plant Taxonomy (flora)	Akman, 1979	Ayaş Mountain
Plant Biology	Plant Taxonomy (flora)	Akman and Yurdakulol, 1981	Semen Mountains (Bolu)
Plant Biology	Plant Taxonomy (flora)	Türker, 1990	Ayaş, Güdül, Beypazarı and Polatlı
Plant Biology	Plant Taxonomy (flora)	Yavuz, 1992	Kazan, Ayaş, and Çanıllı village
Plant Biology	Plant Taxonomy (flora)	Yılmaz (Özkan), 1996	Sarıçal Mountain (Nallıhan)
Plant Biology	Plant Taxonomy (flora)	Pazarlıkçı, 1998	Sarıyar Dam Lake
Plant Biology	Plant Taxonomy (flora)	Türe and Tokur, 2000	Yirce-Bürmece-Kömürsu and Muratdere (Bilecik-Bursa)
Plant Biology	Plant Taxonomy (flora)	Ocak and Tokur, 2000	Gülümbe Mountain (Bilecik-Turkey)
Plant Biology	Plant Taxonomy (flora)	Güner, 2000	Doğandede Hill (Beypazarı)
Plant Biology	Plant Taxonomy (flora)	Doğan, 2000	Nallıhan Bird Sanctuary
Plant Biology	Plant Taxonomy (flora)	Aksoy, 2001	Karakiriş Dağı (Seben-Nallıhan)
Plant Biology	Plant Taxonomy (flora)	Doğan, 2002	Nallıhan

1.8. Scope and purpose

Diversity indices were supposed to a good starting point in this study. Of course there are many diversity indices that were developed to measure the species diversity in literature, but some of them such as; Shannon Wiener and Simpson have been widely accepted and used (Barbour *et al.*, 1987; Molles, 1999). Depending of their importance, both Shannon Wiener and Simpson diversity indices were chosen to conduct this study. By employing these two indices, it was also guaranteed to get a chance to make comparison between the developed models and to decide which one is better at the end of this study. Consequently, two distinct ‘plant species diversity maps’ were aimed to delineate current position of the area by using these diversity indices. For this purpose, dependent variables were defined as Shannon Wiener and Simpson diversity indices. Plant species data necessary to calculate both indices were decided to collect from the field. Determined independent variables to be investigated can be summarized as; (1) variables related to topography (elevation, slope, aspect), (2) variables related to climate (temperature, precipitation, potential evapotranspiration), (3) variables related to geology (geological formations), (4) variables related to soil (data from soil maps and field).

GIS and remote sensing tools are employed to classify and display all spatial data with grid maps. All of these visual products contain valuable data not only for exploratory analyses but also for the statistical analyses that are necessary for investigation of the relationships and developing models.

Principle component analysis and *multiple regression* statistics are decided for investigation and modeling relationships, respectively. Previously, this kind of data was not been examined by using statistical surface models in this area.

Another important goal of this research is to test the validity of the two models in order to make a correct interpretation. Residual maps and logical interpretations are chosen to get the idea about the validity of the models. Consequently, the objectives of this study are defined below.

Objective 1: To integrate all compiled data in order to analyze the relation between the dependent (Shannon Wiener and Simpson indices) and pertinent independent (topography, geology, soil, climate, and remote sensing classes) variables.

Objective 2: To develop a reliable model by using various (*multiple regression*) statistics for predicting biodiversity in a given area.

Objective 3: Run the developed models to produce ‘plant species diversity maps’ for both indices by using GIS, test the validity of the models that were produced for each indices, and compare the two indices according to the maps derived from models, and decide which one produced more reliable results,

Objective 4: Evaluate and interpret overall results and make some recommendations for future research possibilities to illuminate other researchers working in this area.

1.9. Organization of this thesis

This thesis is organized into seven chapters. Each chapter was briefly summarized below.

Chapter 1 puts forward the importance of plant biodiversity. Within this frame, current situation and problems encountered in the world and Turkey are explained. After giving a summary information about the previous works conducted in Turkey, the former studies completed in the study area are summarized and listed. Moreover, the scope and purpose of the study are briefly presented in this chapter.

Chapter 2 is explains the material and methods used in this study. A brief information about study area and applied methods were given in this section.

Chapter 3 deals with plant data. It explains the field studies starting from the first field surveys and ending field data collection. The issues pursued in this chapter can be summarized as initial field surveys, determination of sampling approach, collection of field data, and identifying plant samples.

Chapter 4 explains the complementary data that can be summarized as remote sensing, climate, geology, soil, topography and main vegetation characteristics. This chapter deals how these complementary data were created and used in order to develop models.

Chapter 5 deals the applied statistics and pertinent results. Results are given in a detailed manner after a background information about the applied statistical method. Two resulting maps produced by the application of developed models are also explained in this chapter.

Chapter 6 discusses the overall study under three headings. Firstly, the quality and reliability of all data sets are scrutinized. Then, the produced biodiversity maps depending on the developed models are questioned by checking residual maps and logical interpretations. Finally, some of the similarities and conflicts between the two biodiversity maps are questioned in this chapter.

Chapter 7 briefly puts the outcomes of this study.

CHAPTER 2

MATERIALS AND METHODS

This chapter contains the information about the study area and applied methods. First, the physiographic setting and general topography of the study area were shortly explained. Then, a brief information about the methods of the study was given. After explaining the methods of the study, a background information about the each methods were also supplied.

2.1. Study area

In a general view, Ankara province is located on the plateau of Central Anatolia, and great deal of this plateau lies between 800 and 1000 m altitude except various mountains.

Nallıhan administrative district is distinctive because of its natural beauty. Accordingly, it has some protected areas such as; 'Nallıhan Bird Paradise' and 'Hoşebe National Park'. The study area is specifically called Erenler forest region, and belongs to Nallıhan Forest Management District.

Erenler forest region is 138 km far from the city of Ankara (Capital), and located on the northwest of the city. Physiographic setting of the study area was given in Figure 2.01. The study area contains the forest management series called Sarıçal Mountain, Erenler and Kavacık, and totally covers 327.31 km^2 (32731.29 ha) area.

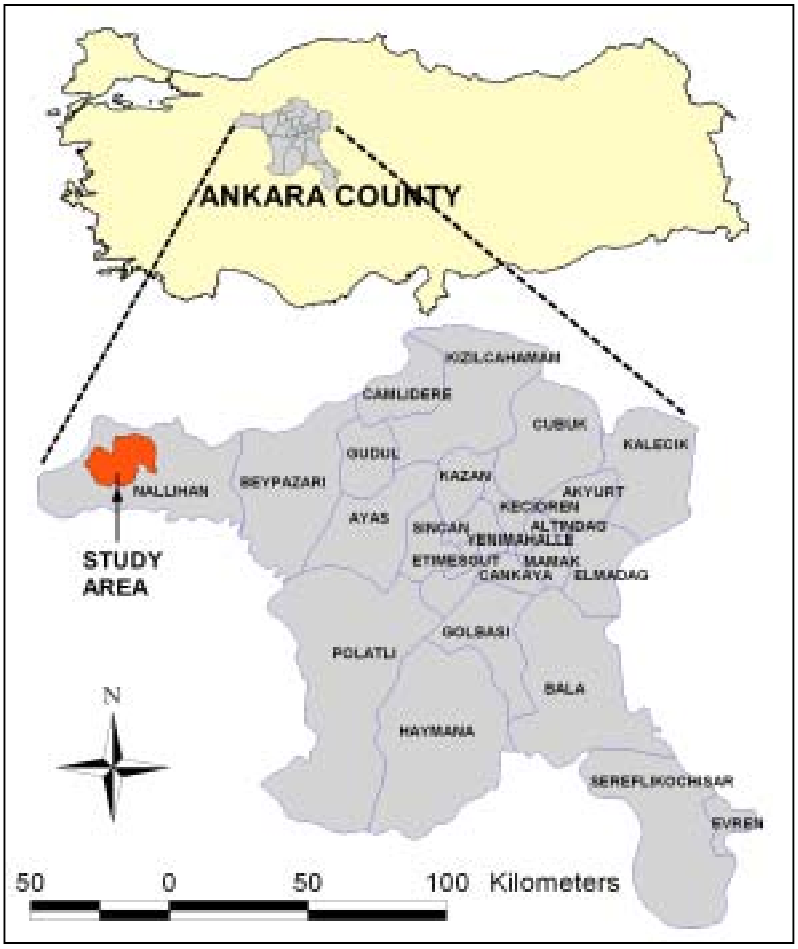


Figure 2.01. Physiographic setting of the study area

The general topography of the study area shows mountainous characters (Figure 2.02). The main mountain ranges can be summarized as; Sarıçal Mountain (1740 m), Epçeler Kayası (1559 m), Kemiklikaya Hill (1568 m), Karakuz Hill (1595 m), and Kavşak Hill (1554 m). Broad valleys were separated by the mountains. Naldere river, Karahisar and Karakaya streams constituted the main drainage with their branches.

Generally, agricultural lands are concentrated along the river basins, while forest dominated in higher elevations. The main settlement is the downtown of Nallıhan a, which is located outside of the Nallıhan Forest Management district's border. Inside the border, there are 28 settlements, and the majority of them are little villages. Irrigated agriculture and orchards can be seen along the rivers, while non-irrigated (dry) agricultural practices take place in the upland. Human effects on the forest can not be underestimated in the area. Moreover, a considerable area is under erosion. About 5.6% forest degraded by natural or anthropogenic causes in the area.

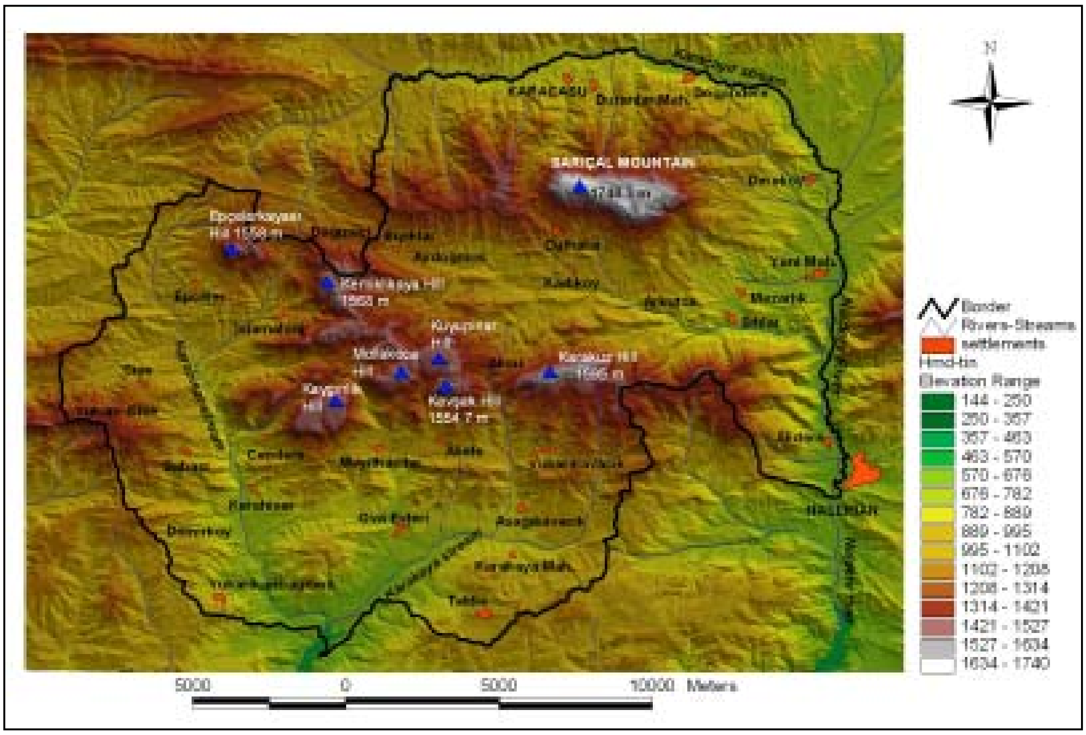


Figure 2.02. Physical geography of the study area (from triangulated network 'TIN' map)

2.2. Methods of the study

In general, the work done in this thesis can be summarized as (1) office work, (2) fieldwork, and (3) identification of plant specimens.

All office work were carried out in the GIS and RS Department of the Central Research Institute for Field Crops, Ankara. Basically, the office work contains the use of several UNIX and PC software developed for statistics, GIS and RS applications. Arc/Info software was used to digitize the paper maps, to correct the digitizing errors, to create topology, to make map projection transformations, to label geographic objects or entities, to determine the attributes, and finally to establish a reasonable digital database (ESRI, 1994; ESRI, 1997).

Some of the database, such as; climate, was produced in LOCCLIM software by evaluating the point data that consists of locations of climate stations and the elevation data from the digital database of topography (Grieser, 2002). A LANDSAT-TM image that belongs to 21 August 2000 was classified by using ERDAS-Imagine software (ERDAS, 1995; ERDAS, 1997). Then, all produced maps were converted to grid themes by using a standard grid size 30 x 30 m (ESRI, 1996). The software used in this study was listed in Table 2.01.

Table 2.01. The software used in the study

Software Name	Program Type	Using Purpose
UNIX-Arc/Info Version 7	GIS	Design and develop a digital spatial database
PC-Arc/Info Version 3.5	GIS	Design and develop a digital spatial database
PC-Arc/View Version 3.2	GIS, Spatial Analysis	Managing, analyzing, and representing spatial data
PC-Arc/GIS	GIS, Spatial Analysis	Managing, analyzing, and representing spatial data
ERDAS-Imagine 8.3.1	Remote Sensing	Image processing and classification
SPSS 11.0 for Windows	Statistical Package	Statistical Analyses, and presentation
SYSTAT 7.0 for Windows	Statistical Package	Statistical Analyses, and presentation
SIO Version 1.0	DOS	Downloading GPS data
Microsoft Excel	Worksheet	Entering field observations in dbf format
LOCCLIM	Climate Estimator	Derive climate data

The fieldwork was carried out for two distinct aims: (1) collecting point data to determine land-use/land-cover characteristics and necessary accuracy assessment, and (2) collecting quadrat data to determine species composition and richness. All required point and quadrat data were collected in the years 2001 and 2002, respectively.

The studies related to identification of plant specimens were carried out in the ANKARA Herbarium of Ankara University. With the contribution of specialists in Biology Department of Ankara University, all plant samples were identified in the herbarium. The Flora of Turkey and East Aegean Islands (Davis, 1965-1988) were used as the main reference throughout the herbarium studies.

The flowchart in Figure 2.03 summarized the methodology that was used in this study. Statistical analyses were conducted in four steps. Descriptive statistics were employed to explore the overall data characteristics as a first step. Central tendency (mean) and measures of dispersion (maximum, minimum, range, standard deviation, and variance) were used as descriptive statistics.

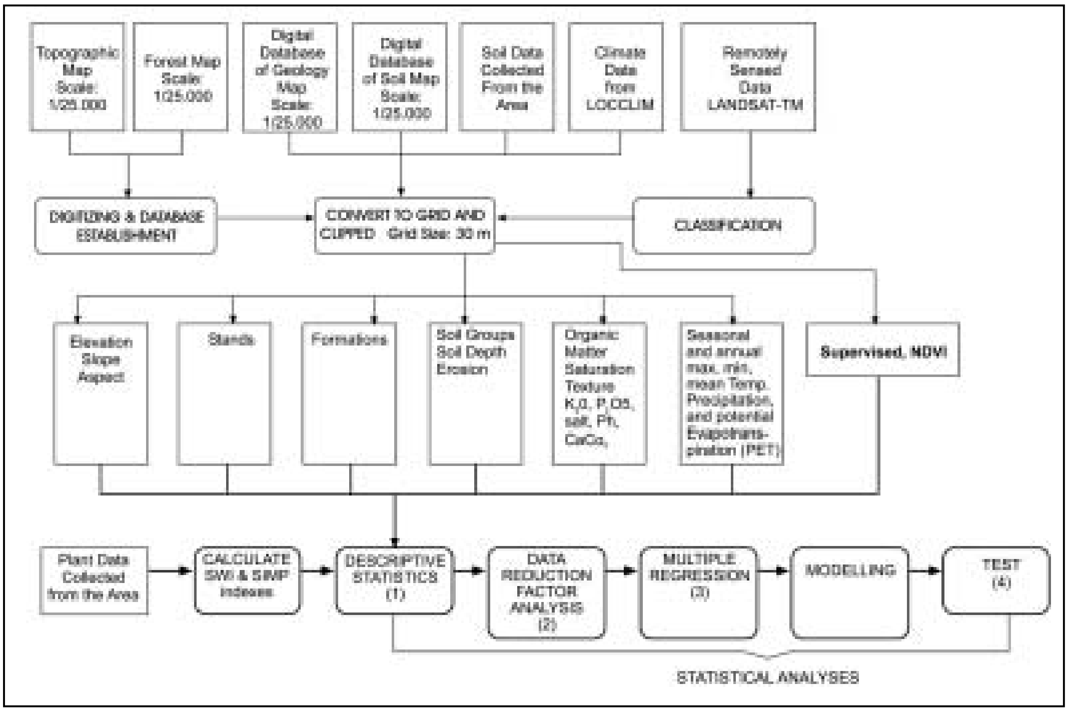


Figure 2.03. The flowchart of the methodology (the rounded rectangles indicate the analyses and processes, rectangles show output products)

Secondly, some statistics were applied for data reduction. For this aim, KMO-Bartlett tests were conducted for both Shannon Wiener and Simpson indices to test the suitability of the data for factor analysis. Then, *principle component analysis* with varimax rotation was applied for data reduction. In this way, redundant (highly correlated) variables were removed from the data file.

Modeling is the third step in this study. In this stage, *multiple regression* was used to formulate overall relationship between the dependent and independent (reduced) variables. After modeling, two distinct species richness maps were produced for Shannon Wiener and Simpson indices. The reliability of the maps was questioned in order to make a correct interpretation as the fourth step. Residual maps and logical interpretation were employed to check the validity of these maps. All details about the methodology were also given in Chapter 3 (plant data) and Chapter 4 (complementary data).

2.3. Background information about the applied methods

2.3.1. Building digital database

This is the most critical and time-consuming part of this thesis, because the completeness and accuracy of the digital database determines the quality of the analysis and final products. In general, building digital database can be summarized under two main headings; (1) establishment of digital database from paper maps (digitizing), and (2) creation digital database from coordinated point data (interpolating surface).

2.3.1.1. Establishment of digital database from paper maps

In this study, the methods developed by Environmental Systems Research Institute (ESRI) were employed in order to establish a reliable digital database (ESRI, 1995). The geology and soil maps had already obtained in digital format from the General Directorate of Mineral Research and Exploration (MTA) and the General Directorate of Rural Affairs (KHGM), respectively. For this reason, only the paper maps (1 /25 000 scale) related to topography and forest were digitized by following the methods defined by ESRI (1995). The method can be summarized under three headings, and these are (1) database design, (2) automating the data, and (3) managing the database.

Database design includes several stages such as; (1) determining the study area boundary, (2) deciding the coordinate system to be used, (3) selecting required data layers, (4) choosing features for each layer, (5) finding out the needed attributes for each feature type, and (6) determine how the attributes are to be coded and organized.

Automating the data also involves several steps. These steps can be summarized as; (1) getting the spatial data into the database (digitizing), (2) making the spatial data usable (verifying-editing errors, and creating topology), and (3) getting the attribute data into the database (entering the attribute data into the computer, and associating the attributes with spatial features). All these steps were completed by using Arc/Info (PC and UNIX) software (ESRI, 1994, 1995, 1997).

2.3.1.2. Setting up digital database from coordinated point data

Visiting every location in a study area to measure the concentration of a phenomenon is usually difficult or expensive. For this reason, the attribute (Z) values of strategically dispersed sample input point locations (X and Y coordinates) and *interpolate surface* techniques have been used to assign an estimated value to all other locations (ESRI, 1996). Input points can be either randomly or regularly spaced points containing different attribute values such as; soil saturation (%), air temperature (^oC), and precipitation (mm).

The resulting grid theme is the best estimate of what the quantity is on the actual surface for each location. The surface interpolators make certain assumptions about how to determine the best estimated values. Based on the phenomena the values represent and on how the sample points are distributed, different interpolators will produce better estimates relative to the actual values. No matter which interpolator is selected, the more input points and the greater their distribution, the more reliable the results (ESRI, 1996).

In this study, point data for the soil variables were created by using the analyze results of collected soil samples from the 56 quadrats in the study area. Moreover, point data for the climate variables were established by using the 576 selected point locations and LOCCLIM software developed by FAO (Grieser, 2000). Arc/View GIS (Version 3.2) software and Inverse Distance Weighted (IDW) method (ESRI, 1996) was used to produce grid maps of the soil (K_2O, P_2O_5, organic matter, pH, salt, $CaCO_3$, saturation and texture) and climate (temperature, precipitation, and potential evapotranspiration) variables.

Basically, IDW interpolator assumes that each input point has a local influence that diminishes with distance. It weights the points closer to the processing cell greater than those farther away. A specified number of points, or optionally all points within a specified radius, can be used to determine the output value for each location. More distant locations have less influence. The power parameter in the IDW interpolation controls the significance of the surrounding points upon the interpolated value. A higher power results in less influence from distant points (ESRI, 1996).

2.3.2. Air photo interpretation and image classification

In this study, both monoscopic air photographs with 1 / 35 000 scale and a LANDSAT-TM satellite image were utilized. Aerial photographs were used to make the visual interpretation in an accurate way, while LANDSAT-TM image was employed to classify the area according to the land-use/land-cover characteristics.

All air photos were combined in ERDAS-Imagine 8.3.1 software by using *photo mosaic* function and *rubbersheeting* method (ERDAS, 1995; ERDAS, 1997). The LANDSAT-TM image, belongs to 21 August 2000, was registered, geo-referenced, and rectified in the same software by using the available point data collected from the field. *Maximum likelihood parametric rule* and *4, 5, 3 band combination* were chosen for the supervised classification of the LANDSAT-TM image. Statistical filtering with 7 x 7 window size was employed to get a neat classified map from the satellite image. Essentially, the maximum likelihood algorithm assumes that the histograms of the bands of data have normal distributions. The maximum likelihood decision rule is based on the probability that a pixel belongs to a particular class. The basic equation assumes that these probabilities are equal for all classes, and that the input bands have normal distributions (ERDAS, 1995; ERDAS, 1997).

The accuracy assessment of the supervised classification was performed with ground truth data collected from 306 points throughout the study area. Overall accuracy with a Kappa coefficient was employed to check the accuracy of image classification. Basically, the Kappa coefficient expresses the proportionate reduction in error generated by a classification process compared with the error of a completely random classification. For example, a value of 0.82 implies that the classification process is avoiding 82 percent of the errors that a completely random classification generates (Congalton, 1991).

The Normalized Difference Vegetative Index (NDVI) model derived from the algorithm of *NDVI = (infrared – red) / (infrared + red)* was also used to get NDVI classes of the area. Basically, NDVI is sensitive to small differences between various vegetation classes that might be valuable for a model. For this reason, it was also employed in this study. The NDVI classes of the LANDSAT-TM image were determined according to unsigned 8 bit values between 0 and 255 in ERDAS-Imagine 8.3.1 software (ERDAS, 1995; ERDAS, 1997).

.

2.3.3. Field studies

Field studies established the most important part of this thesis, and took two years (2001 and 2002). In the first year of the study (2001), initial field surveys were conducted to understand the basic vegetation and land characteristics of the study area. A total of 791 point data, and 91 plant samples were collected within this period. Some of the point data (306) were also collected for accuracy assessment of supervised classification.

In the second year of the study (2002), detailed plant data necessary to calculate Shannon Wiener and Simpson indices were collected from the 56 established quadrats (20 x 50 m) in the study area. Total 752 plant samples were collected from 56 quadrats within this period. All field observations were taken with their geographic coordinates by using Global Positioning System (GPS). Details about the field studies were given in Chapter 3 (Plant Data).

2.3.4. Identification of plant samples

Total 752 plant specimens collected from 56 quadrats were identified by means of using Davis' Flora of Turkey and the East Aegean Island vol. 1-10 (Davis, 1965-1988). Ankara Herbarium of the Biology Department of Ankara University was used for this purpose.

2.3.5. Descriptive statistics

Measure of central tendency (mean) and measures of dispersion (maximum, minimum, range, standard deviation, and variance) were employed as descriptive statistics in this study. Histograms and boxplots were also used for descriptive statistics. All descriptive statistics were determined by using SPSS and SYSTAT software (SPSS, 2001; SPSS, 1997).

The arithmetic mean is undoubtedly the most widely used measure of central tendency. It is usually the most appropriate measure when using interval or ratio data. The arithmetic mean is defined as follows (McGrew and Monroe, 1993):

$$\bar{X} = \sum_{i=1}^{n} X_i \ / \ n$$

$\bar{X}$ is the mean of the variable, X_i is the value of observation i, and n is the number of observations.

Standard deviation is the most common measure of variability. Standard deviation (*s*) defined as (McGrew and Monroe, 1993):

$$s = \sqrt{\frac{\Sigma(Xi - \bar{X})^2}{n}}$$

The variance of a data set, defined as the square of the standard deviation (McGrew and Monroe, 1993):

$$s^2 = \frac{\Sigma(Xi - \bar{X})^2}{n}$$

2.3.6. Factor analysis

In this study, *factor analysis* (principal component method with varimax rotation) was employed to remove redundant (highly correlated) variables, and to determine important components before *multiple regression* statistics (modeling). Before applying *factor analysis*, KMO-Bartlett tests were conducted for both Shannon Wiener and Simpson indices in order to see the suitability of *factor analysis*.

Factor analysis was first invented nearly 100 years ago by psychologist Charles Spearman (Williams et al., 2003), and Harman (1976) prepared the well-known factor-analytic textbook called *Modern Factor Analysis* in 1976. Basically, *factor analysis* is a mathematical tool which can be used to examine a wide range of data sets. It has been used in disciplines as diverse as chemistry, biology, sociology, economics, and psychology.

Many statistical methods can be used to study the relation between independent and dependent variables. *Factor analysis* is different, because it is convenient to study the patterns of relationship among many dependent variables. The goal of factor analysis is to discover something about the nature of the independent variables. A typical factor analysis suggests answers to four major questions (Darlington *et al.*, 1973): (1) How many different factors are needed to explain the pattern of relationships among these variables? (2) What is the nature of those factors? (3) How well do the hypothesized factors explain the observed data? (4) How much purely random or unique variance does each observed variable include?

In particular, factor analysis seeks to discover if the observed variables can be explained largely or entirely in terms of a much smaller number of variables called factors. The main applications of factor analytic techniques are: (1) to reduce the number of variables and (2) to detect structure in the relationships between variables, that is to classify variables. Therefore, factor analysis is applied as a data reduction or structure detection method. Rotation is the step in factor analysis that allows to identify meaningful factor names or descriptions (Gorsuch, 1983; Morrison, 1990).

The *factor analysis* procedure constructs a solution with several extraction methods such as; unweighted least squares, generalized least squares, maximum likelihood, principle axis factoring, alpha factoring, image factoring, etc. *Principle component analysis* (PCA) is one of these methods(SPSS, 2001).

The *principal component* method of extraction begins by finding a linear combination of variables (a component) that accounts for as much variation in the original variables as possible. It then finds another component that accounts for as much of the remaining variation as possible and is uncorrelated with the previous component, continuing in this way until there are as many components as original variables. Usually, a few components will account for most of the variation, and these components can be used to replace the original variables (SPSS, 2001). The decision of when to stop extracting factors basically depends on when there is only very little "random" variability left. The nature of this decision is arbitrary.

2.3.7. Multiple regression

Modeling stage is simply a process to formulate the relationships for the estimation of plant biodiversity pattern in the study area. For this aim *multiple regression* method was employed in this study. "Regressing a variable *Y* on a series of independent variables could be done by successively regressing deviations. Generally, if we suspect several variables of being functionally related to *Y*, we try to regress *Y* on all of them simultaneously. This technique is called *multiple regression* (Sokal and Rohlf, 1995)."

There are two main purposes of *multiple regression* analysis. One is to establish a linear prediction equation that will enable a better prediction of a dependent variable *Y* than would be possible by any single independent variable X_j. The second is to estimate and fit a structural model to explain variation in the observations of *Y* in terms of the independent variables X_j. The *multiple regression* equation can be stated in two distinct but interrelated ways, the conventional and standardized form. The conventional *multiple regression* equation is defined as below (Sokal and Rohlf, 1995).

$$\hat{Y} = a + b_{Y1}.X_1 + b_{Y2}.X_2 + \ldots\ldots\ldots + b_{Yk}.X_k$$

The estimate of the dependent variable $\hat{Y}$ is a function of k independent variables $X_1, X_2, \ldots\ldots, X_k$. A coefficient such as b_{yj} denotes the regression coefficient of Y on variable X_j that one would expect if all the other variables in the regression equation had been held constant experimentally. In the standardized form, the variables are transformed to standard deviates by subtracting means and dividing by the standard deviation as below.

$$y' = (Y - \tilde{Y}) / s_y \qquad x'_j = (X_j - X_j) / s_{xj}$$

Consequently, the following equation is defined as standardized form of *multiple regression*. Here the coefficients b'_{Yj} are standard partial regression coefficients also known as b-primes, beta coefficients, or beta weights, and are in a simple relation to the convential partial regression coefficients b_{Yj}.

$$\hat{y}' = b'_{Y1}.x'_1 + b'_{Y2}.x'_2 + \ldots\ldots\ldots\ldots + b'_{Yk}.x'_k$$

CHAPTER 3

PLANT DATA

No doubt, plant data constituted the most important part of this thesis. This chapter deals with the field studies starting from the first field surveys and ending field data collection that were necessary to calculate diversity indices. Collecting plant data is not an easy process, and contains several important phases. For this reason, this chapter is divided to four sections as the initial field surveys, determination of sampling approach, collection of field data, and identifying plant samples.

3.1. Initial field surveys

Initial field surveys were planned to understand the basic vegetation and land characteristics of the study area. Identification of each feature of interest (main vegetation types, land use and land cover) constituted the first aim of these surveys. Locating representative areas of each feature to generate spectral signatures from the LANDSAT-TM image, and generating adequate additional data to test the accuracy of the image classification formed the second and third aims, respectively.

Appearance or physiological characteristics of plant species is the most significant feature in classification process within the world of GIS and remote sensing. Terms such as ‘forest’ and ‘grassland’ are physiognomic descriptions of the size and spacing of the main components of the vegetation, which may be qualified and subdivided at various levels. If these descriptions are based on general terms, classification and mapping process might be complicated. Therefore, widely applicable and descriptive systems based on plant physiognomy are needed in order to get neat and understandable products. Formation classes developed by UNESCO (1973) were employed to identify and classify the main vegetation types in the area (Table 3.01). These classes were also modified and used successfully by USGS-NPS (United State Geological Survey-National Park Survey) Vegetation Mapping Program (Anon., 1994). Not only vegetation but also land-use characteristics were identified and classified (Table 3.02) according to the standard terms and definitions developed by FAO (Anon., 1990).

By the application of these systems, it is assumed that image classification could be proceeded easily in a standardized way.

Table 3.01. Formation classes developed by UNESCO (1973).

Formation Classes	
I . Closed Forest (CF)	stand of trees > 5 m tall, with interlocking crowns
II. Woodland (W)	stand of trees > 5 m tall, not interlocking crowns but tree cover > 40%
III. Scrub	stand of caespitose woody perennials 0.5 – 5 m tall
• **shrubland (S)**	Scrub with crowns not interlocking
• **thicket (T)**	Scrub with crowns interlocking
IV. Dwarf scrub	Stand of caespitose woody perennials ~0.5 m tall
• **dwarf shrubland (DS)**	Dwarf scrub with woody perennials isolated or in clumps
• **dwarf shrub thicket (DT)**	Dwarf scrub with woody perennials with interlocking crowns
V. Herbaceous communities	
• **tall graminoid (TG)**	Stand of grasses or graminoids > 2 m tall
• **medium tall grassland (MG)**	Stand of grasses 0.5 - 2 m tall
• **short grassland (SG)**	Stand of grasses < 0.5 m tall
• **forb vegetation (FB)**	Stand of broad-leaved herbaceous species
VI. Deserts and other sparsely vegetated areas (DE)	
VII. Aquatic Plant Formations (AQ)	
- Each vegetation type in I-IV may be further qualified as Evergreen, Semideciduous, Deciduous, Xeromorphic - Each vegetation type in III-V may be further qualified as having Trees (> 5 m tall) contributing to 10-40 % of cover, Trees contributing to < 10% of cover, Shrubs, Tuft plants, No woody plants	

To achieve the goals, the first field survey was done on 29 June 2001. Coordinated point data were collected to determine the vegetation and land characteristics of the area. Basically, point data include some specific areas representing different characteristics of the land and ground control points. Some plant samples were also taken and pressed. Moreover, some digital photographs were taken to recognize certain characteristic areas. All observations were written in a detailed manner in a field notebook. After the first field trip, field studies were carried through fourteen additional days between 2 July 2001 and 5 August 2001. A total of 791 point data, and 91 plant samples were collected within this period. Some of the point data (306) were collected for accuracy assessment of supervised classification. All field observations were taken with their geographic coordinates by using Global Positioning System (GPS). In this study, a Scoutmaster GPS receiver (Trimble, 1993) was employed to determine the coordinates in the field.

Table 3.02. Some land use classes developed by FAO (Anon., 1990)

CLASS	CHARACTERISTICS	SYMBOL
Land Use Classes	Settlement/Industry	SE
	Forest Land	F
	Field Farming	FF
	Perennial Crops	PC
	Rotated Agriculture	CRF
	Crop Agriculture	CA
	Irrigated Agriculture	IA
	Mixed Farming	MF
	Rangeland/Grassland	RG
	Natural Protected Areas	PA
	Not Used	NU

After initial field surveys, basic vegetation and land characteristics of the study area were very well documented. All GPS coordinates that show the locations of sampled points were transformed to dBASE data format (dbf) by using SIO software (Trimble, 1994). Then, the field notes were entered manually in this dbf file in an organized way. Consequently, the worksheet data files that consist of XYZ values (Table 3.03) were created. Basically, creating this worksheet data file is the most important step in order to conduct GIS and RS applications.

According to results of the initial survey; the study area shows very distinct vegetation characteristics that is shifting from the closed forest to steppe. During the field observations it is realized that plant diversity is higher in the areas that show woodland (W), shrubland (S), and dwarf shrubland (DS) formation characters. In fact, these three formation classes can be in a simple group that is called degraded forest. Some of the digital photographs that summarize main vegetation types of the study area are illustrated in Figure 3.01.

Table 3.03. Structure of data worksheet files

GPS RECORDS (geographic)				Z VALUES				
No	**X**	**Y**	**DATE**	**Formation Class**	**Land Use Class**	**Plant Sample No**	**Photo No**	**NOTE**
1	31.2461	40.1697	29-Jun-01	W	F	-	-	….
2	31.1760	40.1943	29-Jun-01	W	F	1	1	….
….	….	….	….	….	….	….	….	….
….	….	….	….	….	….	….	….	….
….	….	….	….	….	….	….	….	….

Figure 3.01. Some vegetation characteristics of the study area; a) closed forest, b) mixed woodland, c) medium tall grassland, d) short grassland, e) shrubland, f) thicket (*Juniperus communis* and *Pinus nigra*).

3.2. Determination of sampling approach

"It is rarely feasible, or desirable, to census every individual in a community. Such a strategy would be prohibitively time-consuming and expensive; it would also damage or possibly even destroy the community in question (Magurran, 1988)." In most ecological and geographical research, sampling is an essential component. In fact, sampling is a necessity, if the population being studied is extremely large.

Sampling is an efficient and cost-effective method of collecting data, and can provide highly detailed information. It allows repeated collection of information quickly and inexpensively, and often provides a high degree of accuracy (McGrew and Monroe, 1993). Ecologists, therefore, rely on sampling to provide an accurate picture of community composition (Magurran, 1988).

Before collecting sample data a series of issues should be considered. First of all the definition of 'target population' and 'target area' is important. The target population is the complete set of individuals from which information is collected, whereas the target area is the entire region or set of locations from which information is gathered. After deciding target population and area, researchers should designate sampled population and sampled area within the sampling frame. Selection of sampling design is also important (McGrew and Monroe, 1993). According to Magurran (1988); sampling is a significant consideration in the studies of ecological diversity. Choosing correct sample size, defining the study area, and selecting the appropriate technique for measuring abundance are extremely important. Basically, the most important steps to develop an appropriate sampling approach can be summarized as (1) identifying parameters to be collected, (2) sampling design, (3) sampling size and intensity, and (4) the distribution of sampling sites. Each step was explained in a detailed manner under this section.

3.2.1. Identifying plant parameters to be collected

Before developing a robust sampling approach, the first issue is the determination of the measurement techniques, and the parameters required by these techniques. Ecologists define species diversity on the basis of two factors: species richness, and species evenness. The number of species in the community is called species richness, while the relative abundance of species is described as species evenness. Although there are many developed indices for species diversity by using these two important factors, Shannon-Wiener and Simpson indices have been mostly used (Barbour *et al.*, 1987; Molles, 1999). Consequently, both of these two indices were employed to measure the species diversity in each sampling unit (quadrat). The formulas of Shannon-Wiener and Simpson indices were given in Table 3.04.

Table 3.04. Shannon-Wiener (Molles, 1999) and Simpson Index (Barbour *et al.*, 1987)

Shannon-Wiener Index	**Simpson Index**
$H' = -\sum_{i=1}^{s} p_i \log_e p_i$	$D = \sum_{i=1}^{S} p_i^2$
H' = the value of the Shannon-Wiener diversity index p_i = the proportion of the i^{th} species loge = the natural logarithm of p_i s = the number of species in the community	D = the value of the Simpson diversity index p_i = the proportion of the i^{th} species s = the number of species in the community

To calculate both indices, parameters to be collected from the sampling units (quadrats) were identified as (1) species component, (2) number of species, (3) species cover (%), and (4) species density. The species component is the fundamental structure of a plant community. For this reason, a species list is an essential part of all vegetation studies. The number of species is defined as the total number of each species within a quadrat that is placed on the ground. On the other hand, the cover of a species is the proportion of ground occupied by vertical projection, and it is normally expressed as a percentage. For instance, the maximum cover of any one species is 100 percent. Density is the number of species in a unit area, and usually expressed as number / m^2. The content of field data was determined as in Table 3.05.

Table 3.05. The content of collected field data

FIELD DATA CONTENT	DESCRIPTION	LEVELS OF MEASUREMENT
- Plant Sample number in a quadrat	Number	RATIO
- Density of sampled plant in the quadrat	plants / m^2	RATIO
- Percent area covered by sampled plant	percent (%)	RATIO
-Total number of sampled species	Number	RATIO
-Formation Class of Quadrat	Symbol	NOMINAL
-Sampled strata	Tree/Shrub/Grass	NOMINAL

3.2.2. Sampling design

Sampling design refers to the way in which individuals or locations are selected from the sampling frame. Under certain circumstances, ‘spatial sampling’ is necessary. Spatial sampling is applied when a map of continuously distributed variable (such as vegetative cover, soil type, pH or surface water) is being selected from this map. If a researcher conducting a fieldwork must select sample site locations within defined target area, spatial sampling is also needed. According to McGrew and Monroe (1993); a spatial sampling from maps or other spatial sampling frames may involve point samples, line samples or area samples (Figure 3.02).

According to Dobrowski and Greenberg (2003); if the main goal is to determine the number of objects per unit area which is called density, quadrat sampling design is the best. Basically, a quadrat is plot of a fixed size in which density of objects can be measured. Ancillary information such as type of objects, size and shape can be measured as well. In this study, a quadrat sampling design was determined to collect the necessary field data.

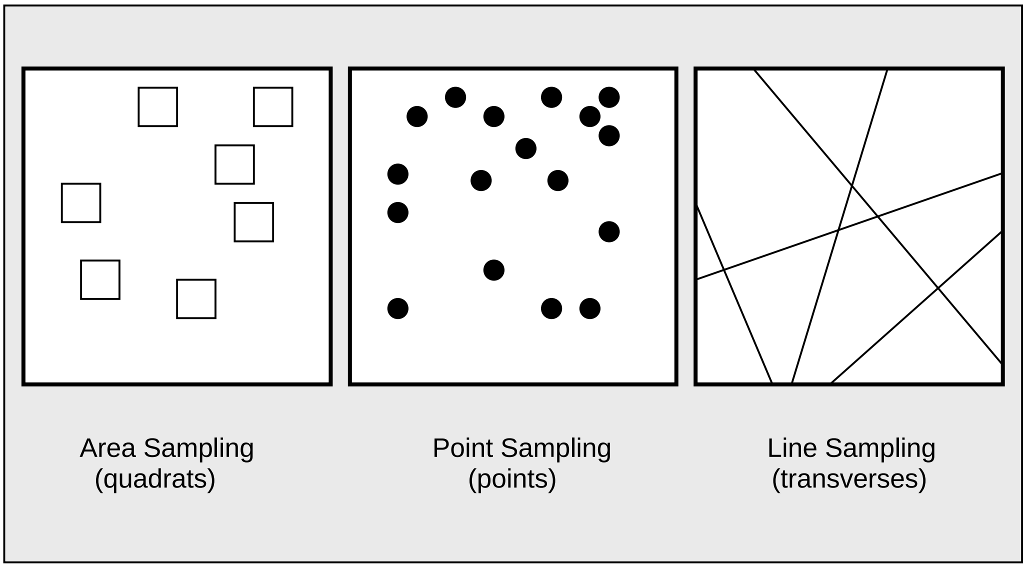


Figure 3.02. Spatial sampling designs (McGrew and Monroe, 1993).

3.2.3. Sampling size and intensity

After deciding the sampling design, sampling size should be defined. In general, a quadrat size up to 0.25 m^2 is suitable for herbaceous vegetation, while very much larger sizes are required for work with most woody species (Causton, 1988). According to Marriot *et al.* (1998); forest and woodland communities were sampled with 20x20 meter quadrats while herbaceous communities were sampled with 10x10 meter quadrats. In some instances rectangular plots of the same area, 10x40 m or 5x20 m, were used in linear or narrow polygons. Dolittle (2002) stated that the plots are usually circular or square in shape but can be other shapes as well. Studies have shown that rectangular quadrats are more accurate than square or circular ones because of the general tendency of vegetation to "clump". Rectangles should be laid-out up and down slope rather than across slope. Grossman *et al.* (2003) proposed guidelines (Table 3.06) for determining plot (quadrat) sizes in USGS-NPS Vegetation Mapping Program. Accordingly, a quadrat size was determined as 20x50 meter (1000 m^2) in this study. For dense herbaceous plant communities, a sampling frame (1x1 m) was used within the quadrats.

Table 3.06. Guidelines for determining plot (quadrat) size (Grossman *et al.*, 2003)

Class	Area	Dimensions
Forest	100 - 1,000 m2	10x10 - 20x50
Woodland	100 - 1,000 m2	10x10 - 20x50
Sparse Woodland	25 - 1,000 m2	5x5 - 20x50
Shrubland	25 - 400 m2	5x5 - 20x20
Sparse Shrubland	25 - 400 m2	5x5 - 20x20
Dwarf shrubland	25 - 400 m2	5x5 -20x20
Sparse Dwarf shrubland	25 - 400 m2	5x5 - 20x20
Herbaceous	25 - 400 m2	5x5 - 20x20
Nonvascular	1 - 25 m2	1x1 - 5x5

Sampling intensity however affects the species richness. Some researchers (Kirby *et al.*, 1986; Bunge and Fitzpatrick, 1993) reported that there is a positive relationship between the survey effort and number of recognized species. In another word, the more the sample is taken, the more the species is recorded. On the other hand, some limits must be defined in order to conduct a feasible fieldwork. In order to obtain a faithful idea about sampling intensity, it is better to look at the study that was conducted by Magurran (1981) in oakwood and conifer plantation at Banagher. Basically, Magurran (1981) conducted a quadrat survey in two contrasting woodlands: an oakwood which is a remnant of primeval forest and also a nature reserve, and a conifer plantation. The numbers of sampled quatrats and diversity values that belong to three indices (Shannon, Simpson, Berger-Parker) and number of species were plotted in X and Y axis, respectively (Figure 3.03). The diversity curves produced by the three indices level off at about 50 quadrats in both sites indicating that this is the minimum sample size on which a diversity estimate should be based. Consequently, it is assumed that the quadrat number between 50 and 60 could be enough to represent the study area well. For this reason 56 quadrats were established and sampled in this study.

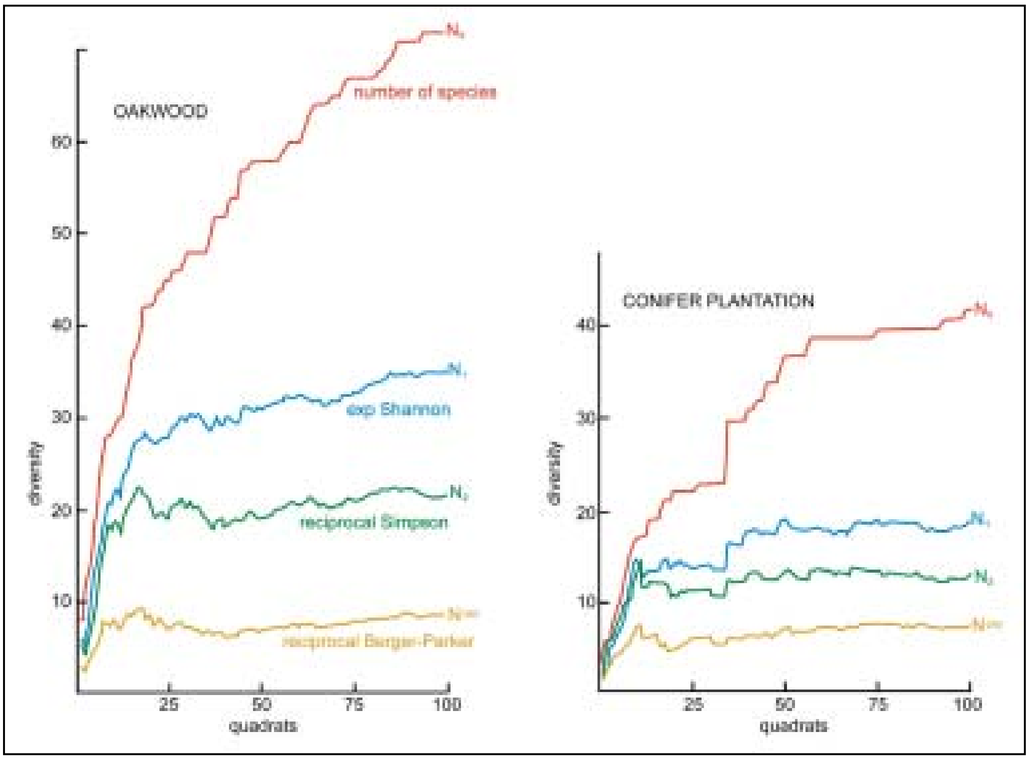


Figure 3.03. Diversity curves of indices in different quadrat numbers (developed for the oakwood and conifer plantation at Banagher, the point where the curve flattens indicates the minimum viable sample size). Magurran (1981).

3.2.4. The distribution of sampling sites

Whatever the sampling types, the distribution of them in the target area could be several types. There are some types of probability sampling which can be summarized as; simple random, systematic, stratified, cluster, and hybrid. If the study area consists of different strata (soils, geography, vegetation etc.), stratified sampling might be the most appropriate for modeling studies (McGrew and Monroe, 1993). To alleviate the defects of standard random sampling, stratified sampling schemes have been used to provide both accuracy and statistical validity. Stratified sampling divides a study area into compartments and locates samples randomly within compartments. This approach has been used successfully over large heterogeneous areas (Grossman *et al.*, 2003). Orloci and Stanek (1979) used a nested stratified random sampling design to characterize vegetation pattern in southern Yukon, Canada. The results of the study indicated that the selected stratifying variables accounted for a large part of the regional variation in vegetation.

In this study, a stratified random sampling design was used to investigate species richness. The selection of quadrat sites were primarily determined from the environmental stratification of the study area and the interpretation of the remotely sensed data (supervised classification, NDVI and air photo). In another words, the study area was stratified according to its topography, geology, vegetation and soil characteristics in order to find the best placement of quadrats (Figure 3.04). The availability of local expertise and collateral information also helped to determine how the sampling should be stratified, where it should be concentrated, and how much new information will be required to meet the objectives of this thesis.

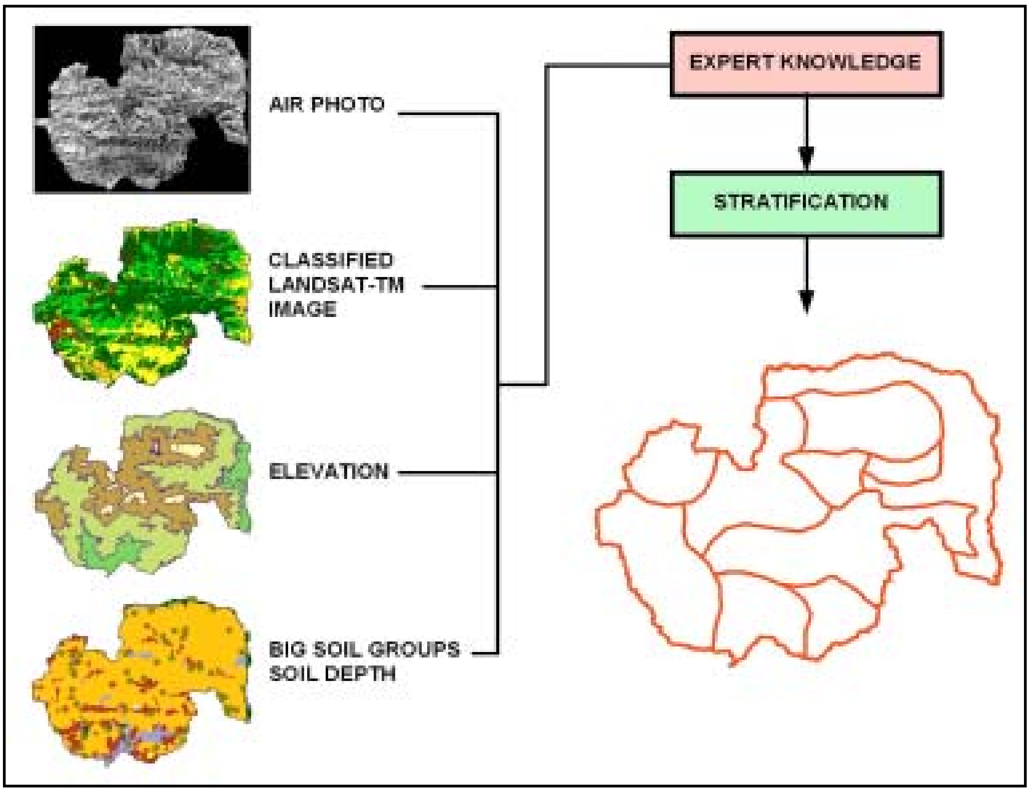


Figure 3.04. Stratification of the study area by using background information

3.3. Collection of field data

3.3.1. Establishing quadrats in the field

Field studies were programmed according to the important vegetation periods because timing and the quality of samples are highly effective on their identifications. Considering this important issue, field studies were conducted in the extended time period between mid April and early August in 2002. During that period, 38 field visits were carried out. Considering the previous stratification, quadrat sites were randomly distributed in these strata. In the field studies, total 56 quadrats were established. The locations of quadrats were determined by using GPS, and given in Figure 3.05. The geometric center of each quadrat was determined as the base point of each GPS reading. Quadrats were settled down according to the general rules explained by Dolittle (2002) and Grossman *et al.* (2003).

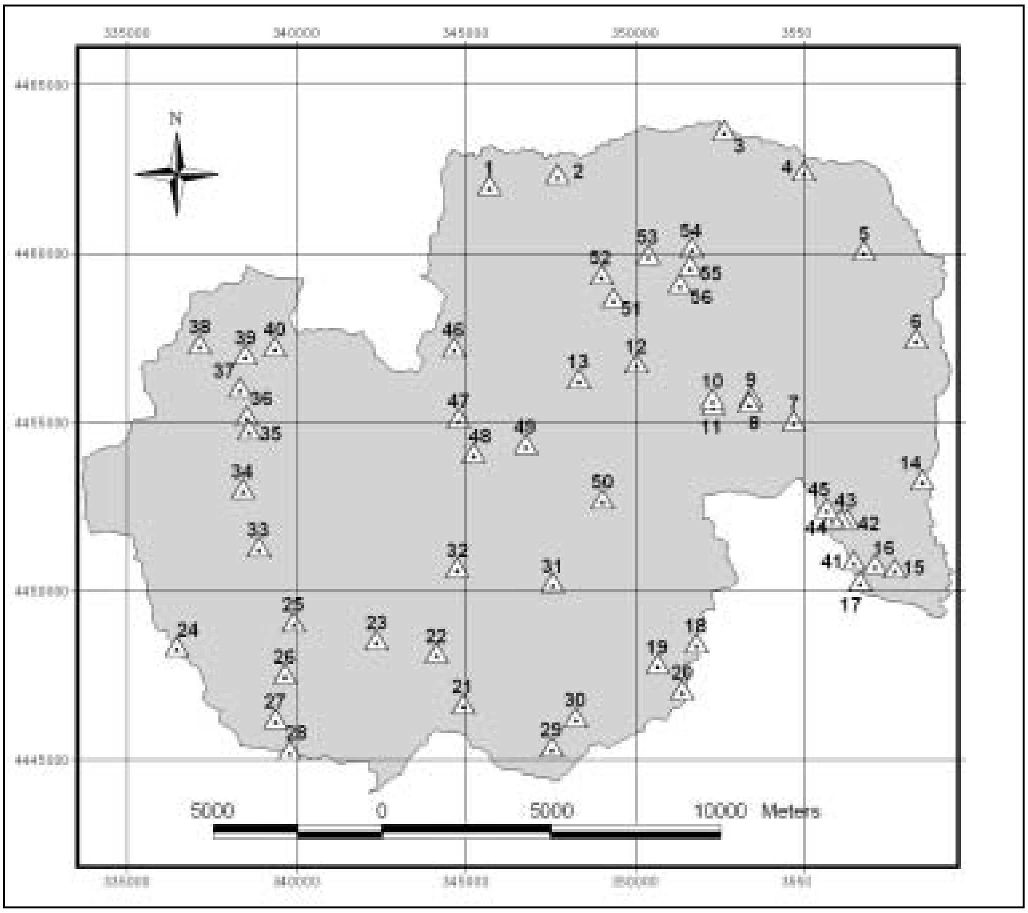


Figure 3.05. Locations of established quadrats

3.3.2. Measurements of plant parameters

Measurements of plant parameters in established quadrats constituted the most critical and laborious work in the field. After establishing a sampling quadrat in the field, the vegetation characteristics were examined carefully. The first issue is investigation of species component. After a close examination of the species composition, number of each species was determined and recorded prudently. Trees, shrubs, big annuals and perennials were counted directly. So the numbers of this kinds of plants were the result of direct counting. On the other hand, direct counting is impossible for the dense and small herbaceous plants such as some species in Gramineae family (*Aegilops umbellulata, Agropyron cristatum,* etc.). For this reason, a sampling frame (1x1 m) was used to measure the density of small plants in a quadrat. After deciding the density and cover values of small plants, their numbers can be calculated by using a certain formula. In the field, just cover (%) and density (number of plant / m^2) values were recorded, and their calculations were left to the office work in order to gain more time. Counting process were illustrated in Figure 3.06.

For classification purposes, the most common practice is estimation of cover in field. Cover is essentially a vegetation parameter that can be estimated by measurement, but it can also be estimated much more easily by eye. The actual measurement of cover in the field is extremely laborious. Estimating cover by eye to the nearest 1% is almost out of the question, but it is more feasible to the nearest 5%. However, if the cover of a species in a stand is estimated to be less than 5%, then a real attempt should be made to 'get it down' to the nearest 1%. If these recommendations are followed, then the visual estimation of species percentage cover is very useful (Causton, 1988). There are a number of 'scales' or ratings based on cover. Braun-Blanquet method developed by Josias Braun-Blanquet (1932) is the one that has been highly used especially in Europe. For this reason, cover classes and sociability scale of Braun-Blanquet (Table 3.07) were employed to determine cover and sociability parameters of each species in a quadrat.

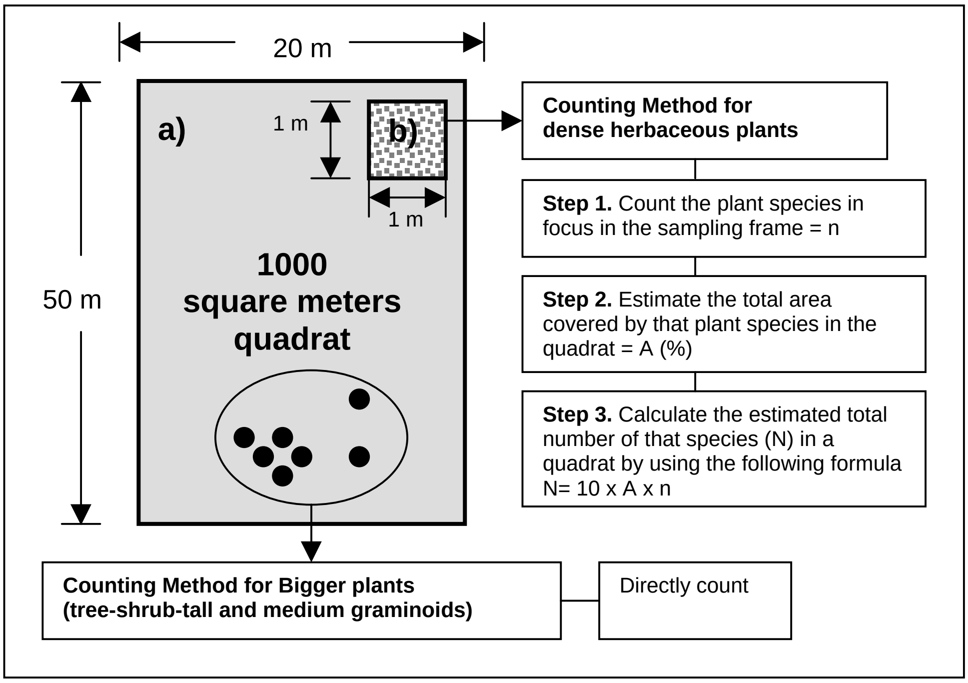


Figure 3.06. Establishment of a quadrat in the field and counting methods a) sampling quadrat, b) sampling frame for dense herbaceous plants to determine density in the sampling quadrat

Following the methods explained above, all measurements were taken and recorded in the field data notebook. Depending on the species composition and complexity, a time period between 4 and 8 hours was spent in each quadrats.

Table 3.07. Cover classes and sociability classes of Braun-Blanquet method (Tables were taken and modified from Barbour *et al.*, 1987)

Cover Classes of Braun-Blanquet		
Class	**Range of Cover (%)**	**Mean**
5	75-100	87.5
4	50-75	62.5
3	25-50	37.5
2	5-25	15.0
1	1-5	2.5
†	<1	0.1
r	<<1	*
* Individuals occurring seldom or only once; cover ignored and assumed to be insignificant		
Sociability scale of Braun-Blanquet		
Value	**Meaning**	
5	Growing in large, almost pure stands	
4	Growing in small colonies or carpets	
3	Forming small patches or cushions	
2	Forming small but dense clumps	
1	Growing singly	

3.3.3. Collecting plant samples

Identifying plants is not an easy task and to do it in a scientifically correct manner requires a high degree of specialization. It is usually not adequate to identify trees, for example, as merely "oaks." The term "oak" after all is simply a common name of one genus (*Quercus*) for which there are several species. In some cases, there are subspecies and varieties. Sometimes, the vegetation community is not so complex, and it is sufficient to record only the more common plants. On the other hand, plant samples that require a high degree of expertise should be collected in order to identify later by a specialist in a herbarium. Consequently, 752 plant samples, representative of their populations, were collected from 56 quadrats for correct identifications. Collecting such samples involves the use of a plant press. Plant presses in the standard size (40.64x 25.4 cm or 16x10 inch) were used for this purpose (Figure 3.07). The rules and definitions explained by Davis and Heywood (1965) were followed for a correct pressing and drying process.

Figure 3.07. Standard plant press (in size 40.64x 25.4 cm or 16x10 inch)

Each plant sample was labeled carefully and numbered. A unique number starting from 1 were given to each plant sample and this rule was followed throughout the field studies. In this way, it is prevented from some mistakes. Labels included the following information that is reported by Davis and Heywood (1965). Basically, the information content of the labels were determined as; (1) Number of plant (a straight serial number so that the specimen can be referred to merely by the collector's name and number), (2) name of plant (as far as known—even if only to family), (3) province or major divisions of the country, (4) locality (geographic coordinates from GPS readings), (5) altitude (GPS readings as meters), (6) habitat (include the type of terrain and preferably some indication of the community), (7) observations (any data may help the taxonomist in the herbarium; flower color, scent, duration and habit (biennial or perennial), height (tree or shrub), observed pollinators, etc.).

Most importantly, digital photographs were also employed to identify the plant species in each quadrat. Digital photographs have several advantages. First of all they can supply very detailed illustrations of plants that are valuable for their identification after fieldwork. Secondly, it is possible to use them as a digital image database. In this way a hot link between the maps and these digital image database can easily be established. When it is simply clicked on a point of a digital map, software displays the plant photo that is found in this area. Thirdly, digital photos are limitless and cheap. Researchers can take many photos and download easily in a laptop during fieldwork. In this way, they can organize some kind of digital herbaria that can be referenced easily in digital environment.

3. 4. Identifying plant samples

According to Davis and Heywood (1965), there are several ways open to us to identify the species, and four sources can be important in this process. Those are (1) floras, (2) monographs and revisions, (3) herbaria, and (4) illustrations. Monographs and revisions are only available for certain groups, and are sometimes only traced with difficulty. For practical reasons, the classification of the world's flora is primarily based on herbarium material and the literature associated with it. Despite its limitations, a herbarium has certain advantages over living collections. It is usually only in the herbarium that we can compare all the related species of a genus in the same place, in the same state and at the same time. It is possible, if we have access to a good herbarium, to "match up" the specimen directly with named herbarium material.

The Biology Department of Ankara University has a regional herbaria (Ankara Herbarium) that concentrates on Turkey's Flora. It is the biggest herbarium in Turkey and serves the dual purpose of research and identification. Contributes which are necessary for plant identification were also supplied by the specialists from the Biology department of Ankara University. Total 752 plant specimens collected from 56 quadrats were identified by means of using Davis' Flora of Turkey and the East Aegean Island vol. 1-10 (Davis, 1965-1988).

Total 239 species belonging to 45 family were determined. According to the records of Turkey's Plant Database (TUBIVES, 2003); 14 species were detected as endemic in the study area. Number of species recognized in each family is stated in Table 3.08. Leguminosae, Compositae, Labiatae, Rosaceae, Cruciferae, and Gramineae families have more species comparing the others. The full list of identified species was given in Appendix A, and endemic species were pointed out with special symbols.

Table 3.08. Number of species recognized in each family

Family	Number of Species	Family	Number of Species	Family	Number of Species
Leguminosae	37	Ranunculaceae	3	Iridaceae	1
Compositae	34	Cistaceae	3	Acanthaceae	1
Labiatae	28	Papaveraceae	3	Anacardiaceae	1
Rosaceae	15	Fagaceae	3	Chenopodiaceae	1
Cruciferae	10	Santalaceae	2	Convolvulaceae	1
Graminae	10	Illecebraceae	2	Coryllaceae	1
Liliaceae	9	Rhamnaceae	2	Crassulaceae	1
Boraginaceae	8	Geraniaceae	2	Equisetaceae	1
Scrophulariaceae	8	Linaceae	2	Euphorbiaceae	1
Caryophyllaceae	7	Berberidaceae	2	Globulariaceae	1
Umbelliferae	7	Cyperaceae	2	Guttiferae	1
Campanulaceae	5	Paeoniaceae	2	Malvaceae	1
Rubiaceae	5	Pinaceae	2	Orchidaceae	1
Cupressaceae	4	Valerianaceae	2	Polygalaceae	1
Plumbaginaceae	4	Dipsacaceae	1	Urticaceae	1

CHAPTER 4

COMPLEMENTARY DATA

Complementary data contain the information that is amenable to be represented in map form in order to predict plant biodiversity distribution pattern in terms of external influences. For this purpose, five categories of information traditionally have been useful. These data categories can be summarized as topography, climate, soil, land-use/land-cover, and vegetation. This chapter explains how complementary data sets were prepared to conduct further analyses and develop models.

4.1. Data preparation

In general, the main goal of the data preparation is to convert all complementary data set to grid themes (grid maps) in order to conduct spatial analysis and run developed models in GIS environment. A grid theme is a theme in which geographic data is stored in array of equally sized square cells arranged in rows and columns. Each cell has an attribute value defined as a piece of information describing a map feature.

In this study, the data incorporated into GIS were obtained from the graphic sources (existing maps in paper and digital format), spatial sources (photographs and remotely sensed image), and non-graphic sources (coordinated point data in tabular form).

Topography, soil, geology and forest maps with 1/25 000 scale were used as graphic sources. Only the maps related to topography and forest were digitized by using the Arc/Info (PC and UNIX) software. The geology and soil maps were obtained in digital format from the General Directorate of Mineral Research and Exploration and the General Directorate of Rural Affairs, respectively. Aerial photographs and a LANDSAT-TM image were used as spatial sources, and ERDAS-Imagine software was employed to process those spatial data. Finally, two sources of information were used as non-graphic data. These are the coordinated point data collected from the field and the coordinated point data derived from LOCCLIM software. Basically, the coordinated point data collected from the field were used to image classification and accuracy assessment, and to develop grid maps for some soil variables. The coordinated data derived from LOCCLIM software were employed to produce grid maps of climatic variables.

To conduct the spatial analysis and to develop a model, each grid theme (topography, climate, soil, land-use/land-cover, and vegetation) must be defined in the same coordinate system. Moreover, the grid sizes of all themes are also important, and they should be in the same size. In another words, the grid size and the map projections of all complementary data set must be the same for analyzing process. For this reason, a standard grid size (30 x 30 m) and Universal Transverse Mercator (UTM) projection were chosen to proceed. Basically, UTM projection system was developed by the U.S. Army in 1947, and it is currently used by the United States and NATO armed forces. With the advent of inexpensive GPS receivers, many other map users are adopting the UTM grid system for coordinates that are simpler to use than latitude and longitude. The UTM system divides the earth into 60 zones each 6 degrees of longitude wide. These zones define the reference point for UTM grid coordinates within the zone. UTM zones extend from a latitude of 80° south to 84° north, and they are numbered 1 through 60, starting at the International Date Line (longitude 180°) and proceeding east (Muehrcke and Muehrcke, 1992). According to UTM projection system, the location of Turkey extends from the zone 35 to zone 38, and the study area is located in zone 36 (Figure 4.01).

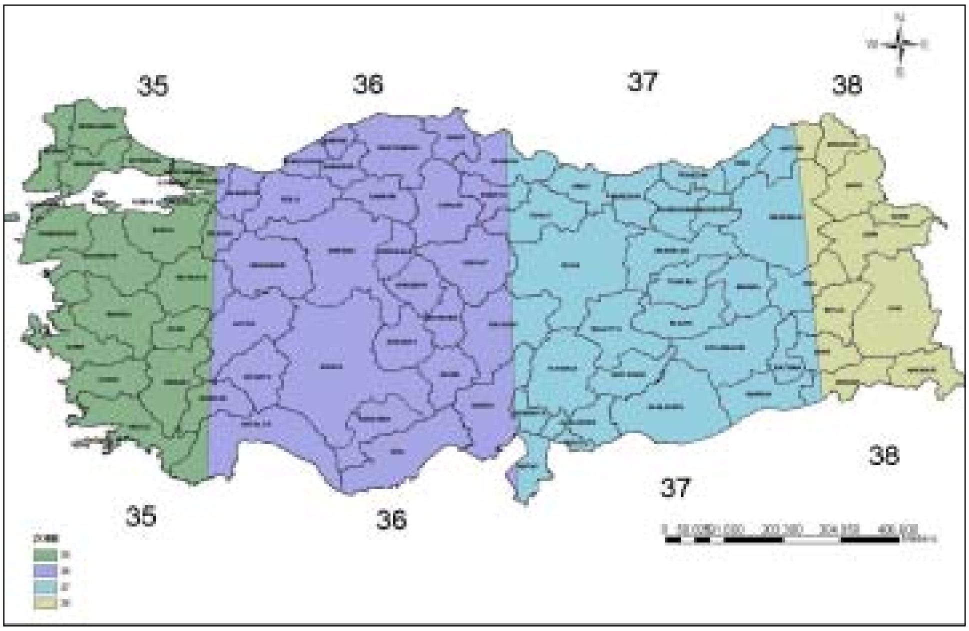


Figure 4.01. UTM zones of Turkey (starting from zone 35 and ending zone 38)

4.2. Grid map layers

4.2.1. Topography

Topography refers to the variation in elevation of the land surface on a broad scale. Topography pattern of hills and valleys, slopes and plateaus, influences climate, water drainage, soil erosion, and plant growth. Local climate is also driven by regional and local topography. Aspect, slope, and surrounding topography determine potential insulation. For this reason, south-facing slopes are warmer and north-facing slopes cooler at a given elevation. Topographic position and slope angle affect sediment accumulation and soil depth. These factors form complex environmental gradients that have profound effects on plant distributions. Consequently, topography is highly important to explain climate and plant distribution characteristics.

There is a complete data set about the topography of Turkey. The detailed topographic maps that contain whole country are currently available. Turkish Army (HGK:General Commandership of Mapping) has been producing detailed topographic maps of Turkey with 1/25 000 scale. Establishments or institutions can purchase these maps by following the certain process.

To develop a digital database of topography, total six topographic map sheets (H26: a4, a3, b4, c1, d1, d2) with 1/25 000 scale were selected from the map archive of GIS and Remote Sensing Department of Central Research Institute for Field Crops (CRIFC). Using these topographical paper map sheets, a digital database of the area was created. A triangulated network (TIN) map was also created to obtain the general idea about the topography (Figure 2.02 in Chapter 2).

The topography of the study area was investigated in three basic categories in this study. Those categories can be summarized as elevation, aspect, and slope. The main coverage is elevation. After producing a grid map of elevation by using digital topographic database, the grid maps of aspect and slope were derived in Arc/View software. The grid maps of elevation, aspect, and slope were given in Figure 4.02. In Figure 4.02, elevation classes were expressed as meters, and summarized in 16 classes with 100 m interval. Aspect classes were described in nine classes, and summarized as (1) North, (2) Northeast, (3) Northwest, (4) South, (5) Southeast, (6) Southwest, (7) West, (8) East, and (9) flat. Slope classes were defined as percent, and summarized in 18 classes with 5% interval.

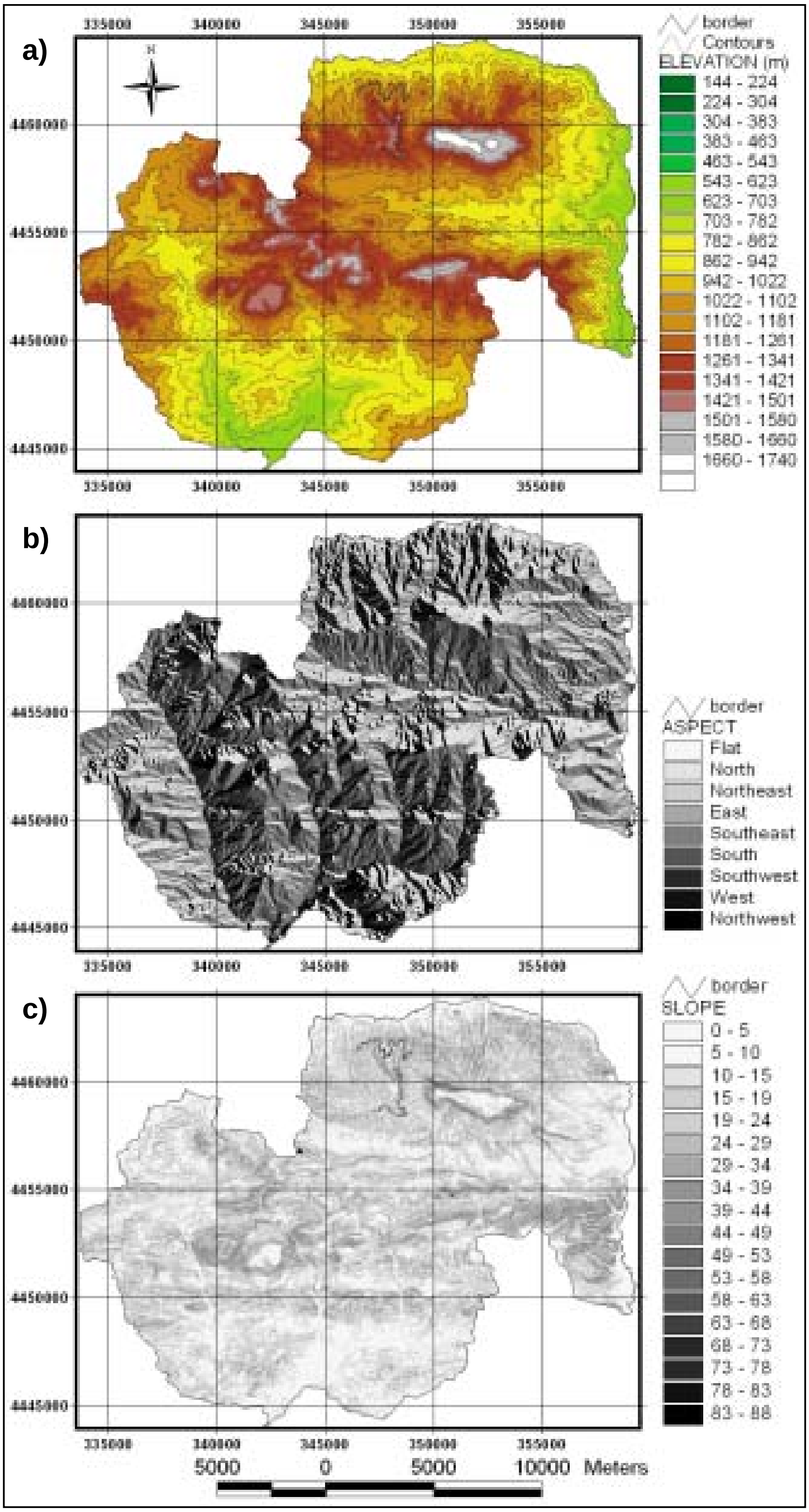


Figure 4.02. Topographical characteristics of the study area a) elevation (meters), b) aspect, c) slope (%).

4.2.2. Climate

The natural vegetation of cover is often a distinctive feature of a climatic region. Climate may be described in terms of individual parameters, such as temperature and rainfall, or by means of synoptic combinations of such parameters. Moreover, the development of soils, as well as the types of processes that shape landforms, is partly dependent on temperature and precipitation. The most relevant climate data, of course, are the annual, seasonal or monthly means, extremes, and ranges of temperature and precipitation for any given vegetation cover. According to Reid (1980); climate data should be long-term norms when someone study on plant populations.

Long-term data of meteorological stations are available in published form in a number of international accounts in addition to national publications. Food and Agriculture Organization of the United Nations (FAO), for instance, maintains a database of climate information (FAOCLIM) which contains climatological records from thousands of weather stations worldwide. FAO also developed a software program called LOCCLIM that is a powerful local climate estimator. The software estimates the climate for any location on Earth, and it is based on the database of the Agrometeorology Group in the Environment and Natural Resources Service (SDRN) in the FAO (Grieser, 2000).

LOCCLIM not only calculates an estimate for different climate variables at given locations; but also provides quite a variety of tools to improve the estimates and gives warnings in case of different kinds of pitfalls. Using the Inverse Distance Weighted Average (IDWA) approach, LOCCLIM offers an estimate of the climate at any location specified either by coordinates, or by a click on a map. For any given location LOCCLIM searches for the nearest stations that fulfill given criteria (absolute number, maximum distance, altitude constraints). If desired LOCCLIM fits a linear altitude function through the observations to reduce all of them to the elevation of the desired location. This minimizes the systematic error resulting from the different elevations of the neighboring stations. The altitude of the desired location can either be defined by the user or taken from a digital elevation model (DEM) with a spatial resolution of 10 km and an altitudinal resolution of 20 m (DEM downgraded from the NOAA/NCDC Global Land One-kilometer Base Elevation).

For the best estimation of climatic variables, total 576 sites were selected from the established grids on the study area (Figure 4.03). A climate database including the coordinates and altitudes of those grids were created in notepad text format which LOCCLIM software can recognize. Instead of LOCCLIM's DEM with a spatial resolution of

10 km, the DEM based on topographic maps (1/25 000 scale) was chosen (user defined) to increase accuracy. Then, best estimates for those grid coordinates were derived using LOCCLIM software.

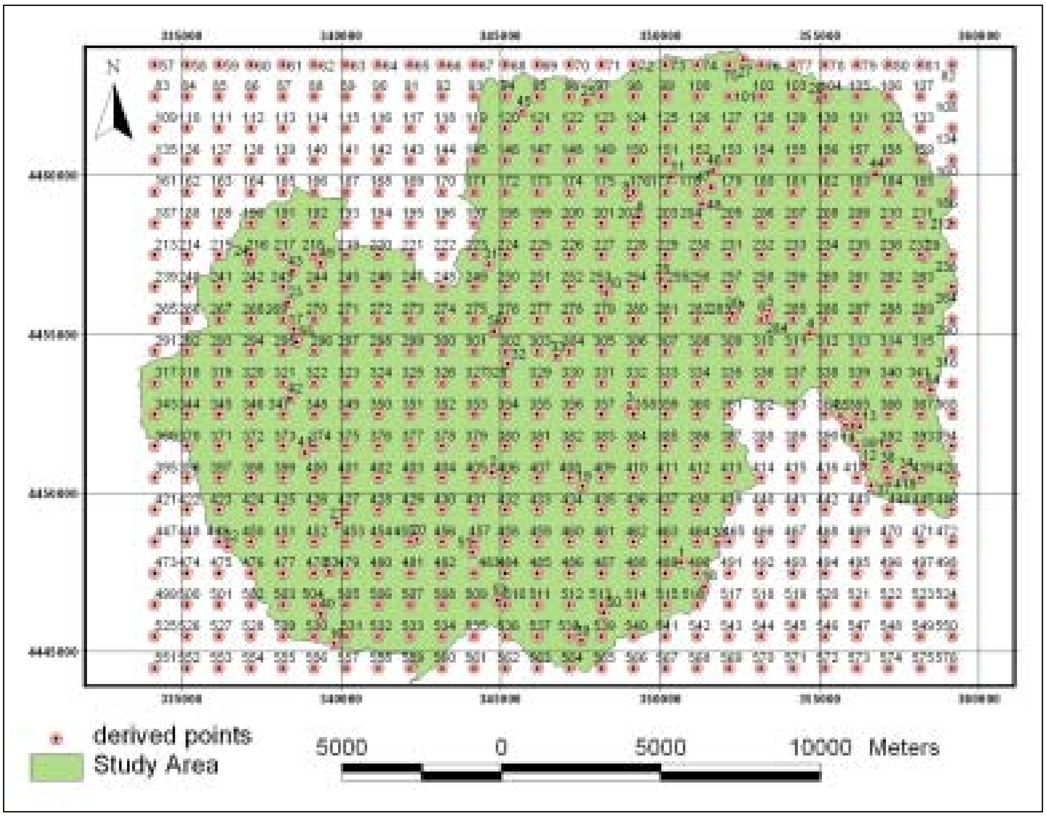


Figure 4.03. Determined points for climatic data estimation and aggregation

To derive the climatic data for selected 576 points, total 20 stations that are located in the nearby area were automatically selected by LOCCLIM, and 8 of them eliminated by considering the distance to the study area and their geographic positions. The positions of the selected 12 stations in the nearby area were given in Figure 4.04. Six climatic variables, important for the plants, were aggregated for 576 sites in monthly basis. The climatic variables can be summarized as; (1) mean temperature, (2) minimum temperature, (3) maximum temperature, (4) precipitation, (5) potential evapotranspiration, and (6) sunshine fraction. The averages of 576 sites were used to summarize and characterize the overall climate of the study area in monthly basis.

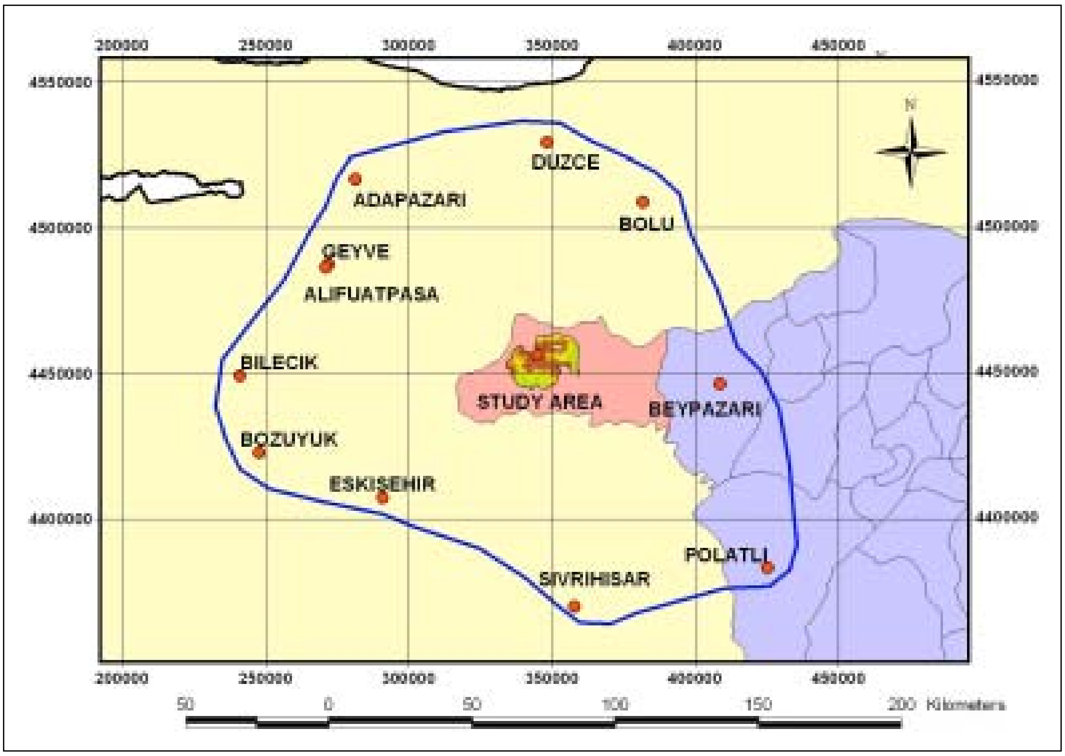


Figure 4.04. Nearest meteorological stations selected and used for best estimations by LOCCLIM software

Some kinds of graphs were employed to present the climate of a location at a glance. One type of these graphs is climograph that is a pictorial device developed to summarize the climate of an area. Basically, a climograph shows the annual cycles of monthly mean air temperature and monthly mean precipitation for a location, along with some useful information such as; sunshine fraction or sunshine degree (Strahler and Strahler, 1996). The climograph of the study area is given in Figure 4.05.

At the top of the climograph, the mean monthly temperature is plotted as a line graph. In the middle, the mean monthly precipitation is shown as a bar graph. Sunshine fraction is given as a line graph at the bottom. According to the Köppen-Geiger Climate System (Strahler and Strahler, 1996), a global climate classification based on mean annual temperature and mean annual precipitation, the study area has typically dry mid-latitude climate characteristics (BSk). On the other hand, it is partly under the effects of Mediterranean climate (CSa).

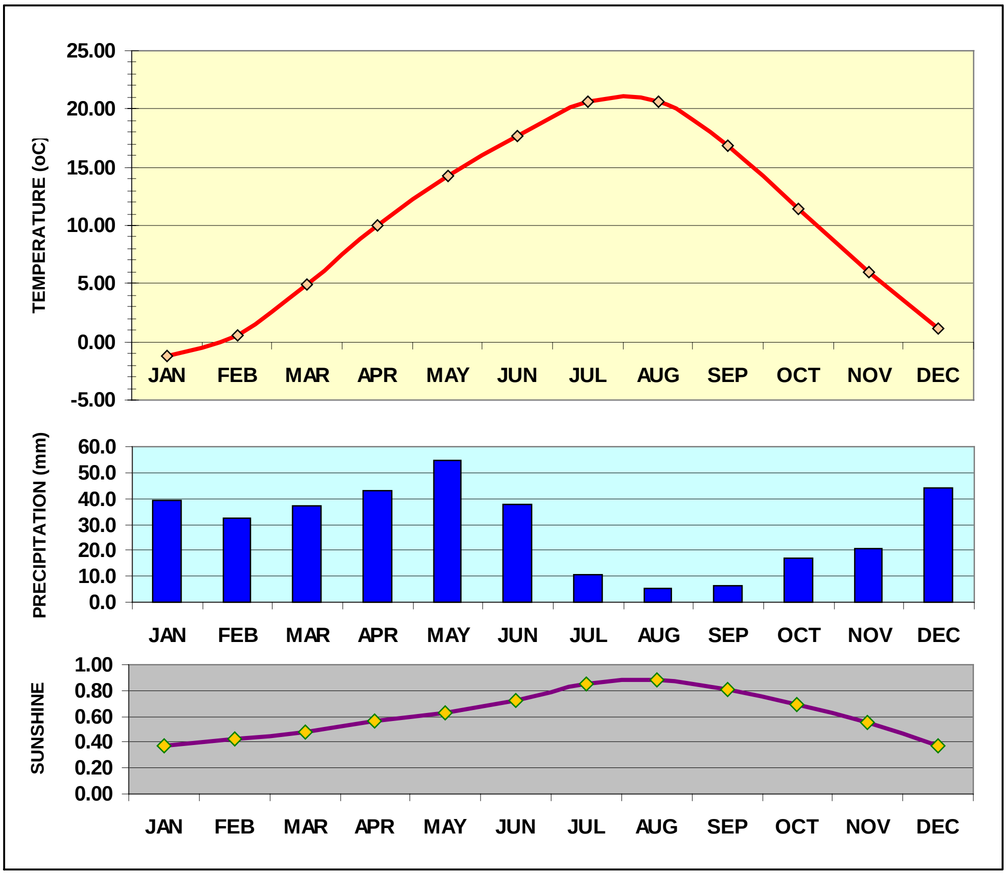


Figure 4.05. The climograph of the study area

The precipitation regime is extremely important in biological point of view, because the vegetation in a region is directly affected by the distribution of precipitation through the years. In another words, dry periods extremely control the vegetation. For this reason, Akman (1999) classified Turkey`s climate by using long term (over 20 years) climatic records belonging 383 climate stations. Akman (1999) used EMBERGER, KÖPPEN, and GIACOBBE methods in his research. EMBERGER method produced very satisfactory results for bioclimatic interpretations of the Mediterranean climate types because it classified the Mediterranean climate layers according to the Q values that is based on annual precipitation, mean maximum temperature (M) of the hottest month, and mean minimum temperature (m) of the coldest month (Table 4.01). Bioclimatic layers of Mediterranean climate were classified by using the Q values and subdivided by using the mean minimum temperature of the coldest month (Table 4.01). According to EMBERGER method; the climate of the study area showed “Semi-arid Upper Mediterranean Bioclimate” characteristics with cold winters.

Table 4.01. Bioclimatic characteristics of the study area according to EMBERGER method (Akman, 1999)

Region	Bio-climatic Variables of Region								Bio-climate
	Mean Elevation (m)	P (mm)	M (°C)	m (°C)	Q	PE	S (PE/M	Precipitation regime	
Nallıhan	650	413	30.9	-1.4	44.4	55.7	1.8	S.W.A.Sm.	Semi-arid Upper Mediterranean

PE: Summer Precipitation (mm); S: EMBERGER Drought Index (PE/M); S.W.A.Sm: Precipitation regime from the highest to lowest (Spring, Winter, Autumn, Summer)

Upper Mediterranean Bio-climate Layers

Q Values		m Values (°C)	
Q> 98	High precipitation bio-climate layer	m< -10	Icy winters
63 - 98	Low precipitation bio-climate layer	-7<m<-10	Extremely cold winters
32 - 63	**Semi-arid bio-climate layer**	-3<m<-7	Very cold winters
20 - 32	Arid bio-climate layer	**0<m<-3**	**Cold winters**
Q<32	Very arid bio-climate layer	0<m<+3	Chilly winters
		+3<m<+4.5	Temperate winters
		+4.5<m<+7	Mild winters
		+7<m<+10	Hot winters
		m>+10	Very hot winters

EMBERGER EQUATION

$Q = (2000 \times P) / (M^2 - m^2)$

P = Annual precipitation (mm)

M = Mean maximum temperature of the hottest month (°C)

m = Mean minimum temperature of the coldest month (°C)

Three important climate characteristics (mean temperature, precipitation, and potential evapotranspiration were represented in Figure 4.06. in monthly basis. Basically, four seasons (autumn, winter, spring and summer) are recognized in the study area. Spring (March, April, May) and early summer (June) are the first periods that the most precipitation is received throughout the year. Precipitation is decreased sharply in the mid/late summer (July, August), and early autumn (September) where mean temperature is in peak. In mid/late autumn (October, November), precipitation goes up little. Winter (December, January, February) is the second period that the most precipitation is recorded. Precipitation is mostly rain throughout the year except winters, and total number of snowy days does not exceed 20 days.

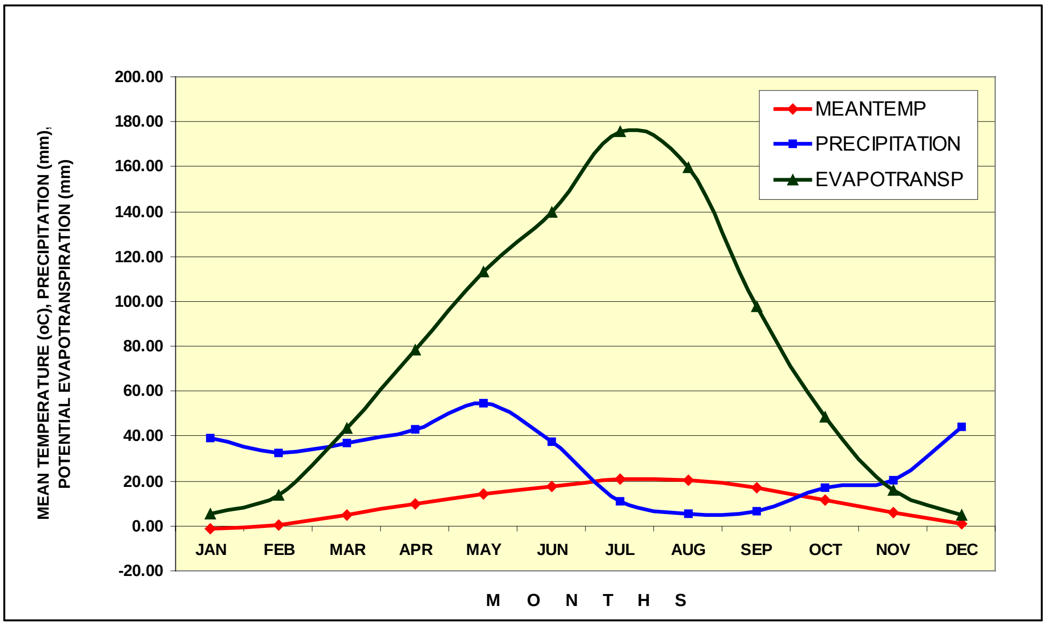


Figure 4.06. Some important climatic attributes of the study area (mean temperature, precipitation, and potential evapotranspiration (PET)

One of the most important characteristics for terrestrial primary production is evapotranspiration that is defined as the combined water loss to the atmosphere by evaporation from the soil and transpiration from plants. Annual actual evapotranspiration (AET) is the total amount of water that evaporates and transpires off a landscape during the course of year and is measured in milimeters of water per year. Both temperature and precipitation are effective on the AET process. The ecosystems that are warm and receive large amounts of precipitation show the highest levels of primary production. Conversely, ecosystems that are cold and receive little precipitation suggest low levels of AET. For instance, both hot deserts and tundra exhibit low levels of AET. There are strong correlations between AET or precipitation and rates of terrestrial primary production (Molles, 1999).

Saturation deficit is a more meaningful measure of the evaporative power of the air, and called potential evapotranspiration (PET). Basically, PET can be defined as the difference between the actual vapor pressure and the saturation vapor pressure at the same temperature. For this reason, mean temperature, precipitation, and potential evapotranspiration (PET) were plotted in the same graph to get better insight about the climate (Figure 4.06).

The most critical period was detected between May and September in Figure 4.06. During this critical period, the gap between precipitation and PET is getting bigger starting from May, and reached its peak in July. Therefore, the climate variables that belong to this critical period were extracted as ’seasonal’. To produce the best model for the study area, both ‘mean annual’ and ‘mean seasonal’ climatic variables were searched. Mean temperature, minimum temperature, maximum temperature, precipitation, and potential evapotranspiration (PET) variables were selected for mapping. Then, the grid maps related to these climatic variables were created in order to delineate their distributions by using Inverse Distance Weighted (IDW) method in Arc/View software (ESRI, 1994, 1996, 1997).

In Figure 4.07, grid maps produced from mean temperature were given in annual and seasonal basis. In the same way, Figure 4.08 and Figure 4.09 delineated the grid maps of maximum and minimum temperature, respectively. Grid maps of precipitation (Figure 4.10) and potential evapotranspiration (Figure 4.11) showed the spatial distributions of these two climatic variables.

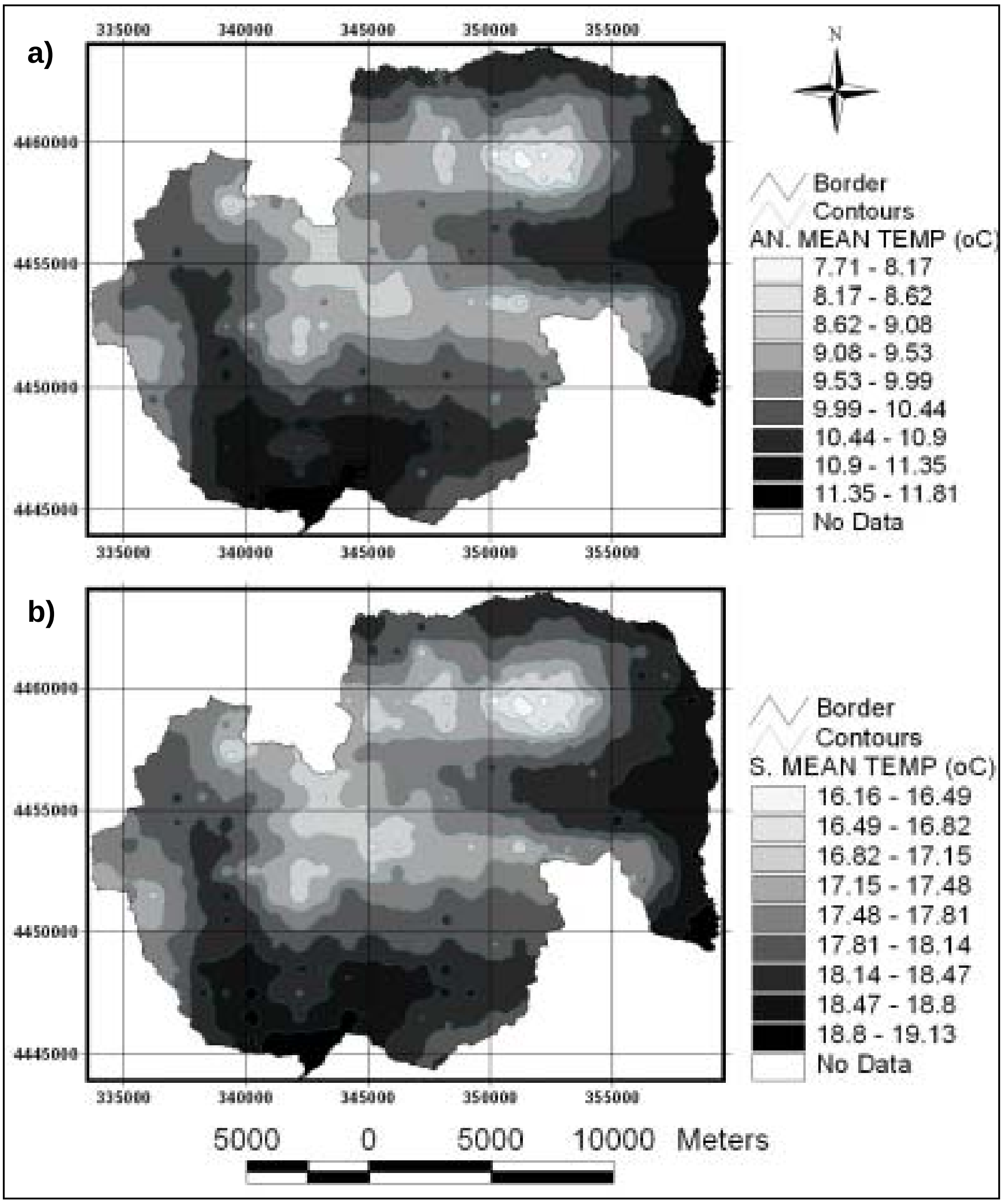


Figure 4.07. Grid maps of mean temperature (°C) values in (a) annual, and (b) seasonal basis (derived from point data by using IDW method in Arc/View software)

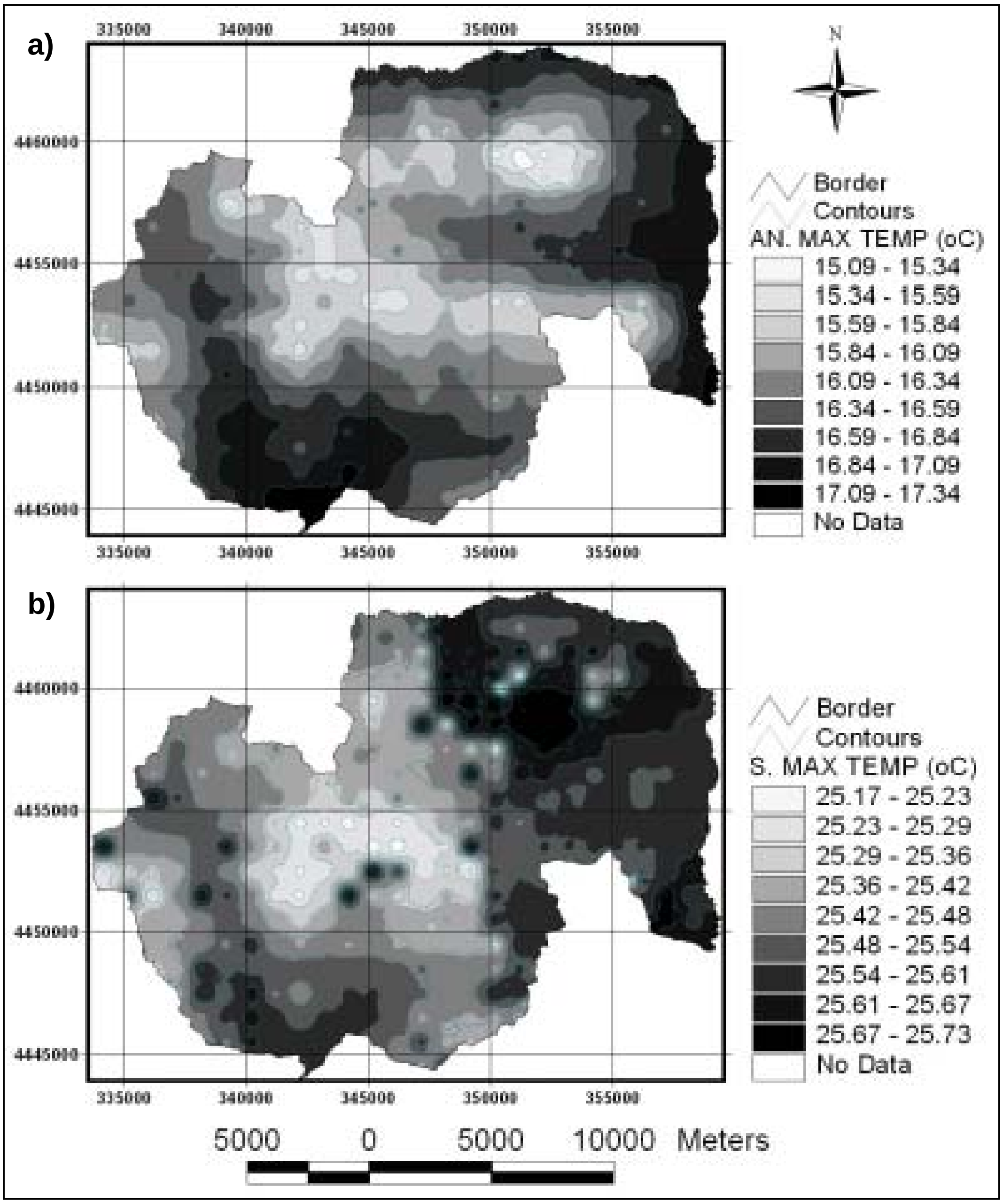


Figure 4.08. Grid maps of maximum temperature (^{o}C) values in (a) annual, and (b) seasonal basis (derived from point data by using IDW method in Arc/View software)

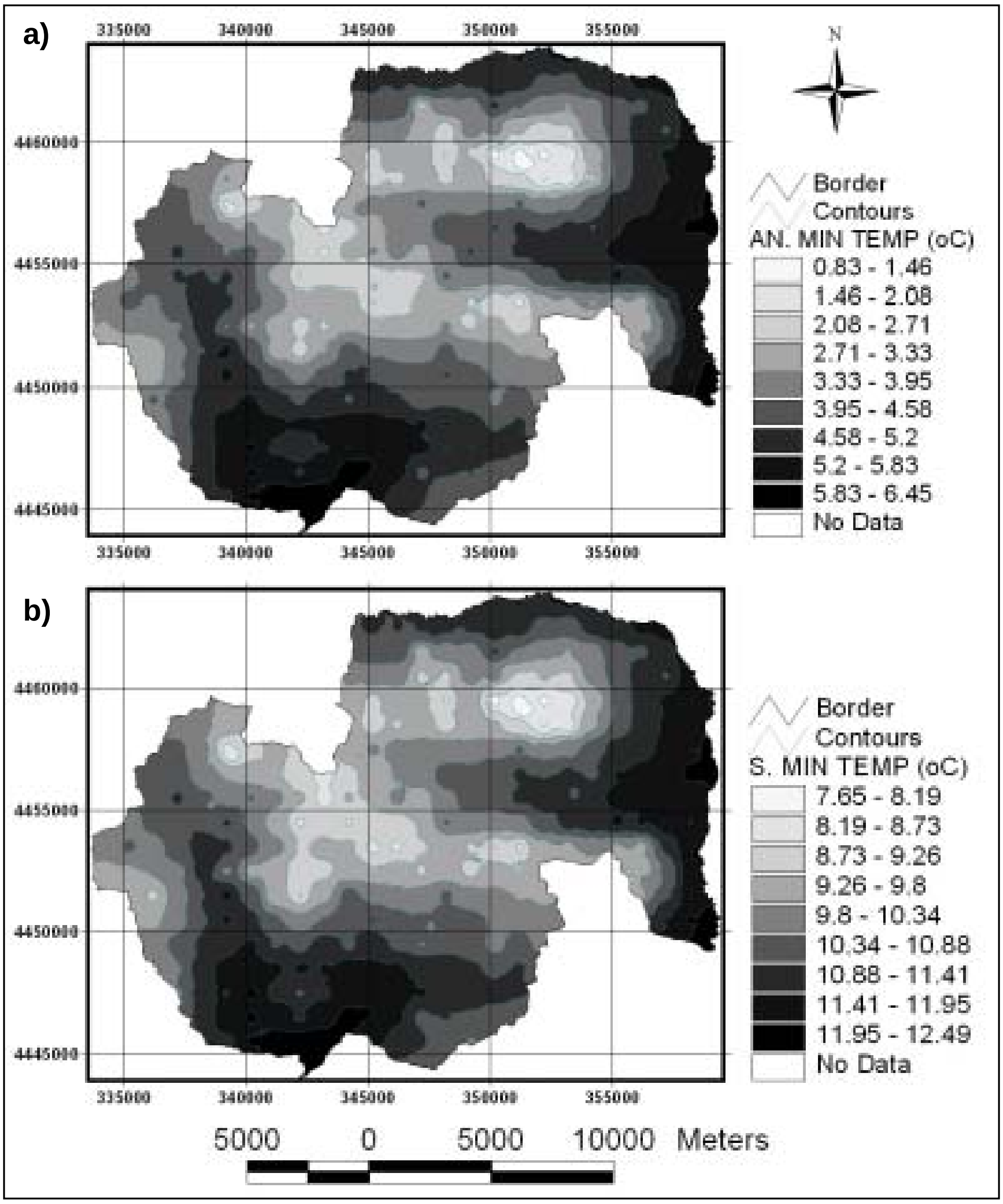


Figure 4.09. Grid maps of minimum temperature (^oC) values in (a) annual, and (b) seasonal basis (derived from point data by using IDW method in Arc/View software)

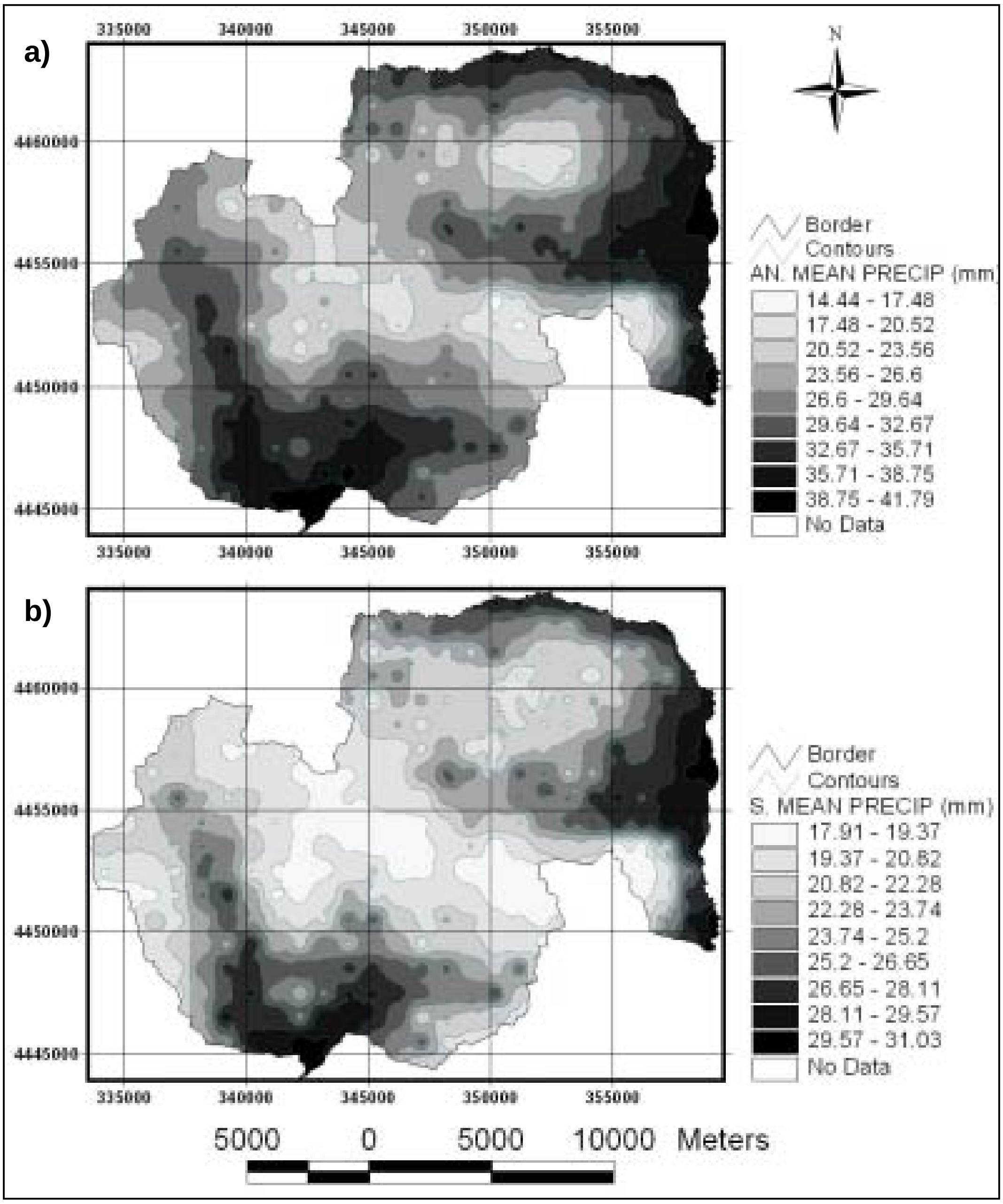


Figure 4.10. Grid maps of precipitation (mm) values in (a) annual, and (b) seasonal basis (derived from point data by using IDW method in Arc/View software)

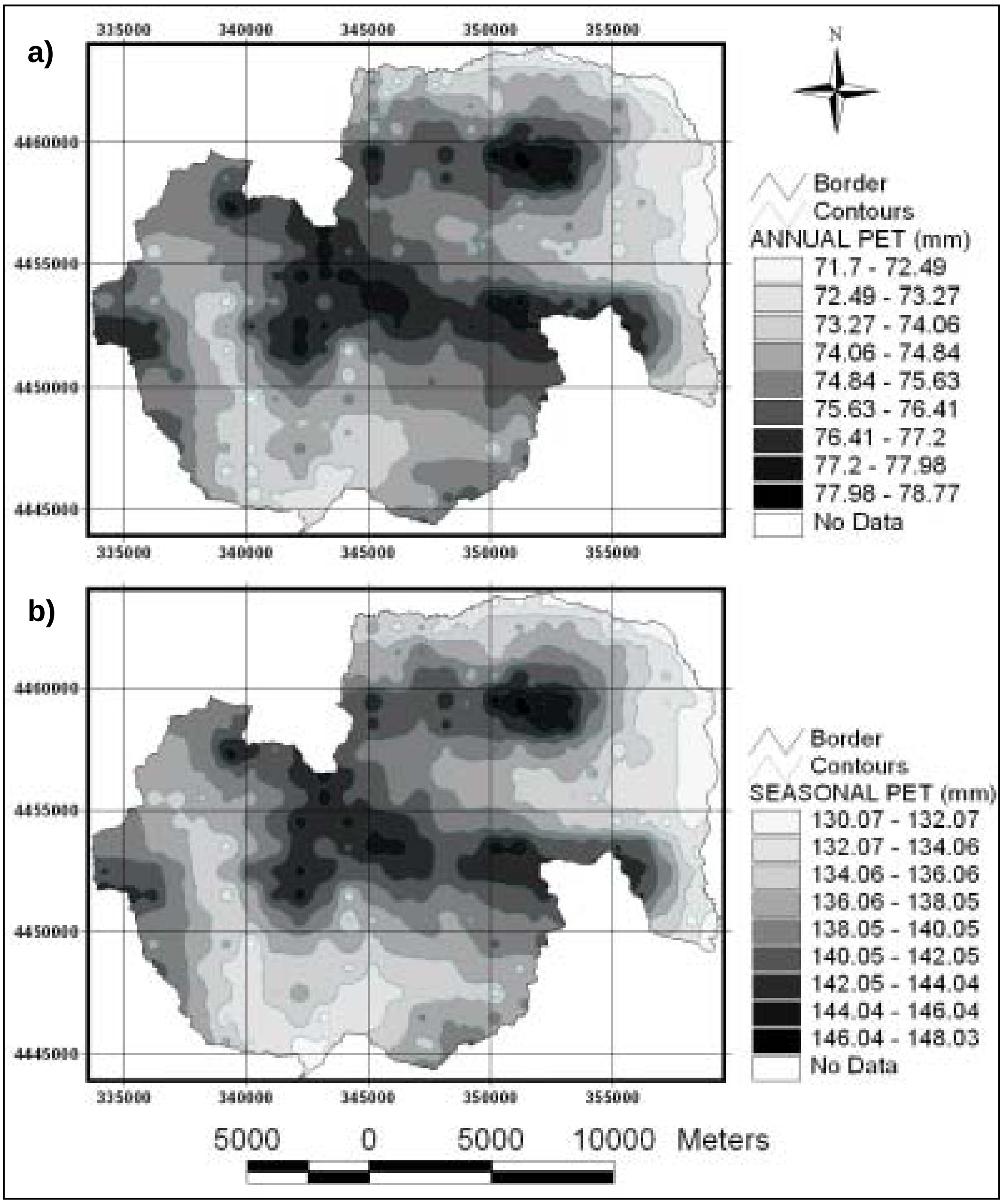


Figure 4.11. Grid maps of potential evapotranspirations (PET:mm) values in (a) annual, and (b) seasonal basis (derived from point data by using IDW method in Arc/View software)

4.2.3 Soil

Soil is a complex mixture of living and nonliving material upon which most terrestrial life depends (Molles, 1999). Soil characteristics such as texture, nutrient status, and depth are well documented as important factors determining the competitive relationships and growth rates of plants in a wide variety of environments.

Significant variation in terrestrial primary production can be described by the differences of edaphic factors. Therefore, several important soil characteristics were drawn to examine their influences on plant biodiversity. Soil data were compiled from the available soil maps and recent soil analyses. The digital soil map of Nallıhan county with 1 / 25 000 scale was obtained from the digital soil database of General Directorate of Rural Affairs (KHGM). The database content of soil maps were consist of (1) big soil groups, (2) erosion classes, (3) soil depth, (4) current land use, (5) land types, (6) land use ability for agriculture, (7) sub classes of land use ability, and (8) geographical data (minings, rivers, lakes, etc.). All soil data layers were examined carefully, and the layers that contain big soil groups, erosion classes, and soil depth were selected for further analysis (Table 4.02).

Three big soil groups (alluvial, colluvial, and brown forest soil) were recognized in the area (Figure 4.12-a). Brown forest soil group covered the majority of the study area, while colluvial soils were not widespread. Alluvial soils normally appeared along the Naldere river, because they were transported and deposited by the stream. Alluvial soil group has been mainly used for agriculture. Besides the big soil groups, four soil depth classes were observed (Figure 4.12-b). Very shallow and shallow depth classes were found substantial amount, while so small areas pertained to medium and deep classes. When we look at the erosion classes (Figure 4.12-c), it can be easily seen that more than half of the area is in severely erosion class.

Table 4.02. Summary information about selected soil layers from digital soil map

SOIL DEPTH (cm)		EROSION (1-4)		BIG SOIL GROUPS	
Classes	**% Area covered**	**Classes**	**% Area covered**	**Groups**	**% Area covered**
Deep (+90)	2.40	no or little	1.50	Colluvial	3.17
Medium (90-50)	4.80	medium	5.65	Alluvial	16.23
Shallow (50-20)	13.60	very severe	35.50	Brown Forest	80.60
Very Shallow (20-0)	79.20	severe	57.35		

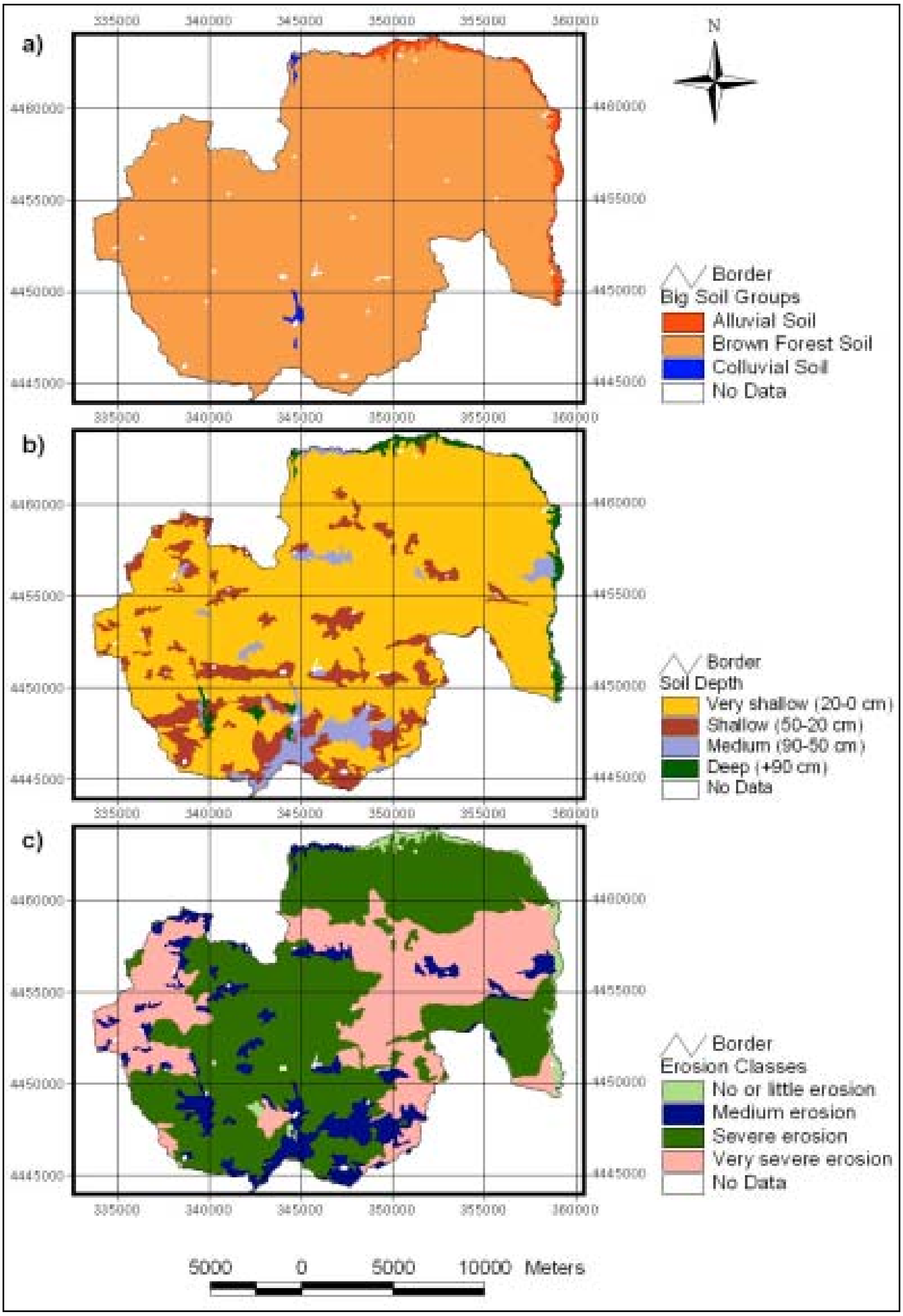


Figure 4.12. Soil maps of the study area a) big soil groups, b) soil depths, and c) erosion classes

Although digital soil map supplied valuable data, it did not comprise any information about nutrients, physical, and chemical characteristics of soil that could be important to explain the distribution of plant species throughout the area. For this reason, representative soil samples were collected from the 56 quadrats according to the certain soil sampling methods (Ateşalp, 1976; Ülgen and Yurtsever, 1995) in this study. Then, a series of soil analyses (Table 4.03) were carried out by Soil and Fertilizer Research Institute of the General Directorate of Rural Affairs. Each soil variables were mapped according to the principles explained in this chapter (data preparation) and Chapter 2 (setting up digital database from coordinated point data). The spatial distributions of the analyzed soil variables were given in Figure 4.13, Figure 4.14, and Figure 4.15.

Table 4.03. Summary information about selected soil variables to be analyzed

S O I L V A R I A B L E S		
Nutrients	**Chemical Characters**	**Physical Characters**
Effective K_2O (kg/da)	pH (values)	Texture (texture classes)
Effective P_2O_5 (kg/da)	Salt (%)	Saturation (%)
Organic Matter (%)	CaCO3 (%)	
NOTE: 1 da = 1000 m^2		

The nutrient conditions of the study area were defined by using potassium (K_2O), phosphorus (P_2O_5) and organic matter (Figure 4.13). In the area, the effective amount of K_2O varies between 280 and 5360 kg/ha, while P_2O_5 changes between 6 and 234 kg/ha. Generally, the soils of the study area contain K_2O and P_2O_5 in the amounts of 840-1400 and 6-107 kg/ha, respectively. The organic matter extends to 6.48% starting from 0.94 %, and usually gets the value between 1 % and 4 %.

The chemical characters focused in this study are pH, salt and calcium carbonate (Figure 4.14). Soil pH is important, because it has many effects on biological and chemical activity of the soil. The pH scale ranging from 0 to 14 is used to indicate acidity and alkalinity. A pH of 7.0 is neutral, values below 7.0 are acid, and those above are alkaline. According to this range, study area soils generally showed alkaline character between 7.5 and 8.0 values. Salt content of the study area, on the other hand, changes between 0 and 0.084%, and usually gets the values between 0.03 and 0.07%. The soils also contain different amounts of CaCO3 between 0 and 27.49%. Texture and saturation are the physical attributes to represent the soil. Two distinct texture classes, clay (C) and clay loam (CL), were recognized in the study area. Depending on the texture, saturation values change between the 32% and 67%, but usually obtain the values around 50%. Texture and saturation characteristics of the area were summarized in Figure 4.15.

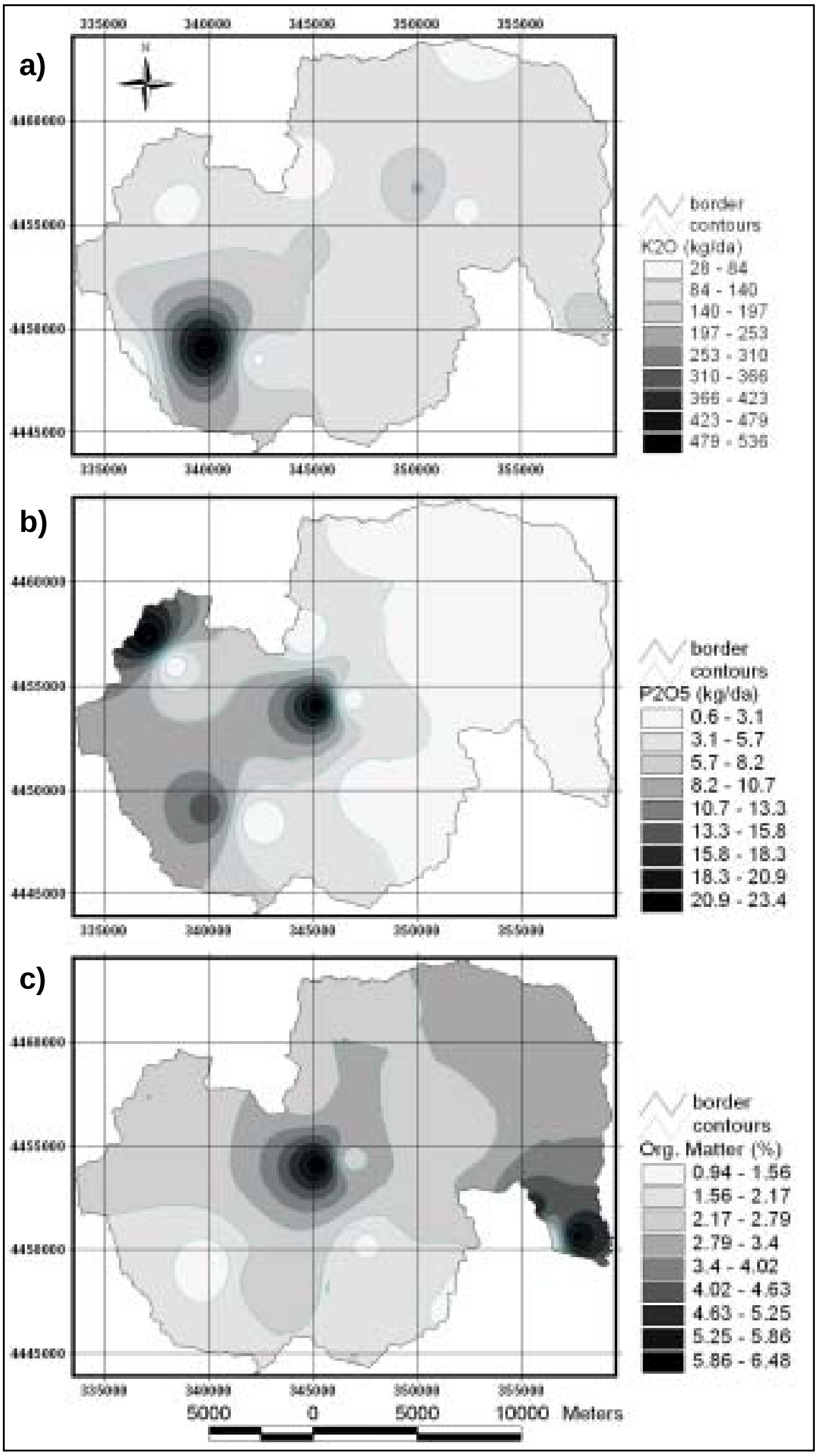


Figure 4.13. Soil nutrient characters of the study area: a) P2O (kg/da), (b) K2O (kg/da), and (c) organic matter (%) NOTE: 1 da = 1000 m^2, and 1ha = 10 da

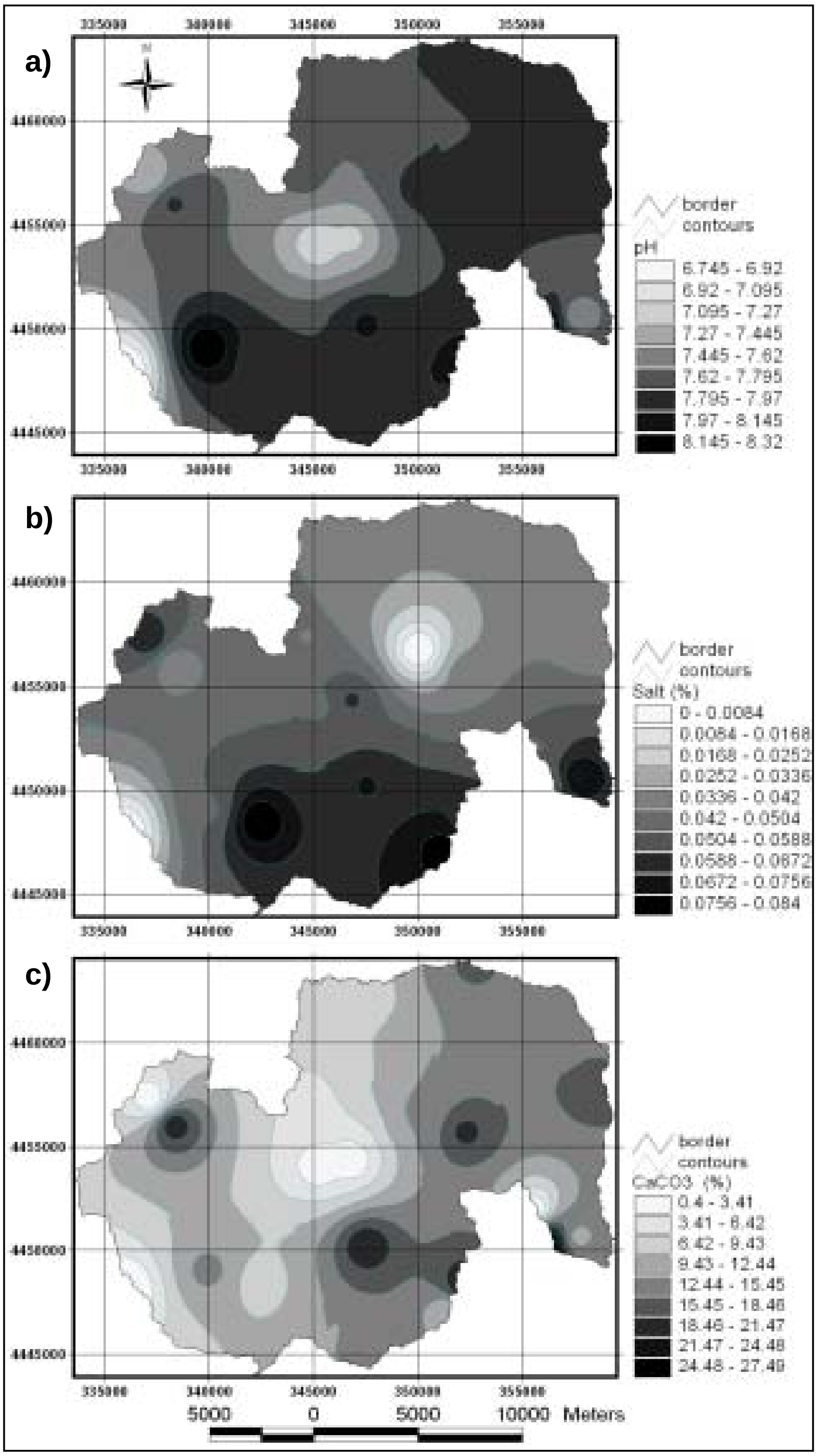


Figure 4.14. Soil chemicals characters of the study area: a) pH, (b) salt (%), and (c) $CaCO_3$

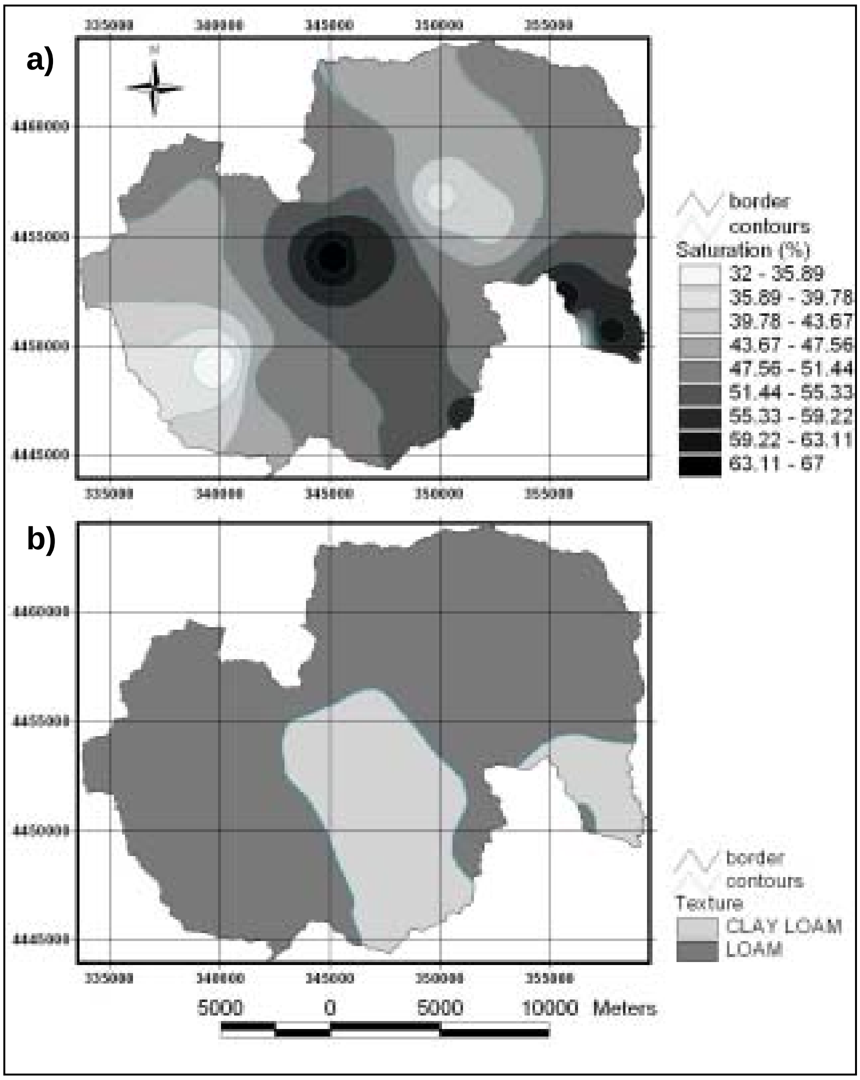


Figure 4.15. Some soil physical characters of the study area: a) saturation (%), and (b) texture

4.2.4. Geology

Molles (1999) reported that landscapes are structured and change in response to geological processes as well as climate, activities of organisms, and fire. Therefore, geology is one of the important factors that is highly effective on the plant distributions. Rock types existing in the area may influence, directly or indirectly, the growth of the plants. They should be, therefore, considered in the evaluation of the spatial distribution of plants. Other geological structures and processes than the rock types, however, may not be genetically linked to the plants in the area. For this reason, the term “geology” in this thesis refers only to the “rock types”; and considering the purpose of the thesis a simplified geological map of the region is prepared (Figure 4.16) show distribution of the rocks existing in the area excluding other geological information.

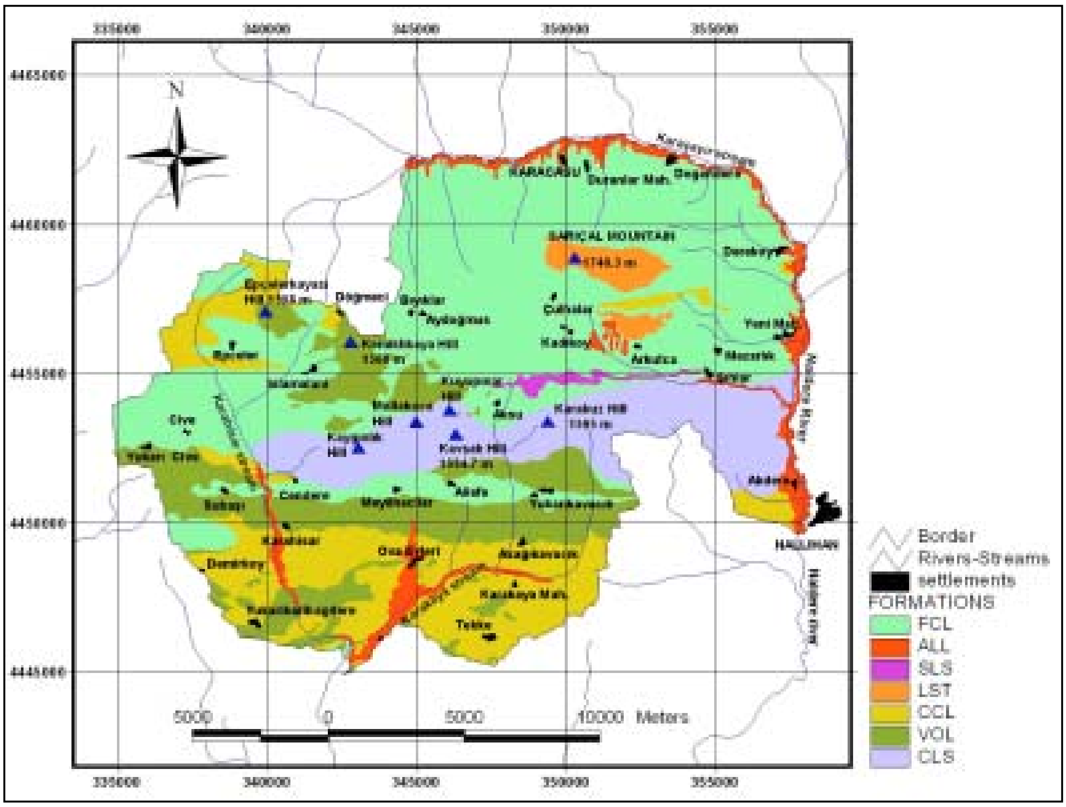


Figure 4.16. Categorized rock types of the study area; FCL: Fine Clastics, ALL: Alluvium, SLS: Shaly Limestone, LST: Limestone, CCL: Coarse Clastics, VOL: Volcanics, CLS: Cherty Limestone

Geological map of the area is provided from MTA (General Directorate of Mineral Research and Exploration) in digital format. The area is included in six sheets of geological map at 1/25.000 scale. Total number of the rock types is 13 in the original MTA maps. The name and age of these rock types are illustrated in Table 4.04. Some of these rock types, however, are similar to each other in their lithological characteristics and the time span they were formed. Therefore, these rock types are reclassified and the number of the classes is reduced to 7 (Table 4.04). One class of these rock types belongs to volcanic rocks, all others to sedimentary rocks. A short description of each rock type is given below.

Table 4.04. Reclassified rock types present within the study area.

Original Rock Type by MTA	Age	Categorized Rock Type	Area Covered (%)	Abbreviation
Alluvium Alluvial Fan Old Alluvium	Quaternary Quaternary Quaternary	Alluvium	4.62	ALL
Andesite, Basalt, Pyroclastic Andesite, Basalt, Pyroclastic	Eocene Paleocene	Volcanics	15.77	VOL
Conglomerate-Sandstone-Mudstone	Paleocene	Coarse Clastics	17.34	CCL
Sandstone-Mudstone-Limestone Sandstone-Mudstone-Limestone	Paleocene Albian (Cretaceous)	Fine Clastics	46.65	FCL
Shaly Limestone Shaly Limestone	Turonian (Cretaceous) Campanian (Cretaceous)	Shaly Limestone	0.59	SLS
Cherty Limestone	Callovian – Aptian (Cret.)	Cherty Limestone	13.00	CLS
Limestone Limestone	Hauterevian (Cretaceous) Callovian (Jurassic)	Limestone	2.03	LST

Alluvium in general refers to the loose material transported by a river and deposited (accumulated) along the plains of the river. It is, therefore, observed at lower elevations of the region. It is commonly composed of sand and gravels. Two classes of this rock type exist in the area and are named as “Alluvium” and “Old alluvium” (Table 4.04). Both have an age of Quaternary. Alluvial fan, on the other, hand is a rock type similar to alluvium. The only difference is that, alluvial fan is deposited at a certain point along the river course where there is a decrease in the gradient. These three river-associated rock types are re-categorized into the same type considering major similarities in their age and lithological characteristics. Alluvium within the study area is mainly exposed along the stream beds of three rivers, namely, Karaçayır to the northeast, Naldere to the east, Karakaya to the south and Karahisar to the southwest (Figure 4.16).

Volcanic rocks are the products of volcanic activity that occur in an area. Two original rock types are mapped in the area as “andesite, basalt and pyroclastics”. The ages of these are very close to each other suggesting that they are the products of the same volcanic phase. Therefore they are classified under the same rock type. Individual rocks identified within these two classes are the same. Andesite is an extrusive volcanic rock erupted in the form of lava flows. It is characterized by presence of fine-grained minerals mostly oligoclase or andesine. Basalt is also an extrusive rock erupted in the form of lava flows. Essential minerals of basalt are plagioclase and pyroxene with or without olivine. Pyroclastic rocks consist of fragmental volcanic material, which is blown into the atmosphere by explosive activity and accumulated on the surface after they settle down. Volcanics rocks in the area are exposed in the vicinity of Subaşı, Karahisar, Meyilhacılar, Aşağıkavacık, Yukarıkavacık, Kemiklikaya Hill, Epçeler Kayası Hill, and Döğmeci, and cover an area of 51.62 km^2 (Figure 4.16).

Coarse clastics refers to the sedimentary rocks composed of sediments (grains or fragments) transported by certain agents and deposited in a water body. Coarse clastics in the study area correspond to Palocene age consolidated material. Dominant rock types of this unit are conglomerates, sandstones and mudstones. These rocks are scattered within the area as separate outcrops. The largest outcrop is exposed in the southern part (Figure 4.16). Coarse clastics cover approximately 17.34 % of the area. Fine clastics are sedimentary rocks that have smaller grain size. Within the area, this group rocks are composed of two separate sequences, which are Paleocene and Albian in age. Although these two sequences have a time gap in between they are put into the same category due to their lithological similarities. The sequences are composed of alternation of sandstone, mudstone and limestone layers. They are the most commonly observed rock type and cover 46.65 % of the area.

Shaly limestone is a sedimentary rock that is essentially composed of limestone (carbonates) with minor content of shale (fine sedimentary clastic rock). Two sequences of different ages (Turonian and Campanian) exist in the area. They are exposed within a limited area as a thin belt extending in E-W direction. Cherty limestone is a sedimentary rock formed by chemical precipitation of soluble material within a water body. The dominant rock type is limestone (carbonates) with certain intercalation of chert (silica) layers. This unit is represented by a single large outcrop in the central part of the area extending in E-W direction. Limestone is a chemical sedimentary rock consisting essentially of carbonates. The most important constituent of limestone is calcite ($CaCO3$). There are two outcrops of this rock unit in the study area that belong to Hauterevian and Callovian ages. Limestone covers 2.03% of the area.

4.2.5. Land cover and land-use

Land cover is defined as the observed (bio) physical cover on the earth's surface, and it describes vegetation and man-made features. Consequently, areas where the surface consists of bare rock or bare soil are describing land itself rather than land cover. Also, it is disputable whether water surfaces are real land cover. However, in practise, the scientific community usually describes those aspects under the term land cover. Land use is characterized by the arrangements, activities and inputs people undertake in a certain land cover type to produce, change or maintain it. Definition of land use in this way establishes a direct link between land cover and the actions of people in their environment. In this study, two sources, remote sensing data and detailed forest stand map, were utilized to delineate the land cover and land use characteristics of the study area.

4.2.5.1. Remote sensing data

Modeling the relationship between the vegetation and land characteristics is a challenging task. Achieving this goal is inconceivable without detailed information that delineates the spatial distribution of important land features especially for plants. Therefore, reliable data that supply these kinds of information are essential. In this study, both aerial photographs and satellite image data were used to delineate the current vegetation position. Available monoscopic aerial photographs with 1 / 35 000 scale were used to make the visual interpretation in an accurate way, while a LANDSAT-TM image was employed to classify the area which a model can build it up. Total 26 pieces of monoscopic air photographs (16 July 1998) were determined and purchased from the archives of the General Commandership of Mapping of Turkish Army. All photos were scanned finely and saved in TIFF format. Then, each photograph was registered and rectified in ERDAS IMAGINE software (ERDAS, 1997) by using ground control points determined on both the topographic maps (1 / 25 000 scale) and the air photographs. Finally, all air photos were merged together in the software by using ‘photo mosaic’ function and ‘rubbersheeting’ method. After all air photographs unified, metadata records including all necessary information about the data were created and added. The photo mosaic constituted a valuable data in two ways; (1) it helped to determine signatures of the land cover visually throughout the area, and (2) it gave a chance to check the validity of supervised classification and applicability of the developed model. Evaluation of photo mosaic data is based on entirely visual interpretation in this study (Figure 4.17).

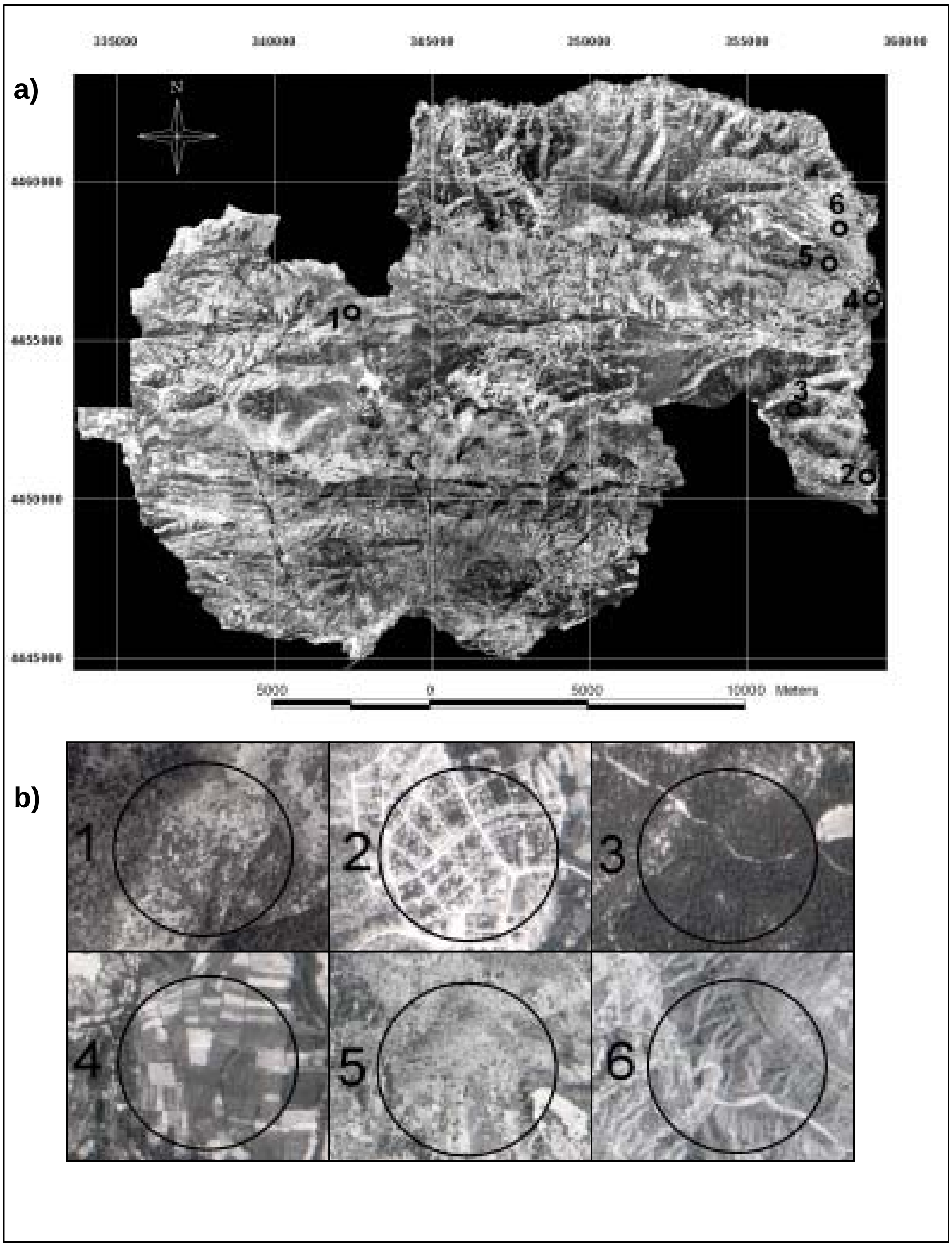


Figure 4.17. Air photo interpretation a) Photo mosaic of the study area b) signature of land covers determined as (1) deciduous forest (mainly oak), (2) setllement, (3) coniferus forest (mainly black pine), (4) agriculture, (5) degregated forest/rangeland, (6) stony/erosion areas

A LANDSAT-TM image, belongs to 21 August 2000, was selected from the image archive of Central Research Institute for Field Crop, GIS and Remote Sensing Department. The image (path/row: 178/32) has 30 x 30 m spatial, 2^8 (256) radiometric, 7 bands spectral, and 16 days temporal resolution. The LANDSAT-TM image was registered, geo-referenced, and rectified in ERDAS IMAGINE software (ERDAS, 1997) by using the available data. After preparing the image, a supervised classification was applied. A signature file created by evaluating available maps, field observations, and photo mosaic data was used in this process. Maximum likelihood parametric rule and 4, 5, 3 band combination were chosen to classify the image properly. Finally, statistical filtering with 7 x 7 window size was employed to get a neat classified map. The accuracy assessment of the supervised classification was performed with ground truth data collected from 306 points throughout the study area. The results of accuracy assessment were summarized in Table 4.05. Supervised classification obtained 92.16 % overall accuracy with a Kappa coefficient of 0.8828, and produced a reliable result.

There are a group of available algorithms that are commonly used for either vegetation or mineral delineation in ERDAS software. In many cases, judiciously chosen indices can highlight and enhance differences which cannot be observed in the display of the original color bands. Normalized difference vegetative index (NDVI) is one of these indices that has been used extensively in vegetation analyses. NDVI brings out small differences between various vegetation classes that might be valuable for a model. Depending on its importance, an NDVI layer was created to get better insight in this study. The NDVI model is derived from the algorithm of NDVI = (infrared – red) / (infrared + red). Although, NDVI basically gets the values between +1 and –1, it is also possible to stretch this index to unsigned 8 bit values between 0 and 255 in the software. To conduct a meaningful statistical evaluation, unsigned 8 bit values were chosen to produce NDVI layer of the study area (ERDAS, 1997). The grid maps of supervised and NDVI classification were given in Figure 4.18.

Table 4.05. Accuracy assessment of supervised classification

ACCURACY TOTALS						
Class Name	**Reference Totals**	**Classified Totals**	**Number Correct**	**Producers Accuracy %**	**Users Accuracy %**	**Kappa (K^) Statistics**
Unclassified	0	0	0	-	-	0.0000
Agriculture	84	79	73	86.90	92.41	0.8953
Coniferus (black pine)	153	144	143	93.46	99.31	0.9861
Degraded forest	38	46	36	94.74	78.26	0.7518
Deciduous (oak)	15	18	14	93.33	77.78	0.7663
Stone-	16	19	16	100.00	84.21	0.8334
TOTALS	306	306	282			
Overall Classification Accuracy = 92.16 % Overall Kappa Statistics = 0.8828						

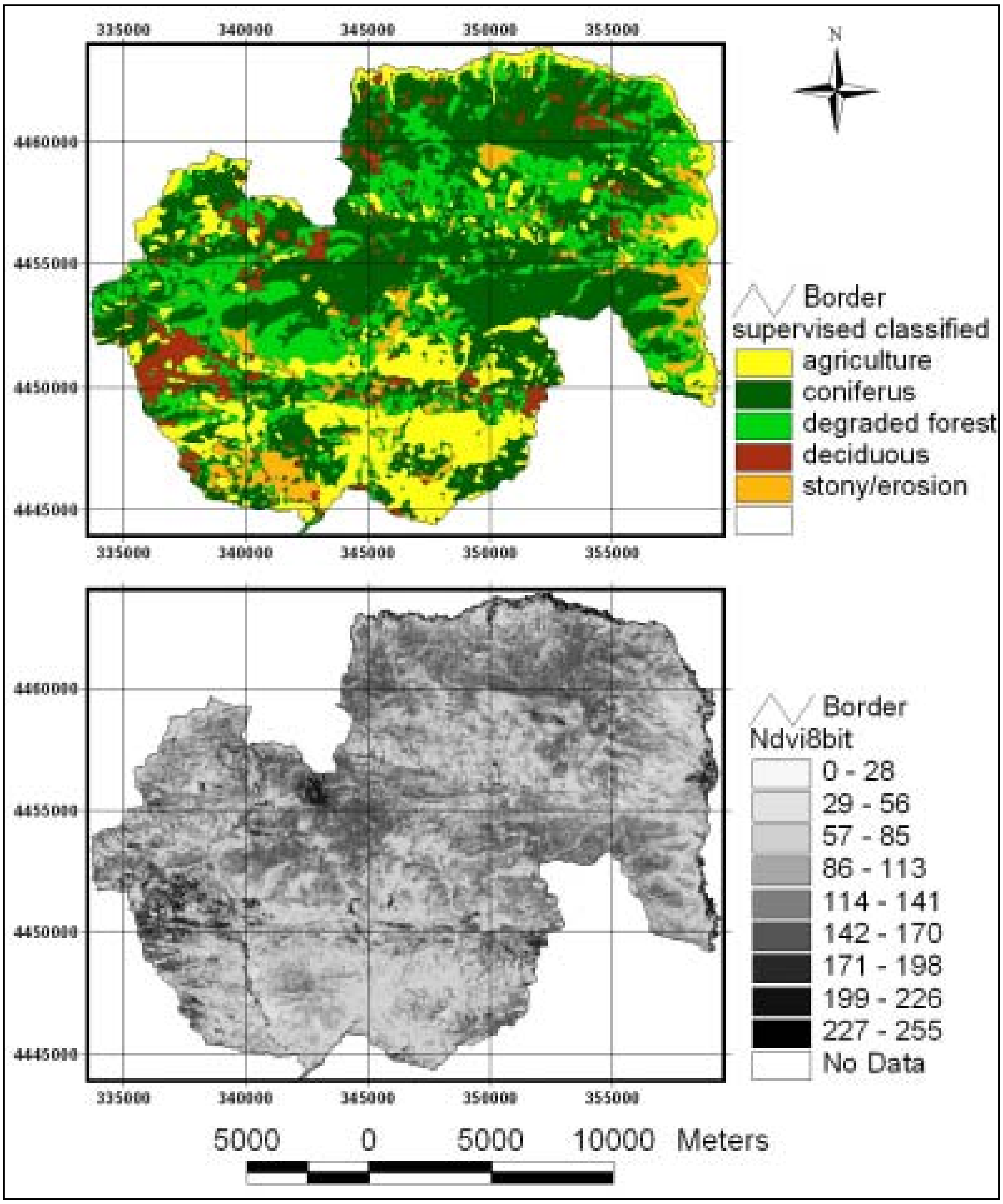


Figure 4.18. Supervised classification and NDVI maps produced from LANDSAT-TM image (Total five reliable classes determined in supervised classification are (1) agriculture, (2) coniferus forest (mainly black pine), (3) degregated forest, (4) deciduous forest (mainly oak), and (5) stony-erosion-shrub areas)

4.2.5.2. Forest data

The recent available forest stand map (1996) with 1/25 000 scale was delivered from the General Directorate of Forest. This map was transformed to the digital database format according to the methods that were explained in ‘data preparation’ section (Figure 4.19). The digital forest map supplied very detailed information about the forest stands and their distribution. According to the forest stand map, the main forest plant species are black pine (*Pinus nigra*), juniper (*Juniperus spp*.), red pine (*Pinus brutia*), and oak (*Quercus spp*.).

The land cover characteristics of main forest species were summarized in Table 4.06. According to the Table 4.06; *Pinus nigra* is the dominant species that covers 31.47% of the area. Although forest map database was used extensively in the field studies and image classification process, it was not used in the modeling stage. Instead of forest map, supervised classified and NDVI images were used to reflect the recent land cover and land use characteristics in this study.

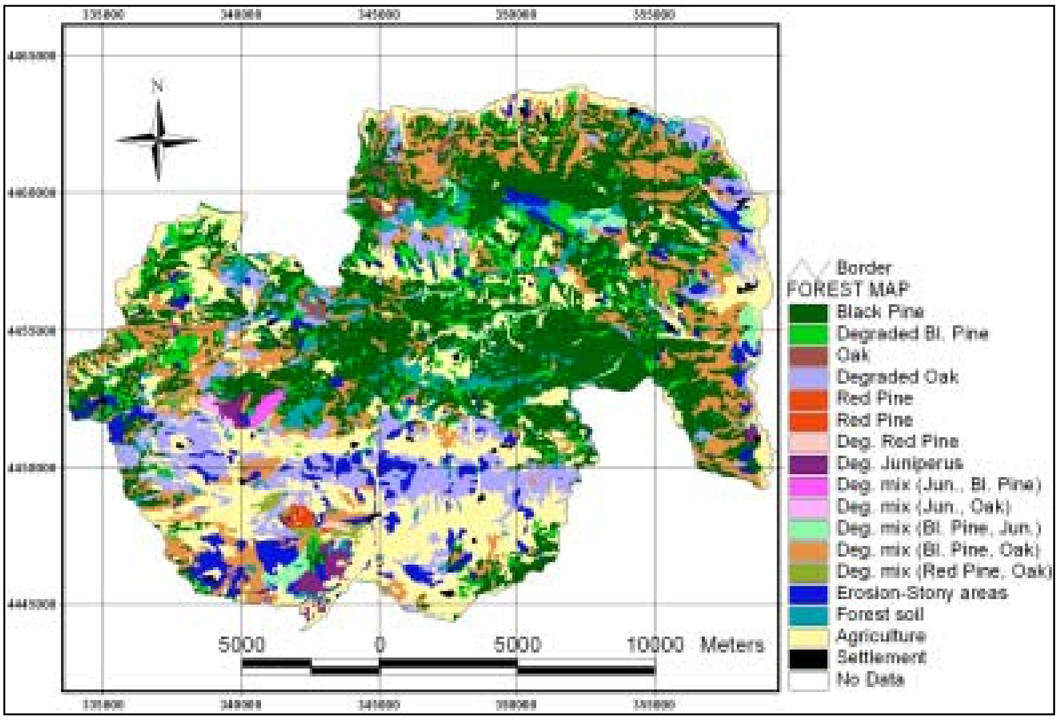


Figure 4.19. Main land cover characteristics of the study area (derived from forest stand map with 1 / 25 000 scale)

Table 4.06. The summary of land cover characteristics of the study area (according to the forest stand map with 1 / 25 000 scale)

Forest Stand Characteristics	Area covered (ha)	% area covered
Black Pine	10300.95	31.47
Agriculture	7520.04	22.97
Degraded Mixed (Black Pine,Oak)	4800.51	14.67
Degraded Oak	3165.03	9.67
Degraded Black Pine	2036.07	6.22
Erosion	1833.3	5.60
Forest Soil	1587.78	4.85
Degraded Mixed (Black Pine, Juniperus)	467.73	1.43
Degraded Juniperus	285.21	0.87
Settlement	183.15	0.56
Oak	155.07	0.47
Degraded Mixed(Red Pine,Oak)	139.32	0.43
Degraded Mixed(Juniperus, Black Pine)	82.08	0.25
Red Pine	70.29	0.21
Degraded Mixed (Juniperus, Oak)	57.24	0.17
Degraded Red Pine	47.52	0.15
TOTAL	**32731.29**	**100.00**

CHAPTER 5

DATA ANALYSIS

This chapter concerns with the applied statistical analyzes, and interpretation of the results. Basically, statistical analyzes were explained as descriptive statistics, data reduction, and multiple regression. A brief introduction was given for each statistical method for easy understanding. The developed models and resulting plant biodiversity maps of these models were also explained in this chapter.

5.1. Data Enumeration

The nature of the data is important to decide the most suitable statistical method. Generally, parametric and non-parametric statistical methods are available for the quantitative and qualitative data, respectively. Comparing to non-parametric methods, parametric methods are generally preferable to produce more faithful results (McGrew and Monroe, 1993).

The nature of the data can be explained by the measurement levels. Basically, variables are classified into four measurement levels: (1) nominal, (2) ordinal, (3) interval, and (4) ratio. The details about these measurement levels were given in literature review (Chapter 2). In this study, majority of the data has quantitative characteristics at ratio and interval levels, and they are suitable for parametric statistical methods. On the other hand, some of the data has qualitative characteristics at nominal level, and they are not appropriate for parametric statistical methods. Geological formations, soil texture, big soil groups, soil depth, erosion, aspect, and supervised classes are the data at nominal level. To apply parametric methods, nominal data should be explained in numeric form. For this reason, the nominal data were enumerated to obtain the integrity of the data analysis in this study. The enumerated nominal data and the criteria that had been considered through this process were summarized in Table 5.01.

Table 5.01. Enumeration of nominal data

Geology	Numeric Value	Criteria
FCL	1	Wild Plant Cover Increase ↓
ALL	2	
LST	3	
CCL	4	
VOL	5	
SLS	6	
CLS	7	
Texture	**Numeric Value**	**Criteria**
L	1	Productivity increase ↓
CL	2	
Big soil groups	**Numeric Value**	**Criteria**
Alluvial-Colluvial Soils	1	Wild Plant Cover Increase ↓
Brown Forest Soil	2	
Soil depth	**Numeric Value**	**Criteria**
No Soil	1	Soil Depth Increase ↓
Very shallow	2	
Shallow	3	
Medium	4	
Deep	5	
Erosion	**Numeric Value**	**Criteria**
Very severe erosion	1	Wild Plant Cover Increase ↓
Severe erosion	2	
Medium erosion	3	
No or little erosion	4	
Aspect	**Numeric Value**	**Criteria**
N	1	Sunshine fraction increase ↓
NE, NW	2	
E, W, F (Flat)	3	
SE, SW	4	
S	5	
Supervised classes	**Numeric Value**	**Criteria**
Stony/erosion	1	Wild plant Cover increase ↓
Agriculture	2	
Degraded forest	3	
Deciduous	4	
Coniferus	5	

5.2. Exploring Data

The first step of data analysis should always be detailed examination of the data regardless of the simplicity or complexity of the problem to be solved and statistical method to be used. There are several important reasons for examining data before the beginning of any statistical analysis. Identifying mistakes is one of the important reasons for examining data because errors and mistakes can be made during any step of any research including measurement and data gathering stages. Without checking the data for mistakes, errors could effect all of the statistical analysis to be conducted, and produce incorrect results. Examination of the data also supplies enough information to establish some additional hypotheses or to modify existing hypotheses at the beginning of the statistical analyses. Another reason for examining the data is to decide the appropriate statistical methods to be conducted. The distribution of data is also important for evaluating the appropriateness of the statistical techniques for hypothesis testing or model building.

As a first step, descriptive statistics is important to provide a concise and easily understood summary of the characteristics of a data set. Measure of central tendency (mean) and measures of dispersion (maximum, minimum, range, standard deviation, and variance) are important descriptive statistics. Creating graphical representation is another technique that shows us data characteristics in a visual way. Histograms and Boxplots can be used to create the graphical representation of data (Norušis, 1993).

The histogram is commonly used to represent data graphically and based on the frequency distribution. It is basically a graph of vertical bars in which the height of each bar corresponds to the frequency of occurrence of a given X score (Grimm, 1993). Each bar in a histogram represents the number of cases with values within the interval (Norušis 1993). The boxplot also summarizes and displays further information about the distribution of the values (Figure 5.01). Instead of plotting the actual values, a boxplot displays summary statistics for the distribution by plotting the median, the 25th percentile, the 75th percentile, and values that are far removed from the rest (Norušis 1993). If the median is not defined in the center of the box, this shows that the distribution of the data is skewed. The median's position that is closer to the left of the box indicates positively skewed data while the median that is closer to the right of the box displays negatively skewed values.

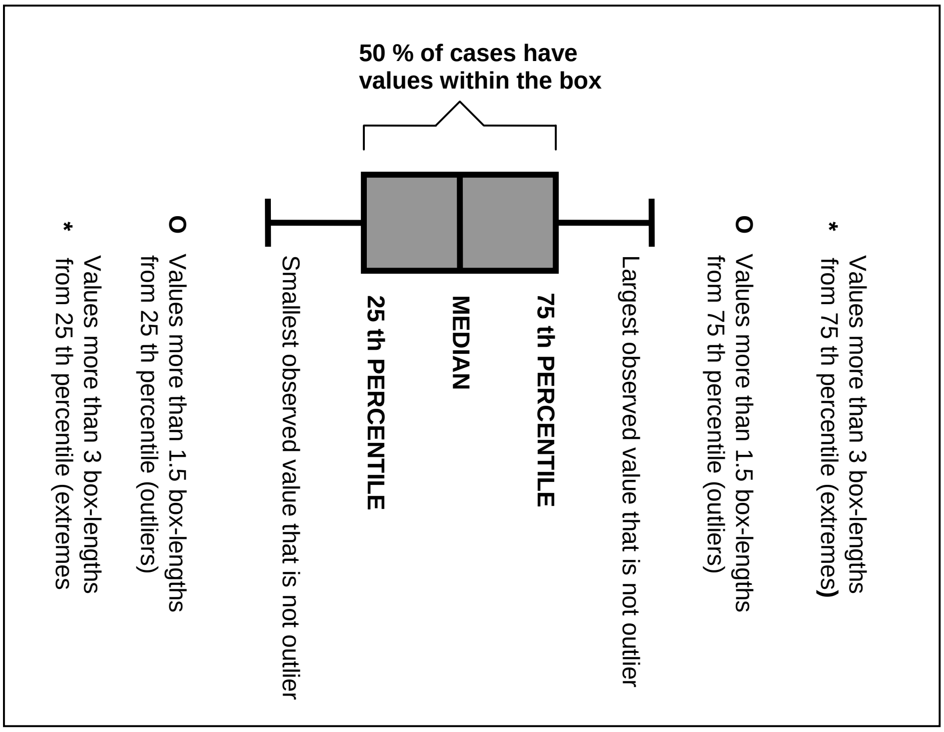


Figure 5.01. Illustrated sketch of a boxplot (Norušis 1993)

Both descriptive statistics and graphical representations were used to explore data characteristics of this study. Range, minimum, maximum, mean, standard deviation, and variance were employed as descriptive statistics, and they were summarized in Table 5.02. Histograms and boxplots were also utilized to gain more insight about the overall data visually. The histograms of the data were shown in Figure 5.02 and Figure 5.03, and the boxplots were given in Figure 5.04 and Figure 5.05. All statistics were conducted by using SPSS and SYSTAT software (SPSS, 2001; SPSS, 1997) in this study.

Descriptive statistics have produced valuable information about the data characteristics. Although all of the topographic variables (elevation, slope, and aspect) slightly skewed (positive skewness), they nearly showed normal distribution characteristics without outliers and extremes (Figure 5.04 and Figure 5.05). Diversity indices also display the normal distribution characteristics without outliers and extremes. Shannon Wiener index variable, for instance, is nearly identical to normal distribution, while Simpson index variable is slightly skewed (positive skewness). The distribution of $CaCO_3$ and salt delineate normal distribution characteristics with one exteme. On the other hand, soil nutrients (organic matter, P_2O_5, and K_2O) positively skewed with some extremes and outliers. When we looked at the climatic variables, they also delineate nearly normal distribution characteristics without outliers and extremes. Only seasonal maximum temperature variable has one extreme, but it also shows normal distribution characteristics.

The histograms and boxplots suggested that the overall data has nearly normal distribution characteristics and there is no need to normality test for any variable. Consequently, it is decided that the overall data characteristics are suitable for any parametric statistical solutions.

Table 5.02. Descriptive statistics of the data

	N	Range	Min	Max	Mean	Std. Deviation	Variance
ELEV	56	1060	650	1710	1040.84	243.80	59440.27
ASPECT	56	4	1	5	3.34	1.30	1.68
SLOPE	56	41.36	.00	41.36	16.11	10.04	100.76
ORGM	56	5.54	.94	6.48	2.80	1.13	1.27
P2O5	56	22.79	.60	23.39	4.64	4.89	23.88
K2O	56	504.83	30.74	535.57	124.16	77.29	5973.17
CACO3	56	25.58	.54	26.12	11.90	5.73	32.84
PH	56	1.56	6.76	8.32	7.76	0.24	0.06
SALT	56	.084	.000	.084	0.05	0.02	0.00
STR	56	34.98	32.02	67.00	48.88	7.04	49.56
TEXTR	56	1	1	2	1.30	0.46	0.22
SOILG	56	1	1	2	1.96	0.19	0.04
ERS	56	3	1	4	1.87	0.81	0.66
SLDPT	56	3	2	5	2.34	0.79	0.63
GEO	56	6	1	7	3.50	2.44	5.93
NDVI	56	141	58	199	106.79	33.39	1115.08
SPVSD	56	4	1	5	3.43	1.31	1.70
SWI	56	2.750	.250	3.000	1.67	0.76	0.59
SIMP	56	.872	.056	.928	0.37	0.26	0.07
META	56	3.87	7.71	11.58	10.21	0.86	0.73
METS	56	2.80	16.16	18.96	17.97	0.62	0.39
MAXTA	56	2.11	15.09	17.20	16.38	0.51	0.26
MAXTS	56	.50	25.23	25.73	25.51	0.13	0.02
MINTA	56	5.35	.83	6.18	4.20	1.21	1.46
MINTS	56	4.63	7.65	12.28	10.50	1.05	1.11
PRCPA	56	24.63	16.50	41.13	28.63	6.49	42.07
PRCPS	56	12.96	18.07	31.03	22.84	3.37	11.35
PETAN	56	6.90	71.87	78.77	74.92	1.65	2.72
PETSE	56	17.23	130.81	148.04	137.53	4.18	17.49

ABBREVIATIONS: **ELEV:** elevation, **ASP:** aspect, **STR:** saturation, **TEXTR:** texture, **SOILG:** big soil group, **ERS:** erosion, **SLDPT:** soil depth, **GEO:** geological formations, **SPVSD:** supervised classes, **SWI:** Shannon Wiener index, **SIMP:** Simpson index, **META:** annual mean temperature, **METS:** seasonal mean temperature, **MAXTA:** annual maximum temperature, **MAXTS:** seasonal maximum temperature, **MINTA:** annual minimum temperature, **MINTS:** seasonal minimum temperature, **PRCPA:** annual precipitation, **PRCPS:** seasonal precipitation, **PETAN:** annual potential evapotranspiration, **PETSE:** seasonal potential evapotranspiration

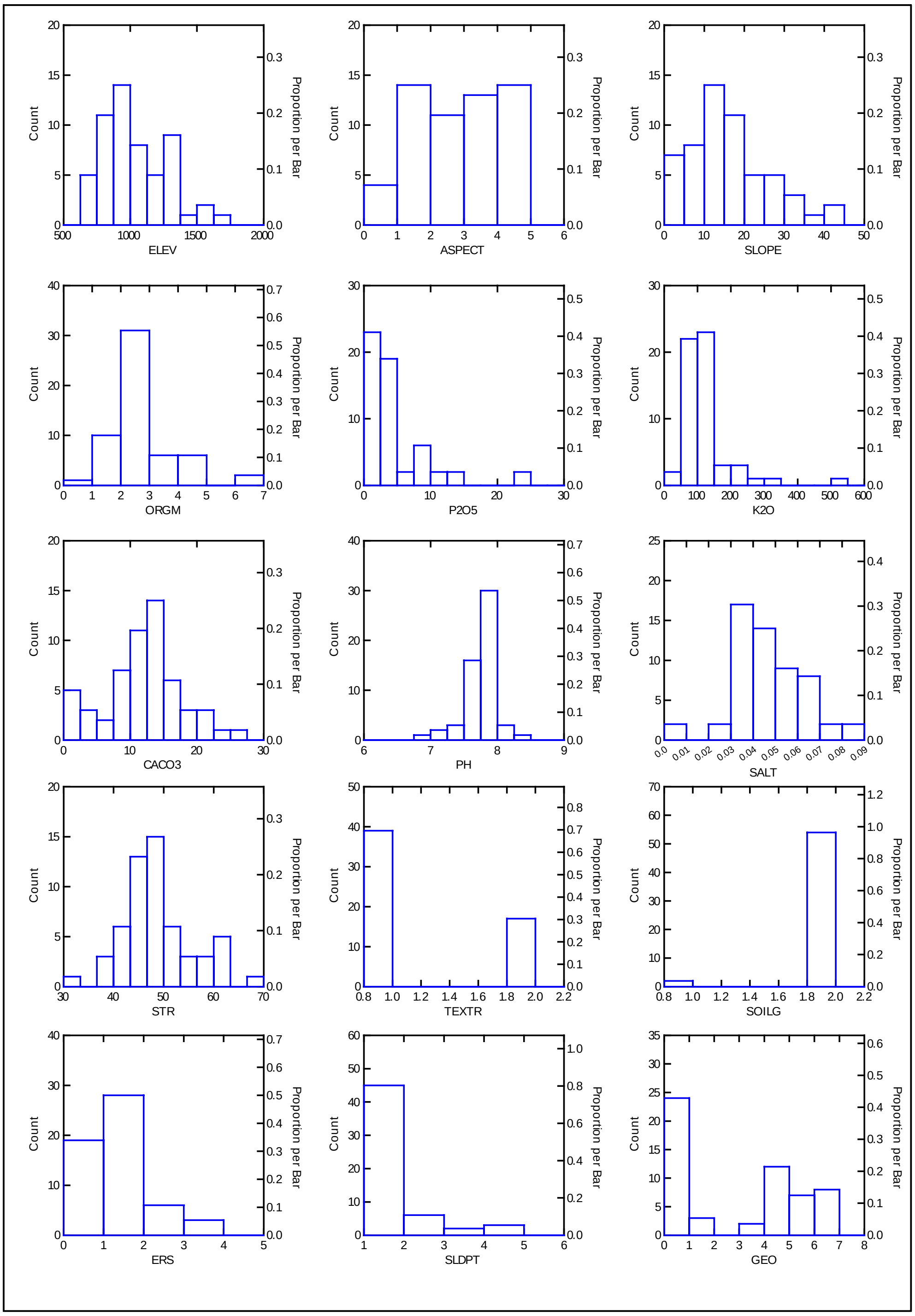


Figure 5.02. The histograms of topography, soil, and geology variables (ELEV (elevation): m, ASPECT:1-8, SLOPE: %, ORGM (organic matter): %, P2O5: kg/da, K2O: kg/da, CaCO3: kg/da, pH: 4-9 scale, SALT: %, STR (saturation): %, TEXTR (texture): (1) loam -(2) clay loam, SOILG (soil group): (1)alluvial-colluvial soils, (2) brown forest soil, ERS (erosion): 1-4, SLDPT (soil depth): 1-5, GEO (geology): 1-7) NOTE: 1 da = 1000 m^2

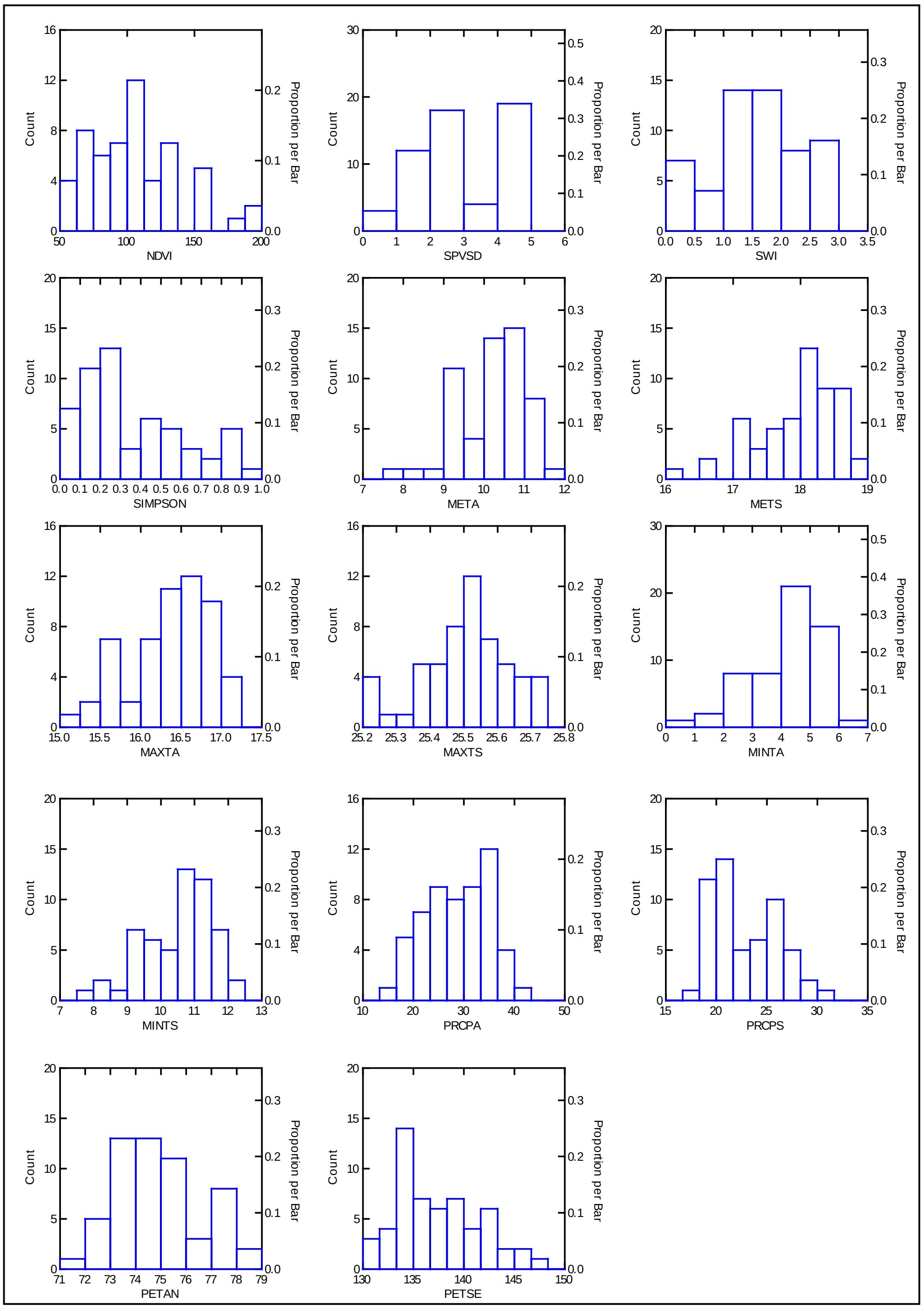


Figure 5.03. The histograms of NDVI-Supervised classes, indices, and climatic variables (NDVI (Normalized Difference Vegetative Index): 0-256, SPVSD (supervised classes from LANDSAT-TM image): 1-5, SWI (Shannon Wiener Index): 0.250-2.750, SIMP (Simpson Index): 0.056-0.928, temperature (META: Annual Mean Temperature, METS: seasonal mean temperature, MAXTA: annual maximum temperature, MAXTS: seasonal maximum temperature, MINTA: annual minimum temperature, MINTS: seasonal minimum temperature): ^{o}C, precipitation (PRCPA: annual precipitation, PRCPS: seasonal precipitation): mm, evapotranspiration (PETAN: annual potential evapotranspiration, PETSE: seasonal potential evapotranspiration): mm)

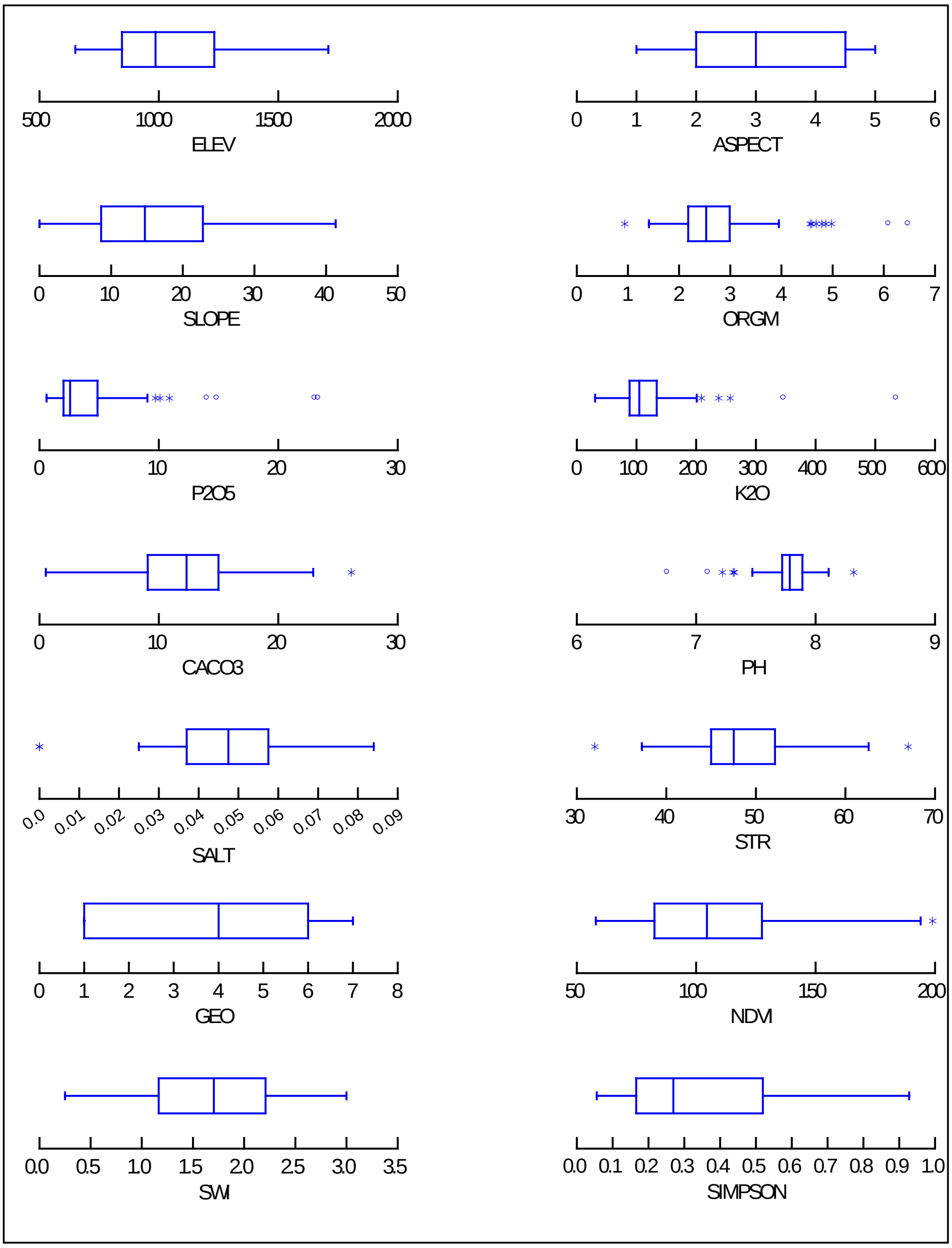


Figure 5.04. The boxplots of the topographic, soil, geology, NDVI, and index variables (ELEV (elevation): m, ASPECT:1-8, SLOPE: %, ORGM (organic matter): %, P2O5: kg/da, K2O: kg/da, CaCO3: kg/da, pH: 4-9 scale, SALT: %, STR (saturation): %, GEO (geology): 1-7, NDVI (Normalized Difference Vegetative Index): 0-255, SPVSD (supervised classes from LANDSAT-TM image): 1-5, SWI (Shannon Wiener Index): 0.250-2.750, SIMP (Simpson Index): 0.056-0.928) NOTE: 1 da = 1000 m^2

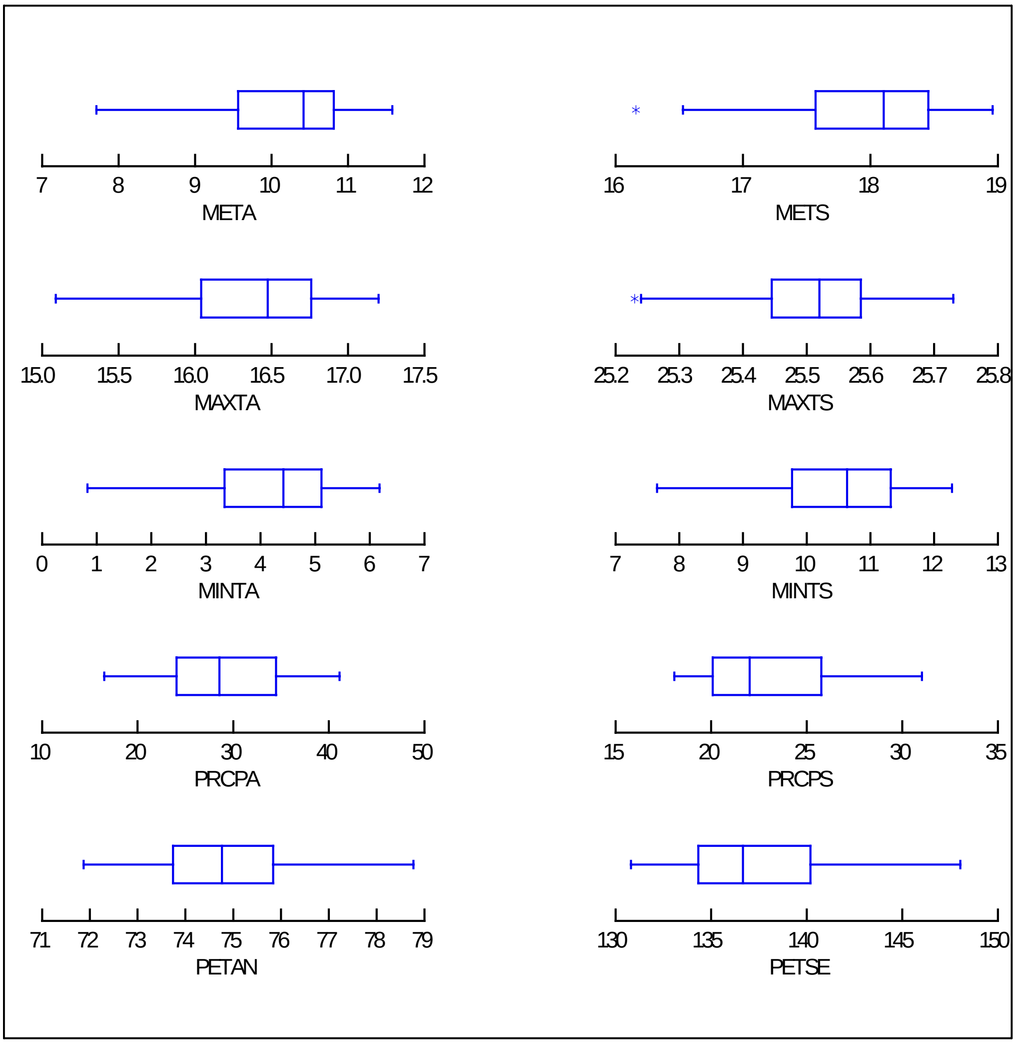


Figure 5.05. The boxplots of the climatic variables (temperature (META: Annual Mean Temperature, METS: seasonal mean temperature, MAXTA: annual maximum temperature, MAXTS: seasonal maximum temperature, MINTA: annual minimum temperature, MINTS: seasonal minimum temperature): ^{o}C, precipitation (PRCPA: annual precipitation, PRCPS: seasonal precipitation): mm, evapotranspiration (PETAN: annual potential evapotranspiration, PETSE: seasonal potential evapotranspiration): mm)

5.3. Data reduction

Although, classical *correlation analysis* approach can be used to investigate the relation between independent and dependent variables, *factor analysis* was chosen in this study. It was thought that *factor analysis* might be more convenient to study the patterns of relationship among many dependent variables. The *Pearson Correlation Matrix* of the data set were given as a supplementary information in Appendix B.

The purpose of data reduction is to remove redundant (highly correlated) variables from the data file, perhaps replacing the entire data file with a smaller number of uncorrelated variables. The *factor analysis* procedure has several extraction methods for constructing a solution. One of these methods is *principle component* analysis. The *principal component* method of extraction begins by finding a linear combination of variables (a component) that accounts for as much variation in the original variables as possible. It then finds another component that accounts for as much of the remaining variation as possible and is uncorrelated with the previous component, continuing in this way until there are as many components as original variables. Usually, a few components will account for most of the variation, and these components can be used to replace the original variables. This method is most often used to reduce the number of variables in the data file.

In the initial stage, all of the 28 variables were included in the *factor analysis*, and principal component method with varimax rotation was selected as the *factor analysis* method. A number of tests can be performed for the validity of *factor analysis* with the given variables. KMO-Bartlett tests were conducted for both Shannon Wiener and Simpson indices in order to see the suitability of *factor analysis* (Table 5.03). KMO-Bartlett test results show two tests that indicate the suitability of the data for *factor analysis*. The Kaiser-Meyer-Olkin (KMO) Measure of Sampling Adequacy is a statistic that reveals the proportion of variance in the variables which is common variance, i.e. which might be caused by underlying factors. High values (close to 1.0) generally indicate that a *factor analysis* may be useful with the available data. If the value is less than 0.50, the results of the *factor analysis* probably will not be very useful (SPSS, 2001). In this study, they were 0.720 and 0.719 for Shannon Wiener and Simpson indices, respectively (Table 5.03). Bartlett's test of sphericity indicates whether the correlation matrix is an identity matrix, which would indicate that the variables are unrelated. The significance level gives the result of the test. Very small values (less than 0.05) indicate that there are probably significant relationships among the variables (Table 5.03).

Table 5.03. KMO and Bartlett's test with initial 28 variables

Shannon Wiener Index		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		0.720
Bartlett's Test of Sphericity	Approx. Chi-Square	2922.703
	Df	378
	Sig.	0.000
Simpson Index		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		0.719
Bartlett's Test of Sphericity	Approx. Chi-Square	2874.778
	Df	378
	Sig.	0.000

A value higher than about 0.10 or so may indicate that the data are not suitable for *factor analysis*. The significance levels were found 0.00 for both Shannon Wiener and Simpson indices (Table 5.03). Based on both KMO and Bartlett's test of sphericity values, *factor analysis* is supposed to produce very successful results in this study.

The next step in the analyses is to explain which variables are not fitting to the model, which could be done by exploring the anti-image matrices. The anti-image matrices contain the negative partial covariance and correlations. They can give an indication of correlations which aren't due to the common factors. Small values indicate that the variables are relatively free of unexplained correlations. Most or all values off the diagonal should be small (close to zero). Each value on the diagonal of the anti-image correlation matrix shows the Measure of Sampling Adequacy (MSA) for the respective item. Values less than 0.5 may indicate variables that do not seem to fit with the structure of the other variables, and can be considered to drop from the analysis.

The anti image covariance matrices of the initial 28 variables for Shannon Wiener and Simpson indices are presented in Table 5.04 and Table 5.06, and correlation matrices of these are given in Table 5.05 and Table 5.07, respectively. According to the anti-image correlation matrices; aspect, slope, P_2O_5, K_2O, salt, erosion and maximum seasonal temperature variables do not fit into the structure of the remaining variables. As this is the first iterative pass of the system, these unfitting parameters were included in order to see the effects of them before and after their exclusion. Consequently, the amount of the total variance explained with the initial 28 variables for Shannon Wiener and Simpson indices were also shown in Table 5.08 and Table 5.09. These tables give the amount of cumulative variance explained with the initial solution and initial rotation of the factor analyses.

Table 5.04. The anti-image covariance matrice of initial 28 variables for Shannon Wiener Index

	ELEV	ASP	SLOPE	ORGM	P2O5	K2O	CACO3	PH	SALT	STR	TEXTR	SOILG	ERS	SLDPT
ELEV	0.0004	-0.0018	0.0020	-0.0018	0.0007	0.0015	0.0008	-0.0004	-0.0013	0.0013	-0.0025	0.0001	0.0018	-0.0003
ASP	-0.0018	0.5040	0.0728	-0.0399	-0.0315	0.0210	-0.0596	0.0518	-0.0455	0.0198	0.0131	-0.0255	-0.0009	-0.0016
SLOPE	0.0020	0.0728	0.3390	-0.0516	0.0087	-0.0164	-0.0426	0.0125	-0.0172	0.0150	0.0133	-0.0009	-0.0192	0.0261
ORGM	-0.0018	-0.0399	-0.0516	0.0788	-0.0143	-0.0353	-0.0262	0.0131	0.0491	-0.0381	0.0286	-0.0076	0.0069	-0.0180
P2O5	0.0007	-0.0315	0.0087	-0.0143	0.1950	-0.0728	-0.0017	0.0813	-0.0387	0.0104	0.0090	0.0393	0.0174	-0.0009
K2O	0.0015	0.0210	-0.0164	-0.0353	-0.0728	0.1090	0.0586	-0.0663	-0.0130	0.0258	-0.0483	-0.0164	-0.0336	0.0421
CACO3	0.0008	-0.0596	-0.0426	-0.0262	-0.0017	0.0586	0.1500	-0.0709	-0.0221	0.0234	-0.0268	0.0421	-0.0067	0.0306
PH	-0.0004	0.0518	0.0125	0.0131	0.0813	-0.0663	-0.0709	0.1070	-0.0290	-0.0006	0.0133	-0.0005	0.0220	-0.0227
SALT	-0.0013	-0.0455	-0.0172	0.0491	-0.0387	-0.0130	-0.0221	-0.0290	0.0824	-0.0369	0.0234	-0.0266	-0.0185	-0.0041
STR	0.0013	0.0198	0.0150	-0.0381	0.0104	0.0258	0.0234	-0.0006	-0.0369	0.0270	-0.0339	0.0016	0.0037	0.0069
TEXTR	-0.0025	0.0131	0.0133	0.0286	0.0090	-0.0483	-0.0268	0.0133	0.0234	-0.0339	0.1180	0.0624	-0.0300	0.0191
SOILG	0.0001	-0.0255	-0.0009	-0.0076	0.0393	-0.0164	0.0421	-0.0005	-0.0266	0.0016	0.0624	0.2970	-0.0222	0.0704
ERS	0.0018	-0.0009	-0.0192	0.0069	0.0174	-0.0336	-0.0067	0.0220	-0.0185	0.0037	-0.0300	-0.0222	0.1220	-0.0885
SLDPT	-0.0003	-0.0016	0.0261	-0.0180	-0.0009	0.0421	0.0306	-0.0227	-0.0041	0.0069	0.0191	0.0704	-0.0885	0.0993
GEO	0.0014	-0.0433	-0.0085	0.0126	-0.0215	0.0079	0.0315	-0.0185	-0.0074	-0.0002	-0.0134	-0.0075	0.0121	-0.0127
NDVI	-0.0057	0.1310	-0.0157	0.0198	-0.0108	-0.0472	-0.0727	0.0450	-0.0003	-0.0202	0.0622	0.0673	-0.0236	-0.0066
SPVSD	0.0034	-0.0655	-0.0281	-0.0022	-0.0438	0.0670	0.1030	-0.0813	-0.0039	0.0079	-0.0241	-0.0169	0.0193	0.0106
SWI	-0.0014	0.0635	-0.0100	-0.0184	0.0072	0.0088	-0.0190	0.0202	-0.0179	0.0085	0.0038	0.0045	-0.0151	0.0194
META	0.0001	-0.0013	0.0010	-0.0005	0.0014	-0.0016	-0.0006	0.0012	-0.0008	0.0003	-0.0005	-0.0005	0.0024	-0.0020
METS	0.0000	0.0004	0.0003	0.0002	-0.0023	0.0031	0.0004	-0.0026	0.0009	-0.0002	-0.0002	0.0012	-0.0031	0.0031
MAXTA	0.0002	0.0022	0.0045	-0.0024	0.0012	0.0015	0.0020	0.0001	-0.0018	0.0016	-0.0009	0.0013	-0.0023	0.0024
MAXTS	-0.0012	-0.0101	-0.0086	0.0058	-0.0035	-0.0159	-0.0203	0.0069	0.0007	-0.0050	0.0098	-0.0123	0.0198	-0.0201
MINTA	-0.0001	0.0035	0.0060	-0.0026	-0.0056	0.0032	0.0005	-0.0050	0.0029	-0.0004	0.0028	0.0007	-0.0053	0.0036
MINTS	0.0002	-0.0035	-0.0077	0.0010	0.0041	-0.0006	0.0012	0.0023	-0.0018	0.0006	-0.0032	-0.0014	0.0036	-0.0022
PRCPA	0.0003	-0.0009	-0.0047	0.0008	-0.0014	0.0042	0.0004	-0.0007	-0.0016	0.0009	-0.0037	-0.0015	0.0013	0.0004
PRCPS	-0.0004	0.0027	0.0142	-0.0039	0.0032	-0.0084	-0.0011	0.0008	0.0027	-0.0012	0.0086	0.0067	-0.0040	0.0006
PETAN	0.0001	0.0044	0.0093	0.0005	-0.0044	-0.0012	-0.0045	0.0000	0.0016	-0.0010	0.0024	0.0009	0.0017	-0.0010
PETSE	-0.0001	-0.0023	-0.0058	-0.0004	0.0033	0.0003	0.0026	0.0007	-0.0015	0.0008	-0.0017	-0.0009	-0.0008	0.0003

ABBREVIATIONS: **ELEV:** elevation, **ASP:** aspect, **STR:** saturation, **TEXTR:** texture, **SOILG:** big soil group, **ERS:** erosion, **SLDPT:** soil depth, **GEO:** geological formations, **SPVSD:** supervised classes, **SWI:** Shannon Wiener index, **META**: annual mean temperature, **METS**: seasonal mean temperature, **MAXTA:** annual maximum temperature, **MAXTS:** seasonal maximum temperature, **MINTA:** annual minimum temperature, **MINTS:** seasonal minimum temperature, **PRCPA:** annual precipitation, **PRCPS:** seasonal precipitation, **PETAN:** annual potential evapotranspiration, **PETSE:** seasonal potential evapotranspiration

Table 5.04. Continue (the anti-image covariance matrice of initial 28 variables for Shannon Wiener Index)

	GEO	NDVI	SPVSD	SWI	META	METS	MAXTA	MAXTS	MINTA	MINTS	PRCPA	PRCPE	PETAN	PETSE
ELEV	0.0014	-0.0057	0.0034	-0.0014	0.0001	0.0000	0.0002	-0.0012	-0.0001	0.0002	0.0003	-0.0004	0.0001	-0.0001
ASP	-0.0433	0.1310	-0.0655	0.0635	-0.0013	0.0004	0.0022	-0.0101	0.0035	-0.0035	-0.0009	0.0027	0.0044	-0.0023
SLOPE	-0.0085	-0.0157	-0.0281	-0.0100	0.0010	0.0003	0.0045	-0.0086	0.0060	-0.0077	-0.0047	0.0142	0.0093	-0.0058
ORGM	0.0126	0.0198	-0.0022	-0.0184	-0.0005	0.0002	-0.0024	0.0058	-0.0026	0.0010	0.0008	-0.0039	0.0005	-0.0004
P2O5	-0.0215	-0.0108	-0.0438	0.0072	0.0014	-0.0023	0.0012	-0.0035	-0.0056	0.0041	-0.0014	0.0032	-0.0044	0.0033
K2O	0.0079	-0.0472	0.0670	0.0088	-0.0016	0.0031	0.0015	-0.0159	0.0032	-0.0006	0.0042	-0.0084	-0.0012	0.0003
CACO3	0.0315	-0.0727	0.1030	-0.0190	-0.0006	0.0004	0.0020	-0.0203	0.0005	0.0012	0.0004	-0.0011	-0.0045	0.0026
PH	-0.0185	0.0450	-0.0813	0.0202	0.0012	-0.0026	0.0001	0.0069	-0.0050	0.0023	-0.0007	0.0008	0.0000	0.0007
SALT	-0.0074	-0.0003	-0.0039	-0.0179	-0.0008	0.0009	-0.0018	0.0007	0.0029	-0.0018	-0.0016	0.0027	0.0016	-0.0015
STR	-0.0002	-0.0202	0.0079	0.0085	0.0003	-0.0002	0.0016	-0.0050	-0.0004	0.0006	0.0009	-0.0012	-0.0010	0.0008
TEXTR	-0.0134	0.0622	-0.0241	0.0038	-0.0005	-0.0002	-0.0009	0.0098	0.0028	-0.0032	-0.0037	0.0086	0.0024	-0.0017
SOILG	-0.0075	0.0673	-0.0169	0.0045	-0.0005	0.0012	0.0013	-0.0123	0.0007	-0.0014	-0.0015	0.0067	0.0009	-0.0009
ERS	0.0121	-0.0236	0.0193	-0.0151	0.0024	-0.0031	-0.0023	0.0198	-0.0053	0.0036	0.0013	-0.0040	0.0017	-0.0008
SLDPT	-0.0127	-0.0066	0.0106	0.0194	-0.0020	0.0031	0.0024	-0.0201	0.0036	-0.0022	0.0004	0.0006	-0.0010	0.0003
GEO	0.1040	-0.0560	0.0887	-0.0651	0.0002	-0.0001	-0.0001	-0.0061	-0.0023	0.0018	0.0030	-0.0080	-0.0001	-0.0004
NDVI	-0.0560	0.3690	-0.1790	0.0551	-0.0006	-0.0005	-0.0045	0.0341	0.0008	-0.0024	-0.0039	0.0080	0.0008	-0.0001
SPVSD	0.0887	-0.1790	0.2860	-0.0640	-0.0002	0.0015	0.0007	-0.0132	-0.0003	0.0020	0.0055	-0.0124	-0.0023	0.0006
SWI	-0.0651	0.0551	-0.0640	0.0706	-0.0004	0.0000	0.0000	0.0068	0.0007	-0.0006	-0.0009	0.0027	-0.0006	0.0008
META	0.0002	-0.0006	-0.0002	-0.0004	0.0003	-0.0004	-0.0002	0.0024	-0.0001	0.0002	-0.0002	0.0004	0.0000	0.0000
METS	-0.0001	-0.0005	0.0015	0.0000	-0.0004	0.0008	0.0004	-0.0040	0.0002	-0.0003	0.0006	-0.0009	0.0001	-0.0001
MAXTA	-0.0001	-0.0045	0.0007	0.0000	-0.0002	0.0004	0.0009	-0.0061	0.0001	-0.0002	0.0001	0.0000	0.0000	0.0000
MAXTS	-0.0061	0.0341	-0.0132	0.0068	0.0024	-0.0040	-0.0061	0.0553	-0.0004	0.0009	-0.0022	0.0027	0.0003	-0.0003
MINTA	-0.0023	0.0008	-0.0003	0.0007	-0.0001	0.0002	0.0001	-0.0004	0.0013	-0.0007	-0.0004	0.0010	0.0002	-0.0002
MINTS	0.0018	-0.0024	0.0020	-0.0006	0.0002	-0.0003	-0.0002	0.0009	-0.0007	0.0007	0.0003	-0.0009	-0.0004	0.0002
PRCPA	0.0030	-0.0039	0.0055	-0.0009	-0.0002	0.0006	0.0001	-0.0022	-0.0004	0.0003	0.0012	-0.0024	0.0000	0.0000
PRCPS	-0.0080	0.0080	-0.0124	0.0027	0.0004	-0.0009	0.0000	0.0027	0.0010	-0.0009	-0.0024	0.0053	0.0002	-0.0001
PETAN	-0.0001	0.0008	-0.0023	-0.0006	0.0000	0.0001	0.0000	0.0003	0.0002	-0.0004	0.0000	0.0002	0.0013	-0.0008
PETSE	-0.0004	-0.0001	0.0006	0.0008	0.0000	-0.0001	0.0000	-0.0003	-0.0002	0.0002	0.0000	-0.0001	-0.0008	0.0006

ABBREVIATIONS: **ELEV:** elevation, **ASP:** aspect, **STR:** saturation, **TEXTR:** texture, **SOILG:** big soil group, **ERS:** erosion, **SLDPT:** soil depth, **GEO:** geological formations, **SPVSD:** supervised classes, **SWI:** Shannon Wiener index, **META**: annual mean temperature, **METS**: seasonal mean temperature, **MAXTA:** annual maximum temperature, **MAXTS:** seasonal maximum temperature, **MINTA:** annual minimum temperature, **MINTS:** seasonal minimum temperature, **PRCPA:** annual precipitation, **PRCPS:** seasonal precipitation, **PETAN:** annual potential evapotranspiration, **PETSE:** seasonal potential evapotranspiration

Table 5.05. The anti-image correlation matrice of initial 28 variables for Shannon Wiener Index

	ELEV	ASP	SLOPE	ORGM	P2O5	K2O	CACO3	PH	SALT	STR	TEXTR	SOILG	ERS	SLDPT
ELEV	0.8560	-0.1260	0.1680	-0.3150	0.0821	0.2250	0.0984	-0.0641	-0.2250	0.3730	-0.3460	0.0052	0.2500	-0.0499
ASP	-0.1260	0.3870	0.1760	-0.2000	-0.1000	0.0895	-0.2170	0.2230	-0.2230	0.1700	0.0537	-0.0658	-0.0037	-0.0070
SLOPE	0.1680	0.1760	0.4680	-0.3150	0.0336	-0.0853	-0.1890	0.0658	-0.1030	0.1570	0.0662	-0.0030	-0.0942	0.1420
ORGM	-0.3150	-0.2000	-0.3150	0.5250	-0.1150	-0.3810	-0.2410	0.1430	0.6090	-0.8250	0.2960	-0.0495	0.0702	-0.2030
P2O5	0.0821	-0.1000	0.0336	-0.1150	0.4080	-0.4980	-0.0098	0.5620	-0.3050	0.1430	0.0590	0.1630	0.1130	-0.0063
K2O	0.2250	0.0895	-0.0853	-0.3810	-0.4980	0.3410	0.4580	-0.6130	-0.1370	0.4740	-0.4240	-0.0910	-0.2910	0.4040
CACO3	0.0984	-0.2170	-0.1890	-0.2410	-0.0098	0.4580	0.6100	-0.5590	-0.1990	0.3680	-0.2010	0.2000	-0.0494	0.2510
PH	-0.0641	0.2230	0.0658	0.1430	0.5620	-0.6130	-0.5590	0.5760	-0.3090	-0.0103	0.1180	-0.0028	0.1920	-0.2200
SALT	-0.2250	-0.2230	-0.1030	0.6090	-0.3050	-0.1370	-0.1990	-0.3090	0.4360	-0.7820	0.2360	-0.1700	-0.1850	-0.0449
STR	0.3730	0.1700	0.1570	-0.8250	0.1430	0.4740	0.3680	-0.0103	-0.7820	0.5740	-0.5980	0.0183	0.0648	0.1330
TEXTR	-0.3460	0.0537	0.0662	0.2960	0.0590	-0.4240	-0.2010	0.1180	0.2360	-0.5980	0.6090	0.3330	-0.2500	0.1760
SOILG	0.0052	-0.0658	-0.0030	-0.0495	0.1630	-0.0910	0.2000	-0.0028	-0.1700	0.0183	0.3330	0.7830	-0.1170	0.4100
ERS	0.2500	-0.0037	-0.0942	0.0702	0.1130	-0.2910	-0.0494	0.1920	-0.1850	0.0648	-0.2500	-0.1170	0.4420	-0.8050
SLDPT	-0.0499	-0.0070	0.1420	-0.2030	-0.0063	0.4040	0.2510	-0.2200	-0.0449	0.1330	0.1760	0.4100	-0.8050	0.5690
GEO	0.2160	-0.1890	-0.0455	0.1390	-0.1510	0.0739	0.2520	-0.1750	-0.0802	-0.0038	-0.1210	-0.0430	0.1080	-0.1250
NDVI	-0.4590	0.3040	-0.0445	0.1160	-0.0402	-0.2350	-0.3090	0.2260	-0.0016	-0.2030	0.2980	0.2030	-0.1110	-0.0346
SPVSD	0.3120	-0.1730	-0.0902	-0.0146	-0.1850	0.3790	0.4960	-0.4640	-0.0252	0.0901	-0.1310	-0.0578	0.1030	0.0629
SWI	-0.2630	0.3370	-0.0644	-0.2470	0.0614	0.1010	-0.1840	0.2320	-0.2340	0.1940	0.0418	0.0310	-0.1630	0.2310
META	0.3140	-0.1150	0.1100	-0.1070	0.1890	-0.2990	-0.0880	0.2320	-0.1720	0.1250	-0.0832	-0.0555	0.4290	-0.3910
METS	-0.0331	0.0205	0.0200	0.0240	-0.1880	0.3420	0.0409	-0.2850	0.1180	-0.0457	-0.0157	0.0816	-0.3230	0.3480
MAXTA	0.3060	0.1040	0.2570	-0.2910	0.0893	0.1530	0.1750	0.0085	-0.2120	0.3150	-0.0890	0.0794	-0.2180	0.2580
MAXTS	-0.2470	-0.0605	-0.0631	0.0881	-0.0338	-0.2040	-0.2220	0.0895	0.0104	-0.1280	0.1200	-0.0958	0.2420	-0.2720
MINTA	-0.0681	0.1360	0.2840	-0.2560	-0.3500	0.2670	0.0390	-0.4210	0.2800	-0.0612	0.2220	0.0346	-0.4200	0.3190
MINTS	0.2840	-0.1810	-0.4890	0.1340	0.3470	-0.0677	0.1200	0.2630	-0.2270	0.1390	-0.3430	-0.0924	0.3880	-0.2610
PRCPA	0.3920	-0.0367	-0.2350	0.0843	-0.0915	0.3740	0.0324	-0.0663	-0.1670	0.1520	-0.3180	-0.0788	0.1100	0.0397
PRCPS	-0.2740	0.0528	0.3350	-0.1910	0.0991	-0.3490	-0.0401	0.0346	0.1290	-0.0977	0.3410	0.1690	-0.1570	0.0272
PETAN	0.1900	0.1740	0.4490	0.0537	-0.2820	-0.0986	-0.3270	0.0007	0.1540	-0.1720	0.1920	0.0456	0.1340	-0.0861
PETSE	-0.2120	-0.1370	-0.4210	-0.0581	0.3130	0.0383	0.2870	0.0894	-0.2190	0.2060	-0.2100	-0.0669	-0.0912	0.0401

ABBREVIATIONS: **ELEV:** elevation, **ASP:** aspect, **STR:** saturation, **TEXTR:** texture, **SOILG:** big soil group, **ERS:** erosion, **SLDPT:** soil depth, **GEO:** geological formations, **SPVSD:** supervised classes, **SWI:** Shannon Wiener index, **META**: annual mean temperature, **METS**: seasonal mean temperature, **MAXTA:** annual maximum temperature, **MAXTS:** seasonal maximum temperature, **MINTA:** annual minimum temperature, **MINTS:** seasonal minimum temperature, **PRCPA:** annual precipitation, **PRCPS:** seasonal precipitation, **PETAN:** annual potential evapotranspiration, **PETSE:** seasonal potential evapotranspiration

Table 5.05. Continue (the anti-image correlation matrice of initial 28 variables for Shannon Wiener Index)

	GEO	NDVI	SPVSD	SWI	META	METS	MAXTA	MAXTS	MINTA	MINTS	PRCPA	PRCPE	PETAN	PETSE
ELEV	0.2160	-0.4590	0.3120	-0.2630	0.3140	-0.0331	0.3060	-0.2470	-0.0681	0.2840	0.3920	-0.2740	0.1900	-0.2120
ASP	-0.1890	0.3040	-0.1730	0.3370	-0.1150	0.0205	0.1040	-0.0605	0.1360	-0.1810	-0.0367	0.0528	0.1740	-0.1370
SLOPE	-0.0455	-0.0445	-0.0902	-0.0644	0.1100	0.0200	0.2570	-0.0631	0.2840	-0.4890	-0.2350	0.3350	0.4490	-0.4210
ORGM	0.1390	0.1160	-0.0146	-0.2470	-0.1070	0.0240	-0.2910	0.0881	-0.2560	0.1340	0.0843	-0.1910	0.0537	-0.0581
P2O5	-0.1510	-0.0402	-0.1850	0.0614	0.1890	-0.1880	0.0893	-0.0338	-0.3500	0.3470	-0.0915	0.0991	-0.2820	0.3130
K2O	0.0739	-0.2350	0.3790	0.1010	-0.2990	0.3420	0.1530	-0.2040	0.2670	-0.0677	0.3740	-0.3490	-0.0986	0.0383
CACO3	0.2520	-0.3090	0.4960	-0.1840	-0.0880	0.0409	0.1750	-0.2220	0.0390	0.1200	0.0324	-0.0401	-0.3270	0.2870
PH	-0.1750	0.2260	-0.4640	0.2320	0.2320	-0.2850	0.0085	0.0895	-0.4210	0.2630	-0.0663	0.0346	0.0007	0.0894
SALT	-0.0802	-0.0016	-0.0252	-0.2340	-0.1720	0.1180	-0.2120	0.0104	0.2800	-0.2270	-0.1670	0.1290	0.1540	-0.2190
STR	-0.0038	-0.2030	0.0901	0.1940	0.1250	-0.0457	0.3150	-0.1280	-0.0612	0.1390	0.1520	-0.0977	-0.1720	0.2060
TEXTR	-0.1210	0.2980	-0.1310	0.0418	-0.0832	-0.0157	-0.0890	0.1200	0.2220	-0.3430	-0.3180	0.3410	0.1920	-0.2100
SOILG	-0.0430	0.2030	-0.0578	0.0310	-0.0555	0.0816	0.0794	-0.0958	0.0346	-0.0924	-0.0788	0.1690	0.0456	-0.0669
ERS	0.1080	-0.1110	0.1030	-0.1630	0.4290	-0.3230	-0.2180	0.2420	-0.4200	0.3880	0.1100	-0.1570	0.1340	-0.0912
SLDPT	-0.1250	-0.0346	0.0629	0.2310	-0.3910	0.3480	0.2580	-0.2720	0.3190	-0.2610	0.0397	0.0272	-0.0861	0.0401
GEO	0.5870	-0.2860	0.5140	-0.7600	0.0313	-0.0165	-0.0077	-0.0799	-0.1980	0.2050	0.2740	-0.3410	-0.0071	-0.0570
NDVI	-0.2860	0.4920	-0.5520	0.3410	-0.0606	-0.0316	-0.2470	0.2390	0.0381	-0.1490	-0.1880	0.1810	0.0387	-0.0055
SPVSD	0.5140	-0.5520	0.4970	-0.4500	-0.0232	0.1040	0.0414	-0.1050	-0.0130	0.1420	0.3000	-0.3180	-0.1230	0.0470
SWI	-0.7600	0.3410	-0.4500	0.8110	-0.0942	0.0048	-0.0013	0.1080	0.0721	-0.0804	-0.0972	0.1380	-0.0622	0.1190
META	0.0313	-0.0606	-0.0232	-0.0942	0.7870	-0.8890	-0.4480	0.6220	-0.2200	0.3470	-0.4220	0.3750	0.0141	-0.0133
METS	-0.0165	-0.0316	0.1040	0.0048	-0.8890	0.7730	0.4630	-0.6120	0.2030	-0.4070	0.5880	-0.4470	0.0783	-0.0780
MAXTA	-0.0077	-0.2470	0.0414	-0.0013	-0.4480	0.4630	0.8410	-0.8630	0.0578	-0.2560	0.1340	-0.0191	0.0239	0.0289
MAXTS	-0.0799	0.2390	-0.1050	0.1080	0.6220	-0.6120	-0.8630	0.4680	-0.0511	0.1500	-0.2690	0.1600	0.0410	-0.0484
MINTA	-0.1980	0.0381	-0.0130	0.0721	-0.2200	0.2030	0.0578	-0.0511	0.8330	-0.7380	-0.2950	0.3910	0.1540	-0.2110
MINTS	0.2050	-0.1490	0.1420	-0.0804	0.3470	-0.4070	-0.2560	0.1500	-0.7380	0.7900	0.3140	-0.4520	-0.3650	0.3650
PRCPA	0.2740	-0.1880	0.3000	-0.0972	-0.4220	0.5880	0.1340	-0.2690	-0.2950	0.3140	0.7980	-0.9490	-0.0296	0.0349
PRCPS	-0.3410	0.1810	-0.3180	0.1380	0.3750	-0.4470	-0.0191	0.1600	0.3910	-0.4520	-0.9490	0.7670	0.0843	-0.0821
PETAN	-0.0071	0.0387	-0.1230	-0.0622	0.0141	0.0783	0.0239	0.0410	0.1540	-0.3650	-0.0296	0.0843	0.8450	-0.9810
PETSE	-0.0570	-0.0055	0.0470	0.1190	-0.0133	-0.0780	0.0289	-0.0484	-0.2110	0.3650	0.0349	-0.0821	-0.9810	0.8500

ABBREVIATIONS: **ELEV:** elevation, **ASP:** aspect, **STR:** saturation, **TEXTR:** texture, **SOILG:** big soil group, **ERS:** erosion, **SLDPT:** soil depth, **GEO:** geological formations, **SPVSD:** supervised classes, **SWI:** Shannon Wiener index, **META**: annual mean temperature, **METS**: seasonal mean temperature, **MAXTA:** annual maximum temperature, **MAXTS:** seasonal maximum temperature, **MINTA:** annual minimum temperature, **MINTS:** seasonal minimum temperature, **PRCPA:** annual precipitation, **PRCPS:** seasonal precipitation, **PETAN:** annual potential evapotranspiration, **PETSE:** seasonal potential evapotranspiration

Table 5.06. The anti-image covariance matrice of initial 28 variables for Simpson Index

	ELEV	ASP	SLOPE	ORGM	P2O5	K2O	CACO3	PH	SALT	STR	TEXTR	SOILG	ERS	SLDPT
ELEV	0.0005	-0.0010	0.0021	-0.0024	0.0008	0.0018	0.0005	-0.0003	-0.0016	0.0015	-0.0026	0.0001	0.0017	0.0000
ASP	-0.0010	0.5350	0.0717	-0.0320	-0.0252	0.0147	-0.0546	0.0524	-0.0450	0.0180	0.0170	-0.0227	0.0083	-0.0119
SLOPE	0.0021	0.0717	0.3340	-0.0534	0.0031	-0.0154	-0.0417	0.0058	-0.0121	0.0139	0.0104	-0.0040	-0.0190	0.0253
ORGM	-0.0024	-0.0320	-0.0534	0.0828	-0.0153	-0.0352	-0.0317	0.0144	0.0478	-0.0389	0.0295	-0.0083	0.0042	-0.0157
P2O5	0.0008	-0.0252	0.0031	-0.0153	0.1900	-0.0719	-0.0036	0.0804	-0.0412	0.0116	0.0113	0.0412	0.0164	0.0011
K2O	0.0018	0.0147	-0.0154	-0.0352	-0.0719	0.1100	0.0626	-0.0646	-0.0108	0.0253	-0.0485	-0.0168	-0.0328	0.0415
CACO3	0.0005	-0.0546	-0.0417	-0.0317	-0.0036	0.0626	0.1530	-0.0676	-0.0224	0.0251	-0.0279	0.0416	-0.0097	0.0352
PH	-0.0003	0.0524	0.0058	0.0144	0.0804	-0.0646	-0.0676	0.1000	-0.0306	0.0002	0.0154	0.0035	0.0218	-0.0219
SALT	-0.0016	-0.0450	-0.0121	0.0478	-0.0412	-0.0108	-0.0224	-0.0306	0.0794	-0.0362	0.0199	-0.0283	-0.0194	-0.0034
STR	0.0015	0.0180	0.0139	-0.0389	0.0116	0.0253	0.0251	0.0002	-0.0362	0.0273	-0.0331	0.0024	0.0048	0.0062
TEXTR	-0.0026	0.0170	0.0104	0.0295	0.0113	-0.0485	-0.0279	0.0154	0.0199	-0.0331	0.1170	0.0629	-0.0306	0.0203
SOILG	0.0001	-0.0227	-0.0040	-0.0083	0.0412	-0.0168	0.0416	0.0035	-0.0283	0.0024	0.0629	0.2950	-0.0230	0.0732
ERS	0.0017	0.0083	-0.0190	0.0042	0.0164	-0.0328	-0.0097	0.0218	-0.0194	0.0048	-0.0306	-0.0230	0.1240	-0.0902
SLDPT	0.0000	-0.0119	0.0253	-0.0157	0.0011	0.0415	0.0352	-0.0219	-0.0034	0.0062	0.0203	0.0732	-0.0902	0.1020
GEO	0.0006	-0.0037	-0.0180	-0.0021	-0.0430	0.0301	0.0368	-0.0209	-0.0274	0.0091	-0.0262	-0.0158	0.0024	0.0004
NDVI	-0.0056	0.1180	-0.0194	0.0339	-0.0061	-0.0582	-0.0703	0.0445	0.0024	-0.0252	0.0684	0.0748	-0.0172	-0.0171
SPVSD	0.0030	-0.0418	-0.0269	-0.0161	-0.0551	0.0844	0.1070	-0.0858	-0.0063	0.0126	-0.0300	-0.0230	0.0119	0.0235
SWI	0.0010	-0.0802	0.0355	0.0154	-0.0358	-0.0018	0.0219	-0.0488	0.0382	-0.0128	-0.0175	-0.0215	0.0138	-0.0227
META	0.0001	-0.0012	0.0011	-0.0006	0.0013	-0.0016	-0.0006	0.0012	-0.0008	0.0004	-0.0005	-0.0005	0.0024	-0.0020
METS	0.0000	0.0003	0.0004	0.0002	-0.0023	0.0032	0.0005	-0.0025	0.0010	-0.0002	-0.0002	0.0012	-0.0032	0.0031
MAXTA	0.0002	0.0027	0.0042	-0.0026	0.0013	0.0015	0.0019	0.0003	-0.0019	0.0016	-0.0008	0.0014	-0.0024	0.0026
MAXTS	-0.0011	-0.0217	-0.0056	0.0088	-0.0061	-0.0170	-0.0176	0.0019	0.0045	-0.0065	0.0083	-0.0138	0.0225	-0.0239
MINTA	0.0000	0.0034	0.0059	-0.0026	-0.0054	0.0032	0.0007	-0.0047	0.0028	-0.0004	0.0028	0.0007	-0.0053	0.0037
MINTS	0.0002	-0.0030	-0.0077	0.0009	0.0041	-0.0005	0.0011	0.0024	-0.0019	0.0007	-0.0031	-0.0013	0.0036	-0.0021
PRCPA	0.0003	-0.0004	-0.0046	0.0007	-0.0014	0.0044	0.0003	-0.0006	-0.0017	0.0009	-0.0037	-0.0015	0.0012	0.0006
PRCPS	-0.0004	0.0017	0.0139	-0.0037	0.0035	-0.0089	-0.0008	0.0009	0.0026	-0.0013	0.0087	0.0069	-0.0038	0.0003
PETAN	0.0001	0.0040	0.0094	0.0006	-0.0046	-0.0011	-0.0043	-0.0005	0.0019	-0.0011	0.0021	0.0006	0.0017	-0.0011
PETSE	-0.0001	-0.0023	-0.0059	-0.0004	0.0034	0.0002	0.0026	0.0009	-0.0016	0.0008	-0.0015	-0.0007	-0.0007	0.0003

ABBREVIATIONS: **ELEV:** elevation, **ASP:** aspect, **STR:** saturation, **TEXTR:** texture, **SOILG:** big soil group, **ERS:** erosion, **SLDPT:** soil depth, **GEO:** geological formations, **SPVSD:** supervised classes, **SWI:** Shannon Wiener index, **META**: annual mean temperature, **METS**: seasonal mean temperature, **MAXTA:** annual maximum temperature, **MAXTS:** seasonal maximum temperature, **MINTA:** annual minimum temperature, **MINTS:** seasonal minimum temperature, **PRCPA:** annual precipitation, **PRCPS:** seasonal precipitation, **PETAN:** annual potential evapotranspiration, **PETSE:** seasonal potential evapotranspiration

Table 5.06. Continue (the anti-image covariance matrice of initial 28 variables for Simpson Index)

	GEO	NDVI	SPVSD	SWI	META	METS	MAXTA	MAXTS	MINTA	MINTS	PRCPA	PRCPE	PETAN	PETSE
ELEV	0.0006	-0.0056	0.0030	0.0010	0.0001	0.0000	0.0002	-0.0011	0.0000	0.0002	0.0003	-0.0004	0.0001	-0.0001
ASP	-0.0037	0.1180	-0.0418	-0.0802	-0.0012	0.0003	0.0027	-0.0217	0.0034	-0.0030	-0.0004	0.0017	0.0040	-0.0023
SLOPE	-0.0180	-0.0194	-0.0269	0.0355	0.0011	0.0004	0.0042	-0.0056	0.0059	-0.0077	-0.0046	0.0139	0.0094	-0.0059
ORGM	-0.0021	0.0339	-0.0161	0.0154	-0.0006	0.0002	-0.0026	0.0088	-0.0026	0.0009	0.0007	-0.0037	0.0006	-0.0004
P2O5	-0.0430	-0.0061	-0.0551	-0.0358	0.0013	-0.0023	0.0013	-0.0061	-0.0054	0.0041	-0.0014	0.0035	-0.0046	0.0034
K2O	0.0301	-0.0582	0.0844	-0.0018	-0.0016	0.0032	0.0015	-0.0170	0.0032	-0.0005	0.0044	-0.0089	-0.0011	0.0002
CACO3	0.0368	-0.0703	0.1070	0.0219	-0.0006	0.0005	0.0019	-0.0176	0.0007	0.0011	0.0003	-0.0008	-0.0043	0.0026
PH	-0.0209	0.0445	-0.0858	-0.0488	0.0012	-0.0025	0.0003	0.0019	-0.0047	0.0024	-0.0006	0.0009	-0.0005	0.0009
SALT	-0.0274	0.0024	-0.0063	0.0382	-0.0008	0.0010	-0.0019	0.0045	0.0028	-0.0019	-0.0017	0.0026	0.0019	-0.0016
STR	0.0091	-0.0252	0.0126	-0.0128	0.0004	-0.0002	0.0016	-0.0065	-0.0004	0.0007	0.0009	-0.0013	-0.0011	0.0008
TEXTR	-0.0262	0.0684	-0.0300	-0.0175	-0.0005	-0.0002	-0.0008	0.0083	0.0028	-0.0031	-0.0037	0.0087	0.0021	-0.0015
SOILG	-0.0158	0.0748	-0.0230	-0.0215	-0.0005	0.0012	0.0014	-0.0138	0.0007	-0.0013	-0.0015	0.0069	0.0006	-0.0007
ERS	0.0024	-0.0172	0.0119	0.0138	0.0024	-0.0032	-0.0024	0.0225	-0.0053	0.0036	0.0012	-0.0038	0.0017	-0.0007
SLDPT	0.0004	-0.0171	0.0235	-0.0227	-0.0020	0.0031	0.0026	-0.0239	0.0037	-0.0021	0.0006	0.0003	-0.0011	0.0003
GEO	0.1980	-0.0381	0.0994	0.0890	-0.0002	-0.0001	-0.0006	0.0053	-0.0036	0.0023	0.0046	-0.0122	0.0001	-0.0004
NDVI	-0.0381	0.3970	-0.1820	-0.0634	-0.0004	-0.0007	-0.0045	0.0275	0.0006	-0.0021	-0.0037	0.0075	0.0005	0.0000
SPVSD	0.0994	-0.1820	0.3220	0.0828	-0.0005	0.0019	0.0003	-0.0033	0.0001	0.0016	0.0056	-0.0127	-0.0020	0.0005
SWI	0.0890	-0.0634	0.0828	0.2050	0.0005	0.0003	-0.0010	0.0115	-0.0009	-0.0003	0.0007	-0.0035	0.0029	-0.0021
META	-0.0002	-0.0004	-0.0005	0.0005	0.0003	-0.0004	-0.0002	0.0024	-0.0001	0.0001	-0.0002	0.0005	0.0000	0.0000
METS	-0.0001	-0.0007	0.0019	0.0003	-0.0004	0.0008	0.0004	-0.0040	0.0002	-0.0003	0.0006	-0.0009	0.0001	-0.0001
MAXTA	-0.0006	-0.0045	0.0003	-0.0010	-0.0002	0.0004	0.0009	-0.0061	0.0001	-0.0002	0.0001	0.0000	0.0000	0.0000
MAXTS	0.0053	0.0275	-0.0033	0.0115	0.0024	-0.0040	-0.0061	0.0553	-0.0006	0.0010	-0.0020	0.0023	0.0005	-0.0005
MINTA	-0.0036	0.0006	0.0001	-0.0009	-0.0001	0.0002	0.0001	-0.0006	0.0013	-0.0007	-0.0004	0.0010	0.0002	-0.0002
MINTS	0.0023	-0.0021	0.0016	-0.0003	0.0001	-0.0003	-0.0002	0.0010	-0.0007	0.0007	0.0003	-0.0009	-0.0004	0.0002
PRCPA	0.0046	-0.0037	0.0056	0.0007	-0.0002	0.0006	0.0001	-0.0020	-0.0004	0.0003	0.0012	-0.0024	0.0000	0.0000
PRCPS	-0.0122	0.0075	-0.0127	-0.0035	0.0005	-0.0009	0.0000	0.0023	0.0010	-0.0009	-0.0024	0.0054	0.0002	-0.0001
PETAN	0.0001	0.0005	-0.0020	0.0029	0.0000	0.0001	0.0000	0.0005	0.0002	-0.0004	0.0000	0.0002	0.0012	-0.0008
PETSE	-0.0004	0.0000	0.0005	-0.0021	0.0000	-0.0001	0.0000	-0.0005	-0.0002	0.0002	0.0000	-0.0001	-0.0008	0.0006

ABBREVIATIONS: **ELEV:** elevation, **ASP:** aspect, **STR:** saturation, **TEXTR:** texture, **SOILG:** big soil group, **ERS:** erosion, **SLDPT:** soil depth, **GEO:** geological formations, **SPVSD:** supervised classes, **SWI:** Shannon Wiener index, **META**: annual mean temperature, **METS**: seasonal mean temperature, **MAXTA:** annual maximum temperature, **MAXTS:** seasonal maximum temperature, **MINTA:** annual minimum temperature, **MINTS:** seasonal minimum temperature, **PRCPA:** annual precipitation, **PRCPS:** seasonal precipitation, **PETAN:** annual potential evapotranspiration, **PETSE:** seasonal potential evapotranspiration

Table 5.07. The anti-image correlation matrice of initial 28 variables for Simpson Index

	ELEV	ASP	SLOPE	ORGM	P2O5	K2O	CACO3	PH	SALT	STR	TEXTR	SOILG	ERS	SLDPT
ELEV	0.8640	-0.0642	0.1680	-0.3900	0.0812	0.2600	0.0647	-0.0378	-0.2590	0.4220	-0.3550	0.0047	0.2250	-0.0046
ASP	-0.0642	0.4670	0.1690	-0.1520	-0.0792	0.0605	-0.1910	0.2260	-0.2180	0.1490	0.0679	-0.0572	0.0324	-0.0510
SLOPE	0.1680	0.1690	0.4670	-0.3210	0.0122	-0.0803	-0.1840	0.0315	-0.0741	0.1460	0.0527	-0.0128	-0.0933	0.1370
ORGM	-0.3900	-0.1520	-0.3210	0.5170	-0.1220	-0.3680	-0.2820	0.1580	0.5890	-0.8200	0.2990	-0.0531	0.0411	-0.1700
P2O5	0.0812	-0.0792	0.0122	-0.1220	0.3810	-0.4970	-0.0209	0.5840	-0.3360	0.1610	0.0757	0.1740	0.1070	0.0079
K2O	0.2600	0.0605	-0.0803	-0.3680	-0.4970	0.3320	0.4820	-0.6140	-0.1150	0.4610	-0.4270	-0.0933	-0.2800	0.3910
CACO3	0.0647	-0.1910	-0.1840	-0.2820	-0.0209	0.4820	0.6120	-0.5460	-0.2030	0.3880	-0.2080	0.1960	-0.0703	0.2810
PH	-0.0378	0.2260	0.0315	0.1580	0.5840	-0.6140	-0.5460	0.5620	-0.3430	0.0043	0.1420	0.0201	0.1950	-0.2160
SALT	-0.2590	-0.2180	-0.0741	0.5890	-0.3360	-0.1150	-0.2030	-0.3430	0.4180	-0.7770	0.2060	-0.1850	-0.1950	-0.0376
STR	0.4220	0.1490	0.1460	-0.8200	0.1610	0.4610	0.3880	0.0043	-0.7770	0.5570	-0.5860	0.0273	0.0831	0.1170
TEXTR	-0.3550	0.0679	0.0527	0.2990	0.0757	-0.4270	-0.2080	0.1420	0.2060	-0.5860	0.5940	0.3380	-0.2530	0.1850
SOILG	0.0047	-0.0572	-0.0128	-0.0531	0.1740	-0.0933	0.1960	0.0201	-0.1850	0.0273	0.3380	0.7730	-0.1200	0.4210
ERS	0.2250	0.0324	-0.0933	0.0411	0.1070	-0.2800	-0.0703	0.1950	-0.1950	0.0831	-0.2530	-0.1200	0.4560	-0.8000
SLDPT	-0.0046	-0.0510	0.1370	-0.1700	0.0079	0.3910	0.2810	-0.2160	-0.0376	0.1170	0.1850	0.4210	-0.8000	0.5710
GEO	0.0677	-0.0114	-0.0697	-0.0167	-0.2210	0.2040	0.2110	-0.1480	-0.2180	0.1240	-0.1720	-0.0653	0.0152	0.0027
NDVI	-0.4180	0.2560	-0.0532	0.1870	-0.0222	-0.2780	-0.2850	0.2230	0.0134	-0.2420	0.3170	0.2190	-0.0775	-0.0849
SPVSD	0.2450	-0.1010	-0.0819	-0.0988	-0.2230	0.4480	0.4820	-0.4780	-0.0396	0.1340	-0.1550	-0.0746	0.0596	0.1300
SWI	0.1020	-0.2420	0.1360	0.1180	-0.1810	-0.0120	0.1240	-0.3400	0.2990	-0.1710	-0.1130	-0.0875	0.0863	-0.1570
META	0.3060	-0.1010	0.1120	-0.1270	0.1810	-0.2930	-0.0989	0.2250	-0.1730	0.1330	-0.0861	-0.0580	0.4250	-0.3860
METS	-0.0303	0.0133	0.0236	0.0288	-0.1900	0.3430	0.0453	-0.2850	0.1250	-0.0512	-0.0187	0.0790	-0.3240	0.3480
MAXTA	0.3070	0.1250	0.2450	-0.3060	0.1010	0.1540	0.1670	0.0331	-0.2300	0.3280	-0.0801	0.0853	-0.2260	0.2730
MAXTS	-0.2140	-0.1260	-0.0411	0.1300	-0.0594	-0.2170	-0.1910	0.0256	0.0674	-0.1680	0.1030	-0.1080	0.2710	-0.3180
MINTA	-0.0565	0.1290	0.2790	-0.2510	-0.3400	0.2620	0.0459	-0.4050	0.2740	-0.0660	0.2240	0.0372	-0.4170	0.3160
MINTS	0.2690	-0.1540	-0.4950	0.1150	0.3520	-0.0598	0.1040	0.2800	-0.2480	0.1600	-0.3370	-0.0880	0.3780	-0.2440
PRCPA	0.3840	-0.0156	-0.2340	0.0676	-0.0933	0.3870	0.0205	-0.0587	-0.1720	0.1640	-0.3180	-0.0800	0.0994	0.0559
PRCPS	-0.2570	0.0316	0.3290	-0.1740	0.1090	-0.3650	-0.0277	0.0378	0.1280	-0.1080	0.3470	0.1740	-0.1450	0.0114
PETAN	0.1950	0.1540	0.4590	0.0599	-0.3020	-0.0936	-0.3140	-0.0467	0.1890	-0.1890	0.1710	0.0309	0.1390	-0.1000
PETSE	-0.2040	-0.1320	-0.4320	-0.0526	0.3330	0.0285	0.2830	0.1270	-0.2460	0.2160	-0.1880	-0.0519	-0.0888	0.0440

ABBREVIATIONS: **ELEV:** elevation, **ASP:** aspect, **STR:** saturation, **TEXTR:** texture, **SOILG:** big soil group, **ERS:** erosion, **SLDPT:** soil depth, **GEO:** geological formations, **SPVSD:** supervised classes, **SWI:** Shannon Wiener index, **META**: annual mean temperature, **METS**: seasonal mean temperature, **MAXTA:** annual maximum temperature, **MAXTS:** seasonal maximum temperature, **MINTA:** annual minimum temperature, **MINTS:** seasonal minimum temperature, **PRCPA:** annual precipitation, **PRCPS:** seasonal precipitation, **PETAN:** annual potential evapotranspiration, **PETSE:** seasonal potential evapotranspiration

Table 5.07. Continue (the anti-image correlation matrice of initial 28 variables for Simpson Index)

	GEO	NDVI	SPVSD	SWI	META	METS	MAXTA	MAXTS	MINTA	MINTS	PRCPA	PRCPE	PETAN	PETSE
ELEV	0.0677	-0.4180	0.2450	0.1020	0.3060	-0.0303	0.3070	-0.2140	-0.0565	0.2690	0.3840	-0.2570	0.1950	-0.2040
ASP	-0.0114	0.2560	-0.1010	-0.2420	-0.1010	0.0133	0.1250	-0.1260	0.1290	-0.1540	-0.0156	0.0316	0.1540	-0.1320
SLOPE	-0.0697	-0.0532	-0.0819	0.1360	0.1120	0.0236	0.2450	-0.0411	0.2790	-0.4950	-0.2340	0.3290	0.4590	-0.4320
ORGM	-0.0167	0.1870	-0.0988	0.1180	-0.1270	0.0288	-0.3060	0.1300	-0.2510	0.1150	0.0676	-0.1740	0.0599	-0.0526
P2O5	-0.2210	-0.0222	-0.2230	-0.1810	0.1810	-0.1900	0.1010	-0.0594	-0.3400	0.3520	-0.0933	0.1090	-0.3020	0.3330
K2O	0.2040	-0.2780	0.4480	-0.0120	-0.2930	0.3430	0.1540	-0.2170	0.2620	-0.0598	0.3870	-0.3650	-0.0936	0.0285
CACO3	0.2110	-0.2850	0.4820	0.1240	-0.0989	0.0453	0.1670	-0.1910	0.0459	0.1040	0.0205	-0.0277	-0.3140	0.2830
PH	-0.1480	0.2230	-0.4780	-0.3400	0.2250	-0.2850	0.0331	0.0256	-0.4050	0.2800	-0.0587	0.0378	-0.0467	0.1270
SALT	-0.2180	0.0134	-0.0396	0.2990	-0.1730	0.1250	-0.2300	0.0674	0.2740	-0.2480	-0.1720	0.1280	0.1890	-0.2460
STR	0.1240	-0.2420	0.1340	-0.1710	0.1330	-0.0512	0.3280	-0.1680	-0.0660	0.1600	0.1640	-0.1080	-0.1890	0.2160
TEXTR	-0.1720	0.3170	-0.1550	-0.1130	-0.0861	-0.0187	-0.0801	0.1030	0.2240	-0.3370	-0.3180	0.3470	0.1710	-0.1880
SOILG	-0.0653	0.2190	-0.0746	-0.0875	-0.0580	0.0790	0.0853	-0.1080	0.0372	-0.0880	-0.0800	0.1740	0.0309	-0.0519
ERS	0.0152	-0.0775	0.0596	0.0863	0.4250	-0.3240	-0.2260	0.2710	-0.4170	0.3780	0.0994	-0.1450	0.1390	-0.0888
SLDPT	0.0027	-0.0849	0.1300	-0.1570	-0.3860	0.3480	0.2730	-0.3180	0.3160	-0.2440	0.0559	0.0114	-0.1000	0.0440
GEO	0.6710	-0.1360	0.3930	0.4410	-0.0282	-0.0065	-0.0440	0.0510	-0.2230	0.1900	0.2990	-0.3730	0.0052	-0.0427
NDVI	-0.1360	0.5330	-0.5090	-0.2220	-0.0435	-0.0401	-0.2390	0.1860	0.0265	-0.1220	-0.1720	0.1630	0.0213	-0.0028
SPVSD	0.3930	-0.5090	0.5310	0.3220	-0.0496	0.1210	0.0198	-0.0248	0.0026	0.1060	0.2880	-0.3060	-0.0998	0.0407
SWI	0.4410	-0.2220	0.3220	0.8370	0.0625	0.0255	-0.0726	0.1080	-0.0562	-0.0211	0.0478	-0.1040	0.1800	-0.2010
META	-0.0282	-0.0435	-0.0496	0.0625	0.7840	-0.8890	-0.4530	0.6410	-0.2170	0.3400	-0.4310	0.3840	0.0194	-0.0146
METS	-0.0065	-0.0401	0.1210	0.0255	-0.8890	0.7710	0.4600	-0.6100	0.2010	-0.4090	0.5920	-0.4520	0.0820	-0.0826
MAXTA	-0.0440	-0.2390	0.0198	-0.0726	-0.4530	0.4600	0.8350	-0.8680	0.0618	-0.2550	0.1310	-0.0114	0.0103	0.0431
MAXTS	0.0510	0.1860	-0.0248	0.1080	0.6410	-0.6100	-0.8680	0.4540	-0.0650	0.1570	-0.2540	0.1340	0.0664	-0.0821
MINTA	-0.2230	0.0265	0.0026	-0.0562	-0.2170	0.2010	0.0618	-0.0650	0.8340	-0.7340	-0.2920	0.3880	0.1470	-0.2050
MINTS	0.1900	-0.1220	0.1060	-0.0211	0.3400	-0.4090	-0.2550	0.1570	-0.7340	0.7910	0.3070	-0.4420	-0.3700	0.3750
PRCPA	0.2990	-0.1720	0.2880	0.0478	-0.4310	0.5920	0.1310	-0.2540	-0.2920	0.3070	0.7970	-0.9470	-0.0267	0.0364
PRCPS	-0.3730	0.1630	-0.3060	-0.1040	0.3840	-0.4520	-0.0114	0.1340	0.3880	-0.4420	-0.9470	0.7660	0.0732	-0.0768
PETAN	0.0052	0.0213	-0.0998	0.1800	0.0194	0.0820	0.0103	0.0664	0.1470	-0.3700	-0.0267	0.0732	0.8390	-0.9830
PETSE	-0.0427	-0.0028	0.0407	-0.2010	-0.0146	-0.0826	0.0431	-0.0821	-0.2050	0.3750	0.0364	-0.0768	-0.9830	0.8430

ABBREVIATIONS: **ELEV:** elevation, **ASP:** aspect, **STR:** saturation, **TEXTR:** texture, **SOILG:** big soil group, **ERS:** erosion, **SLDPT:** soil depth, **GEO:** geological formations, **SPVSD:** supervised classes, **SWI:** Shannon Wiener index, **META**: annual mean temperature, **METS**: seasonal mean temperature, **MAXTA:** annual maximum temperature, M**AXTS:** seasonal maximum temperature, **MINTA:** annual minimum temperature, **MINTS:** seasonal minimum temperature, **PRCPA:** annual precipitation, **PRCPS:** seasonal precipitation, **PETAN:** annual potential evapotranspiration, **PETSE:** seasonal potential evapotranspiration

Table 5.08 and Table 5.09 give the amount of cumulative variance explained with the initial solution and initial rotation of the factor analyses for both indices. The “total” column gives the amount of variance in the observed variables accounted for by each component or factor. The “% of variance” column gives the percent of variance accounted for by each specific factor or component, relative to the total variance in all the variables. The “cumulative %” column gives the percent of variance accounted for by all factors or components up to and including the current one. For instance, the cumulative % for the second factor is the sum of the % of variance for the first and second factors. The “extraction sums of squared loadings” group gives information regarding the extracted factors or components. For principal components extraction, these values will be the same as those reported under initial eigenvalues (SPSS, 2001).

In a good *factor analysis*, there are a few factors that explain a lot of the variance, and the rest of the factors explain relatively small amounts of variance. Depending on this rule, eigenvalues of smaller than 1 are not included in the *factor analysis* (SPSS, 2001). Consequently, the first 7 factors are taken into consideration for both indices (Table 5.08 and Table 5.09). Explained total variances of these 7 factors are defined as 83.753% and 83.183% for Shannon Wiener and Simpson indices, respectively.

The scree plots, graphical representation of the eigenvalues of each component, can also be useful to determine the optimal number of components. The eigenvalues of each component in the initial solution is plotted for both indices (Figure 5.06). In scree plots, the components that have steep slope is determined to be extracted, while components that have shallow slope contribute little to the solution.

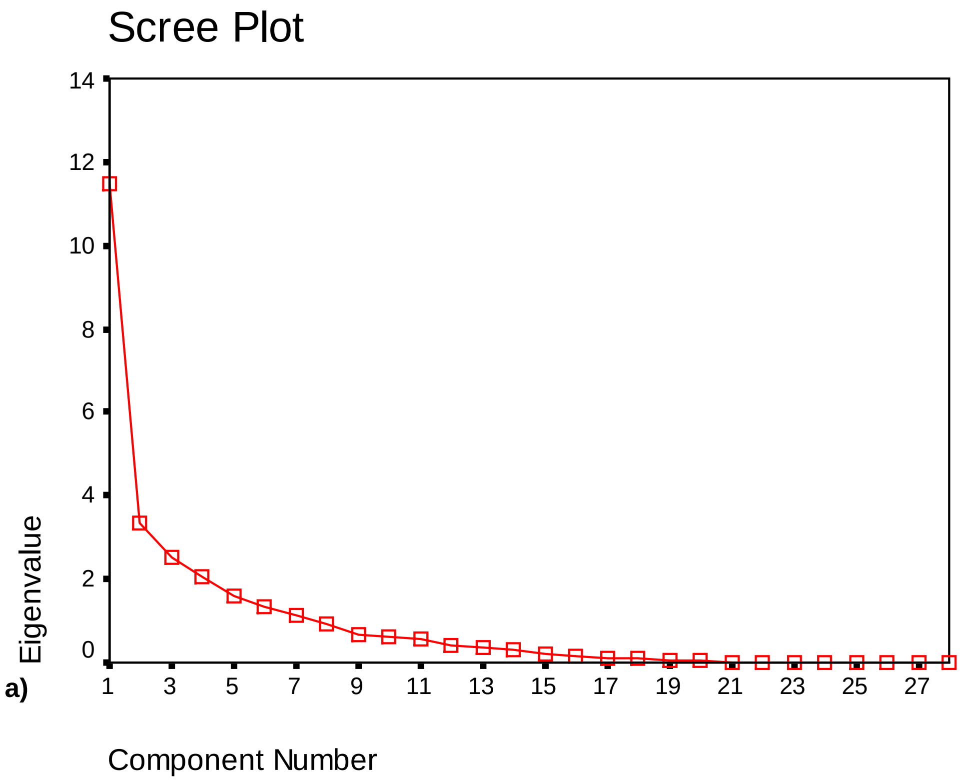
The components’ contents can be explained by analyzing the rotated component matrix after varimax rotation. The factor loadings for each variable on the components after rotation were summarized in Table 5.10 and Table 5.11. In rotated matrices, each number represents the partial correlation between the item and the rotated component. The bold numbers represent the maximum correlations within these components. For instance, elevation, and climatic variables (temperature, precipitation, and evapotranspiration) constituted the first component for each index. The second component is derived from salt, saturation, texture, and geological formations for Shannon Wiener index, while the same component is explained as salt, saturation, and texture for Simpson index. The third component is composed of erosion and soil depth for Shannon Wiener index, and so on.

Table 5.08. Total variance explained via components for Shannon Wiener index (the first 7 factors are taken into consideration, because their initial eigenvalues are bigger than 1) NOTE: The explanation of the components were given in Table 5.10.

	Initial Eigenvalues			**Extraction Sums of Squared Loadings**			**Rotation Sums of Squared Loadings**		
Component	**Total**	**% of Variance**	**Cumulative %**	**Total**	**% of Variance**	**Cumulative %**	**Total**	**% of Variance**	**Cumulative %**
1	11.4770	40.9890	40.989	11.477	40.989	40.989	10.151	36.253	36.253
2	3.3330	11.9050	52.893	3.333	11.905	52.893	3.270	11.680	47.933
3	2.5300	9.0340	61.928	2.530	9.034	61.928	2.720	9.715	57.648
4	2.0510	7.3250	69.252	2.051	7.325	69.252	2.545	9.091	66.738
5	1.6210	5.7880	75.041	1.621	5.788	75.041	1.780	6.358	73.096
6	1.3210	4.7180	79.759	1.321	4.718	79.759	1.713	6.118	79.214
7	1.1180	3.9940	**83.753**	1.118	3.994	**83.753**	1.271	4.539	**83.753**
8	0.9400	3.3570	87.111						
9	0.6540	2.3370	89.448						
10	0.6320	2.2570	91.705						
11	0.5770	2.0620	93.767						
12	0.4140	1.4790	95.246						
13	0.3360	1.1990	96.445						
14	0.2900	1.0350	97.480						
15	0.2120	0.7570	98.237						
16	0.1450	0.5190	98.756						
17	0.1240	0.4430	99.200						
18	0.0858	0.3070	99.506						
19	0.0647	0.2310	99.737						
20	0.0344	0.1230	99.860						
21	0.0225	0.0804	99.940						
22	0.0102	0.0364	99.977						
23	0.0035	0.0124	99.989						
24	0.0014	0.0051	99.994						
25	0.0007	0.0026	99.997						
26	0.0003	0.0012	99.998						
27	0.0003	0.0011	99.999						
28	0.0002	0.0006	100.000						

Table 5.09. Total variance explained via components for Simpson index (the first 7 factors are taken into consideration, because their initial eigenvalues are bigger than 1) NOTE: The explanation of the components were given in Table 5.11.

	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
Component	**Total**	**% of Variance**	**Cumulative %**	**Total**	**% of Variance**	**Cumulative %**	**Total**	**% of Variance**	**Cumulative %**
1	11.3110	40.3980	40.398	11.311	40.398	40.398	10.015	35.768	35.768
2	3.2860	11.7370	52.135	3.286	11.737	52.135	3.123	11.153	46.921
3	2.5750	9.1960	61.331	2.575	9.196	61.331	2.791	9.966	56.887
4	2.0510	7.3250	68.656	2.051	7.325	68.656	2.584	9.230	66.117
5	1.6220	5.7930	74.449	1.622	5.793	74.449	1.785	6.377	72.494
6	1.3190	4.7120	79.161	1.319	4.712	79.161	1.706	6.093	78.586
7	1.1260	4.0220	**83.183**	1.126	4.022	**83.183**	1.287	4.597	**83.183**
8	0.9510	3.3970	86.580						
9	0.7100	2.5360	89.116						
10	0.6290	2.2460	91.362						
11	0.5870	2.0950	93.457						
12	0.4030	1.4400	94.897						
13	0.3400	1.2140	96.111						
14	0.2860	1.0230	97.134						
15	0.2360	0.8420	97.976						
16	0.1660	0.5940	98.570						
17	0.1340	0.4790	99.050						
18	0.1170	0.4180	99.467						
19	0.0640	0.2280	99.696						
20	0.0425	0.1520	99.848						
21	0.0259	0.0923	99.940						
22	0.0104	0.0371	99.977						
23	0.0034	0.0123	99.989						
24	0.0014	0.0051	99.994						
25	0.0007	0.0025	99.997						
26	0.0004	0.0013	99.998						
27	0.0003	0.0011	99.999						
28	0.0002	0.0006	100.000						



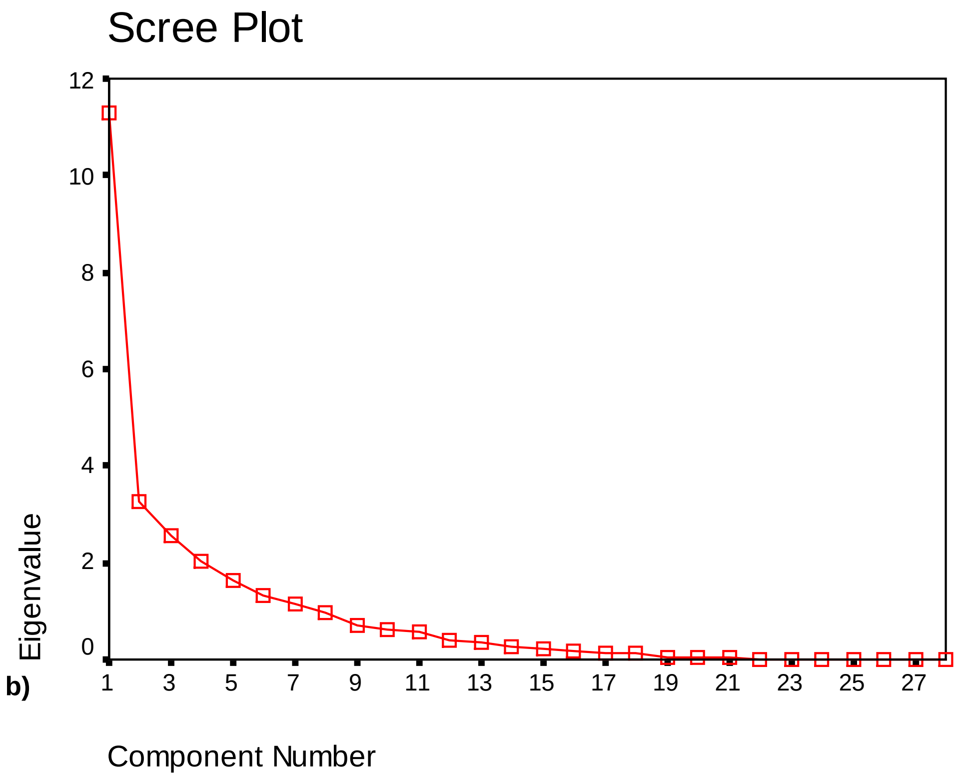


Figure 5.06. Scree plots for (a) Shannon Wiener index, and (b) Simpson index

Table 5.10. Rotated Component Matrix for Shannon Wiener index

VARIABLE	COMPONENT						
	1	2	3	4	5	6	7
ELEV	**-0.978**	0.027	-0.065	-0.111	0.091	-0.074	0.078
ASP	0.051	0.024	-0.066	0.002	**-0.729**	-0.224	-0.147
SLOPE	-0.208	0.123	-0.365	0.053	0.081	0.009	**0.657**
ORGM	-0.169	0.339	0.100	**-0.625**	-0.010	-0.230	0.488
P2O5	-0.016	0.025	-0.075	**-0.552**	0.081	**0.628**	-0.278
K2O	0.248	-0.005	0.007	0.049	0.080	**0.886**	0.081
CACO3	0.237	-0.174	-0.096	**0.767**	-0.211	-0.207	0.090
PH	0.298	0.031	0.025	**0.864**	0.081	0.064	0.105
SALT	0.254	**0.834**	0.038	0.180	0.088	0.037	-0.145
STR	-0.196	**0.733**	0.071	-0.412	0.131	-0.355	0.183
TEXTR	-0.159	**0.802**	0.061	-0.237	0.040	-0.116	0.144
SOILG	-0.248	0.039	-0.762	0.045	-0.071	0.235	-0.065
ERS	0.048	0.174	**0.875**	-0.024	0.053	0.213	0.022
SLDPT	0.224	-0.045	**0.887**	0.004	-0.004	-0.035	-0.205
GEO	-0.210	**0.808**	0.000	-0.033	-0.143	0.269	-0.037
NDVI	-0.260	0.039	0.298	-0.038	**0.688**	-0.025	-0.059
SPVSD	-0.243	0.039	-0.268	-0.094	**0.690**	-0.132	-0.208
SWI	**-0.638**	**0.567**	-0.276	-0.092	0.007	-0.001	0.014
META	0.971	0.030	0.045	0.122	-0.114	0.059	-0.094
METS	**0.946**	0.125	0.035	0.153	-0.156	0.030	-0.042
MAXTA	**0.972**	-0.053	0.085	0.131	-0.110	0.072	-0.031
MAXTS	0.251	-0.204	0.197	0.294	-0.260	0.007	**0.510**
MINTA	**0.975**	0.012	0.059	0.109	-0.102	0.065	-0.054
MINTS	**0.974**	0.030	0.062	0.123	-0.115	0.052	-0.016
PRCPA	**0.941**	-0.196	0.144	0.081	-0.032	0.095	-0.033
PRCPS	**0.837**	-0.238	0.266	0.046	-0.003	0.121	0.126
PETAN	**-0.919**	0.205	-0.119	-0.144	0.112	0.068	-0.084
PETSE	**-0.957**	0.150	-0.099	-0.135	0.104	0.022	-0.026

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.
a Rotation converged in 9 iterations.
NOTE: For variable abbreviations see Table 5.04

Table 5.11. Rotated Component Matrix for Simpson index

VARIABLE	COMPONENT						
	1	2	3	4	5	6	7
ELEV	**-0.977**	0.021	-0.067	-0.111	0.090	-0.073	0.086
ASP	0.050	0.036	-0.070	0.000	**-0.727**	-0.234	-0.167
SLOPE	-0.209	0.112	-0.364	0.078	0.067	0.022	**0.659**
ORGM	-0.168	0.308	0.111	**-0.609**	-0.020	-0.231	**0.521**
P2O5	-0.015	0.028	-0.080	**-0.569**	0.079	**0.613**	-0.275
K2O	0.248	-0.001	0.011	0.038	0.075	**0.893**	0.074
CACO3	0.237	-0.163	-0.100	**0.776**	-0.209	-0.199	0.054
PH	0.299	0.035	0.031	**0.859**	0.086	0.078	0.083
SALT	0.253	**0.849**	0.033	0.159	0.094	0.031	-0.132
STR	-0.196	**0.714**	0.079	-0.418	0.134	-0.357	0.235
TEXTR	-0.158	**0.786**	0.070	-0.254	0.047	-0.114	0.199
SOILG	-0.244	0.042	**-0.763**	0.032	-0.066	0.232	-0.063
ERS	0.042	0.180	**0.873**	-0.021	0.046	0.216	0.013
SLDPT	0.219	-0.032	**0.883**	0.006	-0.006	-0.036	-0.218
GEO	-0.207	**0.806**	-0.010	-0.051	-0.140	0.251	-0.041
NDVI	-0.263	0.043	0.295	-0.037	**0.684**	-0.026	-0.062
SPVSD	-0.242	0.041	-0.271	-0.101	**0.692**	-0.138	-0.198
SWI	**0.547**	-0.483	0.381	-0.011	0.039	0.043	0.043
META	**0.970**	0.036	0.047	0.120	-0.112	0.057	-0.101
METS	**0.945**	0.131	0.036	0.152	-0.155	0.027	-0.050
MAXTA	**0.971**	-0.046	0.086	0.135	-0.111	0.072	-0.045
MAXTS	**0.244**	-0.189	0.190	0.332	-0.283	0.014	0.454
MINTA	**0.975**	0.017	0.061	0.110	-0.102	0.063	-0.062
MINTS	**0.973**	0.034	0.064	0.125	-0.115	0.052	-0.025
PRCPA	**0.940**	-0.193	0.148	0.086	-0.034	0.097	-0.043
PRCPS	**0.835**	-0.239	0.273	0.058	-0.010	0.126	0.112
PETAN	**-0.918**	0.205	-0.124	-0.156	0.116	0.064	-0.071
PETSE	**-0.956**	0.148	-0.103	-0.143	0.106	0.020	-0.013

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.
A Rotation converged in 11 iterations.
NOTE: For variable abbreviations see Table 5.05

Upon the completion of the initial iterative pass, a step by step variable removal scheme is applied based on the rules of Anti-Image Matrices. The best solution is found after the second pass with the removal of (1) aspect, (2) slope, (3) P_2O_5, (4) K_2O, (5) salt, (6) erosion, and (7) seasonal maximum temperature. After the removal of these six variables the Kaiser-Meyer-Olkin Measure of Sampling Adequacy is increased from 0.720 to 0.808 for Shannon Wiener index, and from 0.719 to 0.811 for Simpson index (Table 5.03 and Table 5.12).

Table 5.12. KMO and Bartlett's test with last 22 variables

Shannon Wiener Index		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.808
Bartlett's Test of Sphericity	Approx. Chi-Square	2502.219
	df	210
	Sig.	.000
Simpson Index		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.811
Bartlett's Test of Sphericity	Approx. Chi-Square	2454.079
	df	210
	Sig.	.000

The final factor loadings are given in Table 5.13. The rotation scheme and the factor loading scheme are slightly changed due to the different variables encountered into the analyses. Consequently, a more stable model is produced. The amount of the total variance explained with the variables are increased from 83.753% to 86.448 % for Shannon Wiener index (Table 5.08 and Table 5.13), and from 83.183% to 85.633 % for Simpson index (Table 5.09 and Table 5.13).

The stability of the model can also seen in the generic differentiation of the factors and their responsible variables (Table 5.14). For instance, the first component did not change and consists of elevation and climatic variables for each index. This is reasonable, because elevation is highly effective on climatic factors, and they are extremely important for plant species. Moreover, organic matter promoted to the second component, and establish a reasonable group with other related soil characters such as geology, texture, and saturation in both index. The third component totally changed and turned to $CaCO_3$ and pH in Shannon Wiener index. This is also more meaningful because of the apparent relationship between $CaCO_3$ and pH. Soil groups and soil depths constitute the fourth component of Shannon Wiener index, while they are in the third component of Simpson index. Normally, Classes derived from satellite images (NDVI and Supervised classes) take part in the fifth component of both indices.

Table 5.13. Total variance explained of two indices after removal of seven variables (the first 5 factors are taken into consideration for both indices, because their initial eigenvalues are bigger than 1) NOTE: The explanation of the components were given in Table 5.14 for each index.

Component	Shannon Wiener Index								
	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	11.1090	52.9020	52.902	11.109	52.902	52.902	9.492	45.200	45.200
2	2.7270	12.9850	65.887	2.727	12.985	65.887	3.054	14.541	59.741
3	1.8780	8.9410	74.828	1.878	8.941	74.828	2.105	10.023	69.765
4	1.2670	6.0310	80.859	1.267	6.031	80.859	2.008	9.563	79.327
5	1.1740	5.5890	**86.448**	1.174	5.589	86.448	1.495	7.121	86.448
6	0.8350	3.9740	90.423						
7	0.5120	2.4370	92.859						
8	0.4290	2.0420	94.901						
9	0.3380	1.6090	96.511						
10	0.2710	1.2900	97.801						
11	0.1700	0.8110	98.612						
12	0.1260	0.6020	99.214						
13	0.0898	0.4280	99.642						
14	0.0438	0.2080	99.850						
15	0.0186	0.0887	99.939						
16	0.0058	0.0275	99.966						
17	0.0042	0.0199	99.986						
18	0.0013	0.0061	99.992						
19	0.0008	0.0039	99.996						
20	0.0005	0.0024	99.998						
21	0.0003	0.0015	100.000						
Simpson Index									
1	10.9440	52.1130	52.113	10.944	52.113	52.113	9.377	44.653	44.653
2	2.6810	12.7640	64.877	2.681	12.764	64.877	2.873	13.680	58.333
3	1.9210	9.1470	74.025	1.921	9.147	74.025	2.120	10.097	68.430
4	1.2570	5.9860	80.011	1.257	5.986	80.011	2.117	10.082	78.512
5	1.1810	5.6220	**85.633**	1.181	5.622	85.633	1.495	7.121	85.633
6	0.8480	4.0370	89.670						
7	0.5440	2.5910	92.261						
8	0.4240	2.0200	94.281						
9	0.3960	1.8830	96.164						
10	0.2790	1.3280	97.493						
11	0.1850	0.8800	98.373						
12	0.1470	0.6980	99.071						
13	0.0977	0.4650	99.537						
14	0.0652	0.3100	99.847						
15	0.0192	0.0912	99.938						
16	0.0058	0.0278	99.966						
17	0.0042	0.0199	99.986						
18	0.0013	0.0061	99.992						
19	0.0008	0.0039	99.996						
20	0.0005	0.0024	99.998						
21	0.0003	0.0016	100.000						

Table 5.14. Rotated component matrices of two indices after removal of seven variables

VARIABLE	Components for Shannon Wiener Index				
	1	2	3	4	5
ELEV	**-0.971**	0.117	-0.133	-0.077	0.098
ORGM	-0.146	**0.531**	-0.475	0.141	-0.040
CaCO3	0.169	-0.193	**0.818**	0.029	-0.234
PH	0.280	-0.055	**0.846**	0.026	0.055
STR	-0.164	**0.794**	-0.384	0.100	0.100
TEXTR	-0.120	**0.846**	-0.175	0.073	-0.036
SOILG	-0.199	-0.058	-0.021	**-0.891**	0.014
SLDPT	0.227	-0.087	-0.039	**0.815**	0.018
GEO	-0.100	**0.813**	0.065	-0.146	0.020
NDVI	-0.259	0.056	-0.010	0.372	**0.752**
SPVSD	-0.210	0.003	-0.120	-0.255	**0.853**
SWI	**-0.570**	**0.625**	-0.019	-0.336	0.096
META	**0.969**	-0.056	0.157	0.059	-0.106
METS	**0.954**	0.056	0.196	0.043	-0.123
MAXTA	**0.966**	-0.134	0.146	0.091	-0.104
MINTA	**0.974**	-0.067	0.139	0.069	-0.100
MINTS	**0.974**	-0.046	0.147	0.073	-0.108
PRCPA	**0.917**	-0.287	0.073	0.165	-0.076
PRCPS	**0.816**	-0.293	0.011	0.270	-0.050
PETAN	**-0.886**	0.251	-0.140	-0.165	0.161
PETSE	**-0.932**	0.210	-0.143	-0.137	0.137
VARIABLE	**Components for Simpson Index**				
	1	**2**	**3**	**4**	**5**
ELEV	**-0.971**	0.110	-0.088	-0.135	0.097
ORGM	-0.152	**0.525**	0.145	-0.582	-0.047
CaCO3	0.171	-0.181	0.025	**0.824**	-0.232
PH	0.279	-0.058	0.044	**0.833**	0.048
STR	-0.169	**0.797**	0.095	-0.389	0.096
TEXTR	-0.125	**0.848**	0.071	-0.184	-0.040
SOILG	-0.189	-0.077	**-0.884**	-0.034	0.009
SLDPT	0.219	-0.070	**0.807**	-0.026	0.026
GEO	-0.097	**0.799**	-0.168	0.056	0.029
NDVI	-0.263	0.062	0.364	-0.008	**0.754**
SPVSD	-0.207	-0.005	-0.256	-0.127	**0.851**
SIMPSON	0.477	-0.499	0.462	-0.074	-0.103
META	**0.969**	-0.050	0.067	0.159	-0.104
METS	**0.954**	0.064	0.048	0.198	-0.122
MAXTA	**0.966**	-0.127	0.100	0.148	-0.103
MINTA	**0.974**	-0.062	0.079	0.139	-0.099
MINTS	**0.974**	-0.039	0.082	0.148	-0.108
PRCPA	**0.916**	-0.283	0.181	0.074	-0.078
PRCPS	**0.813**	-0.292	0.291	0.009	-0.054
PETAN	**-0.885**	0.245	-0.179	-0.142	0.163
PETSE	**-0.931**	0.203	-0.150	-0.145	0.138

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.
A Rotation converged in 6 iterations.
NOTE: For variable abbreviations see Table 5.04 and Table 5.05

5.4. Modeling

Two sets of multiple regression were employed to develop two distinct models for each indices (Shannon Wiener and Simpson). In multiple regression, each index values were defined as dependent variable (Y), while the extracted variables for each indices in *factor analysis* stage were described as independents (X_n).

The outcomes of multiple regression were summarized in three important domains. Those are ANOVA, Model Summary, and Coefficients. The ANOVA table tests the acceptability of the model from a statistical perspective. In the ANOVA table, the regression row displays information about the variation accounted for by the model, while the residual row displays information about the variation that is not accounted for by the model. While the ANOVA table is a useful test of the model's ability to explain any variation in the dependent variable, it does not directly address the strength of that relationship. The model summary table reports the strength of the relationship between the model and the dependent variable. R, the multiple correlation coefficient, is the linear correlation between the observed and model-predicted values of the dependent variable. Its large value indicates a strong relationship. R square, the coefficient of determination, is the squared value of the multiple correlation coefficient.

The results of multiple regression were summarized in Table 5.15 and Table 5.16 for Shannon Wiener and Simpson indices, respectively. In the tables the regression sum of squares is found bigger than the residual sum of squares, which indicates that about 90% of the variation is explained by the model for Shannon Wiener index, and about 80% of the variation is explained by the model for Simpson index. Moreover, the significance values of the F statistic are less than 0.05, which means that the variation explained by the models are not due to chance. The unstandardized coefficients are the coefficients of the estimated regression model. According to these values, two distinct models were established and given in Table 5.17.

Table 5.15. The summary of multiple regression for Shannon Wiener index

ANOVA[b]

Model		Sum of Squares	Df	Mean Square	F	Sig.
1	Regression	29.173	20	1.459	17.055	.000[a]
	Residual	2.993	35	.086		
	Total	32.167	55			

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.952[a]	.907	.854	.292452

a Predictors: (Constant), SPVISED, GEOLOGY, PH, SOILDPTH, ORGMT, NDVI, MINTEMAN, TEXTURE, SOILGR, CACO3, PRECIPSE, STR, PETAN, METEMSE, PRECIPAN, MAXTEMAN, MINTEMSE, ELEV, PETSE, METEMAN
b Dependent Variable: SWI

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	22.296	28.951		0.770	0.446
	ELEV	0.008	0.006	2.434	1.246	0.221
	META	1.688	1.894	1.888	0.891	0.379
	METS	-0.981	1.670	-0.796	-0.588	0.561
	MAXTA	-1.068	1.072	-0.711	-0.996	0.326
	MINTA	-0.289	0.702	-0.457	-0.412	0.683
	MINTS	0.944	0.959	1.298	0.985	0.332
	PRCPA	0.097	0.148	0.822	0.654	0.517
	PRCPS	-0.143	0.132	-0.631	-1.088	0.284
	PETAN	-0.168	0.524	-0.362	-0.321	0.750
	PETSE	-0.032	0.316	-0.176	-0.102	0.919
	STR	-0.004	0.016	-0.041	-0.287	0.776
	TEXTR	-0.008	0.193	-0.005	-0.043	0.966
	GEO	0.226	0.027	0.720	8.386	0.000
	ORGM	0.071	0.081	0.104	0.870	0.390
	CACO3	0.020	0.014	0.151	1.486	0.146
	PH	-0.378	0.301	-0.121	-1.255	0.218
	SOILG	0.005	0.366	0.001	0.015	0.988
	SLDPT	-0.081	0.080	-0.084	-1.014	0.318
	NDVI	-0.002	0.002	-0.099	-1.449	0.156
	SPVSD	0.139	0.043	0.238	3.241	0.003

a Dependent Variable: SWI

Table 5.16. The summary of multiple regression for Simpson index

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	2.669	20	.133	5.027	.000[a]
	Residual	.929	35	.027		
	Total	3.598	55			

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.861[a]	.742	.594	.162938

a Predictors: (Constant), SPVISED, GEOLOGY, PH, SOILDPTH, ORGMT, NDVI, MINTEMAN, TEXTURE, SOILGR, CACO3, PRECIPSE, STR, PETAN, METEMSE, PRECIPAN, MAXTEMAN, MINTEMSE, ELEV, PETSE, METEMAN
b Dependent Variable: SIMPSON

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	10.424	16.130		0.646	0.522
	ELEV	-0.002	0.003	-2.053	-0.631	0.532
	META	-0.009	1.055	-0.031	-0.009	0.993
	METS	-0.220	0.931	-0.533	-0.236	0.815
	MAXTA	-0.293	0.597	-0.583	-0.490	0.627
	MINTA	0.437	0.391	2.062	1.117	0.272
	MINTS	-0.428	0.534	-1.759	-0.801	0.429
	PRCPA	-0.064	0.083	-1.632	-0.779	0.441
	PRCPS	0.087	0.073	1.146	1.186	0.244
	PETAN	-0.027	0.292	-0.177	-0.094	0.926
	PETSE	0.028	0.176	0.462	0.161	0.873
	STR	-0.003	0.009	-0.074	-0.310	0.758
	TEXTR	0.079	0.108	0.144	0.738	0.466
	GEO	-0.064	0.015	-0.607	-4.244	0.000
	ORGM	0.004	0.045	0.018	0.092	0.927
	SOILG	-0.001	0.204	-0.001	-0.005	0.996
	SLDPT	0.030	0.045	0.093	0.673	0.505
	PH	0.269	0.168	0.257	1.602	0.118
	CACO3	-0.011	0.008	-0.236	-1.392	0.173
	NDVI	0.001	0.001	0.117	1.031	0.310
	SPVSD	-0.053	0.024	-0.269	-2.202	0.034

a Dependent Variable: SIMPSON

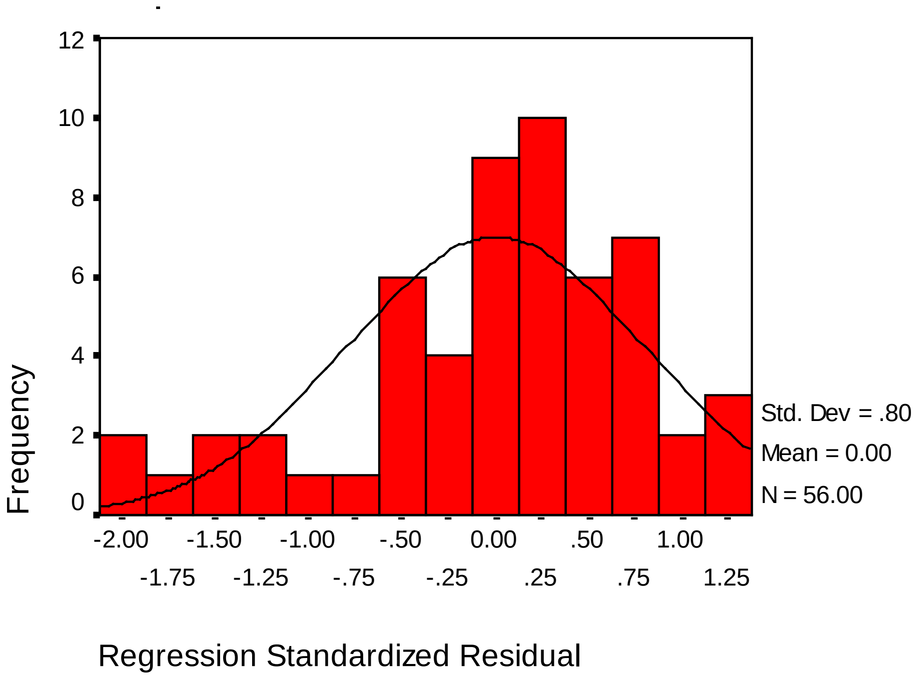
Table 5.17. Estimated models for Shannon-Wiener and Simpson indices (NOTE: For variable abbreviations see Table 5.04 and Table 5.05)

MODELS
SHANNON WIENER INDEX = 22.296 + (0.008 * ELEV) + (1.688 * META) + (0.944 * MINTS) + (0.097 * PRCPA) + (0.226 * GEO) + (0.071 * ORGM) + (0.020 * CACO3) + (0.005 * SOILG) + (0.139 * SPVSD) – (0.981* METS) – (1.068 * MAXTA) – (0.289 * MINTA) – (0.143 * PRCPS) – (0.168 * PETAN) – (0.032 * PETSE) - (0.004 * STR) –(0.008 * TEXTR) – (0.378 * PH) – (0.081 * SLDPT) – (0.002 * NDVI)
SIMPSON INDEX = 10.424 + (0.437 * MINTA) + (0.087 * PRCPS) + (0.028 * PETSE) + (0.079 * TEXTR) + (0.004 * ORGM) + (0.030 * SLDPT) + (0.269 * PH) + (0.001 * NDVI) – (0.002 * ELEV) – (0.009 * META) – (0.220 * METS) – (0.293 * MAXTA) –(0.428 * MINTS) – (0.064 * PRCPA) – (0.027 * PETAN) – (0.003 * STR) – (0.064 * GEO) – (0.001 * SOILG) – (0.011 * CACO3) – (0.053 * SPVSD)

Another important component of the multiple regression is residuals. Basically, the residual is the difference between the observed value of the dependent variable and the value predicted by the model. Residual statistics are summarized in Table 5.18 for each index. In Table 5.18, the minimum, maximum, mean, standard deviation and sample size are displayed for predicted value, residual, standardized predicted value, and standardized residual. For each case, the predicted value is the value predicted by the regression model. If the model is appropriate for the data, the residuals should follow a normal distribution. Standardized predicted values are the values having 0 mean and 1 standard deviation. A histogram or P-P plot of the residuals is also helpful to check the assumption of normality of the error term. The shape of the histogram should approximately follow the shape of the normal curve, and the P-P plotted residuals should follow the 45-degree line (SPSS, 2001). The histograms and P-P plotted residuals are given in Figure 5.07 and Figure 5.08. The histograms are acceptably close to the normal curve for both indices. Neither the histogram nor the P-P plot indicates that the normality assumption is violated.

Table 5.18. Residual statistics of two models

SHANNON WIENER MODEL					
	Minimum	**Maximum**	**Mean**	**Std. Deviation**	**N**
Predicted Value	.24243	3.05349	1.67230	.728301	56
Residual	-.60747	.40209	.00000	.233296	56
Std. Predicted Value	-1.963	1.896	.000	1.000	56
Std. Residual	-2.077	1.375	.000	.798	56
SIMPSON MODEL					
	Minimum	**Maximum**	**Mean**	**Std. Deviation**	**N**
Predicted Value	-.03671	.91534	.36804	.220290	56
Residual	-.28914	.36157	.00000	.129980	56
Std. Predicted Value	-1.837	2.484	.000	1.000	56
Std. Residual	-1.775	2.219	.000	.798	56



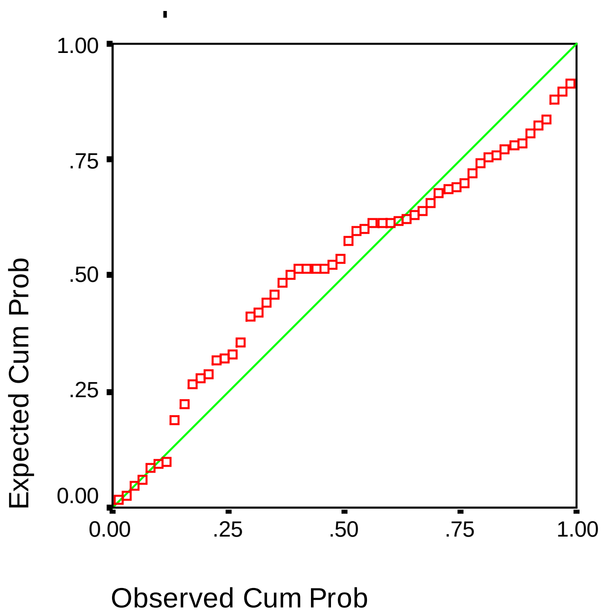
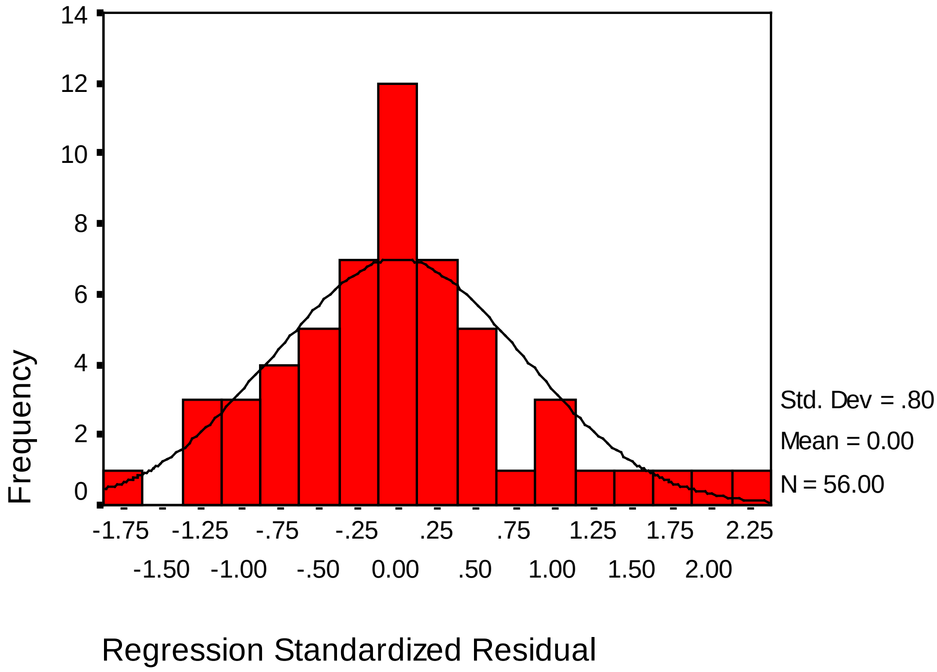


Figure 5.07. Histogram and Normal P-P plot of regression standardized residual for Shannon Wiener Index



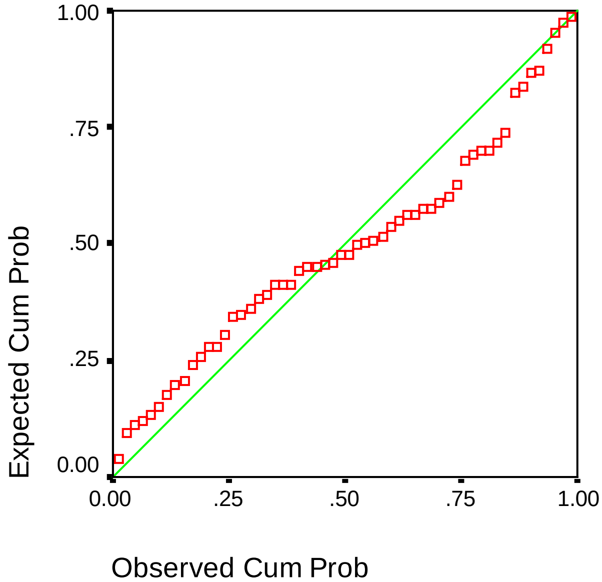


Figure 5.08. Histogram and Normal P-P plot of regression standardized residual for Simpson Index

5.5. Mapping

Mapping process is simply the application of the two developed models (regression equations) by using GIS functions. Application of the models contains map calculations by using developed grid maps in standard grid size (30 x 30 m). With the power of GIS, continuous surfaces such as elevation, soil geology etc. can be represented by grid maps as already stated in Chapter 4, and some mathematical operations can easily be conducted on these grid themes. In this study, the developed grid maps (complementary data set) and map calculator functions were employed throughout the application process of two models .

In Arc/View software, the map calculator dialog aids in the creation of an expression that produces a new output grid theme. The expression can be based on a single grid theme or multiple grid themes. The Map Calculator can create simple expressions with a single function or complex with many operations and functions. Mathematical operators apply a mathematical operation to the values in two or more input grid themes; mathematical functions apply a mathematical function to the values in a single input grid theme. There are three groups of mathematical operators in the Map Calculator: Arithmetic, Boolean, and Relational (ESRI, 1994,1996,1997).

The Arithmetic operators (*, /, -, +) allow for the addition, subtraction, multiplication, and division of two grid themes, or numbers, or combination of the two. The Boolean operators (And, Not, Or, and Xor) use Boolean logic (TRUE or FALSE) on the input values. Output values of TRUE are written as 1 and FALSE as 0. The Relational operators (<, <=, <>, =, >, and >=) evaluate specific relational conditions. If the condition is TRUE, the output is assigned 1; if the condition is FALSE, the output is assigned 0. There are also four subgroups of mathematical functions: Logarithms, Arithmetic, Trigonometric, and Powers (ESRI, 1994,1996,1997).

The output grid theme from Map Calculator is the result of any operation explained above. Resulting species diversity maps for each indices were developed after running of the models, and they given in Figure 5.09 and Figure 5.10.

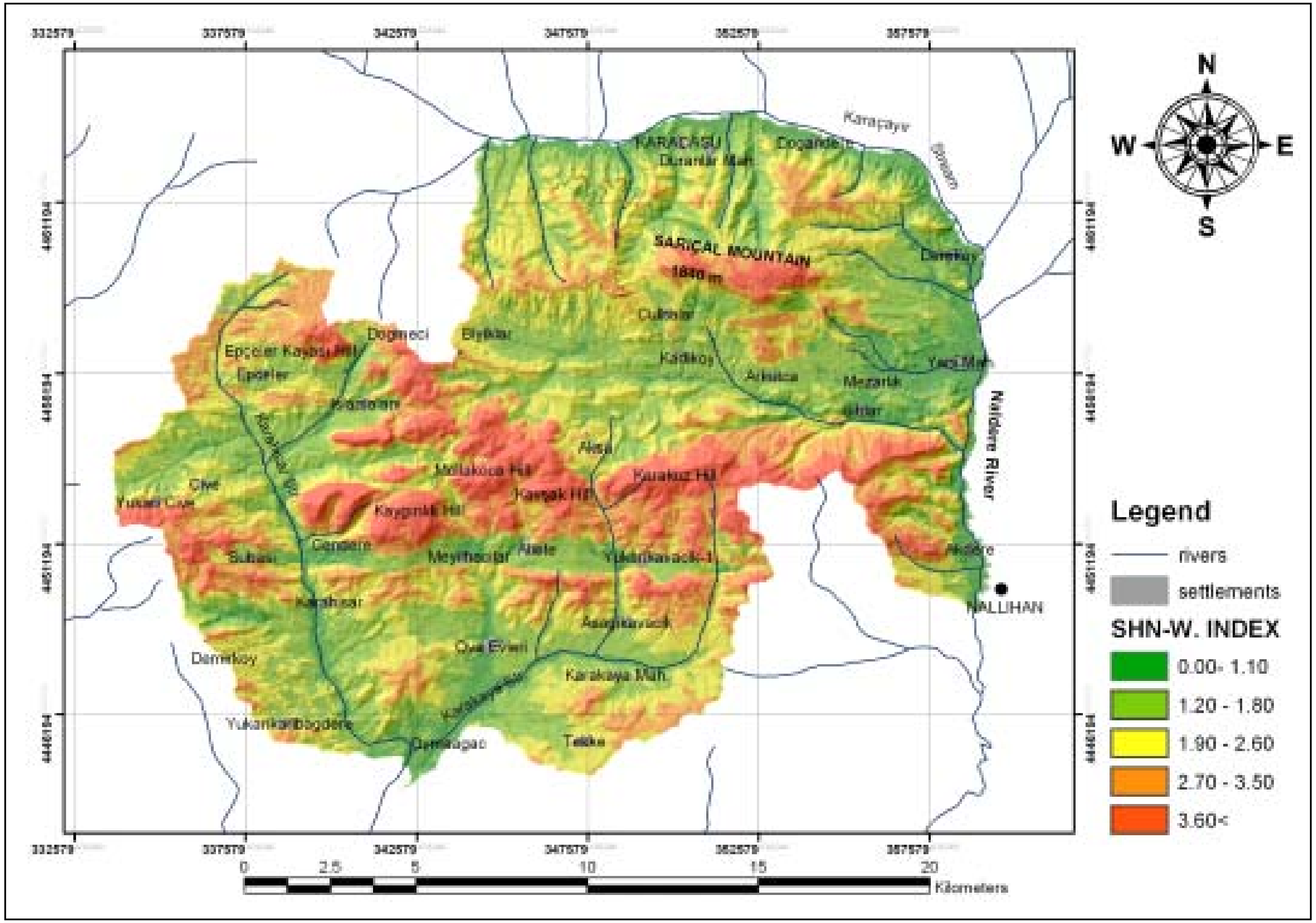


Figure 5.09. Plant Species Diversity Map according to the developed model for Shannon Wiener Index

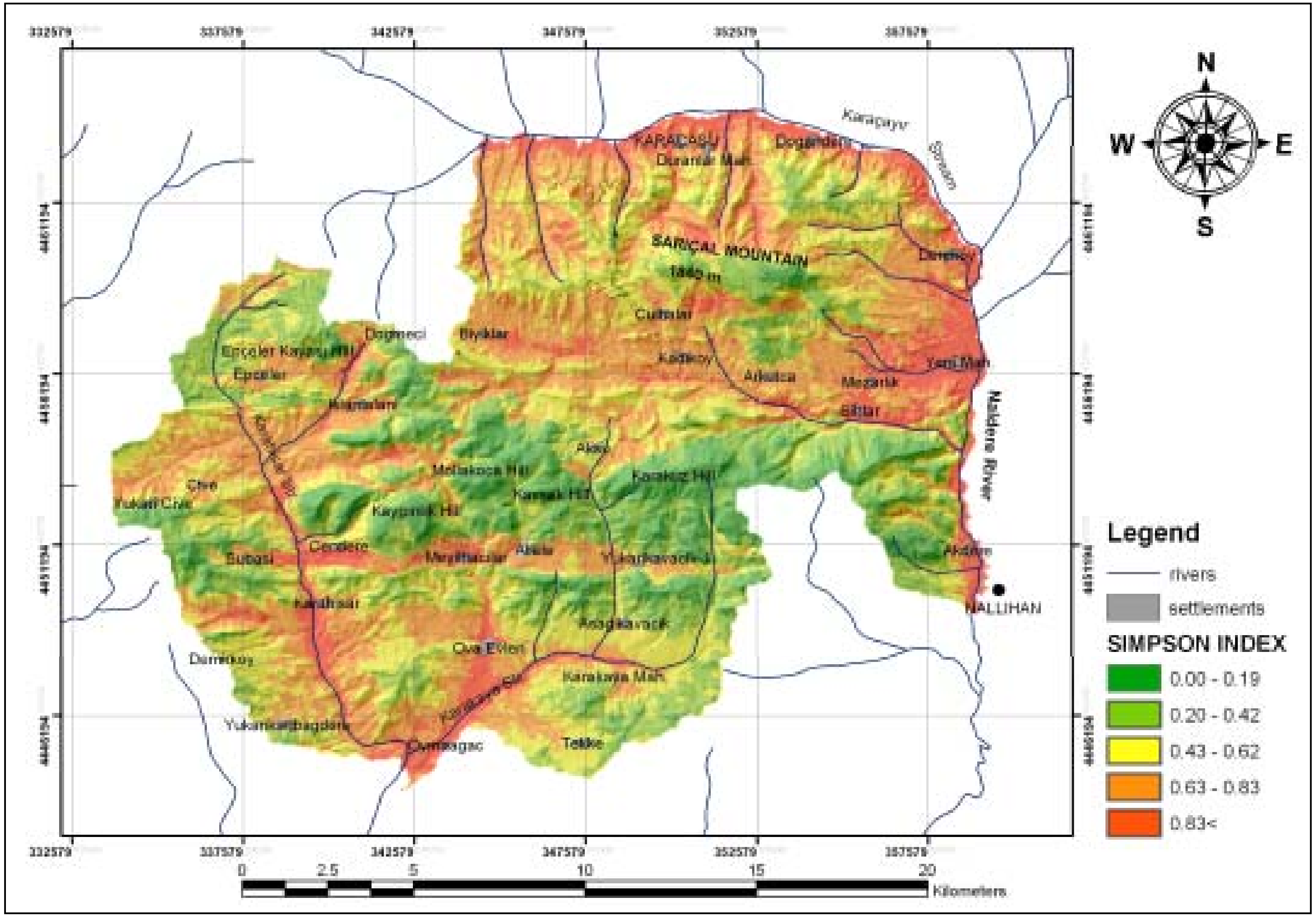


Figure 5.10. Plant Species Diversity Map according to the developed model for Simpson Index

CHAPTER 6

DISCUSSION

This chapter discusses the results of the study. First, the quality of the data used in this study was examined. Then, the results of the study, and reliability of the two developed models questioned.

6.1. Data quality

6.1.1. Remote sensing data

The supervised classification and Normalized Difference Vegetative Index (NDVI) maps delineated the current situation of the land cover. By using these visual tools, the nature and characteristics of the main plant establishments are understood very well. Consequently, a lot of distinguished characteristics that are impossible to be shown in otherwise are detected and displayed. The 4, 5, 3 band combination is produced very satisfactory results for vegetation investigation. The supervised classification obtained 92.16 % overall accuracy with a Kappa coefficient of 0.8828, and produced a reliable result. On the other hand, the detailed information about the distribution of plant species is still remained unidentified because of the spatial resolution of LANDSAT image and species characteristics. For instance, both black pine and red pine species took place in the coniferous class. Nevertheless, supervised classification and NDVI played important roles especially in the modeling stage by supplying the recent reliable information about land use and main plant covers. Although monoscopic air photos are not so new, they also associated with the recent data very well, and supplied valuable visual information during the field surveys and supervised classification process. Signatures of the land cover, for instance, are determined easily by using air photos. Moreover, the air photos helped a lot to check the validity of supervised classification and applicability of the developed model in a visual way.

6.1.2. Soil data

The soil data used in this study is originated from two distinct resources. The first one is the digital soil map database produced by the General Directorate of Rural Affairs, and the second one is the recent soil samples collected from the study area. Both these two resources have several weak points that should be explained here. The principle weakness of the digital soil map database depends on its date. Although this digital database has been created recently, all of the information contents came from the obsolete soil maps in paper format. The first national soil surveys were conducted by Harvey Oaks from the U.S.A. and his Turkish colleagues between 1952 and 1954 in Turkey, and a general soil map with 1 / 800 000 scale was produced in national level. In the period of planning development of Turkey, this map could not meet the increased demands. For this reason, General Directorate of Soil and Water (TOPRAKSU) started "Developed Soil Mapping Surveys of Turkey" in 1966. This project was finished in 1971 and produced very detailed soil maps (1 / 25 000 scale) in paper format. This soil database was transformed to a digital soil database between 1999 and 2002, and contains important information about (1) big soil groups, (2) erosion classes, (3) soil depth, (4) current land use, (5) land types, (6) land use ability for agriculture, (7) sub classes of land use ability, and (8) geographical data (rivers, lakes, etc.).

The most significant drawback of this digital soil database is spatio-temporal change in a 32-year period. Many land-use and soil characteristics changed in constructive or destructive ways during this period. For this reason, some parts of digital soil database, which are sensitive to this kind of change, were excluded to eliminate this disadvantage in this study. Consequently, big soil group and soil depth layers were selected for further analysis. Although erosion classes may change over time, it was also chosen for further analysis by depending on its importance. Accordingly, big soil groups and soil depth variables were found important and, they took place in the final models of two indices. On the contrary, erosion classes could not get a chance to enter the models.

The second weakness of the digital soil database is related to its content, because they do not contain enough information about some physical and chemical soil characters that could be so important for plant diversity. So, collected soil samples were analyzed and used to satisfy these kinds of data needs. Despite of the detailed information, the collected soil samples represented only soil surface characteristics, and they do not reflect the information about soil horizons. Moreover, the grid maps of these soil variables were produced by using an interpolating technique (Inverse Distance Weighted). So these are the weak points of sampled soil data.

6.1.3. Forest data

The digital forest map (1/25 000 scale) set up a good starting point for the investigations at species level. It gives the essential evidences not only about forest species (black pine, red pine, oak spp.) but also about land-use (agriculture, settlement, forest etc.). Moreover, the ages, stages and conditions of each forest stand are available within the content of this digital database. Typically, General Directorate of Forestry renews the forest stand maps in some certain time periods (5-10 years) which are determined by forest management districts. The most recent forest map of the study area was produced in 1996, and the next one will probably be produced in 2006. So, the only available recent forest map belongs to the year 1996. Although this map was produced in 1996, essential parts of it are still valid. In addition, some important changes in the area are well known by the forest management staff, and these changes were easily marked on the forest map to update. Therefore, the digital forest map may be the most detailed and important database in this study, and it was very helpful in many stages of the study such as; site selection for sampling, classification, and accuracy assessment.

6.1.4. Geologic data

Geological surveys have been done since 1935 in Turkey, and the original geological maps of the study area were prepared between 1981 and 1987 by the General Directorate of Mineral Research and Exploration (MTA). These maps were digitized in 1997. Digital geological maps contain the data about geological formations and their ages. Depending on their relative importance, the ages of the geological formations were underestimated in this study. Instead of ages, geological formations were focused to investigate their effects on plant species. Consequently, some geological formations such as alluvium and old alluvium were evaluated in only one class "alluvium". In a geological map, spatio-temporal changes might not be so important, because geological changes generally occur in long time periods. Although digital geological map of the study area was derived from old paper maps, it still contains valuable information about geological formations that could be related to plant distribution in the study area.

6.1.5. Topographic data

Topographical maps with 1 / 25 000 scale have been produced by General Commandership of Mapping (HGK) in Turkey. Those maps have been renewed, when the big topographical changes occurred in an area such as; construction of a dam. The only available topographic maps of the study area were produced in 1977. Since then, they were not renewed, because of unimportant changes in topography. The digital database of topography was created from the paper maps of H26 (a4, a3, b4, c1, d1, d2). This database supplied very detailed and reliable information about elevation, aspect, and slope that are so important for the plant characteristics of the study area. Moreover it was used to derive detailed climate data in LOCCLIM software.

6.2. Statistical analyses

The reliability of overall statistical analyses should be questioned in order to make a correct interpretation. This can be done in two ways. Those are (1) mapping residuals to predict the locations where the models work perfectly, and (2) logical interpretations in ecological point of view. Consequently, the residuals from regressions were investigated and mapped to check the validity of regression equations in this study. Basically, residual from regression is defined as the Y_{cn}-Y_n formula where Y_n represents the observed value of Y for the *nth* unit area and Y_{cn} represents the computed value of Y for the *nth* unit area (Berry and Marble 1968, McGrew and Monroe 1993). According to McGrew and Monroe (1993); in multivariate regression, the independent variable explains a portion of the total variation in the dependent variable and leaves a remainder unexplained (residual error). If this error is interpreted as a new dependent variable, other variables can be identified to explain more of the remaining variation. Thus, residuals from regression should not be viewed as the end of the research process, but rather as an intermediate step in uncovering further influences on the dependent variable. Maps of the residual may be particularly useful to formulate some hypotheses under certain circumstances. Specifically, maps of residuals from regression are useful for the geographer to formulate new hypotheses and to identify new variables for inclusion in an investigation (Berry and Marble 1968). The magnitude of residuals can be interpreted two ways. First, it is possible to check the existence of another variable with a spatial pattern similar to that of the residuals. A residual map can be a useful tool to give researchers some clues about a new variable in this case. Second, it shows if there is a logical, rational reason for this variable to influence the dependent variable (McGrew and Monroe, 1993).

The two residual maps were developed by using the residuals of Shannon Wiener and Simpson indices (Figure 6.01). Residual values of the two indices were converted to maps according to Inverse Distance Weighted (IDW) method. In both maps, the areas that have small residual values (between –1 and +1 standard deviations) indicated that models are strongly predictive. Medium residual values (between -1 and –2 or +1 and +2 standard deviations) suggested that models are medium predictive. On the other hand, the areas have bigger residual values (<-3 or >+3 standard deviations) pointed out that models are less predictive, and another independent variable hidden in that area may be effective on index values. The areas that have bigger values need further research to investigate any other independent variables effects on index values.

The percent area covered by each distinct residual (less, medium, and strong) class could also delineate the credibility of two models (Table 6.01). The areas that were found less predictive for both models cover small percentages of the total area (7.82% for Shannon Wiener, 6.60% for Simpson). On the contrary, the areas that were found moderate and strongly predictive encompass significant part of the total area (92.18% for Shannon Wiener, 93.4% for Simpson). According to figures in Table 6.01; it seems that the models developed for both indices run very well.

Table 6.01. Percent area values of predictivity classes of the two models

PREDICTIVITY LEVELS	% area covered SWI	% area covered SIMPSON
less predictive	7.82	6.60
moderate predictive (1)	27.33	25.28
strongly predictive (2)	64.85	68.12
TOTAL of 1 and 2	92.18	93.40

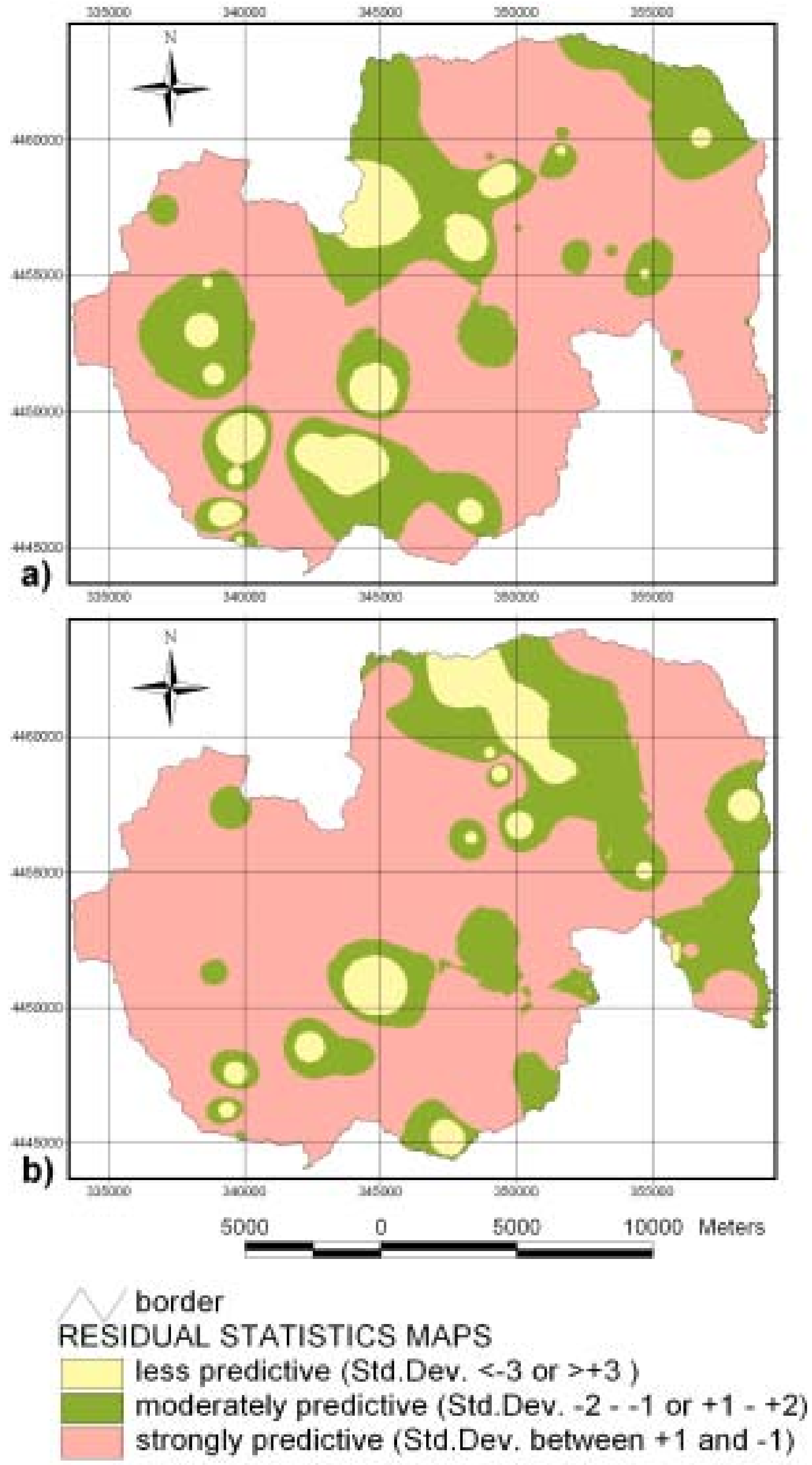


Figure 6.01. Residual maps of the two model a) Shannon Wiener Index Model, and b) Simpson Index Model

When we investigated the map developed by using Shannon Wiener Model (Figure 5.09), very low (0 - 1.19) and low (1.20 - 1.89) index values were found along the rivers where irrigated agriculture, orchards, and settlements are common. Specifically, these low index values were detected along the Karaçayır stream in north, Naldere river and its branches in east, Karakaya stream in south, and Karahisar stream in southwest. All these areas were accompanied with the settlements that were called Karacasu, Duranlar Mahallesi, Doğandere, Dereköy, Yeni Mahalle, Şıhlar, Akdere, Arkutça, Kadıköy, Cendere, Karahisar, Ovaevleri, and Karakaya Mahallesi. On the contrary, high (0.63 - 0.83) and very high (more than 0.84) index values were detected for the same areas in the map developed by using Simpson Index Model (Figure 5.10).

When we came back to the Shannon Wiener Model (Figure 5.09), high (2.70 – 3.59) and very high (more than 3.60) index values were detected in the areas where settlements are rare, and the elevation is high. For instance, Sarıçal Mountain in north, Epçeler Kayası and Kemiklikaya Hills in northwest, Kaygınlık, Mollakoca, Kuyupınar, Kavşak and Karakuz hills in the middle of the study area gained high index values that means plant biodiversity higher in those areas. On the other hand, Simpson index values were found very low (0 - 0.19) and low (0.20 - 0.42) in those areas (Figure 5.10).

The results showed that two models worked inversely. The difference between the two models comes from their inherent characteristics. Because their formulas are different, the absolute values of Shannon Wiener and Simpson indices are also different for the same community (Barbour *et al*., 1987). Basically, Shannon Wiener Index gives importance to both species evenness (relative abundance of species) and species richness (the number of species in a community) (Molles, 1999), while Simpson index reflects dominance because it weights the most abundant species more heavily than the rare species. In another words, if an area has really high plant biodiversity, both species evenness and species richness components must be higher according to the Shannon Wiener Index. On the other hand, Simpson Index is more sensitive to the most abundant species (Barbour *et al*., 1987).

To show the differences of these two indices, an example about two different plant communities was given in Table 6.02. In Table 6.02, both Shannon Wiener and Simpson indices were calculated for two plant communities (community-a and community-b) having the same number of species (species richness) but different species evenness. Although both plant communities have totally 5 species and 25 individual plants, their numerical distributions are different. For this reason, index values of these communities were not alike. Shannon Wiener Index value was found lower (0.662) in community-a, and higher (1.610) in community-b. On the contrary, Simpson Index value was found higher (0.712) in community-a, and lower (0.200) in community-b. As it is clearly seen in Table 6.02, Shannon Wiener Index is sensitive to species richness as well as species evenness, but Simpson Index is sensitive to the most abundant species only.

Table 6.02. The inherent differences between Shannon Wiener and Simpson indices (The data of the two communities were taken from Molles (1999), and Simpson indices were calculated and added to the table to make a comparison)

Community a					
species	number	Proportion (p_i)	logep_i	$p_i$logep_i	pi^2
1	21	0.84	- 0.174	- 0.146	0.7056
2	1	0.04	- 3.219	- 0.129	0.0016
3	1	0.04	- 3.219	- 0.129	0.0016
4	1	0.04	- 3.219	- 0.129	0.0016
5	1	0.04	- 3.219	- 0.129	0.0016
TOTAL	25	1.00		**0.662 Shannon W.**	**0.7120 Simpson**
Community b					
species	number	Proportion (pi)	logep_i	$p_i$logep_i	pi^2
1	5	0.20	-1.609	- 0.322	0.040
2	5	0.20	-1.609	- 0.322	0.040
3	5	0.20	-1.609	- 0.322	0.040
4	5	0.20	-1.609	- 0.322	0.040
5	5	0.20	-1.609	- 0.322	0.040
TOTAL	25	1.00		**1.610 Shannon W.**	**0.200 Simpson**

Depending on the field surveys and observations, some high biodiversity areas were detected near the agricultural lands where can be exemplified by low elevation, warmer climate and productive soil. On the other hand, these kinds of diverse areas might not reflect the certain situation, because they contain small sections between the agricultural areas that can not represent the majority of wild plant cover. Although the number of species in those small sections is tend to be higher, species evenness or relative abundance of species are low because of the expanding human activities.

There is some literature to clarify this situation. First of all, the relationship between the disturbance and diversity should be investigated to decide which model is running well and most suitable for these kinds of applications. ”Generally, there is a gradient of increasing species diversity from the high elevations to low elevations. These gradients follow complex environmental gradients of increasing warmth, among other factors. Maintenance of high diversity appears to require episodic, random (stochastic) disturbance. Very stable, regionally extensive, and homogenous communities exhibit lower species diversity than communities composed of mosaic of patches disturbed at various times in the past by wind throw, fire, disease etc. Following disturbance, diversity increases with time up to a point where dominance by a few, long-lived, large-sized species reverses the trend, and diversity falls thereafter (Barbour *et al.* 1987)."

According to Connell (1978); disturbance is a prevalent feature of nature that significantly influences the diversity of communities. He proposed that high diversity is a consequence of continually changing conditions, not of competitive accommodation at equilibrium. Connell (1978) predicted that intermediate levels of disturbance promote higher levels of diversity, while both high and low levels of disturbance would lead to reduced diversity. A wide array of species can colonize open habitats, but there is not enough time for the most effective competitors to exclude other species in the intermediate levels of disturbance. Although intense human disturbance reduces the species diversity, moderate levels of disturbance may increase the diversity of some communities.

Moss and Guarino (1995) said that fires, floods, landslides, high winds and drought may be termed 'natural' disturbance factors in that they can occur without human help. Purely 'artificial' disturbance ranges from management to complete destruction of the habitat, for example for mineral exploitation or construction or by pollution. According to Moss and Guarino (1995); artificial disturbance is best noted under land-use.

Agriculture is kind of disturbance that reduces the species diversity. Although, agricultural lands are highly productive and suitable for many plant species, this suitability has generally interrupted by agricultural practices itself. Consequently, the areas that are near or within the agricultural lands are highly supposed to be low plant biodiversity. In this point of view, the model developed for Simpson Index could not reflect the real situation in the area.

Potential evapotranspiration (PET) might be another important indicator to evaluate the two models, because changes in PET are likely to have large impacts on terrestrial vegetation. High PET values indicate high biomass production that might be a clue for high biodiversity. When grid maps of PET in annual and seasonal basis (Figure 4.11) and plant species diversity maps (Figure 5.09 and Figure 5.10) were examined carefully, the agreement between the Shannon Wiener biodiversity map and PET maps appeared clearly. In another words, high PET values matches with the high Shannon

Wiener Index values which is very reasonable. On the contrary, there is an inverse relationship between the Simpson biodiversity map and PET maps which is unsound.

Considering the disturbance and PET factors, it might be said that the model developed for Shannon Wiener index runs better than the model developed for Simpson index.

CHAPTER 7

CONCLUSIONS

Ecosystems that are defined as the sum of the plant community, animal community, and physical environment in particular regions or habitat are subject to rapid change. Today, many valuable forest ecosystems are under pressure of the rapid development of settlements, agriculture, and industry. These rapid changes bring together some environmental and managing problems. Consequently, lack of information leads to practices that use resources in ways that can not be sustained over time.

This study put forward a new approach to the conventional diversity (Shannon-Wiener and Simpson) indices by using new and substantial geographic information systems (GIS) and remote sensing (RS) tools. Using this new approach and tools, plant biodiversity of Nallıhan forest ecosystem was modeled and mapped.

The study revealed several important outcomes. First of all, it showed the importance of GIS for “Conservation Biology” that is an emerging discipline dedicated to the preservation of endangered species and habitats. Within this frame, mapping the areas with high plant biodiversity is very important for decision makers. Basically, these kinds of applications let researchers to determine the specific sites such as; hot spots, protecting areas, national parks, and gene management zones. Effective management plans and actions can only be applied after determining these important sites. Although traditional mapping methods have been available, this study brought a fresh and objective point of view, maybe, at the first time in Turkey.

The outcomes of this study showed that biological sciences especially plant ecology and sociology are highly suitable for GIS and RS applications. The study indicated that the spatial analysis is the most important contribution of new GIS technology. The methodology of the study delineated how a link between map-based analysis of spatial patterns and well-developed rigorous quantitative analytical methods can be established by using the spatial analysis. Consequently, this study demonstrated that the interpretation of spatial patterns is no longer subjective for the disciplines mentioned above. With appropriate measurements of map features, cumbersome and time-consuming analysis of complicated spatial relationships become accessible. This supplies a deeper understanding of the issues facing, and let researchers bring more information and less conjecture to the problem solving process. For mapping and analyzing purpose, scientists working in these areas can rely on GIS and RS.

This study also revealed that remotely sensed data especially LANDSAT-TM images are highly suitable to determine the current land cover and land use characteristics in a cheap, fast, and accurate way. The results of the study indicated that both land cover and land use characteristics are highly effective on the plant species distribution and their diversity. Applied supervised classification with maximum likelihood parametric rule and the 4-5-3 band combination of the LANDSAT-TM image were found suitable to display the land cover and land use characteristics of the area. Therefore, scientists who are studying on plant ecology can produce very satisfactory results by using these methods and band combination.

The importance of monitoring process, detecting the changes of an area through time, was also implied in this study. Monitoring is highly important for the identifications of spatio-temporal dynamics of important changes such as; shrub invasion, grassland fragmentation, erosion and desertification. Analyzing, understanding, and modeling of the desertification and erosion are especially important for Turkey. Monitoring of an area is possible by using the power of GIS and RS techniques. On the other hand, a well developed compatible standard digital database is important as well as these techniques. This study showed that a complete digital database set is a good starting point for monitoring studies. In future research, this database can easily be updated and used in order to detect the mostly changed areas in an objective and quantitative way.

This study emphasized the importance of quantitative field data. Although Braun Blanquet method has been successfully applied in the studies related to plant sociology, more quantitative methods are necessary for the spatial analysis. This study proved that some of the subjectivities could be prevented by using more objective methods. In this point of view, quadrat sampling method is highly applicable, and Braun Blanquet values can easily been quantified by using some additional data such as plant density. Moreover, the study showed that data gathering, storaging, analyzing and reporting are very important tasks in GIS. These procedures require reasonable methodologies. Basically, data must be gathered in a suitable and standardized way. Data gathering methods may be different from one discipline to another but their compatibility with the

GIS and RS techniques must be provided. Scientist studying in the field need to record the coordinates of sampling sites in order to establish a link between the data and the geographic entities focused. This can be done by using the coordinated field data. Global Positioning System (GPS) is useful to determine the exact locations of sampling sites on the earth. For this purpose, GPS devices that are getting cheaper and more accessible can be used effectively.

The complementary data about topography, geology, soil, forest, climate, remote sensing classes, and NDVI classes supplied very important information, and played the back bone role at the spatial analysis and modeling stages in this study. To proliferate the modeling studies in large areas, these kinds of data should be available, accessible and cheap. Institutions in the position of primary data producers should establish their digital database in a standardized way. The database in digital format should be compatible with the common GIS and RS software in worldwide. In Turkey, digital forest database has not been available yet. Topographical, geological, and soil maps have been produced by General Commandership of Mapping (HGK), General Directorate of Mineral Research and Exploration (MTA), and General Directorate of Rural Affairs (KHGM), respectively. On the other hand the prices are high, and accessibility is difficult. It is recommended and believed that this situation will be changed by the establishment of National Database of Turkey. Establishment such a national database is fundamental to conduct in the near future by the widespread GIS and RS applications in universities, research institutes, and companies.

The results of the study implied that further research about habitat characterization at species level can be achieved by the development of different models. The new satellites with high spatial resolution (IKONOS or QUICKBIRD) can improve the quality of these kinds of studies. With the contribution of detailed satellite images individual species can be mapped by using new sampling methods, GIS and RS.

REFERENCES

Achard, F., H. Eva, and P. Mayaux, 2001. Tropical forest mapping from coarse spatial resolution satellite data: production and accuracy assessment issues. International Journal of Remote Sensing, 22/14: 2741–2762.

Ager, D. V., 1993. The new catastrophism: the importance of the rare event in geological history. Cambridge University Press, Cambridge.

Akman, Y., 1974. Contribution à l'étude de la flore de la region de Beypazarı Karaşar et de Nallıhan. Communication, De La Faculté des Sciences de L'univesité d'Ankara, Serie C, 18/2: 7-50.

Akman, Y., 1974. Etüde phyto-ecologique de la region de Beypazan-Karaşar et Nallıhan, Communication, De La Faculté des Sciences de L'univesité d'Ankara. Serie C, 18/3: 50-113.

Akman, Y., and O. Ketenoğlu, 1976. The phytosocilogical and phytoecological investigetion on the Ayaş Mountain. Communication, De La Faculté des Sciences de L'univesité d'Ankara, Serie C, 20/2: 1-43.

Akman, Y., and O. Ketenoğlu, 1978. The phytosocilogical investigation of Köroğlu Mountain. Communication, De La Faculté des Sciences de L'univesité d'Ankara, Serie C, 22/1: 1-24.

Akman, Y., 1979. Contribution a l'etüde de la flore des montagnes d'Ayaş. Communication, De La Faculté des Sciences de L'univesité d'Ankara, Serie C, 23/2: 2.

Akman, Y., and E. Yurdakulol, 1981. Contribution to the Flora of Semen Mountains (Bolu). Communication, De La Faculté des Sciences de L'univesité d'Ankara, Serie C, 24/2: 1-43.

Aksoy, N., 2001. The flora of Karakiriş Mountain (Seben-Nallıhan). Master Thesis, Graduate School of İstanbul University, İstanbul, Turkey.

Allen, T. F. H., and T. B. Starr, 1982. Hierarchy: Perspectives for ecological complexity. University of Chicago Press, Chicago.

Alvera, B., Puigdefabregas, J., and Diez, C., 1997. Response of high mountain landscape to topographic variables: Central Pyrenees. Landscape Ecology, 12: 95–116.

Anon., 1990. Guidelines for Soil Description. Food and Agriculture Organization (FAO), Rome, Italy.

Anon., 1992. Agenda 21: Rio Declaration on Environment and Development. The United Nations Conference on Environment and Development, 3-14 June 1992. Rio de Janerio. U.N. Publication (http://www.igc.apc.org/habitat/agenda21/rio-dec.html).

Anon., 1994. Field methods for vegetation mapping” NBS/NPS Vegetation Mapping Program, Report Prepared For: United States Department of Interior National Biological Survey and National Park Service. Prepared by The Nature Conservancy & Environmental System Research Institute (http://biology.usgs.gov/npsveg/fieldmethods/toc.html).

Anon., 1996. Report on the state of the world's plant genetic resources for food and agriculture. Prepared for the International Technical Conference on Plant Genetic Resources, Leipzig, Germany, 17-23 June. Food and Agriculture Organization (FAO), Rome, Italy.

Anon., 1997. Antalya declaration. XI. world forestry congress forestry for sustainable development, 13-22 October. Antalya, Turkey (www.wasmeinfo.org/icsme2001.htm).

Anon., 1999. Benefits of biodiversity. Council for Agricultural Science and Technology. Task Force Report, U.S.A., No:133.

Anon., 2002. Living planet report 2002. Edited by Jonathan Loh. WWF–World Wide Fund For Nature.

Apan, A. A., 1997. Land cover mapping for tropical forest rehabilitation planning using remotely-sent data. International Journal of Remote Sensing, 18:1029–1049.

Ateşalp, M. 1976. Soil and water sampling for analysis. Soil and Fertilizer Research Institute, General No: 68, Farmer Pub. No:3.

Atkinson, P. M., 1997. Selecting the spatial resolution of airborne MSS imagery for smallscale agricultural mapping. International Journal of Remote Sensing, 18:1903–1917.

Atkinson, P. M., and Curran, P. J., 1997. Choosing an appropriate spatial resolution for remote sensing investigations. Photogrammetric Engineering and Remote Sensing, 63:1345–1351.

Atkinson, M. D., Jervis, A. P., and Sangha, R. S., 1997. Discrimination between *Betula pendula*, *Betula pubescens* and their hybrids using near infrared reflectance spectroscopy. Canadian Journal of Forest Research, 27:1896–1900.

Barbour, M. G., J. H. Burk, and W. D. Pitts, 1987. Terrestrial plant ecology. Benjamin Cumming Publishing Company, Inc. California.

Barrio, G. D., Alvera, B., Puigdefabregas, J., and Diez, C., 1997. Response of high mountain landscape to topographic variables: Central Pyrenees. Landscape Ecology, 12: 95–116.

Baugh, W., L. Klinger, A. Guenther, and C. Geron, 2001. Measurement of oak tree density with LANDSAT TM data for estimating biogenic isoprene emissions in Tennessee, USA. International Journal of Remote Sensing, 22/14: 2793–2810.

Biging, G. S., Dobbertin, M., and Murphy, E. C., 1995. A test of airborne multispectral videography for assessing the accuracy of wildlife habitat maps. Canadian Journal of Remote Sensing, 21:357–366.

Blaszczynski, J. S., 1997. Landform characterization with geographic information systems. Photogrammetric Engineering and Remote Sensing, 63: 183–191.

Boissier, E., 1867-1888. Flora orientalis Volume: 1-5, Suplement by Buser, R. Geneve, Switzerland.

Bolstad, P. V., W. Swank, and J. Vose, 1998. Predicting southern Appalachian overstory vegetation with digital terrain data. Landscape Ecology, 13: 271–283.

Bowker, G. C., 2000. Mapping biodiversity. International Journal of Geographical Information Science, 14/ 8: 739-754.

Braun-Blanquet, J., 1932. Plant sociology: the study of plant communities. McGraw-Hill, New York.

Breininger, D. R., V. L. Larson, B. W. Duncan, R. A. Smith, D. M. Oddy, and M. F. Goodchild, 1995. Landscape patterns of Florida Scrub Jay habitat use and demographic success. Conservation Biology, 9: 1442–1453.

Brown, W. M., M. Astley, T. Baker, and H. Mitasova, 1995. GRASS as an Integrated GIS and visualization system for spatio-temporal modeling. Twelfth International Symposium on Computer- Assisted Cartography, Charlotte, North Carolina 4: 89-99.

Bunge, J. and Fitzpatrick, M., 1993. Estimating the number of species; a review. Am. Statist. Assoc. 88: 364-373.

Burrough, P. A., 1986. Principles of geographical information systems for land resources assessment. Monographs on Soil and Resources Survey No:12, Oxford Science Publications, London.

Buttenfield, B. P. and J. H. Ganter, 1990. A visualization and GIS: what should we see? What might we miss? In Proceedings of the 4th International Symposium on Spatial Data Handling, Edited by K. Brassel and H. Kishimoto. Zurich, Switzerland, 1/2: 307-316.

Campbell, J., 1993. Map use and analysis. Wm.C. Brown Publishers, Dubuque, Iowa.

Causton, D. R., 1988. Introduction to vegetation analysis. Academic Division of Unwin Hyman Ltd. London.

Chou, Y. H., 1997. Exploring spatial analysis in geographic information systems. Onword Press. Santa Fe, New Mexico, U.S.A. pp. 107-110.

Clench, H., 1979. How to make regional lists of butterflies: some thoughts. Journal Lepidopterists Society, 33: 216-231.

Cohen, W. B., T. A. Spies, and G. A. Bradshaw, 1990. Semivariograms of digital imagery for analysis of conifer canopy structure. Remote Sensing of Environment, 34: 67-178.

Cohen, W. B., T. K. Maiersperger, T. A. Spies, and D. R. Oetter, 2001. Modeling forest cover attributes as continuous variables in a regional context with Thematic Mapper data. International Journal of Remote Sensing, 22/12: 2279–2310.

Cole, N. H. A., 1994. Conserving Africa's biodiversity: issues, impacts, and priorities. In Widening Perspectives on Biodiverstiy, Edited by Krattiger, A.F., McNeely, J.A., Lesser W.H., Miller K.R., Hill, Y. St., and Senanayake, R., IUCN, The World Conservation Union, The Burlington Press Ltd. Cambridge, UK.

Congalton, R. 1991. A review of assessing the accuracy of classifications of remotely sensed data. Remote Sensing of Environment, 37: 35-46.

Connell, J.H., 1978. Diversity in tropical rain forests and coral reefs. Science, 199: 1302-1303.

Constanza, R., F. H. Sklar, and M. L. White, 1990. Modeling coastal landscape dynamics. BioScience, 40: 91-107.

Constanza, R. and T. Maxwell, 1991. Spatial ecosystem modeling using parallel processors. Ecological Modelling. 58: 159-183.

Cornell, H.V., 1999. Unsaturation and regional influences on species richness in ecological communities: a review of the evidence. Ecoscience, 6: 303-315.

Darlington, R. B., W. Sharon, H. Walberg (1973). Canonical variate analysis and related techniques. Review of Educational Research, 453-454.

Davis, P. H. and V. H. Heywood, 1965. Principles of angiosperm taxonomy. Von Nostrand, New York.

Davis, P.H., 1965-1988. Flora of Turkey and the East Aegean Islands. Vol. 1-10, Edinburgh University Press. Edinburgh.

Davis, P.H., 1971. Distribution patterns in Anatolia with particular reference to endemism. Plant Life of South-West Asia. Edited by Davis Harper and Hedge. Botanical Society of Edinburgh, Great Britain.

Davis, P.H. and J. Edmondson, 1974. The flora of Turkey past present and future. Condollea 30: 331-351.

Debinski, D. M., and P. S. Humphrey, 1997. An integrated approach to biological diversity assessment. Natural Areas Journal, 17:355-365.

Demiriz, H., 1993. The bibliography of Turkey’s flora and vegetation. TBAG. State of Planning Organization, Ankara, Turkey.

Diamond, J., 1997. Location, location, location: The first farmers. Science 278:1243-1244.

Dietz, H., and Steinlein, T., 1996. Determination of plant species cover by means of image analysis. Journal of Vegetation Science, 7:131–136.

Dobrowski, S., and J. Greenberg, 2003. Field sampling methods, error, and limitations in remote sensing. Environmental Remote Sensing Lecture Notes, University of California, Davis (http://www.cstars.ucdavis.edu/classes/ers186-w03/index.html.).

Doğan, E., 2000. The flora of Nallıhan Bird Sanctuary. Master Thesis, Graduate School of Gazi University, Ankara, Turkey.

Doğan, H. M., 1998. Visualizing spatio-temporal mesquite variation on desert grassland under different grazing management applications. Master thesis, Graduate School of New Mexico State University, Master of Applied Geography, New Mexico, Las Cruces, U.S.A.

Doğan, H. M., 2001. The applications of Geographic Information Systems and Remote Sensing in Agriculture. Global Space Activities and Potential in Turkey, 1st International Space Symposium, 30-31 May. The General Commandership of Air Force, Proceedings, Ankara, Turkey, pp. III-265-271.

Doğan, H. M., A. Mermer, and E. Ünal, 2000. Monitoring agricultural lands of Turkey by using satellite images (NOAA) and normalized difference vegetative index (NDVI) values. Journal of Agriculture and Village, Ministry of Agriculture and Rural Affairs, Ankara, Turkey, 135: 38-41.

Doğan, H. M., A. Mermer, Ö. Urla, and E. Ünal, 2000. Investigating the relationship between the topography and rangeland distribution by using geographic information system (GIS) and global positioning system (GPS) in Kargalı and Gölköy pilot areas. AGROENVIRON 2000, 2nd International Symposium on New Technologies for Environment and Monitoring and Agro-Applications, 18-20 October. Proceedings, Tekirdağ, Turkey, pp. 503-505.

Doğan, H. M., 2000a. Geographic information systems and remote sensing in agriculture. Cine-Dergi, Adana, Turkey, 3/23: 20-22.

Doğan, H. M., 2000b. Visualizing spatio-temporal mesquite variation on desert grasslands. Proceedings of International Symposium on Desertification June 2000. Proceedings, General Directorate of Rural Services, Konya, Turkey, pp. 283-288.

Doğan, M., 1998. Turkey's flora, protecting plant biodiversity and the role of flora researches. The Ministry of Environment, Journal of Environment and Human, Ankara, Turkey, 41: 46-49.

Doğan, M., 2002. Efforts in sustainable development in Turkey. The Ministry of Environment, Journal of Environment and Human, Ankara, Turkey, 55: 42-49.

Doğan, M. and G. Akaydın, 2002. A new species of *Acontholimon* Boiss. (Plumbaginaceae) from Ankara, Turkey. Botanical Journal of the Linnean Society 140: 443-448.

Erik, S., 1976. Searching the plant sociology of transition zones between forest and steppe around Karagöl. Journal of Forest Research Inst. 22/2.

ERDAS, 1995. ERDAS Imagine, imagine production tour guides. ERDAS, Inc., Atlanta, Georgia.

ERDAS, 1997. ERDAS field guide, fourth edition. ERDAS, Inc., Atlanta, Georgia.

ESRI, 1994. PC Arc/Info, command references. Redlands, CA.

ESRI, 1996. Arc/View spatial analysis, Advanced Spatial Analysis Using Raster and Vector Data. San Diego, CA.

ESRI, 1997. Understanding GIS, The Arc/Info method, version 7.1 for Unix and Windows NT. John Viley & Sons Inc. New York.

Elichirigoity, F., 1999, Planet management: limits to growth, computer simulations, and the emergence of global spaces. Northwestern University Press.

Everitt, J. H., D. E. Escobar, C. H. Blasquez, M. A. Hussey, and P. R. Nixon, 1987. Evaluation of the mid infrared (1.45 to 2.0 mm) with a black-and-white infrared video camera. Photogrammetric Engineering and Remote Sensing, 52: 1655–1660.

Gorsuch, R. 1983. Factor analysis. Lawrence Erlbaum Associates.

Gotelli, N. J. and R. K. Colwell, 2001. Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. Ecology Letters, 4: 379-391.

Gray, B., 1980. Popper and the 7th Approximation: the problem of taxonomy. Dialectica, 34: 129–153.

Grieser J., 2002. Local climate estimator based on FAOCLIM ver. 2 database of the Agrometeorology Group. FAO Environment and Natural Resources Service SDRN.

Grimm, L. G., 1993. Statistical applications for the behavioral sciences. John Wiley, New York.

Grossman, D., K. L. Goodin, X. Li, and C. Swinehart, 2003. USGS - NPS Vegetation Mapping Program. Contributed by Langendoen D. F., M. Anderson, L. Sneddon, A. Weakley, P. Bourgeron, M. Reid, K. Metzler, R. Vaughan, C. Convis, A. Curtis, M.

Goodchild, and F. Davis. The Nature Conservancy, and Environmental Systems Research Institute, Arlington, Virginia (http://biology.usgs.gov/npsveg/fieldmethods/toc.html).

Güner, M. B., 2000.The Flora of the Doğandede Tepe and it's around (Beypazarı-Ankara.). Master Thesis, Graduate School of Gazi University, Ankara, Turkey.

Güner, A., N. Özhatay, T. Ekim, K. Hüsnü, and C. Baser, 2000. Flora of Turkey and the East Aegean Islands Volume:11. Edinburgh University Press.

Habb, F. J., 1995. Computer Generation and Rendering of Terrain. Masters of Science Thesis, University of Neveda, Las Vegas, Nv.

Hanski, J., 1999. Metapopulation Ecology. Oxford University Press.

Harman, H. H. (1976). Modern factor analysis (3rd ed., rev.). Chicago: University of Chicago Press.

Helmer E. H., S. Brown, and W. B. Cohen, 2000. Mapping Montane tropical forest successional stage and land use with multi-date Landsat imagery. International Journal of Remote Sensing, 21/11: 2163–2183.

Hoersch, B., G. Braun, and U. Schmidt, 2002. Relation between landform and vegetation in alpine regions of Wallis, Switzerland. A multiscale remote sensing and GIS approach. Computers, Environment and Urban Systems, 26: 113–139.

Holt, R.D., G.R. Robinson, and M.S. Gaines, 1995. Vegetation dynamics in an experimentally fragmented landscape. Ecology, 76: 1610-1624.

Hutchinson, G.E. 1957. Concluding remarks. Cold Spring Symposia on Quantitative Biology 22: 415-427.

Hyppanen, H., 1996. Spatial autocorrelation and optimal spatial resolution of optical remote sensing data in boreal forest environment. International Journal of Remote Sensing, 17: 3441-3452.

Innes, J. L., and B. Koch, 1998. Forest biodiversity and its assessment by remote sensing. Global Ecology and Biogeography Letters, 7: 397-419.

Journet, D., 1991. Ecological theories as cultural narratives: F. E. Clements's and H.A. Gleason's 'Stories' of Community Succession. Written Communication, 8, 446–472.

Kalliola, R., and K. Syrjänen, 1991. To what extent are vegetation types visible in satellite imagery? Annales Botanici Fennici, 28: 45–57.

Kirby, K. J., T. Bines, A. Burn, P. Pitkin, and I. Smith, 1986. Seasonal and observer differences in vascular plant records from British woodlands. Journal of Ecology, 74: 123-132.

Konukçu, M., 1998. Statistical profile of Turkish forestry. State Planning Organization, Ankara, Turkey (http://ekutup.dpt.gov.tr/ormancil/konukcum/forest.html).

Krattiger, A. F., 1994. Preface. In widening perspectives on biodiversity, IUCN, The World Conservation Union, International Academy of the Environment, Geneva Switzerland.

Lillesand, T. M., R.W. Kiefer, 1994. Remote sensing and image interpretation. Third Edition, John Wiley & Sons, Inc. Canada.

Lobo, A., O. Chic, and A. Casterad, 1996. Classification of Mediterranean crops with multisensor data: per-pixel versus per-object statistics and image segmentation. International Journal of Remote Sensing, 17: 2385–2400.

Lubchenco, J., A. M. Olson, L. B. Brubaker, S. R. Carpenter, M. M. Holland, S. P. Hubell, S. A. Levin, J. A. MacMahon, P. A. Matson, J. M. Melillo, H. A. Mooney, C. H. Peterson, H. R. Pullam, L. A. Real, P. J. Regal, and P. G. Risser, 1991. The sustainable biosphere initiative: an ecological research agenda. Ecology, 72: 371-412.

Ma, Z., M. M. Hart, R. L. Redmond, 2001. Mapping vegetation across large geographic areas: integration of remote sensing and GIS to classify multisource data. Photogrammetric Engineering & Remote Sensing, 67/3: 295-307.

Magurran, A. E., 1981. Biological diversity and woodland management. Unpublished Phd. Thesis, New University of Ulster (In Magurran, A. E., 1988. Ecological diversity and its measurement. Princeton University Press. New Jersey).

Magurran, A. E., 1988. Ecological diversity and its measurement. Princeton University Press. New Jersey, U.S.A.

Mausel, P., Y. Wu, Y. Li, E. F. Moran, and E. S. Brondizio, 1993. Spectral identification of successional stages following deforestation in the Amazon. Geocarto International, 4: 61-71.

May, R.M., 1988. How many species on earth? Science, 241: 1441-1449.

May, R. M., 1999. The dimensions of life on earth. Proceedings of National Biodiversity Forum. National Academy Press, Washington, D.C.

Mayoraz, R., C.E. Mann, and A. Parriaux, 1992. Three-modeling of complex geologic structures; new development tools for creating 3-D volumes. Computer Modeling of Geographic Surfaces. AAPG Computer Applications in Geology. I. American Association of Petrolium Geologist. Tulsa, Ok. 261-269.

McGraw, J. B., T. A. Warner, T. L. Key, and W. L. Lamar, 1998. High spatial resolution remote sensing of forest trees. Trends in Ecology and Environment, 13: 300-301.

McGrew, J. C. Jr., C. B. Monroe, 1993. Statistical problem solving in geography. Wm. C. Brown Publishers, Dubuque.

McNeely. J. A., 1994. Critical issues in the implementation of the convention on biological diversity. In Widening Perspectives on Biodiversity, IUCN, The World Conservation Union, International Academy of the Environment, Geneva Switzerland.

Meyer, P., K. Staenz, and K. I. Itten, 1996. Semi-automated procedures for tree species identification in high spatial resolution data from digitised colour infrared-aerial photography. ISPRS Journal of Photogrammetry and Remote Sensing, 51: 5-16.

Molles, M.C. Jr., 1999. Ecology, concepts and applications. Internatioal Edition, WCB/McGraw Hill Companies Inc. Boston, U.S.A.

Morrison, DF. 1990. Multivariate Statistical Methods. McGraw-Hill, New York.

Moss, H. and L. Guarino, 1995. Gathering and recording data in the field. In collecting plant diversity, technical guidelines. Edited by Guarino, L., Rao V. R., Reid R., CAB International on behalf of the IPGRI, FAO, IUCN, and UNEP

Muehrcke, P.C., J. O. Muehrcke, 1992. Landform portrayal: map use; reading, analysis interpretation. JP. Publications, Madison, WI.

Nagendra, H., and M. Gadgil, 1999. Biodiversity assessment at multiple scales: linking remotely sensed data with field information. Proceedings of the National Academy of Sciences, USA, 96: 9154–9158.

Nagendra, H. 2001. Using remote sensing to assess biodiversity. International Journal of Remote Sensing, 22/12: 2377–2400.

Nixon, P. R., D. E. Escobar, and R. M. Menges, 1985. Use of a multi-band video system for quick assessment of vegetal condition and discrimination of plant species. Remote Sensing of Environment, 17: 203-208.

Norušis, M. J., 1993. SPSS for Windows base system: user’s guide. Release 6.0. SPSS Inc. Chicago.

Noss, R. F., 1996. Ecosystems as conservation targets. Trends in Ecology and Evolution, 11: 351.

Novotny, V. and Y. Basset, 2000. Rare species in communities of tropical insect herbivores: pondering the mystery of singletons. Oikos, 89: 564-572.

Ocak, A. and S. Tokur, 2000. The flora of Gülümbe Dağı (Bilecik-Turkey). Türk. J. Bot. 24: 121-142.

O’Neill, R. V., C. T. Hunsaker, S. P. Timmins, B. L. Jackson, K. B. Jones, K. H. Riitters, and J. D. Wickham, 1996. Scale problems in reporting landscape pattern at the regional scale. Landscape Ecology, 11:169–180.

Orloci, L., 1978. Multivariate analysis in vegetation research. Second Editon. Dr. W. Junk Publishers. The Hague, Netherlands.

Pimm, S. L., G. J. Russell, J. L. Gittleman, and T. M. Brooks, 1995. The future of biodiversity. Science 269:347-350.

Preston, F. W., 1948. The Commoness, and rarity of species. Ecology, 29: 254-283.

Preston, F. W., 1962a. The canonical distribution of commonness and rarity: part I. Ecology 43:185-215.

Preston, F. W., 1962b. The Canonical Distribution of Commonness and Rarity: Part II. Ecology 43: 410-432.

Price, J. C., 1994. How unique are spectral signatures? Remote Sensing of Environment, 49: 181–186.

Reid, R., 1980. The collection and use of climatic data in pasture plant introduction. In: Clements, R.J. and D.G.Cameron (eds) Collecting and Testing Tropical Forage Plants. CSIRO, Canberra.

Rexstad, E.A., D. D. Miller, C. H. Flather, E. M. Anderson, J. W. Hupp, and D. R. Anderson, 1988. Questionable multivariate statistical inference in wildlife habitat and community studies. J. Wildlife Management 52:794-798.

Robbins, C.S., J.R. Sauer, R.S. Greenberg, and S. Droege, 1989. Population declines in North American birds that migrate to the neotropics. Proc. Natl Acad. Set. (USA), 86: 7658-7662.

Roughgarden, J., S. W. Running, and P. A. Matson, 1991. What does remote sensing do for ecology? Ecology, 72: 1918-1922.

Sader, S. A., T. A. Stone, and A. T. Joyce, 1990. Remote sensing of tropical forests: an overview of research and applications using non-photographic sensors. Photogrammetric Engineering and Remote Sensing, 56: 1343-1351.

Salisbury, J. W., and N. M. Milton, 1987. Preliminary measurements of spectral signatures of tropical and temperate plants in the thermal infrared. Proceedings of the 21st International Symposium on Remote Sensing of Environment, Ann Arbor, Michigan, 26-30 October 1987. Environmental Research Institute of Michigan, pp. 131-143.

Simberloff, D., 1986. Are we on the verge of a mass extinction in tropical rain forests? In: Dynamics of Extinction, ed. Elliott D.K., John Wiley & Sons, New York, pp. 165-180.

Simmons, M. A., V. I. Cullinan, and J. M. Thomas, 1992. Satellite imagery as a tool to evaluate ecological scale. Landscape Ecology, 7: 77-85.

Sokal, R. R., and F. J. Rohlf, 1995. Biometry the principles and practice of statistics in biological research. Third edition, W.H. Freeman and Company, New York, pp. 609.

Solomon Dobrowski & Jonathan Greenberg. 2003. Field sampling methods, error, and limitations in remote sensing. Environmental Remote Sensing Lecture Notes, University of California, Davis (www.cstars.ucdavis.edu/classes/ers186).

SPSS., 1993. SPSS 6.0 for Windows. SPSS Inc., Chicago.

SPSS., 1997. SYSTAT 7.0 for Windows: Unparalleled research quality statistics and graphics. SPSS Inc., Chicago.

SPSS., 2001. SPSS 11.0 for Windows. SPSS Inc., Chicago.

Stenback, J. M., and R. G. Congalton, 1990. Using thematic mapper imagery to examine forest understorey. Photogrammetric Engineering and Remote Sensing, 56: 1285-1290.

Stevens, G. C., 1989. The latitudinal gradient in geographical range: how so many species coexist in the tropics. American Naturalist, 133: 240-256.

Stoms, D. M., and J. E. Estes, 1993. A remote sensing research agenda for mapping and monitoring biodiversity. International Journal of Remote Sensing, 14: 1839-1860.

Stone, L., 1995. Biodiversity and habitat destruction: a comparative study for model forest and coral reef ecosystems. Proc. R, Soc. London, 261/B: 381-388.

St-Onge, B. A., and F. Cavavas, 1997. Automated forest structure mapping from high resolution imagery based on directional semi-variogram estimates. Remote Sensing of Environment, 61: 82-95.

Tappeiner, U., E. Tasser, and G. Tappeiner, 1998. Modeling vegetation patterns using natural and anthropogenic influence factors: preliminary experience with a GIS based model applied to an Alpine area. Ecological Modeling, 113: 225-237.

Taylor, J. E., 1993. Factors causing variation in reflectance measurements from bracken in eastern Australia. Remote Sensing of Environment, 43: 217-229.

Taylor, D. R. F., 1994. Cartographic visualization and spatial data handling. Sixth International Symposium on Spatial Data Handling, Edinburgh, Scotland, 1/2: 16-28.

Tilman, D., R.M. May, C. L. Lehman, and M. Nowack, 1994. Habitat destruction and the extinction debt. Nature, 371 : 65-66.

Trimble,, 1993. A global navigation system in the palm of your hand. Users manual. Trimble Navigation Ltd.Texas, Austin.

Trimble., 1994. Scout it out (SIO) special PC software for the Trimble Scoutmaster GPS receiver . Users manual, Trimble Navigation Ltd., Texas, Austin.

Trisurat, Y., A.. Eiumnoh, S. Murai, M. Z. Hussain, R. P. Shrestha, 2000. Improvement of tropical vegetation mapping using a remote sensing technique: a case of Khao Yai National Park, Thailand. International Journal of Remote Sensing, 21/10: 2031-2042.

Türe, C., S. Tokur. The flora of the forest series of Yirce-Bürmece-Kömürsu and Muratdere (Bilecik-Bursa, Turkey). Turk J. Bot, 24: 47-66.

Türker, H., 1990. The flora of the region among the towns of Ayaş, Güdül, Beypazarı and Polatlı. Master Thesis, Graduate School of Gazi University, Ankara.

UNESCO, 1973. International Classification and mapping of vegetation. Ecology and Conservation. Unesco, Paris.

Ülgen, N., Yurtsever, N. 1995. Fertilizers and fertilizing guide of Turkey. Soil and Fertilizer Research Institute, General No: 209, Technical Pub. No: 66.

Verbyla, D. L., 1995. Satellite remote sensing of natural resources, 1st ed., Lewis Publishers, New York.

Vitousek, P. M., H. A. Mooney, J. Lubchenco, and J. M. Melillo, 1997. Human domination of earth's ecosystems. Science 277:494-499.

Watson, D.F., 1992. Contouring. A Guide to the analysis and display of spatial data. Pergamon Press. Oxford, England.

Weishampel, J. F., G. Sun, J. Ranson, K. D. Lejeune, and H. H. Shigart, 1994. Forest textural properties from simulated microwave backscatter: the influence of spatial resolution. Remote Sensing of Environment, 47: 120-131.

Whittaker, R. H., and W. A. Niering., 1975. Vegetation of the Santa Catalina Mountains, Arizona: V. Biomass, production and diversity along the elevation gradient. Ecology, 56: 771-790.

Wiegleb, G., 2002. Biodiversity - A Central European view. BTU Cottbus, Dept. of General Ecology. (http://www.erm.tu-cottbus.de/Board/scripts/Wiegleb/index.html).

William E. Doolittle., 2002. Vegetation lecture notes, Texas University (www.uts.cc.utexas.edu/~wd/courses/373F/notes/lec19veg.html)

Williams, R.H., W. Z. Donald, D. Z. Bruno, R. Donald. 2003. Charles Spearman: British behavioral scientist. Human Nature Review, 3:114-118 (http://human-nature.com/nibbs/03/spearman.html)

Wolter, P. T., D. J. Mladenoff, G. R. Host, and T. R. Crow, 1995. Improved forest classification in the northern lake states using multi-temporal LANDSAT imagery. Photogrammetric Engineering and Remote Sensing, 61: 1129-1143.

Yavuz, H., 1992. Flora of the region between Ayaş, Çanıllı Village and Kazan. Master Thesis, Graduate School of Ankara University, Ankara.

Yılmaz (Özkan), R., 1996. The flora of Sarıçal Mountain (Nallıhan-Ankara). Master Thesis, Graduate School of Gazi University, Ankara

Zohary, M., 1971. Phytogeographical foundation of the Middle East; in plant life of South-West Asia. Edited by David, Harper, Hedge. The botanical Science of Edinburgh, pp. 43-52.

APPENDIX A

LIST OF IDENTIFIED SPECIES IN THE STUDY AREA

NO	SPECIES NAME	FAMILY
1	*Alhagi pseudalhagi* (Bieb.) Desv.	LEGUMINOSAE
2	*Anthyllis vulneraria* L. subsp. *boissieri* (Sag.) Bornm.	LEGUMINOSAE
3	*Astragalus angustifolius* Lam. subsp. *angustifolius*	LEGUMINOSAE
4	*Astragalus densifolius* Lam.	LEGUMINOSAE
5	*Astragalus glycyphyllos* L. subsp. *glycyphylloides* (DC.) Matthews	LEGUMINOSAE
6	*Astragalus lycius* Boiss.	LEGUMINOSAE
7	*Astragalus macrocephalus* Willd. subsp. *Macrocephalus*	LEGUMINOSAE
8	*Astragalus microcephalus* Willd.	LEGUMINOSAE
9	*Astragalus micropterus* Fischer	LEGUMINOSAE
10	*Astragalus squalidus* Boiss.& Noë *****	LEGUMINOSAE
11	*Astragalus trichostigma* Bunge *****	LEGUMINOSAE
12	*Chamaecytisus pygmaeus* (Willd.) Rothm.	LEGUMINOSAE
13	*Cicer pinnatifidum* Jaub.& Spach	LEGUMINOSAE
14	*Conorilla varia* L. subsp. *Varia*	LEGUMINOSAE
15	*Dorycnium pentaphyllum* Scop. subsp. *anatolicum* (Boiss.) Gams	LEGUMINOSAE
16	*Hedysarum varium* Willd.	LEGUMINOSAE
17	*Lathyrus aureus* (Stev.) Brandza	LEGUMINOSAE
18	*Lotus aegaeus* (Gris.) Boiss.	LEGUMINOSAE
19	*Lotus corniculatus* L. var. *corniculatus*	LEGUMINOSAE
20	*Lotus corniculatus* L. var. *tenuifolius* L.	LEGUMINOSAE
21	*Medicago polymorpha* L. var. *vulgaris* (Benth.) Shinners	LEGUMINOSAE
22	*Medicago sativa* L. subsp. *Sativa*	LEGUMINOSAE
23	*Medicago varia* Martyn	LEGUMINOSAE
24	*Melilotus alba* Desr.	LEGUMINOSAE
25	*Melilotus officinalis* (L.) Desr.	LEGUMINOSAE
26	*Onobrychis argyrea* Boiss. Subsp. *argyrea*	LEGUMINOSAE
27	*Onobrychis armena* Boiss.& Huet.	LEGUMINOSAE
*** endemic species**		

NO	SPECIES NAME	FAMILY
28	*Onobrychis hypargyrea* Boiss.	LEGUMINOSAE
29	*Ononis adenotricha* Boiss. var. *adenotricha*	LEGUMINOSAE
30	*Ononis spinosa* L. subsp. *Leiosperma* (Boiss.) Širj.	LEGUMINOSAE
31	*Pisum sativum* L. subsp. *Elatius* var. *elatius*	LEGUMINOSAE
32	*Trifolium arvense* L. var. *arvense*	LEGUMINOSAE
33	*Trifolium barbulatum* (Freyn & Sint.) Zoh. *****	LEGUMINOSAE
34	*Trifolium repens* L. var. *repens*	LEGUMINOSAE
35	*Vicia cracca* L. subsp. *Stenophylla* Vel.	LEGUMINOSAE
36	*Vicia grandiflora* Scop. var. *grandiflora*	LEGUMINOSAE
37	*Vicia narborensis* L. var. *narborensis*	LEGUMINOSAE
30	*Ononis spinosa* L. subsp. *leiosperma* (Boiss.) Širj.	LEGUMINOSAE
31	*Pisum sativum* L. subsp. *elatius* var. *elatius*	LEGUMINOSAE
32	*Trifolium arvense* L. var. *arvense*	LEGUMINOSAE
33	*Trifolium barbulatum* (Freyn & Sint.) Zoh.*****	LEGUMINOSAE
34	*Trifolium repens* L. var. *repens*	LEGUMINOSAE
35	*Vicia cracca* L. subsp. *stenophylla* Vel.	LEGUMINOSAE
36	*Vicia grandiflora* Scop. var. *grandiflora*	LEGUMINOSAE
37	*Vicia narborensis* L. var. *narborensis*	LEGUMINOSAE
38	*Achillea biebersteinii* Afan.	COMPOSITAE
39	*Achillea setacea* Waldst.& Kit.	COMPOSITAE
40	*Acroptilon repens* (L.) DC.	COMPOSITAE
41	*Anthemis tinctoria* L. var. *discoidea* (All.) DC.	COMPOSITAE
42	*Cardopodium corymbosum* (L.) Pers.	COMPOSITAE
43	*Carlina corymbosa* L.	COMPOSITAE
44	*Centaurea deprassa* Bieb.	COMPOSITAE
45	*Centaurea solstitialis* L. subsp. *solstitialis*	COMPOSITAE
46	*Centaurea triumfettii* All.	COMPOSITAE
47	*Centaurea urvillei* DC. subsp. *Urvillei* *****	COMPOSITAE
48	*Centaurea virgata* Lam.	COMPOSITAE
49	*Chardinia orientalis* (L.) O. Kuntze	COMPOSITAE
50	*Chondrilla juncea* L. var. *juncea*	COMPOSITAE
51	*Cichorium intybus* L.	COMPOSITAE
52	*Cirsium arvense* (L.) Scop. subsp. *vestitum*	COMPOSITAE
53	*Cirsium hypoleucum* DC.	COMPOSITAE
54	*Crepis sancta* (L.) Babcock	COMPOSITAE
55	*Doronicum orientale* Hoffm.	COMPOSITAE
56	*Echinops ritro* L.	COMPOSITAE
57	*Inula oculus-christi* L.	COMPOSITAE
58	*Lactuca serriola* L.	COMPOSITAE
*** endemic species**		

NO	SPECIES NAME	FAMILY
59	*Leontodon asperrimus* (Willd.) J. Ball.	COMPOSITAE
60	*Petasites hybridus* (L.) Gaertner	COMPOSITAE
61	*Pilosella echioides* (Lumn.) C.H.& F.W.Schultz subsp. *procera* (Fries) Sell & West	COMPOSITAE
62	*Pilosella hoppeana* (Schultes) C. H. & F.W. Schultz subsp. *testimonialis* (NP.) Sell &West	COMPOSITAE
63	*Scorzonera cana* (C.A.Meyer) Hoffm.	COMPOSITAE
64	*Scorzonera laciniata* L.	COMPOSITAE
65	*Senecio vernalis* Waldst.& Kit.	COMPOSITAE
66	*Sonchus asper* L. Hill subsp. *glaucescens* (Jordan) Ball.	COMPOSITAE
67	*Tanacetum poteriifolium* (Ledeb.)	COMPOSITAE
68	*Tanacetum vulgare* L.	COMPOSITAE
69	*Taraxacum seronitum* (Waldst.& Kit.) Poiret in Lam.	COMPOSITAE
70	*Tragopogon latifolius* Boiss. var. *angustifolius* Boiss.	COMPOSITAE
71	*Xeranthemum annuum* L.	COMPOSITAE
72	*Acinos rotundifolius* Pers.	LABIATAE
73	*Ajuga chamaepitys* (L.) Schreber, subsp. *chia* (Schreber) Arcangeli, var. *chia*	LABIATAE
74	*Lamium macradon* Boiss.& Huet	LABIATAE
75	*Marrubium parviflorum* Fisch.& Mey. subsp. *oligodon* (Boiss.) Seybold *****	LABIATAE
76	*Mentha spicata* L. subsp. *tomentosa* (Briq.) Harley	LABIATAE
77	*Nepeta nuda* L. subsp.*albiflora* (Boiss.) Gams	LABIATAE
78	*Phlomis armeniaca* Willd. *****	LABIATAE
79	*Phlomis nissolii* L.	LABIATAE
80	*Prunella vulgaris* L.	LABIATAE
81	*Salvia aethiopis* L.	LABIATAE
82	*Salvia hypargeia* Fisch.& Mey.	LABIATAE
83	*Salvia sclarea* L.	LABIATAE
84	*Salvia tomentosa* Miller (Syn: S. *grandiflora* Etl.)	LABIATAE
85	*Salvia verticillata* L. subsp. *amasiaca* (Freyn & Bornm.) Bornm.	LABIATAE
86	*Salvia viridis* L.	LABIATAE
87	*Scutellaria orientalis* L. subsp. *macrostegia* (Hausskn. ex Bornm.) Edmondson	LABIATAE
88	*Sideriris montana* L. subsp. *montana*	LABIATAE
89	*Sideritis galatica* Bornm.	LABIATAE
90	*Stachys annua* (L.) L. subsp. *ammophila* (Boiss. & Bl.) Samuelss	LABIATAE
*** endemic species**		

NO	SPECIES NAME	FAMILY
91	*Stachys annua* (L.) L. subsp. *annua* var. *annua* *****	LABIATAE
92	*Stachys cretica* L. subsp. *anatolica* Rech. fil. *****	LABIATAE
93	*Teucrium chamaedrys* L. subsp. *chamaedrys*	LABIATAE
94	*Teucrium parviflorum* Schreber	LABIATAE
95	*Teucrium polium* L.	LABIATAE
96	*Thymus leucostomus* Hausskn.&Velen. var. *leucostomus*	LABIATAE
97	*Thymus longicaulis* C. Presl subsp. *longicaulis*	LABIATAE
98	*Thymus sipyleus* Boiss. subsp. *sipyleus*	LABIATAE
99	*Ziziphora capitata* L.	LABIATAE
100	*Cotoneaster nummularia* Fisch.& Mey.	ROSACEAE
101	*Crataegus monogyna* Jacq. subsp. *monogyna*	ROSACEAE
102	*Crataegus orientalis* Pallas ex Bieb. var. *orientalis*	ROSACEAE
103	*Crataegus tanacetifolia* (Lam.) Pers. *****	ROSACEAE
104	*Potentilla recta* L.	ROSACEAE
105	*Prunus avium* (L.) L.	ROSACEAE
106	*Prunus divaricata* Ledeb. subsp. *divaricata*	ROSACEAE
107	*Prunus spinosa* L. subsp. *dasyphylla* (Schur) Domin	ROSACEAE
108	*Pyracantha coccinea* Roemer	ROSACEAE
109	*Pyrus elaeagnifolia* Pallas subsp. *elaeagnifolia*	ROSACEAE
110	*Rosa canina* L.	ROSACEAE
111	*Rubus ideaus* L.	ROSACEAE
112	*Rubus sanctus* Schreber	ROSACEAE
113	*Sanguisorba minor* Scop. subsp. *muricata* (Spach) Briq.	ROSACEAE
114	*Sorbus umbellata* (Desf.) Fritsch var. *umbellata*	ROSACEAE
115	*Alyssum desertorum* Stapf. var. *desertorum*	CRUCIFERAE
116	*Alyssum murale* Waldst. & Kit. var. *murale*	CRUCIFERAE
117	*Alyssum sibiricum* Willd.	CRUCIFERAE
118	*Arabis nova* Vill.	CRUCIFERAE
119	*Barbera plantaginea* DC.	CRUCIFERAE
120	*Cardaria draba* (L.) Desv. subsp. *draba*	CRUCIFERAE
121	*Erysimum crassipes* Fisch.& Mey.	CRUCIFERAE
122	*Iberis taurica* DC.	CRUCIFERAE
123	*Thlaspi perfoliatum* L.	CRUCIFERAE
124	*Turritis glabra* L.	CRUCIFERAE
125	*Agropyron cristatum* (L.) Geartner, subsp: *pectinatum* (Bieb.) Tzvelev, var: *pectinatum*	GRAMINEAE
126	*Aegilops umbellulata* Zhuk.	GRAMINEAE
*** endemic species**		

NO	SPECIES NAME	FAMILY
127	*Brachypodium sylvaticum* (Hudson) P. Beauv	GRAMINEAE
128	*Dactylis glomerata* L. subsp. *glomerata*	GRAMINEAE
129	*Festuca airoides* Lam.	GRAMINEAE
130	*Festuca anatolica* Markgr.-Dannenb. subsp. *anatolica*	GRAMINEAE
131	*Festuca ilgazensis* Markgr.-Dannenb.	GRAMINEAE
132	*Poa bulbosa* L.	GRAMINEAE
133	*Stipa bromoides* (L.) Dörfler	GRAMINEAE
134	*Stipa lessingiana* Trin.& Rupr.	GRAMINEAE
135	*Allium scorodoprasum* L. subsp. *rotundum* (L.) Stearn	LILIACEAE
136	*Gagea granatellii* (Parl.) Parl.	LILIACEAE
137	*Muscari armeniacum* Leichtlin ex Baker	LILIACEAE
138	*Muscari longipes* Boiss.	LILIACEAE
139	*Muscari neglectum* Guss.	LILIACEAE
140	*Muscari tenuiflorum* Tausch	LILIACEAE
141	*Ornithogalum oligophyllum* E.D.Clarke.	LILIACEAE
142	*Ornithogalum fimbriatum* Willd.	LILIACEAE
143	*Ornithogalum umbellatum* L.	LILIACEAE
144	*Adonis flammea* Jacq.	RANUNCULACEAE
145	*Ranunculus argyreus* Boiss.	RANUNCULACEAE
146	*Ranunculus ficaria* L. subs. *ficariiformis* Rouy & Fouc.	RANUNCULACEAE
147	*Dianthus anatolicus* Boiss.	CARYOPHYLLACEAE
148	*Dianthus ancyrensis* Hausskn.& Bornm. *****	CARYOPHYLLACEAE
149	*Dianthus zonatus* Fenzl var. *zonatus*	CARYOPHYLLACEAE
150	*Herniaria glabra* L.	CARYOPHYLLACEAE
151	*Minuartia hirsuta* (Bieb.) Hand. & Mazz.	CARYOPHYLLACEAE
152	*Saponaria glutinosa* Bieb.	CARYOPHYLLACEAE
153	*Silene supina* Bieb. subsp.*pruinosa* (Boiss) Chowdh	CARYOPHYLLACEAE
154	*Astrodaucus orientalis* (L.) Drude	UMBELLIFERAE
155	*Coriandrum sativum* L.	UMBELLIFERAE
156	*Falcaria vulgaris* Bernh.	UMBELLIFERAE
157	*Laser trilobum* (L.) Borkh.	UMBELLIFERAE
158	*Malabaila secacul* Banks & Sol.	UMBELLIFERAE
159	*Turgenia latifolia* L. Hoffm.	UMBELLIFERAE
160	*Zosima absinthifolia* (Vent.) Link	UMBELLIFERAE
161	*Alkanna orientalis* (L.) Boiss. var. *orientalis*	BORAGINACEAE
162	*Anchusa leptophylla* Roemer & Schultes subsp. *leptophylla*	BORAGINACEAE
163	*Cerinthe minor* L. subsp. *minor*	BORAGINACEAE
164	*Lithospermum officinale* L.	BORAGINACEAE
*** endemic species**		

NO	SPECIES NAME	FAMILY
165	*Onosma aucheranum* DC.	BORAGINACEAE
166	*Onosma bornmuelleri* Hausskn.	BORAGINACEAE
167	*Onosma isauricum* Boiss. & Heldr. *****	BORAGINACEAE
168	*Onosma tauricum* Pallas ex Willd. var. *tauricum*	BORAGINACEAE
169	*Digitalis ferruginea* L. subsp. *ferruginea*	SCROPHULARIACEAE
170	*Digitalis orientalis* Lam.	SCROPHULARIACEAE
171	*Scrophularia scopolii* [Hoppe ex] Pers. var. *scopolii*	SCROPHULARIACEAE
172	*Verbascum cherianthifolium* Boiss var. *cheiranthifolium*. *****	SCROPHULARIACEAE
173	*Verbascum glomeratum* Boiss	SCROPHULARIACEAE
174	*Veronica chamaedrys* L.	SCROPHULARIACEAE
175	*Veronica multifida* L.	SCROPHULARIACEAE
176	*Veronica pectinata* L. var. *pectinata*	SCROPHULARIACEAE
177	*Asyneuma limonifolium* (L.) Janchen subsp. *pestalozzae* (Boiss.) Damboldt.	CAMPANULACEAE
178	*Asyneuma rigidum* (Willd.) Grossh. subsp. *rigidum*	CAMPANULACEAE
179	*Campanula glomerata* L.	CAMPANULACEAE
180	*Campanula persicifolia* L.	CAMPANULACEAE
181	*Legousia speculum-veneris* (L.) Chaix	CAMPANULACEAE
182	*Asperula stricta* Boiss. subsp. *latibracteata* (Boiss.) Ehrend.	RUBIACEAE
183	*Cruciata taurica* (Pallas ex Willd.) Ehrend.	RUBIACEAE
184	*Galium incanum* Sm. subsp. *elatius* (Boiss.) Ehrend.	RUBIACEAE
185	*Galium palustre* L.	RUBIACEAE
186	*Galium verum* subsp. *verum*	RUBIACEAE
187	*Cistus laurifolius* L.	CISTACEAE
188	*Fumana aciphylla* Boiss.	CISTACEAE
189	*Helianthemum nummularium* (L.) Miller. subsp. *ovatum* (Viv.) Schinz & Thell	CISTACEAE
190	*Juniperus communis* L. subsp. *nana*	CUPRESSACEAE
191	*Juniperus excelsa* Bieb.	CUPRESSACEAE
192	*Juniperus foetidissima* Willd.	CUPRESSACEAE
193	*Juniperus oxycedrus* L. subsp. *oxycedrus*	CUPRESSACEAE
194	*Scabiosa argentea* L.	DIPSACACEAE
195	*Quercus cerris* L. var. *cerris*	FAGACEAE
196	*Quercus pubescens* Willd.	FAGACEAE
197	*Quercus robur* L. subsp. *robur*	FAGACEAE
198	*Osyris alba* L.	SANTALACEAE
199	*Thesium billardieri* Boiss	SANTALACEAE
200	*Paronychia dudleyi* Chaudhri	ILLECEBRACEAE
*** endemic species**		

NO	SPECIES NAME	FAMILY
201	*Paronychia kurdica* Boiss. subsp. *kurdica* var. *kurdica*	ILLECEBRACEAE
202	*Iris orientalis* Miller.	IRIDACEAE
203	*Acanthus hirsutus* Boiss.	ACANTHACEAE
204	*Paliurus spina-christi* Miller	RHAMNACEAE
205	*Rhamnus thymifolius* Bornm. *****	RHAMNACEAE
206	*Geranium robertianum* L.	GERANIACEAE
207	*Geranium tuberosum* L. subsp. *tuberosum*	GERANIACEAE
208	*Linum hirsitum* L. subsp. *anatolicum* (Boiss) Hayek *****	LINACEAE
209	*Linum tenuifolium* L.	LINACEAE
210	*Fumaria cilicica* Hausskn.	PAPAVERACEAE
211	*Hypecoum procumbens* L.	PAPAVERACEAE
212	*Papaver commutatum* Fisch & Mey	PAPAVERACEAE
213	*Rhus coriaria* L.	ANACARDIACEAE
214	*Berberis crataegina* DC.	BERBERIDACEAE
215	*Berberis vulgaris* L.	BERBERIDACEAE
216	*Salsola ruthenica* Iljin	CHENOPODIACEAE
217	*Convolvulus arvensis* L.	CONVOLVULACEAE
218	*Corylus avellana* L. var. *avellana*	CORYLLACEAE
219	*Sempervivum armenum* Boiss.& Huet. var. *armenum*	CRASSULACEAE
220	*Carex flacca* Schreber subsp. *serrulata* (Biv.) Greuter	CYPERACEAE
221	*Carex ovalis* Good.	CYPERACEAE
222	*Equisetum palustre* L.	EQUISETACEAE
223	*Euphorbia macroclada* Boiss.	EUPHORBIACEAE
224	*Globularia trichosanta* Fisch.& Mey.	GLOBULARIACEAE
225	*Hypericum perforatum* L.	GUTTIFERAE
226	*Malva neglecta* Wallr.	MALVACEAE
227	*Cephalanthera rubra* (L.) L.C.M. Richard	ORCHIDACEAE
228	*Paeonia mascula* subsp. mascula	PAEONIACEAE
229	*Paeonia peregrina*	PAEONIACEAE
230	*Pinus brutia*	PINACEAE
231	*Pinus nigra* subsp. *pallasiana*	PINACEAE
232	*Acantholimon acerosum* (Willd.) Boiss	PLUMBAGINACEAE
233	*Acantholimon glumaceum* (Jaub.& Spach) Boiss.	PLUMBAGINACEAE
234	*Acantholimon reflexifolium* Bokhari	PLUMBAGINACEAE
235	*Plumbago europaea* L.	PLUMBAGINACEAE
236	*Polygala anatolica* Boiss.& Heldr.	POLYGALACEAE
237	*Urtica dioica* L.	URTICACEAE
238	*Valeriana alliariifolia* Adams	VALERIANACEAE
239	*Valerianella vesicaria* (L.) Moench	VALERIANACEAE
*** endemic species**		

APPENDIX B

PEARSON CORRELATION MATRIX OF THE DATA SET

	ELEV	ASP	SLOPE	ORGM	P2O5	K2O	CaCO3	PH	SALT
ELEV	1.000								
ASP	-0.096	1.000							
SLOPE	0.256	-0.071	1.000						
ORGM	0.297	0.030	0.196	1.000					
P2O5	0.030	-0.082	-0.163	0.109	1.000				
K2O	-0.293	-0.172	-0.051	-0.192	0.465	1.000			
CaCO3	-0.314	0.205	0.043	-0.428	-0.452	-0.095	1.000		
PH	-0.365	-0.036	-0.048	-0.453	-0.454	0.251	0.680	1.000	
SALT	-0.239	0.040	-0.016	0.019	0.035	0.078	-0.003	0.225	1.000
STR	0.317	-0.044	0.134	0.740	-0.004	-0.329	-0.435	-0.331	0.507
TEXTR	0.226	-0.053	0.171	0.485	0.003	-0.052	-0.325	-0.200	0.524
SOILG	0.258	0.051	0.279	-0.123	0.112	0.090	-0.094	-0.060	0.023
ERS	-0.115	-0.132	-0.194	0.086	0.037	0.172	-0.216	0.019	0.204
SLDPT	-0.304	-0.008	-0.397	-0.071	-0.067	-0.016	-0.068	0.050	0.092
GEO	0.186	0.037	0.096	0.272	0.205	0.085	-0.253	-0.102	0.532
NDVI	0.298	-0.319	-0.023	0.120	0.151	-0.037	-0.185	-0.114	0.050
SPVSD	0.331	-0.227	0.046	0.065	0.131	-0.118	-0.306	-0.081	0.036
SWI	0.659	-0.058	0.319	0.358	0.111	-0.239	-0.262	-0.302	0.214
SIMP	-0.548	-0.001	-0.320	-0.196	-0.061	0.251	0.135	0.259	-0.210
META	-0.995	0.122	-0.255	-0.282	-0.041	0.281	0.324	0.371	0.272
METS	-0.969	0.164	-0.220	-0.223	-0.093	0.256	0.338	0.400	0.316
MAXTA	-0.994	0.113	-0.258	-0.286	-0.052	0.288	0.335	0.380	0.219
MAXTS	-0.260	0.225	-0.022	0.012	-0.263	0.064	0.342	0.355	-0.061
MINTA	-0.996	0.105	-0.243	-0.259	-0.040	0.282	0.313	0.370	0.252
MINTS	-0.992	0.123	-0.218	-0.247	-0.074	0.279	0.319	0.381	0.263
PRCPA	-0.953	0.038	-0.271	-0.315	-0.008	0.299	0.297	0.333	0.134
PRCPS	-0.830	0.012	-0.244	-0.191	-0.030	0.319	0.225	0.314	0.067
PETAN	0.924	-0.111	0.203	0.251	0.169	-0.157	-0.397	-0.391	-0.069
PETSE	0.967	-0.106	0.235	0.270	0.120	-0.206	-0.380	-0.390	-0.125

	STR	TEXTR	SOILG	ERS	SLDPT	GEO	NDVI	SPVSD	SWI
STR	1.000								
TEXTR	0.804	1.000							
SOILG	-0.064	-0.082	1.000						
ERS	0.113	0.151	-0.509	1.000					
SLDPT	-0.062	-0.087	-0.652	0.803	1.000				
GEO	0.464	0.523	0.120	0.143	-0.080	1.000			
NDVI	0.178	0.063	-0.211	0.224	0.166	0.051	1.000		
SPVSD	0.179	0.021	0.213	-0.189	-0.213	0.006	0.437	1.000	
SWI	0.493	0.462	0.336	-0.161	-0.407	0.711	0.131	0.292	1.000
SIMP	-0.341	-0.309	-0.374	0.246	0.421	-0.581	-0.072	-0.277	-0.908
META	-0.287	-0.182	-0.247	0.097	0.286	-0.136	-0.303	-0.336	-0.613
METS	-0.215	-0.103	-0.243	0.090	0.254	-0.044	-0.316	-0.342	-0.538
MAXTA	-0.337	-0.254	-0.267	0.122	0.307	-0.198	-0.289	-0.344	-0.669
MAXTS	-0.225	-0.296	-0.152	0.088	0.100	-0.127	-0.131	-0.236	-0.325
MINTA	-0.287	-0.196	-0.255	0.113	0.291	-0.144	-0.298	-0.332	-0.627
MINTS	-0.271	-0.175	-0.262	0.111	0.285	-0.132	-0.304	-0.338	-0.620
PRCPA	-0.394	-0.346	-0.312	0.158	0.355	-0.324	-0.250	-0.315	-0.755
PRCPS	-0.344	-0.348	-0.390	0.246	0.391	-0.307	-0.190	-0.278	-0.738
PETAN	0.358	0.334	0.324	-0.111	-0.296	0.368	0.281	0.383	0.716
PETSE	0.351	0.306	0.309	-0.106	-0.300	0.313	0.289	0.367	0.709

	SIMP	META	METS	MAXTA	MAXTS	MINTA	MINTS	PRCPA	PRCPS
SIMP	1.000								
META	0.505	1.000							
METS	0.436	0.986	1.000						
MAXTA	0.551	0.988	0.966	1.000					
MAXTS	0.206	0.241	0.284	0.349	1.000				
MINTA	0.525	0.996	0.980	0.992	0.275	1.000			
MINTS	0.516	0.993	0.986	0.990	0.309	0.996	1.000		
PRCPA	0.653	0.923	0.862	0.953	0.307	0.934	0.927	1.000	
PRCPS	0.674	0.786	0.732	0.843	0.431	0.811	0.817	0.941	1.000
PETAN	-0.608	-0.906	-0.873	-0.938	-0.403	-0.911	-0.912	-0.926	-0.843
PETSE	-0.596	-0.953	-0.923	-0.975	-0.363	-0.957	-0.956	-0.951	-0.854

	PETAN	PETSE
PETAN	1.000	
PETSE	0.990	1.000

ABBREVIATIONS: **ELEV:** elevation, **ASP:** aspect, **STR:** saturation, **TEXTR:** texture, **SOILG:** big soil group, **ERS:** erosion, **SLDPT:** soil depth, **GEO:** geological formations, **SPVSD:** supervised classes, **SWI:** Shannon Wiener index, **SIMP:** Simpson Index, **META**: annual mean temperature, **METS**: seasonal mean temperature, **MAXTA:** annual maximum temperature, **MAXTS:** seasonal maximum temperature, **MINTA:** annual minimum temperature, **MINTS:** seasonal minimum temperature, **PRCPA:** annual precipitation, **PRCPS:** seasonal precipitation, **PETAN:** annual potential evapotranspiration, **PETSE:** seasonal potential evapotranspiration

VITA

HAKAN METE DOĞAN

Central Research Institute for
Field Crops, GIS & RS Department
P.Box. 78, Yenimahalle-Ankara

90.312.327-3075 (office)
90.536.253-6846 (cellular)

hmdogan@hotmail.com
hmdogan@ankara.tagem.gov.tr

September 26, 1961- Born in Sivas, Turkey
1977-1979- Gazi High Scholl, Konya
1980-1984- B.Sc. Ankara University, Agricultural Faculty, Ankara
1985-1987- M.Sc. Atatürk University, Agricultural Faculty, Erzurum
1986-1992- Researcher, Eastern Anatolia Agricultural Research Institute, Erzurum,
1992-1996- Researcher, Black Sea Agricultural Research Institute, Samsun
1996-1998- M.Sc. New Mexico State University, Geography Department, the U.S.A.
1996-1998- Researcher, Central Research Institute for Field Crops, GIS & RS Dept., Ankara
1999- - Department Head, Central Research Institute for Field Crops, GIS & RS Dept., Ankara

Area of Interest

Plant ecology, agriculture, forestry, water pollution, conservation biology (in-situ), geographic information systems, remote sensing, and modeling,

Others

Jogging, painting, swimming, computer sciences and programming