UNIVERSITY OF NAIROBI

SCHOOL OF MATHEMATICS

Analysis of Elephant Density Distribution in the Amboseli Ecosystem: Application of Generalized Least Square Model

By

Eric Mboya Ochwang'i

This research project is submitted to the School of Mathematics of the University of Nairobi in partial fulfillment of the requirement for the degree of Masters of Science in Social Statistics.

DECLARATION

This project is my original work and has not been presented for a Masters Degree in any other university. No part of this thesis may be produced without the prior permission of the author and/or University of Nairobi**.**

Eric Mboya Ochwang'i **Reg. No: I591/83411/2012.**

Signature: ______________________________ Date: _____/_____/2014

This thesis has been submitted for examination with our approval as University supervisors.

Dr V.N.Mose

African Conservation Centre

Signature: ______________________________ Date: ____/____/2014

Dr J. M. Ndiritu

University of Nairobi

Signature: ______________________________ Date: ____/____/2014

DEDICATION

I dedicate this project to my children, Elvis and Ruth and my wife, Basilica Bii.Sincere gratitude to my mother, Edna Ochwangi who has never left my side.

I also extend my dedication to my friends who have always given me moral support and above all my Creator for the good health I have enjoyed throughout my entire studies.

ACKNOWLEDGEMENT

I would like to express my sincere thanks to my principal supervisor Dr. Victor Mose for his guidance, encouragement, advice, patience, professional and moral support during the course of my studies. Special thanks also go to my other supervisor Dr. John Ndiritu for his continued assistance and support. Wish to thank the Amboseli Conservation Program for providing financial support for my studies and research work.

Sincere gratitude goes to Dr David Western for his positive criticism and close supervision and particularly for making it possible for me to use the long-term Amboseli Conservation Program data.

I would like to record my appreciation to my wife and mother for their cooperation and support throughout the study**.**

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

NDVI-Normalized Difference Vegetation Index

ACP-Amboseli Conservation Program

NGO-Non Governmental Organizations

GIS-Geographical Information Systems

GLS-Generalized Least Squares

ANOVA-Analysis of Variance

EDA-Exploratory Data Analysis

KWS-Kenya Wildlife Service

ABSTRACT

Elephant population has risen steadily in Amboseli National park since the government enhanced its efforts in combating poaching. However, human population increase and settlements have impacted on vegetation composition leading to a shift in vegetation structure. Due to these changes woody vegetation has shrank to less than 10% and is now dominated by grassland and swamp. We used long-term ACP data to investigate whether human settlement, NDVI and vegetation biomass have significant impact on elephant density distributions over time and further determined the long-term spatial population distribution of elephants in Amboseli. Results from a generalized least square model showed that NDVI, human settlement and vegetation biomass significantly influence elephant density distribution.

The spatial density distribution maps for elephants showed that elephant densities were high inside the park during dry season basically because of the swamp which is located inside the park and it's considered as the main source of water and food. However, during wet season elephants move freely in and outside the park. Spatial density distribution maps for decades showed that elephant densities were very low and evenly distributed during the 1990 decade; concentration for other decades was higher inside the park compared to outside.

A generalized spatial least square model developed to determine the effect of seasonal elephant density distributions on vegetation showed that during wet season elephants have no significant effect on vegetation but they were found to be significant during dry season.

CHAPTER ONE: INTRODUCTION

1.1 Background

Considerable progress in conserving the environment has been made by NGOs, the Kenyan government and other stake holders. They have played a key role in protecting wildlife, forests, as well as promoting peaceful coexistence between wildlife and the communities living in Amboseli. Despite these efforts, still, many challenges need to be addressed. These include:

- Conflicts between the communities living in the area and wildlife.
- Persistent increase in elephant population which is believed to be the main cause of habitat change.
- Decline in the number of browsers and grazers.
- Human settlement.
- Drought.

However, this project focuses on the long term spatial distribution of elephant densities in Amboseli ecosystem and how they are affected by human settlement, NDVI and vegetation biomass.

Between 1970 and 1989 poachers killed more than 5000 elephants in a year (David western, 1995) but poaching levels dropped by almost 50 percent when international ivory ban was imposed in 1990.Kenya’s efforts to end poaching and the success of international ivory ban has led to the increase in elephant population, which is now enjoying protection from Kenya Wildlife Service, ACP and other stake holders. However, over the years as Elephant numbers continue to increase, the composition of woody vegetation (Trees and shrubs) continues to lose its diversity.

Woodland loss is a major cause of biodiversity decline in African savanna parks. There are three major theories which explain woodland loss in Amboseli National Park. They include; overgrazing theory, pathogen/pest theory and climatic theory.

A 20 year experiment conducted by David Western in Amboseli National Park indicates that Elephants are the ones preventing Woodland recovery. They impact on plants by breaking branches/stems, stripping bark, uprooting plants and toppling trees. The persistence of plant species eaten by elephants are dependent on whether they can cope with herbivory of this nature , or whether mortality is balanced or exceeded by recruitment and regeneration.

Increase in elephant densities translate to increase in grazing/browsing pressure, resulting to decline of woody vegetation. Loss of woody vegetation negatively impacts the survival and movement of birds, insects and other mammals.

Human settlement has spread and encroached into the elephant habitat. Elephants are now confined in a small area in the park. It is projected that if this trend continues, then woody vegetation will be completely exhausted as grassland spreads in the area which was originally covered by trees and shrubs. Eventually, elephant population will decrease drastically or possibly be rendered extinct.

As a result of these changes we seek to apply a generalized least square model to determine the effects of human settlement, NDVI and vegetation biomass on elephant densities over time.

1.2 The study area

The study area falls to the North of Amboseli National Park, immediately north of Mount Kilimanjaro and the Tanzania-Kenya border. It covers approximately 400km^2 with an average rainfall of 350mm annually. The area has hot and wet season with maximum temperatures varying between 26 and 44^{o}c. (Mose, 2013).

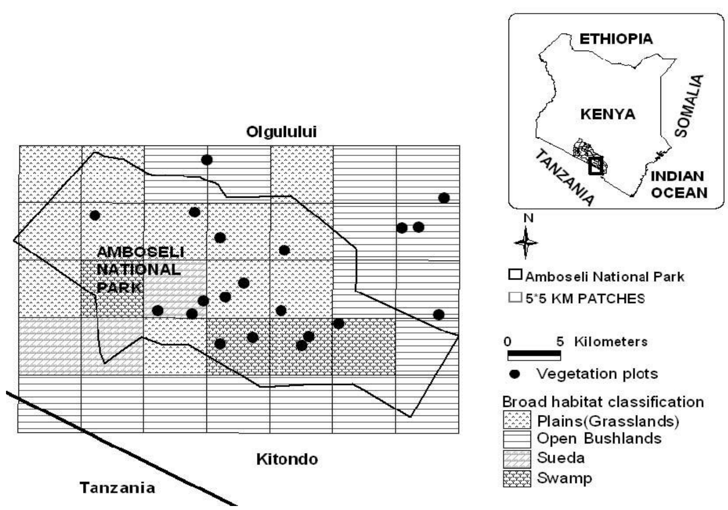


Figure 1.0: Map of Amboseli

1.3 Research problem

Amboseli ecosystem has experienced persistent increase in elephant population over time. However, human settlements have been constantly increasing. This has resulted in loss of woody vegetation diversity and spread of grassland and swamp. If the trend continues then woody vegetation will be completely depleted. Due to limited resources, drought and human-elephant conflicts, elephants might eventually be extinct.

1.4 Objectives of the study

1. To develop a statistical model to determine the effects of vegetation biomass, NDVI and human settlement on elephant densities over time.
2. To generate spatial distribution maps for elephant densities over time using geographical information system.

1.5 Hypothesis

1. Elephant population densities have increased independent of changes in vegetation biomass, NDVI, and human population increase.
2. The spatial distribution of elephants in Amboseli is not affected by seasonal changes.

1.6 Research questions

The main aim of this study is to determine the factors that largely explain seasonal distributions and changes in elephant population densities. The following questions will be answered:

1. Do human settlements, vegetation biomass and NDVI have significant effect on elephant population densities?
2. Do seasonal changes affect the spatial distribution of elephants?

1.7 Justification

The generalized least square model is important in addressing autocorrelation, normality and heteroskedasticity unlike general regression model which assumes that data is normal and no autocorrelation exists. Elephants being the response variable, the model will clearly show the variables which are significant in influencing the elephant population densities over time taking into consideration the effects of autocorrelation and normality.

The GIS maps on the long-term spatial population distribution of elephant densities will be useful in displaying the long term changes of elephant densities in Amboseli National Park and the surrounding. These maps will be helpful in forecasting the long term impact of elephant movements and distribution patterns; which is essential in policy making to reduce human-elephant conflicts while maximizing resources in the ecosystem.

CHAPTER TWO: LITERATURE REVIEW

Elephants being keystone species play a significant role in ecological dynamics. They can influence vegetation structure and composition in ecosystem processes. They are also one of the main tourist attraction species. However, elephants are the main targets of poachers, which make them a species of conservation concern (Omondi and Ngene, 2012) .Over the last 100 years African elephants have reduced dramatically due to ivory trade and drought. It is estimated that elephants declined from around 167,000 in 1973 to 20,000 in 1990.According to the 2014 KWS census they were estimated at 38,000 while in Africa they are estimated to be between 410,000 to 250, 000, which represents more than 50% decline in the last 35 years. However, in Amboseli their numbers have grown steadily over the last two decades which is more worrying since woody vegetation, which is their main forage, has been depleted.

A long-term ecological research in Amboseli showed that woodland and bush land habitats have sharply declined over the past 50 years (Western, 2006) while swamps, scrublands and Grasslands have greatly expanded. The Woodland cover is approximately 10% of Amboseli National Park down from 30% in 1950; it is now dominated by plains and swamp. This is a worrying trend and since woody vegetation is expected to be rendered extinct in the next two decades (Western, 2006), its loss will greatly affect browsing species diversity in the ecosystem.

Large mammal herbivores are the main drivers of African savannah ecosystems and have strong impact on woody vegetation (Lauren, 2013).Mega herbivore exclusion results in increased densities of reproductive trees and tall trees. Consequently, mammalian extinction and changes in land use have complex cascading consequences.

Over the last half a century elephant numbers have drastically increased in Amboseli ecosystem even though there are limited resources, human pressure and competition for resources.

According to Western and Lindsay (1983), elephants select habitats with more abundant but less digestive forage during wet season but during early dry season they heavily use woodlands. African population distributions are influenced by availability of water (Timothy and Child, 2012).Concentration of elephants near water points may lead to vegetation degradation.

Human settlement and poaching also have huge impact on elephant movement; they are forced to alter migration patterns and concentrate in protected areas (Western, 1989)

Elephants are extremely adaptable and are capable of occupying a variety of habitats. Environmental factors that affect elephant movement and distribution include, group size and composition, home range, migration patterns and diet (AWF technical handbook series, 1996).Elephant diet include grass, herbs, bark, tree foliage and fruits.

Elephant mainly affect the size of trees more than density (Claudius et al, 1999) during wet season, as the season progresses they retreat to swamps and swamp edge. They also seek places with high proportions of vegetation and avoid human settlements (Grant and Harris *et al, 2008*).If the area available is large they will probably favour areas near water and high vegetation cover. Artificial water points are mostly affected by elephants (Kanision et al, 2012), canopy cover increases in areas away from the artificial water points. Browsing near areas of water points by large herbivores result in vegetation degradation.

Large herbivores affect the structure and composition of plant communities (Hillary et al, 2013).The response of plant communities to herbivores is dependent on herbivore abundance and herbivore identity. A linear model to determine the effects of herbivore and rainfall on vegetation showed that rainfall had a significant effect on vegetation structure. There was evidence of reduced vegetation cover in areas with high wildlife numbers; grasses also dominated in these areas.

Amahowe et al (2012) used a linear model to assess the woody vegetation structure and plant diversity in areas used by elephants. The model showed that the type of vegetation in an ecosystem could influence the damage caused by elephants. Habitats which are rich in large and tall trees are more exposed to elephant damage. The distribution of elephants in relation to ecological and environmental factors, (William et al, 2013) showed that elephant density was highly correlated with decreasing human density. It was also positively correlated with rainfall and NDVI.Their densities were also found to be correlated with increased human literacy and increased per capita rate.

NDVI is used to measure the density of green on a patch of land. It can also be used as an indicator of climate change. It’s a measure of vegetation which is a major component of land cover (Nduati et al, 2013), a widely used parameter in assessment and characterization of vegetation which is extremely useful for researchers aiming to understand better on how vegetation distribution and dynamics affect diversity, population dynamics, movement patterns and life history traits of animal populations.

High vegetation index indicates that the amount of rainfall is high which results to an increase in elephant densities, unless there are constraint factors such as poaching. Within landscapes, elephants are found on grid-cells with higher NDVI values but the influence of NDVI during dry season may be weak (Young et al, 2009).Also, exists a positive relationship between NDVI and elephant densities in non-forest populations, (Duffy and Pettorelli, 2012).It can also be linked to mega herbivore abundance across Africa. Seasonal rainfall patterns determine the quantity and quality of vegetation and influence large scale movements of herbivores (Mose et al, 2013), but the spread of human settlement, human activities and fragmentations threaten large migratory ungulates in Africa. Elephants move in relation to food abundance and quantity.

When green vegetation is not available elephants adjust their elevation upwards during dry season. When productive vegetation is available at the lower elevations the elephants stick to that home range (Bohrer et al, 2014)

Interaction between elephants and people occur within conservation areas. In most cases it leads to human-elephant conflict which is a highly emotive and politicized issue. In Africa this interaction has occurred for over 1000 years.

Elephants are key to tourism attraction but at the same time vulnerable to poaching due to ivory trade. Unprotected conservation areas attract human settlements which are not only significant in influencing the changes in elephant numbers but also affect the growth and spread of vegetation. The inhabitants use trees for firewood and construction of houses (Western and Thomas, 1979). Human activities have profound effects on the distribution and habitat utilization patterns of elephants in an ecosystem. Increase in human settlement often lead to conflict between human beings and elephants because of the competition for resources.

There has been persistent growth of human population in Africa that has resulted in land subdivision and settlement in areas which were initially reserved for conservation. Africa's population distribution will always be vulnerable to human settlement and expanding agriculture. For instance, Amboseli ecosystem settlement patterns follow various biological and physical characteristics of the landscape (Western and Dunne, 1979).They are located away from bush vegetation and dense trees because of predators.

A linear model developed by Hoare and Toit (1999) showed that elephant density is unrelated to human settlement. Habitat loss has a more serious effect on elephant than poaching while land clearing by rural human population result in a less reversible loss in elephant numbers (Hoare and Toit, 1999).

A generalized least square model to assess whether the movement of elephants across a wide resource gradient was explained by rainfall, population density and primary productivity showed that survival of elephants decreased during dry season but increased during wet season (Young and Aarde, 2010).Elephant densities increased with increase in vegetation productivity but decreased with high human densities.

Merode et al (2000) explored the relationship between the distribution of wildlife populations and human activities. He discovered that the presence of agricultural communities, conservation practices and proximity to urban areas affect the presence, abundance and community structure of large mammal populations. Elephant normal movements and migrations are usually confined to individual home ranges (Viljoen, 1989).They can utilize food resources up to 70kms away from waterholes due to their mobility and ability to go up to four days without drinking water.

Both biotic and abiotic factors constrain the distribution of elephants. Their presence is positively correlated with forest cover and vegetation productivity (Rood et al, 2010).A spatially explicit model showed that elephants mainly utilize forest edges but avoid human dominated areas. Forest encroachment occurs throughout the elephants range and is found within 80% of the elephants’ ecological niche.

CHAPTER THREE: METHODOLOGY

3.1 Data

The data for vegetation, elephant and human settlement was obtained from Amboseli Conservation Program (ACP).The sample design is based on systematic flight lines over the entire study area. The study area is subdivided into grids of 5x5 km.The coordinates are based on the standard geographical system (UTM Grid) numbering from East to West and from South to North. Each cell is located by the column number, followed by the row number.

The area is sampled by flying parallel flight lines, 5km apart, along each column such that the transect runs through the center of each grid. In this way each grid within the study area is sampled and distributions can be established for animals and environmental variables over the region.

Normalized Difference Vegetation Index (NDVI) data was obtained from the Global Land Cover Facility (http://www.glcf.umd.edu/data/gimms/)

We describe a generalized least square model that will be applied to elephant densities in chapter 4. We further outline the advantages and disadvantages of generalized least square model. Finally, we describe the reasons we settled on human settlement, NDVI and vegetation as the ideal variables for our model.

3.2 Choice of variables

3.2.1 Human Settlement

The human population and settlements in Amboseli has been constantly growing, posing a threat to movements and survival of elephants and other herbivores. Human population growth often leads to land subdivision, human-elephant conflicts, poaching and exhaustion of woody vegetation. As a result of being exposed to human beings, elephants spend more time at night than during the day in areas under land use to avoid being attacked by human beings (Graham et al, 2009).This demonstrate that elephants deliberately alter their behavior to avoid risk in human dominated areas. Since elephants move freely within and without the park, human settlements are perceived to be an important variable to be tested to determine whether it has a major impact on elephant densities.

Elephant densities are calculated as:

$$Elephant\ density = \frac{Elephant\ numbers}{Area(5 \times 5km)}$$

3.2.2 Vegetation biomass

Vegetation in Amboseli is classified into woodland, plains, grassland, swamp and bush land; it is quantified in terms of biomass. Elephants being mixed feeders and due to their nature of feeding on bulky, they alter the vegetation structure and composition of an entire ecosystem. In Amboseli, woody vegetation has shrunk to insignificant levels whereas grassland has spread to cover the largest portion. This is a worrying trend even though elephants continue to increase in numbers. These dynamics necessitated us to study whether the changes in structural diversity of vegetation have significant influence on elephant densities.

Vegetation biomass can be calculated as

$$y = 29.03x + 3.34$$

$$Where\ y = Biomass\ per\ m^2$$

$$x = Hits\ per\ pin$$

3.2.3 NDVI

NDVI (Normalized Difference Vegetation Index) is used to measure remote sensing measurements from space platform and is used to assess whether the target being observed contains green vegetation or not. It’s calculated as:

$$NDVI = \frac{(NIR - RED)}{(NIR + RED)}$$

$$NIR = Near - infra\ red$$

$$RED = Red\ reflectance$$

NDVI values are presented as a ratio ranging from -1 to 1 but extremely negative values represent water, values around zero represent bare soil and values over 6 represent dense green vegetation.

There is a direct relationship between NDVI and the amount of stress vegetation experience. Low vegetation index implies that vegetation quality and quantity is also minimum. These changes could influence elephant distribution and abundance.

3.3 Generalized least square model

In this section we begin by describing the classical linear regression model.

Consider the following linear regression model

$$Y = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} +, \dots, + \beta_k X_{i,k-1} + \varepsilon_i \qquad (3.3.1)$$

This model can also be expressed as

$$Y = X\beta + \varepsilon \qquad (3.3.2)$$

Where $\boldsymbol{X}$ is the $\boldsymbol{n} \times \boldsymbol{k}$ matrix of explanatory variables and $\boldsymbol{Y}$ is the$\boldsymbol{n} \times \mathbf{1}$ vector of response variable values. $\boldsymbol{\varepsilon}$ is $\boldsymbol{n} \times \mathbf{1}$ vector of true residuals.

The estimates of $\boldsymbol{\beta}$ and $\boldsymbol{\delta}^{\mathbf{2}}$ are given as

$$\hat{\beta} = (X'X)^{-1}X'y$$

$$\hat{\sigma}^2 = \frac{1}{n}\hat{\varepsilon}'\hat{\varepsilon} = \frac{1}{n}\sum_i \hat{\varepsilon}_i^2$$

Where

$$\hat{\varepsilon} = y - X\hat{\beta} \qquad (3.3.3)$$

The linear model requires relatively strict set of assumptions. These assumptions are required to show that the estimation technique has a number of desirable properties, and also so that the hypothesis tests regarding coefficient estimates could validly be concluded.

Below we present the assumptions classical linear model should meet.

1. $\boldsymbol{E(\varepsilon_i) = 0}$ The errors have zero mean.
2. $\boldsymbol{Var(\varepsilon_i) = \sigma^2} < \infty$ The variance of the errors is constant and finite over all values of $\boldsymbol{X_i}$
3. $\boldsymbol{Cov(\varepsilon_i, \varepsilon_j) = 0}$ The errors are statistically independent of one another.

4. $\boldsymbol{Cov(\varepsilon_i, X_j) = 0}$ There is no relationship between the error and the corresponding $\boldsymbol{x}$
5. $\boldsymbol{\varepsilon_i \sim N(0, \sigma^2) \varepsilon_i}$ is normally distributed

We need to test whether our data satisfies the assumptions of normality and auto regression in order to verify whether any of these assumptions is violated.

3.4 Data validation tests

Validation is the process of assessing how well a model performs against real data. Since we are employing generalized least square model, Durbin Watson, Shapiro-Wilk and Breusch-Pagan tests are used to assess the quality and validity of the data based on assumptions of normality, autocorrelation and heteroskedasticity.

3.4.1 Durbin Watson test for autocorrelation

It's a statistical test used to detect the presence of autocorrelation in the residuals. Let e_i be residual sorted into time order then the Durbin Watson test statistic is

$$\mathrm{d} = \frac{\sum_{i=2}^{n}(e_i - e_{t-1})^2}{\sum_{i=1}^{n} e_i^2}$$

If **d** is less than 2 then there is a positive serial correlation, if d is 2 then there is no serial correlation. If d is more than 2 then there is a negative serial correlation. Because of the dependence of any computed Durbin Watson value on the associated matrix, exact critical values of Durbin Watson statistic are not tabulated for all possible cases. The conventional Durbin Watson tables are not applicable when you do not have constant term in the equation.

According to our model **d=1.5515**, **p value=0.008841**

Where d represents Durbin Watson

Since **d** is less than 2 then there is evidence of positive serial autocorrelation in the residuals.

3.4.2 Shapiro–Wilk test for normality

Shapiro- wilk test can be performed to test whether normality test is statistically valid or not. The test utilizes the null hypothesis to check whether a sample $x_1 \ldots x_n$ came from a normally distributed population.

The test statistic is

$$w = \frac{(\sum_{i=1}^{n} a_i\, x_{(i)})^2}{\sum_{i=1}^{n}(x - \bar{x})^2}$$

Where $\boldsymbol{x_{(i)}}$is the *i*th order statistic

$$\bar{x} = \frac{(x_1, \dots, x_n)}{n}$$

The constants a_i are given by

$$(a_1, \dots, a_n) = \frac{m^T V^{-1}}{(mV^{-1}V^{-1}m^T)^{1/2}}$$

$$m = (m_1, \dots, m_n)^T$$

$\mathrm{m}_1, \dots, \mathrm{m}_\mathrm{n}$ are the expected values of the order statistic of independent and identically distributed random variables sampled from the standard normal population, and V is the covariance matrix of the order statistics.

	W	P-value
Elephant density	0.9364	0.0001494
Vegetation biomass	0.9726	0.03989
Human settlement	0.9807	0.1649
NDVI	0.9159	0.00001145

Table 3.1: Shapiro-Wilk test

From table 3.1, elephant density, vegetation biomass and NDVI are not normally distributed where as human settlement is normally distributed.

3.4.3 Breusch-Pagan test for Heteroskedasticity.

Random variables can be found to be heteroskedastic if some of their sub populations have different variability from others. The possible existence of heterokedasticity is a major concern in the application of regression analysis, including the analysis of variance, because the presence of heteroskedasticity can invalidate statistical tests of significance that assume that the modeling errors are uncorrelated and normally distributed and that their variances do not vary with the effects being modeled.

Breusch-Pagan test is used to test whether the estimated variance of the residual from a generalized least square model depends on the values of the independent variables or not. In a generalized least square model if the p values from the estimated variance is less than 0.05 then we deduce that there is heteroskedasticity or we conclude otherwise if p value is more than 0.05.

The test statistic for Breusch-Pagan test is

$$bp = \frac{1}{v}(u - \bar{\mu}i)'Z(Z'Z)^{-1}Z'(u - \bar{\mu})$$

Where

$$u=(e_1^2, e_2^2, \dots, e_n^2)$$

$$V=\frac{1}{n}\sum_{i=1}^{n}(e_i^2 - \frac{e'e}{n})^2$$

According to our model **BP = 8.6205**, **p value=0.03479**

Where BP represents Breusch-Pagan test

Since the test shows that the p value is 0.03, which is less than 0.05 then we conclude that there is evidence of heterogeneous variance in the data.

Clearly the classical linear regression model will not be ideal since the data violates the assumptions of normality and autocorrelation. Generalized least square model is the most appropriate since it allows for heterogeneous variance in the residuals. The generalized least square estimator$\hat{\beta}$ is an unbiased estimator. If the assumption of normality is relaxed $\hat{\beta}$ becomes the best linear unbiased estimator. The model has a wide range of applications in ecological systems and environment where simple linear regression model cannot be applicable.

Generalized least square model is a technique of estimating unknown parameters in a general linear regression model. This technique is usually applied when the observation data has a certain degree of correlation. Since ordinary least squares can be statistically inefficient generalized least square can give reliable inferences.

Suppose variance ε=$\sigma^2\sum$, where σ^2 isunknown but $\sum$ is known.

Generalized least squares minimizes

$$(y - X\beta)^T\sum^{-1}(y - X\beta) \quad (3.4.1)$$

This is solved by

$$\hat{\beta} = (X^T\sum^{-1}X)^{-1}X^T\sum^{-1}y \quad (3.4.2)$$

$\sum = SS^T$, where S is a triangular matrix.

Using Choleski decomposition we have

$$(y - X\beta)^T S^{-T} S^{-1}(y - X\beta) = (S^{-1}y - S^{-1}X\beta)^T(S^{-1}y - S^{-1}X\beta) \quad (3.4.3)$$

gls is like regressing $S^{-1}X$ on $S^{-1}y$

$$y = X\beta + \varepsilon \quad (3.4.4)$$

$$S^{-1}y = S^{-1}X\beta + S^{-1}\varepsilon$$

Our new regression equation is

$$y' = X'\beta + \varepsilon' \quad (3.4.5)$$

Where the variance of the new errors,ε' are expressed as

$$var\varepsilon' = var(S^{-1}\varepsilon) = S^{-T}(var\varepsilon)S^{-T} = S^{-1}\sigma^2 SS^T S^{-T} = \sigma^2 I$$

$$var\hat{\beta} = (X^T\sum^{-1}X)^{-1}\sigma^2 \quad (3.4.6)$$

$$E(\varepsilon|X) = 0$$

$$Var(\varepsilon|X) = \sigma^2 I$$

The likelihood function of ($\beta_{0,}\sigma^2$) is of the form

$$L(\beta,\delta^2|y) = (2\pi\delta^2)^{-\frac{n}{2}}\exp\left[\frac{1}{2\pi\sigma^2}(y - x\beta)'(y - x\beta)\right] \quad (3.4.7)$$

$\widehat{\beta}$ is unbiased estimator of β and $\hat{\delta}^2$ is unbiased estimate of δ^2

An unbiased estimator for $\hat{\delta}^2$ is the residual variance given by

$$MSE = \frac{1}{n-k}\sum_{i}^{n} \hat{\varepsilon}_i^2$$

T test can be used to test whether the parameter estimates are statistically significant or not

$$T = \frac{\hat{\beta}}{s/\sqrt{S_{xx}}}$$

Where

$$\hat{\beta}=\frac{S_{xy}}{S_{xx}}$$

$$S_{xx}=\sum(X-\bar{X})^2$$

$$S_{xy}=\sum(X-\bar{X})(Y-\bar{Y})$$

Coefficient of determination can be calculated to see how well the model is successful at explaining variability as

$$R^2=\frac{S_{xy}^2}{S_{xx}S_{yy}}$$

Where

$$S_{yy} = \sum(Y-\bar{Y})^2$$

Our generalized least square model equation is described as

$$Elephantdensity(y)=\beta_1 X_1+\beta_2 X_2+\beta_3 X_3+\varepsilon \qquad (3.4.8)$$

Y=Elephant density

X_1=Vegetation biomass

X_2=NDVI

X_3=Human settlement

CHAPTER FOUR: DATA ANALYSIS

4.1 Background

Statistical analyses were conducted using STATA version 12 and R 3.0.1.Elephant spatial population distribution maps were generated using Quantum GIS version 1.7.0 and Arc view version 3.2 a. Graphical representation were used to summarize the data using STATA 12 and R version 3.0.1.Generalized least square model was used to model the effects of vegetation biomass, human settlement and NDVI on elephant densities.Also,generalized spatial least square model was fitted on vegetation biomass to test whether elephant densities had significant effect on vegetation during dry and wet seasons.

4.2 Exploratory data analysis

Exploratory data analysis was applied to vegetation data (bush land, woodland, grassland, and swamp), elephant population density, NDVI and human settlement to determine whether the data satisfies the assumptions of normality and autocorrelation.

QQ plots were used to test for normality, ACF plots were used to test for autocorrelation while histogram and density plots were plotted to test for normality and skewedness.

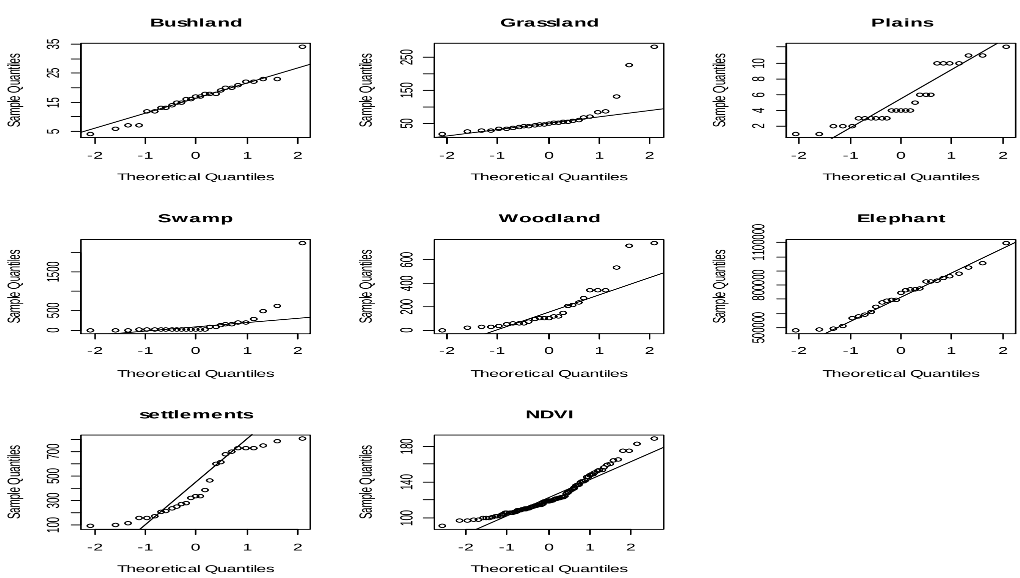


Figure 4.2.1: QQ plots for Elephants, Vegetation biomass and Human settlement

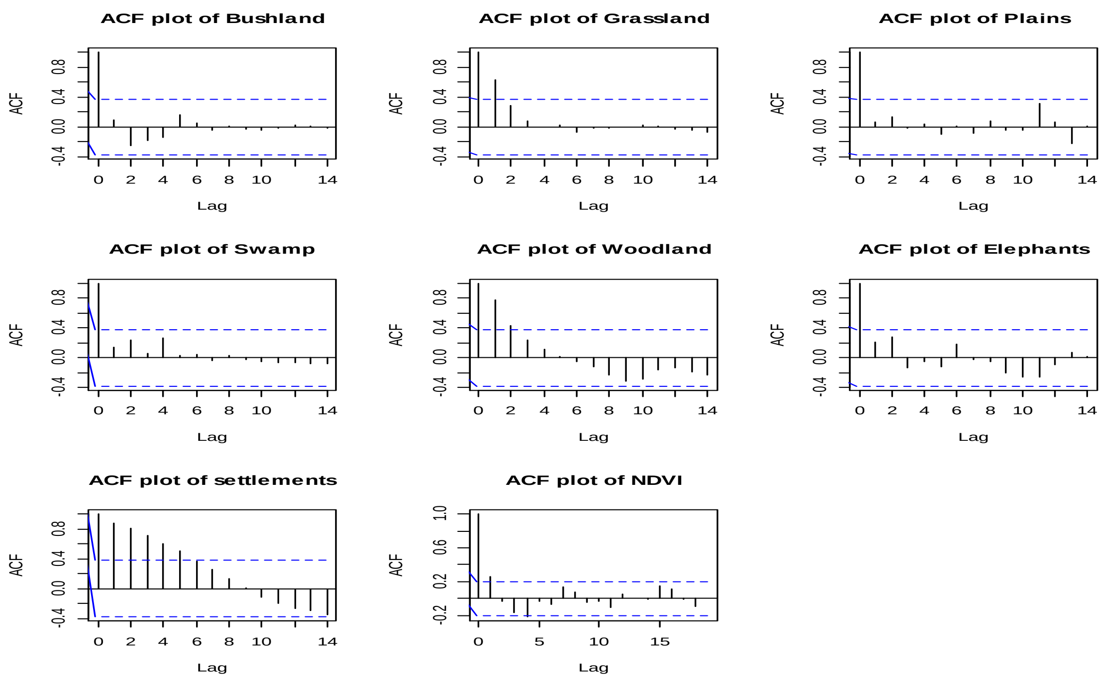


Figure 4.1.2: ACF plots for testing Autocorrelation and partial-correlation.

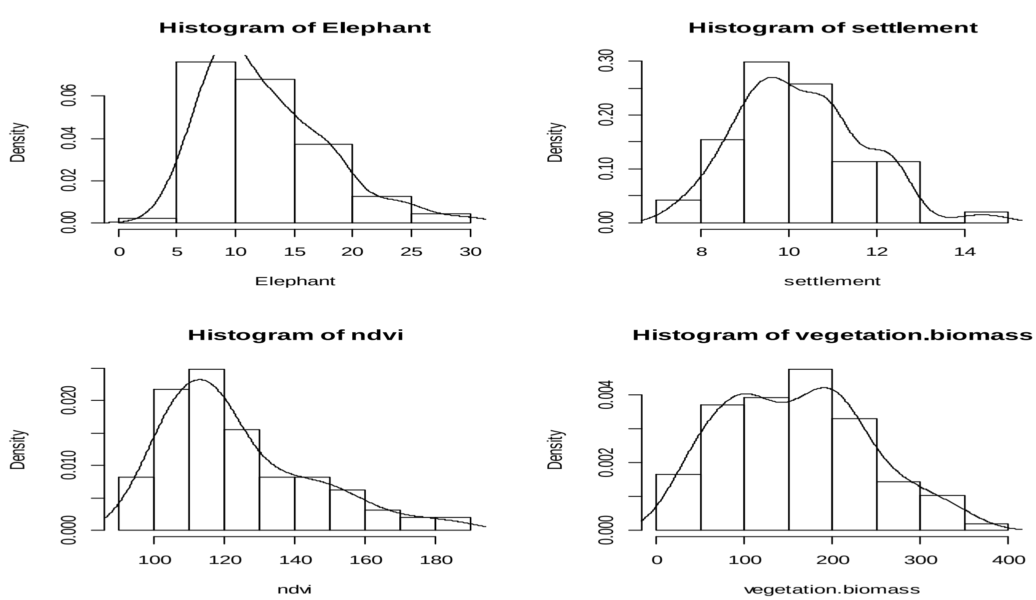
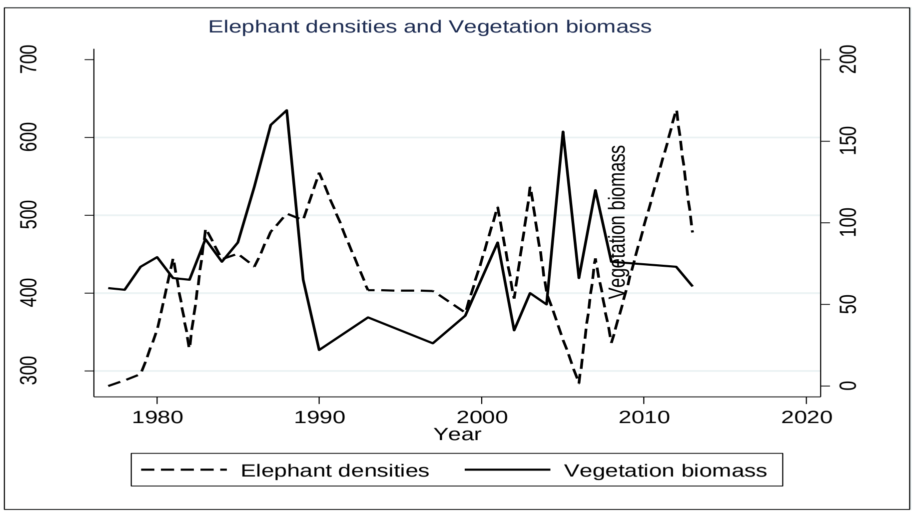


Figure 4.2.2: Histograms and kernel density estimates of vegetation biomass, elephant densities NDVI and human settlement.

There is no evidence of normality from the QQ plots for Grassland, swamp and woodland. (Figure 4.2.1) Autocorrelation is also present in woodland and grassland data (Figure 4.2.2). The histograms for NDVI and elephant densities are skewed to the left. General linear regression will not be suitable for analysis of data which has autocorrelation and lacks normality. Generalized least square model is the most suitable technique since the application of maximum likelihood technique takes care of autocorrelation, normality and heteroskedasticity.

Trend graphs

Trend graphs help in visualizing and describing data and trends. We used line graphs, bar graph and box plot to compare various variables used in the model over time.



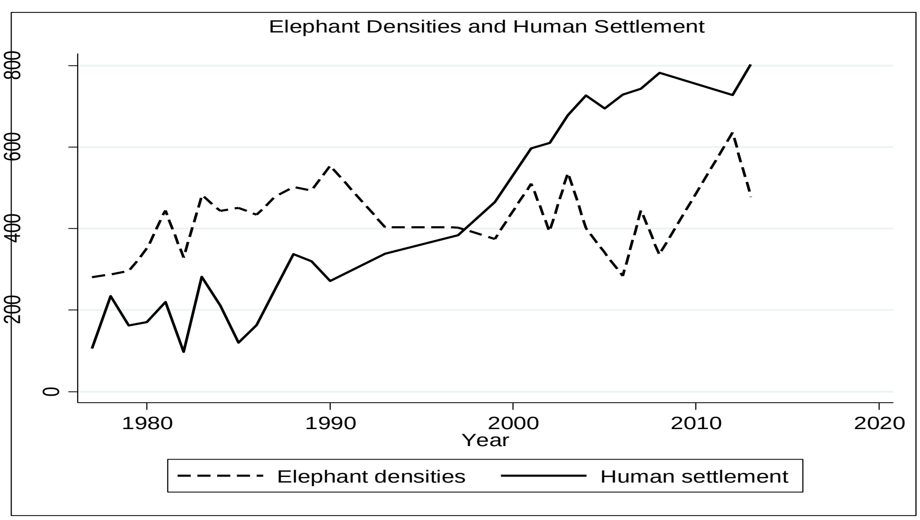


Figure 4.2.3: Line plots of elephant densities against vegetation biomass and human settlement.

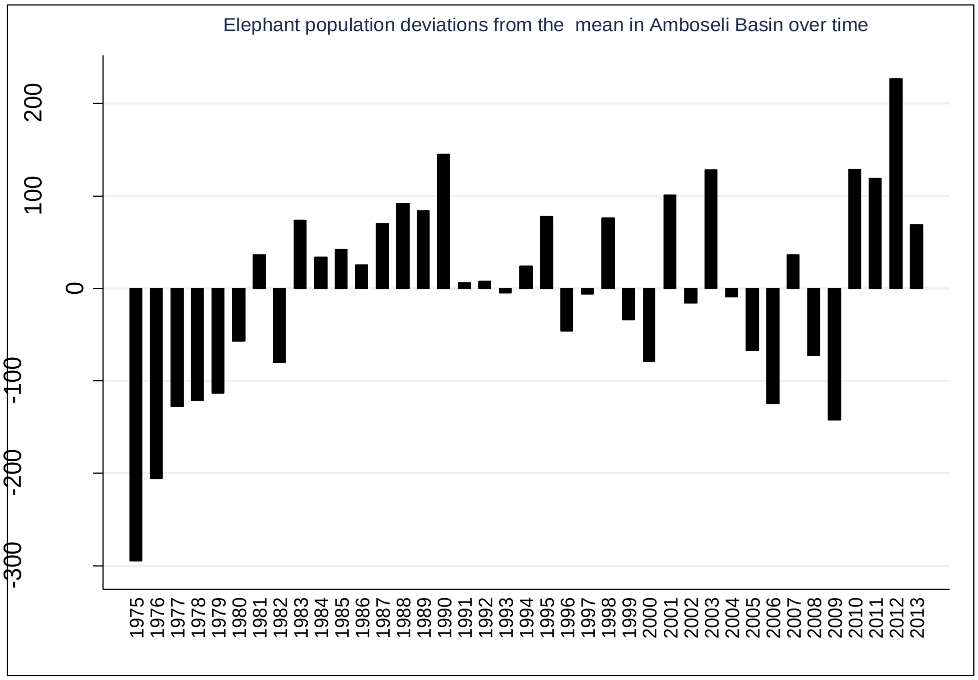


Figure 4.2.4: Mean deviations of elephant densities over time. The graph shows how far each point in the data deviates from the mean densities.

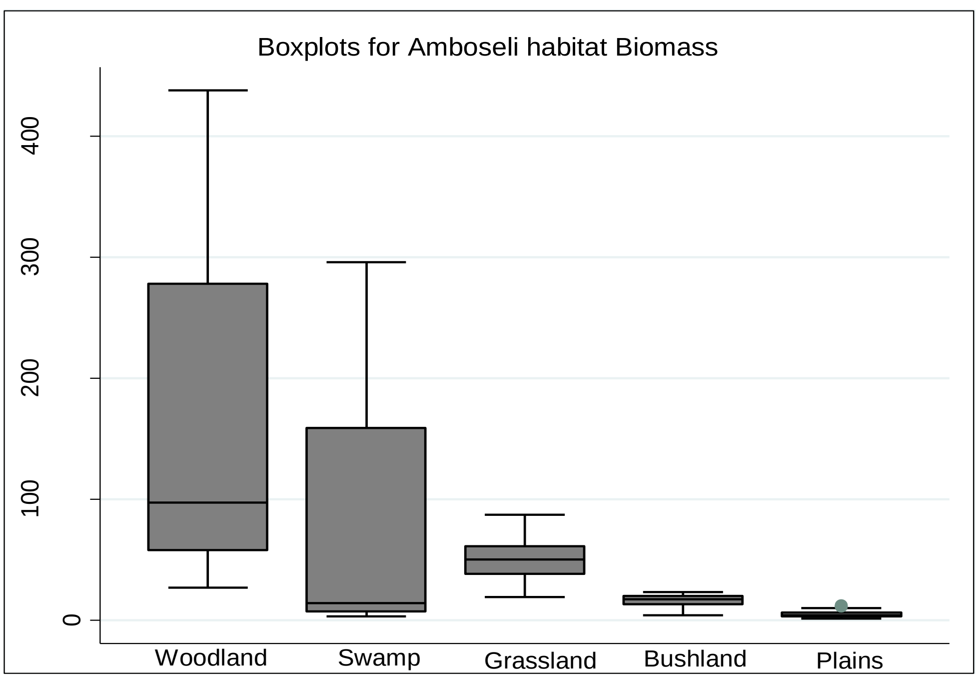


Figure 4.2.5: Box plots for vegetation classes in Amboseli

The rate of increase in human densities is higher than that of elephant densities (figure 4.2.4) but vegetation biomass increases as elephant densities increase. Box plots in figure 4.2.6 show that Woodland has the highest biomass followed by swamp, grassland, bushland and plains respectively.

4.4 Application of generalized least square model

	$Estimate$	$Std.Error$	$t - Value$	$Pr(> \|t\|)$
Vegetation Biomass	0.0112275	0.00553507	2.028420	0.0453
Settlements	-0.5136491	0.15230005	-3.372613	0.0011
NDVI	0.1296039	0.01506687	8.601912	0.0000

Table 4.1: Generalized least square model table of elephant densities against vegetation biomass, NDVI and human settlement.

	$F - Value$	$Pr(> \|t\|)$
Vegetation Biomass	1050.9808	<.0001
Human Settlement	57.5295	<.0001
NDVI	71.7045	<.0001

Table 4.2: ANOVA table for testing the significance of parameters in the gls model.

The generalized least square model equation was found to be

$$Elephant\ density(y) = 0.0112275X_1 - 0.5136491X_2 + 0.1296039X_3 + \varepsilon$$

Human settlements have a negative influence on elephant densities; vegetation biomass and NDVI positively affect elephant densities. A unit increase in vegetation biomass results to 1% increase in elephant densities while a unit increase in NDVI affects the changes of elephant densities by 13%.A unit increase in human settlements results to 51% decrease in elephant densities. All variables are therefore important in determining the population changes and spatial distribution of elephants in an ecosystem.

4.5 Spatial distribution of elephant densities.

Elephant movement in Amboseli ecosystem is based on geographical boundaries. The area is subdivided into grids of 5X5km.Each grid is given a unique code for identification purpose. Elephants move across the grids and within the grids by tracking the resources. We used a spatial generalized least square model to determine whether the movement of elephants across the grids has significant impact on vegetation biomass based on wet and dry seasons.

We first conduct exploratory data analysis to test normality and autocorrelation of variables used in the model.ACF plots and QQ plots are used to test autocorrelation and normality respectively.

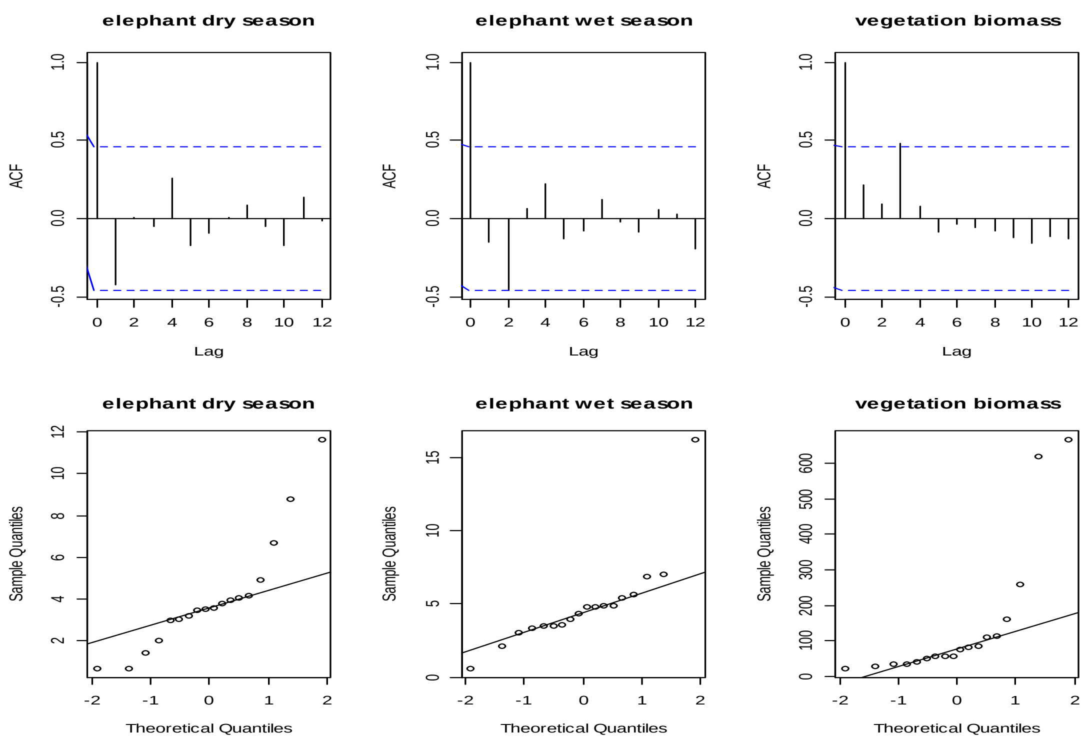


Figure 4.5.1: ACF and QQ plots for testing normality in elephant densities and vegetation biomass data.

There was evidence of autocorrelation in vegetation biomass data. Also, vegetation biomass, dry and wet season elephant densities were not normally distributed.

4.6 Spatial generalized least square model

A spatial generalized least square model was used to test whether elephant densities during wet and dry season had significant effect on vegetation. Grid coordinate system was used as the spatial aspect.

The general spatial least square model can be expressed as

$$y_i = \alpha + \beta x_i + \varepsilon_i \qquad (4.6.1)$$

$i = 1, \ldots, N$

$E(\varepsilon) = 0$

The auxiliary regression is given by

$$z_i^2 = \emptyset + \delta x_i + v_i \qquad (4.6.2)$$

$$s^2 = \frac{\sum \hat{u}_i^2}{N}$$

$$z_i^2 = \frac{\sum \hat{u}_i^2}{s^2}$$

Spatial auto regression

The spatial auto regression model can be expressed as

$$y = \rho W y + x\beta + \varepsilon \qquad (4.6.3)$$

Where

$y = n$ by 1 vector of observations on the dependent variable

$w = n$ by n spatial weights matrix that formalizes

ρ = spatial auto regressive parameter

X= n by k matrix of observations on the exogenous variables, with an associated k by 1 regression coefficient vectorβ

ε = a vector or random error terms

The log-likelihood for the spatial auto regression can be estimated by;

$$\text{In } L = -\ (1/2\sigma^2)\ (y\text{-}\rho Wy\text{-}X\beta)'(y\text{-}\rho Wy\text{-}X\beta) \qquad (4.6.4)$$

The least square estimator and variance for spatial auto regression can be estimated by;

$$\hat{\beta}_{ML}=(X'X)^{-1}X'(y-\lambda Wy)$$

$$\hat{\sigma}_{ML}=\frac{(e_0-\rho e_L)'\,(e_0-\rho e_L)}{N}$$

Where

$$e_0=y-x\hat{\beta}_0$$

$$e_{L=}y-x\hat{\beta}_L$$

We present our model as

$$Vegetation\ biomass(y_i)=\beta_1 Xi_1+\beta_2 Xi_2+\varepsilon_i \qquad (4.6.5)$$

Y_i=Vegetation biomass

X_1=Elephant wet season densities in grid i

X2=Elephant dry season densities in grid i.

	$Estimate$	$Std.Error$	$t-Value$	$Pr(>\|t\|)$
Elephants wet season densities	-16.31656	10.16184	-1.605670	0.1279
Elephants dry season densities	47.13896	12.70136	3.711332	0.0019

Table4.5.2: Generalized least square model for elephant spatial density distribution.

	$F - Value$	$Pr(> \|t\|)$
Elephants wet season densities	0.782412	0.3895
Elephants dry season densities	12.243540	0.0030

Table 4.5.3: ANOVA table for elephant spatial density distribution model

During dry season elephant densities have a significant effect on vegetation biomass (Table 4.4) but during wet season they their movement within Amboseli do not significantly influence vegetation change.

The spatial elephant distribution model equation was found to be

Vegetation biomass(yi)=$-16.31656X_1$+$47.13896X_2 + \varepsilon$

Vegetation biomass increase during wet season but decrease during dry season. During dry season a unit increase in elephant densities results to 16.3% decrease in vegetation biomass but during wet season vegetation biomass increases by 47.1% when elephant densities increase by a single unity

4.7 Density distribution maps for Elephants over time.

Density distribution maps are essential in determining the movement patterns and the level of resource utilization over time based on species densities on specific grids. They can also be useful in tracking species especially in unprotected areas and map generalization which act on a group of elephants or any other species at meso level or a whole thematic class at micro level. Plotting of geographical distribution maps for elephants can be done using ArcView, ArcGIS, R and Quantum gis tools.

Density distribution maps for elephants were generated using Quantum gis 1.7.0 and arc view 3.2 a.Elephant maps were plotted based on graduated sizes of their densities. Dots represent densities per grid, maps portray the geographical distribution of discrete phenomena using an arrangement of identical point symbols. This technique is particularly useful for understanding global distribution of the mapped elephant densities and comparing relative densities of different grids in the map. The aim was to determine whether their distribution within the park and outside the park varies from time to time based on decades and seasonality.

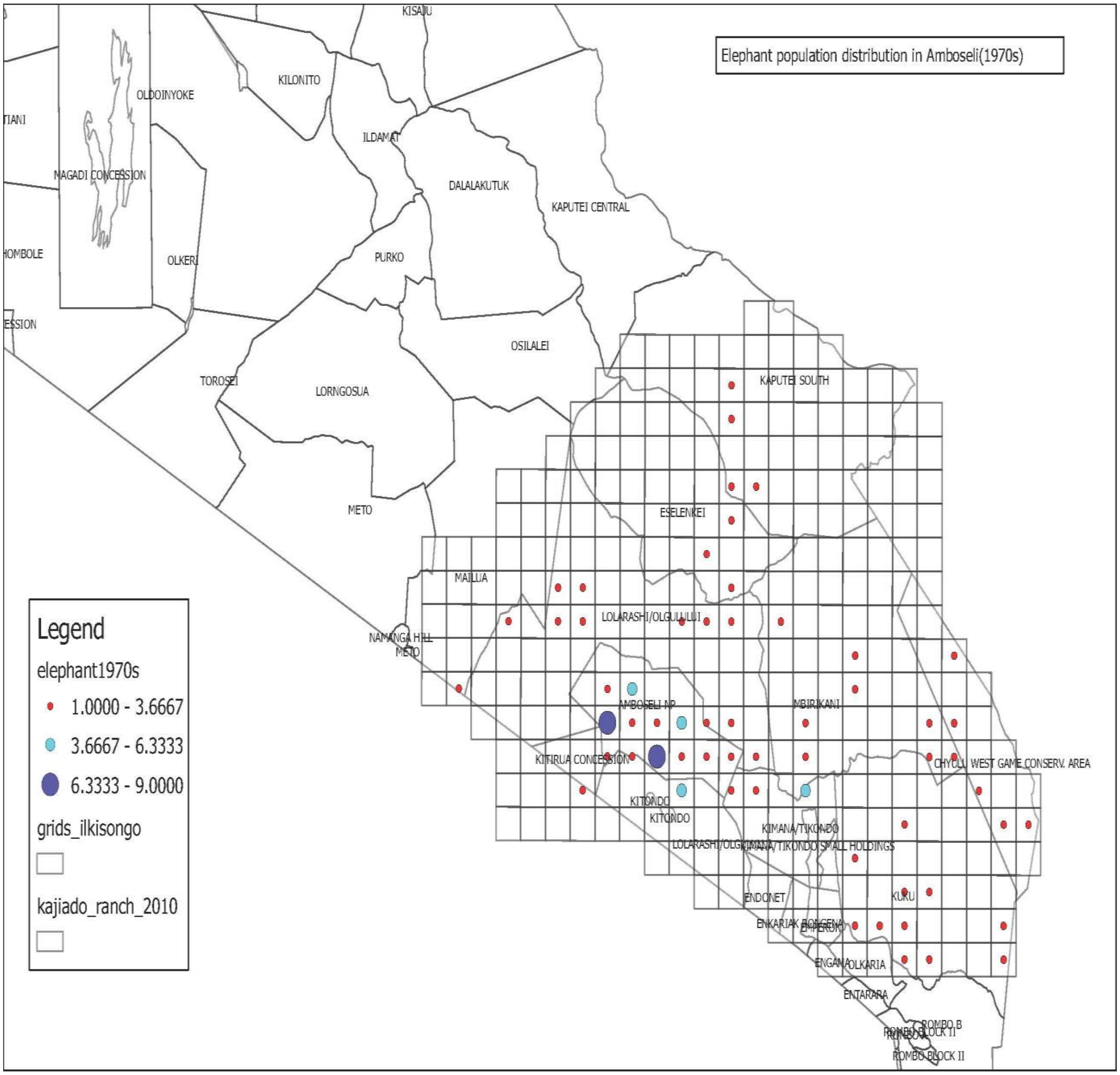


Figure 4.6.1: Elephant population distribution for 1970 decade

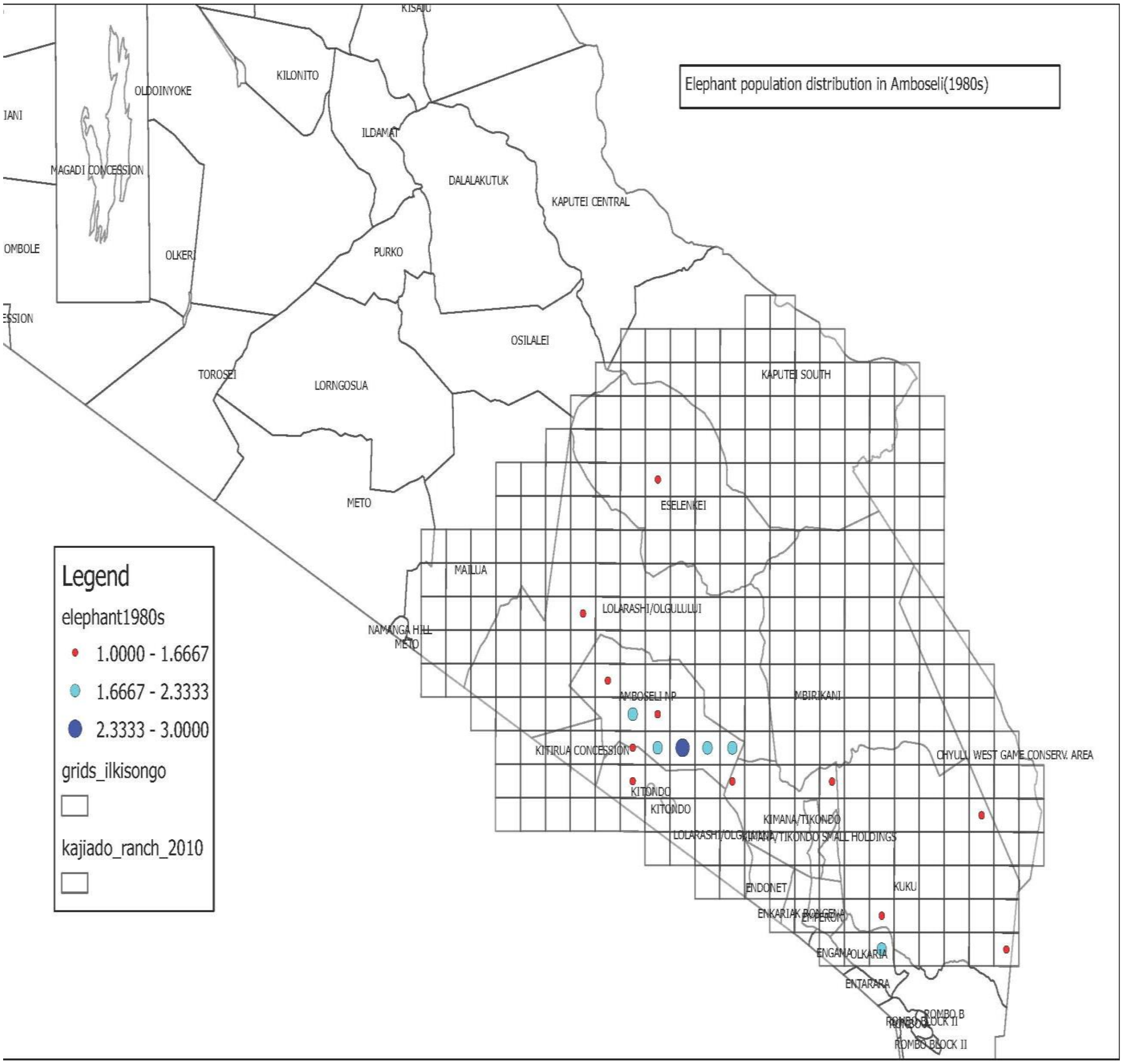


Figure 4.6.2: Elephant population distribution for 1980 decade

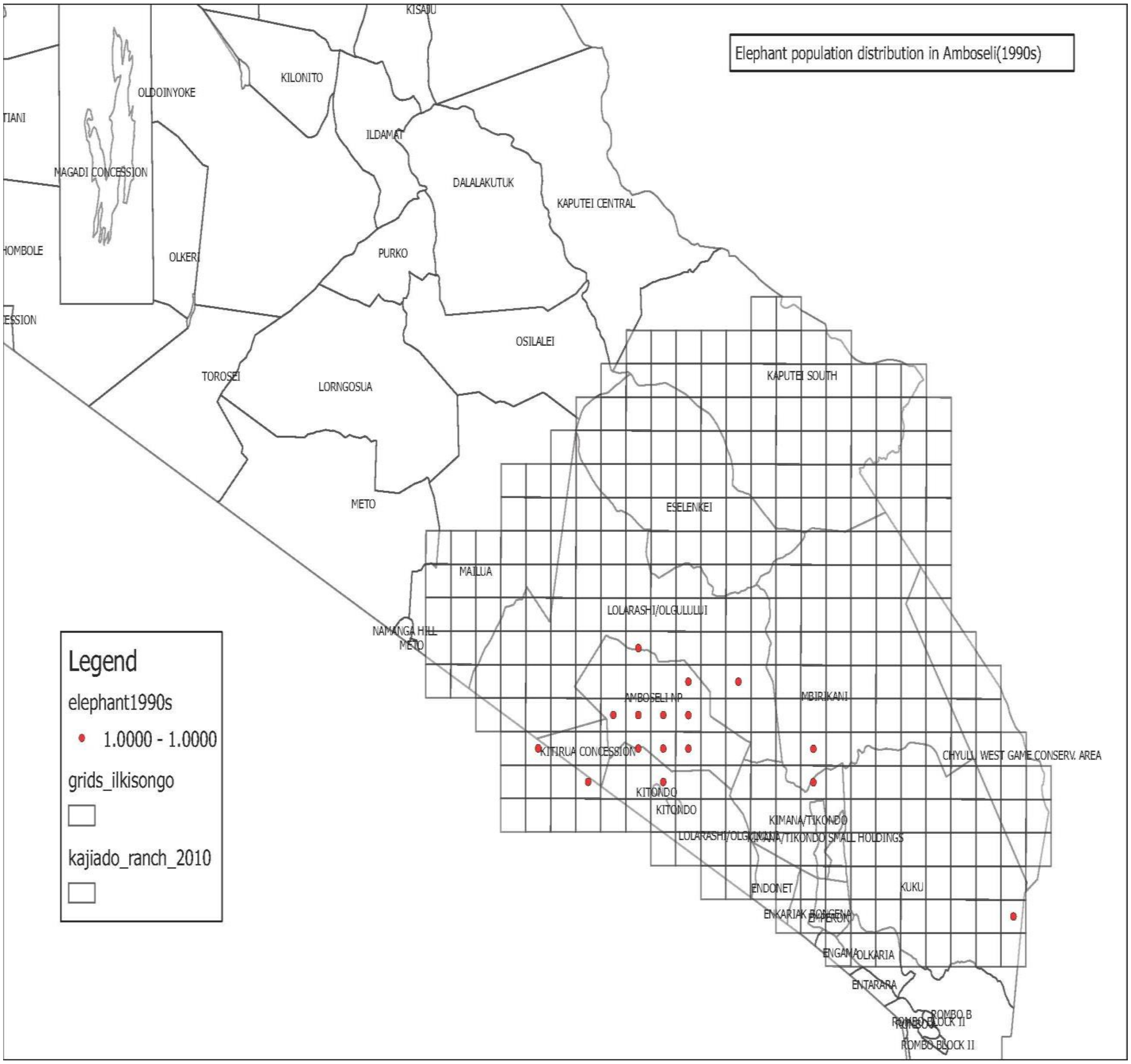


Figure 4.6.3: Elephant population distribution for 1990 decade

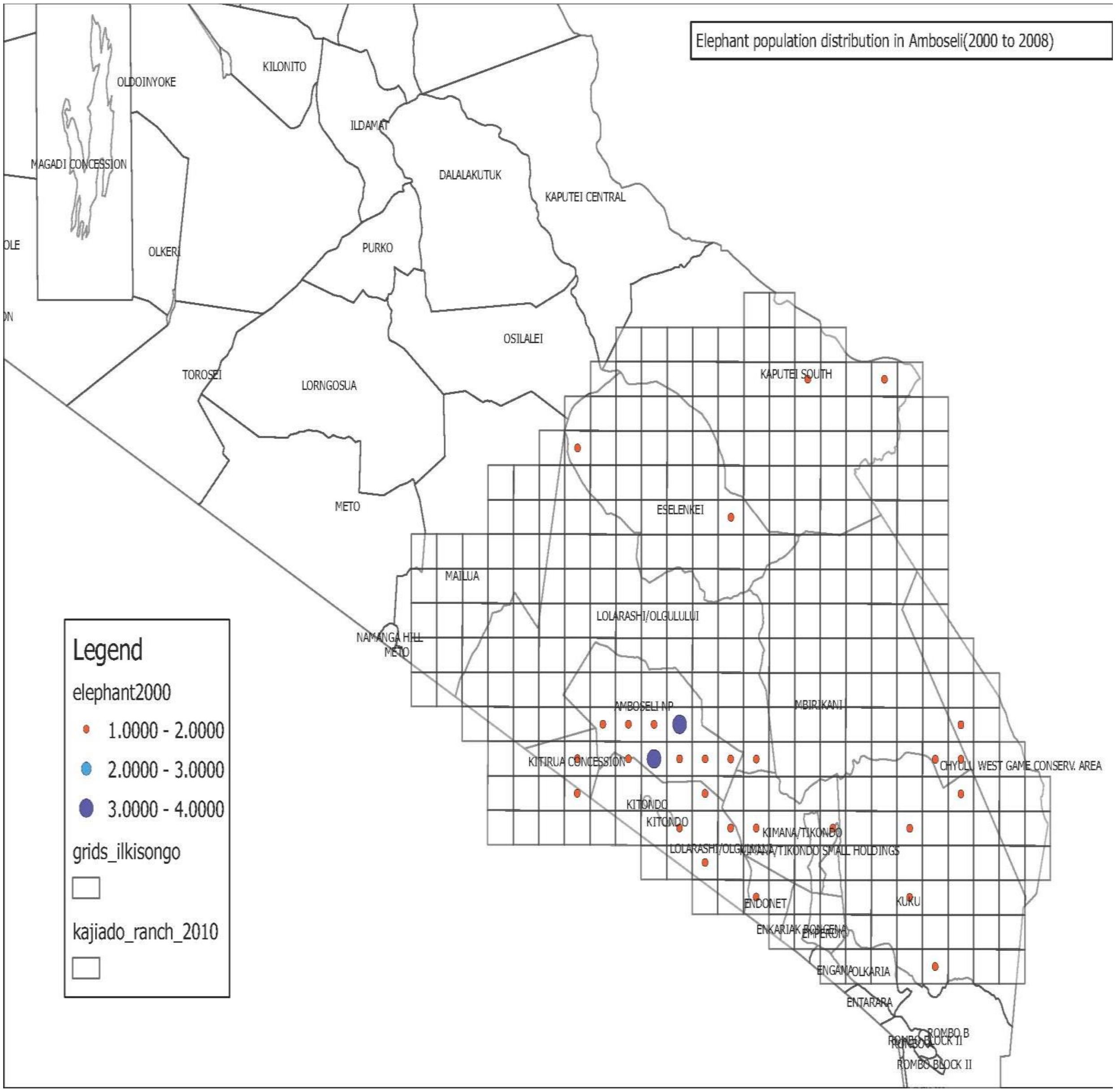


Figure 4.6.4: Elephant population distribution for 2000 decade

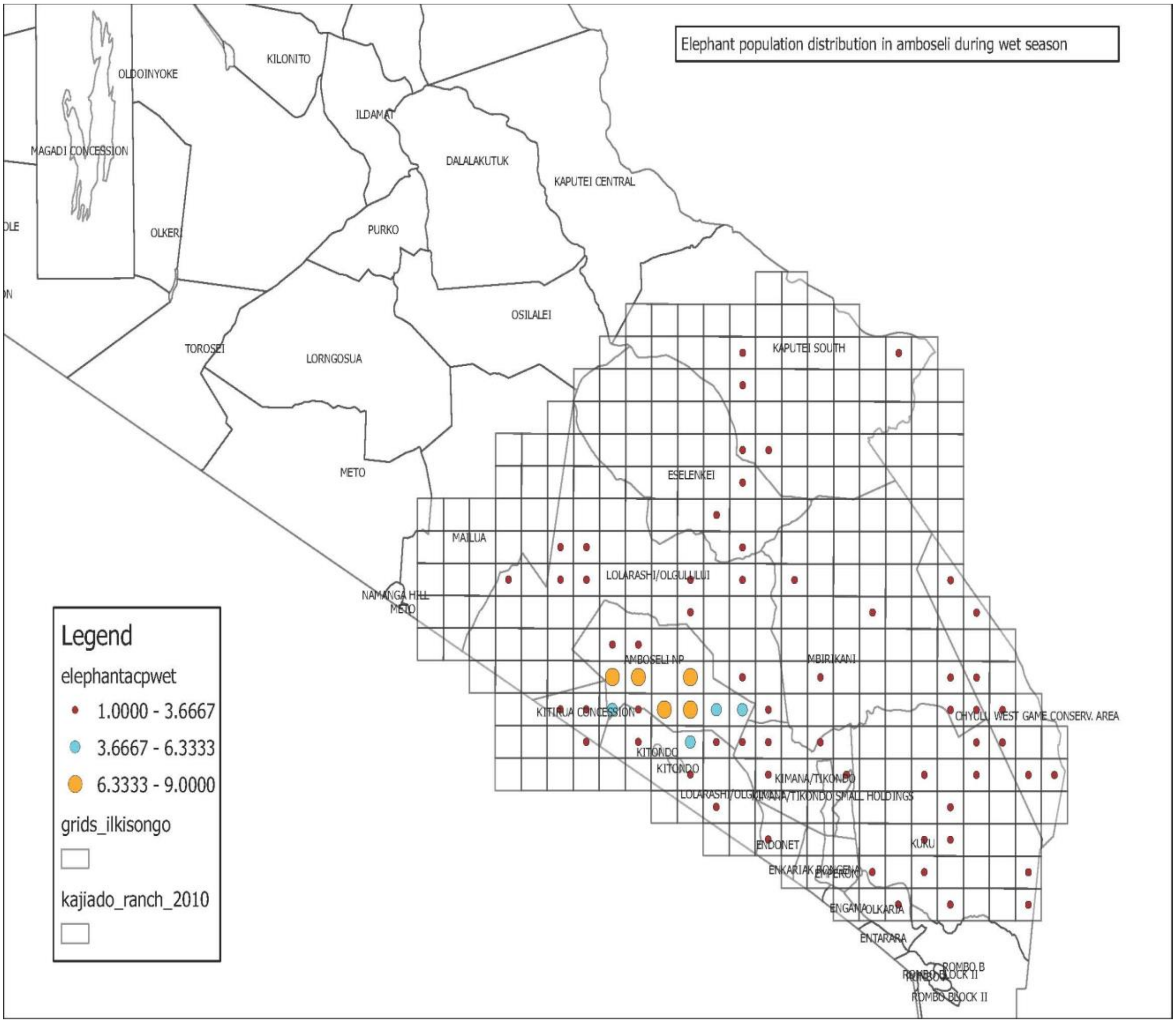


Figure 4.6.5: Wet season density distribution

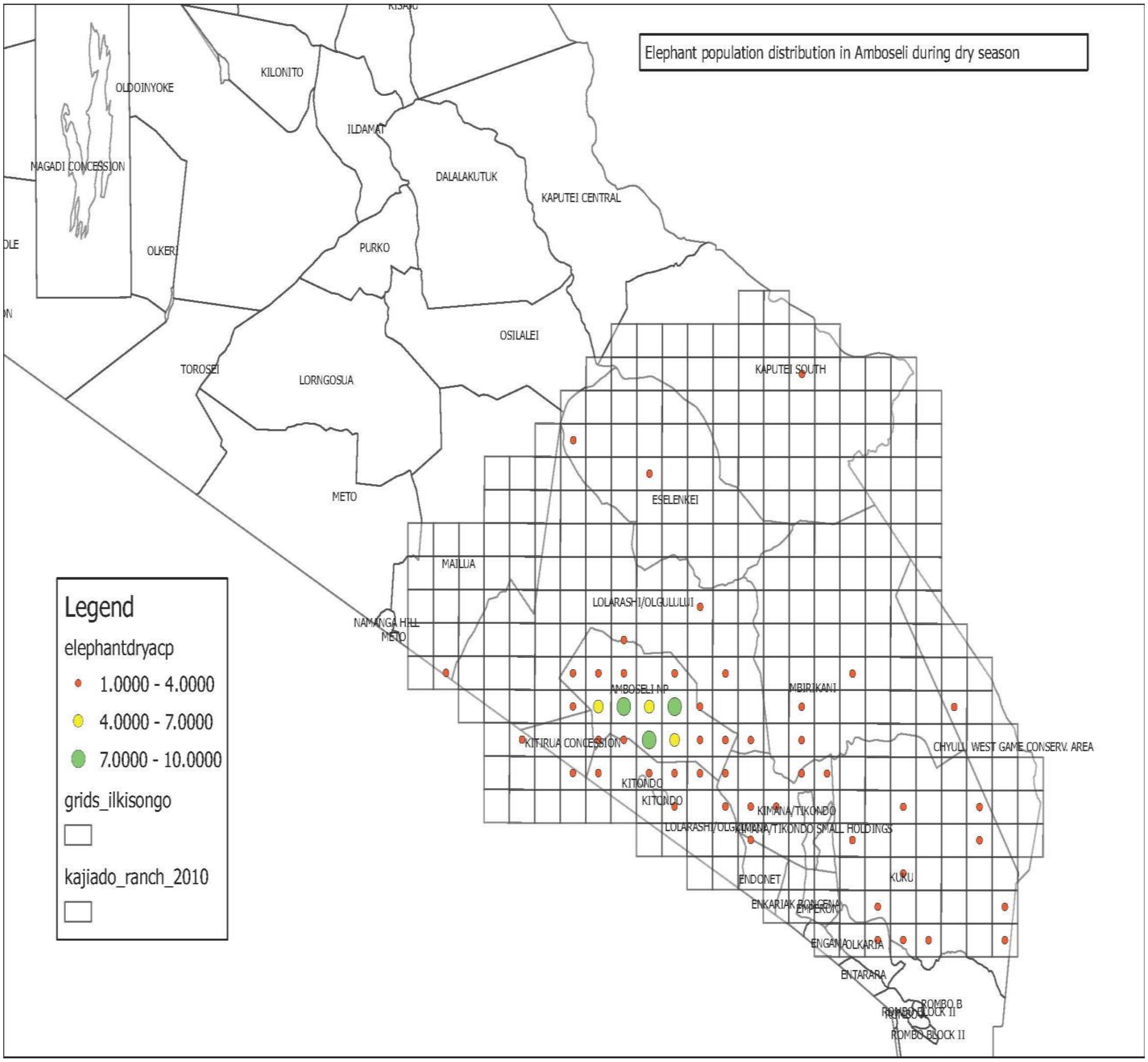


Figure 4.6.6: Dry season density distribution

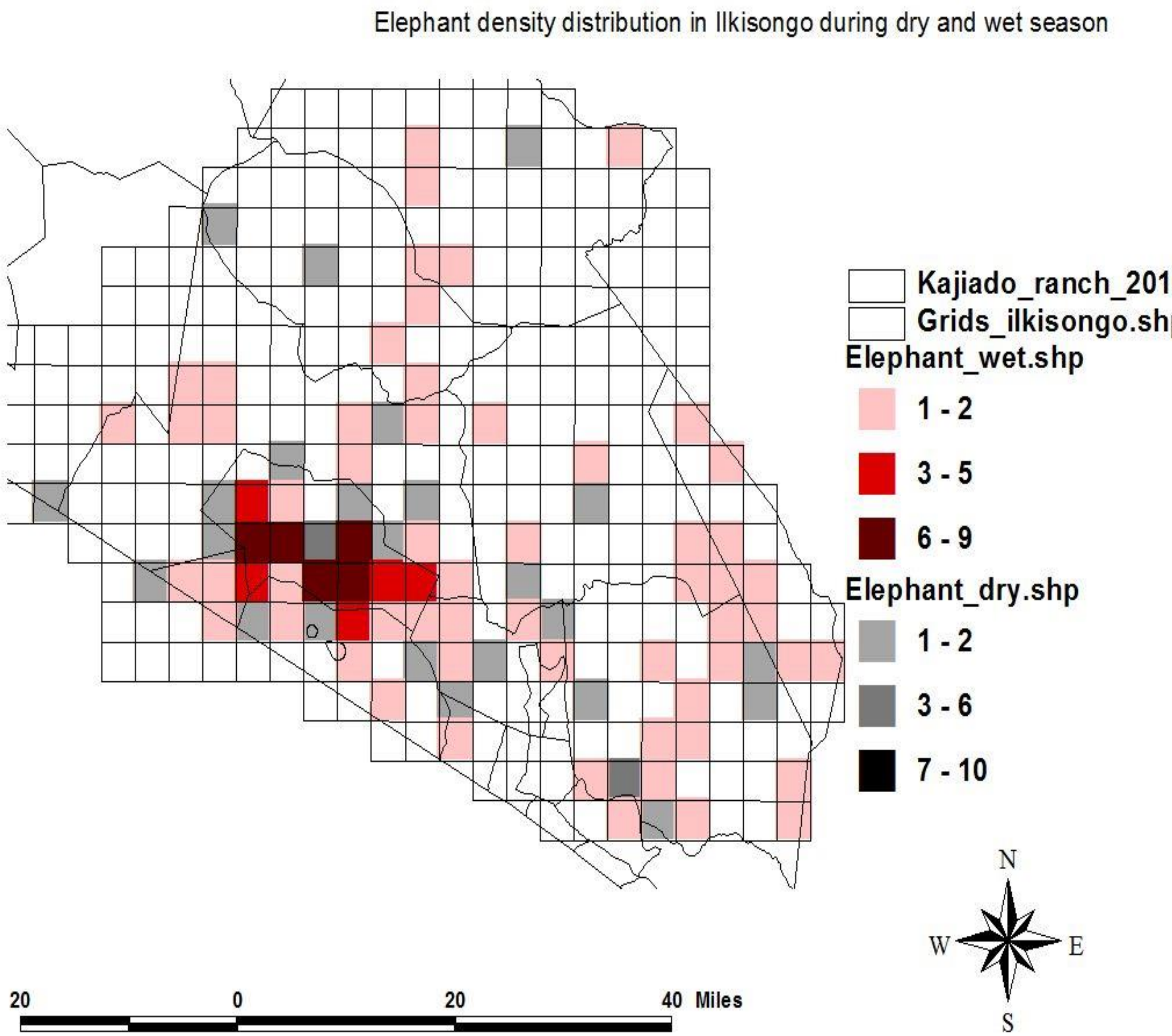


Figure 4.6.7: Elephant population distribution in Amboseli during dry and wet seasons

4.8 Summary of results

The generalized least square model was used to determine whether NDVI, human settlements and vegetation biomass have significant effects on elephant densities. Our findings are consistent with previous studies. NDVI was found to be significant (t_{97}=8.601912, p=0.000), human settlements were also significant (t_{97}=-3.11, p=0.003) while vegetation biomass was significant (t_{97}=2.028420, p=0.0453). A unit increase in vegetation biomass results to 1% increase in elephant density. Human settlements also affects elephant density by 51.4%.The model also showed that a unit increase in NDVI results to 12.9% increase in Elephant density.

The model equation can be expressed as;

$$Elephantdensity(y)=0.0112275X_1 - 0.5136491X_2 + 0.1296039X_3 + \varepsilon$$

The ANOVA test was used to determine the significance of parameters in the generalized least square model. The results show that all the parameters were significant in the model, vegetation biomass (F_{97}=1050.9808, p=0.0001), human settlement (F_{97}=57.5295, p=0.0001), and NDVI (F_{97}=71.7045, p=0.000)

During dry season elephant densities were highest inside the park (Figure 4.6.6). During wet season, although elephant densities were highest in the park, some of the elephants were able to move outside the park (Figure 4.6.5). Elephants are mixed feeders and due to their big size their distribution is motivated by availability of resources. Outside the park there is low vegetation production and low water availability. Due to plenty of food in and outside the park during wet season their distribution is also even but during dry season they are mainly within the park because of the availability of the swamp and the fact that food is limited.

The decadal spatial density distribution showed that elephant distribution was even in 1990sbut the highest concentration was found inside the park in the decades of 1970s, 1980s and 2000s.

CHAPTER FIVE: DISCUSSION, CONCLUSIONS AND RECOMENDATIONS

The main aim of this study was to establish whether human settlements, NDVI and vegetation biomass have significant impact on elephant densities, the result from a generalized least square model confirmed that all variables in the model greatly influence the survival and movement of elephants within and outside the park.

Elephants prefer woody vegetation to grass; vegetation serves as wildlife habitat and the energy source for the vast array of animal species on the planet and, ultimately, to those that feed on these. Vegetation is also critically important to the world economy, particularly in the use of fossil fuels as an energy source, but also in the global production of food, wood, fuel and other materials. Since woody vegetation is currently less diverse in Amboseli elephants feed on alternative forage (grass), which is the dominant pasture in the ecosystem. From the model, vegetation biomass was found to be significant in influencing the survival and movement of elephants, it’s therefore important for stake holders to come up with ways to increase woody vegetation diversity.

Also, the persistent increase of human settlement and land subdivision is worrying.Amboseli ecosystem is an important biological and economic resource for human beings, wildlife, livestock, and insects. Human beings are among the major causes for the loss of biodiversity in Savannah ecosystems through habitat change and destruction (Erhlich, 1988).They mainly convert forests to arable fields and rangeland for grazing of livestock. These land uses have a major effect on the distribution and abundance of world ungulates both at the community level and population level, mostly causing wildlife population decline.

Measures should be put in place to prevent human encroachment and particularly enlighten the communities living around the conservation areas on the importance of peaceful coexistence between human beings and wildlife.

Spatial movement of elephants in the basin confirms that Amboseli National Park is an important area for large mammals. Since their movement is majorly confined within the park, resources should be enough to sustain the ecosystem. Prudent measures should be undertaken to safeguard the ecosystem from further settlements and human-elephant conflicts.

It's evident from the spatial least square model that elephants significantly impact on vegetation biomass during dry season but there was no evidence of any influence on vegetation biomass during wet season.

Since elephants are the main cause of habitat change in Amboseli, a mathematical model should be developed to test whether effects of vegetation change have cascading effect on other species in the ecosystem.

REFERENCES

1. Amahowe I.O, Djossa A.B, Adomou A.C, Kabre A.T and Sinsin B. (2012). Vegetation structure and tree diversity in habitats exploited by elephants in Benin (West Africa).*Journal of Applied Biosciences, 58, 4223-4233.*
2. AWF technical handbook (1996).Studying Elephants.*ISBN: 9966-9915-0-6.*
3. Boer. W, Langevelde. F, Herbert.H.T, Ruiter. P, Blanc. J, Marc. V, Gaston. K, Hamilton. D (2013). Understanding spatial differences in African elephant densities and occurrence, a continent-wide analysis. *Biological Conservation, 159,468-476.*
4. Claudius A. D, Van.D, Charles A.F and Han. O (1999).Changes in the woody component of an East African savanna during 25 years. *Journal of Tropical Ecology, 15, 545–564.*
5. Fullman. T and Child. B (2012). Water distribution at local and landscape scales affects tree utilization by elephants in Chobe National Park, Botswana. *African Journal of Ecology, 51, 235-243.*
6. Gil. B, Pieter S.B, Shadrack M.N, Andrew. S and Hamilton. D (2014). Elephant movement closely tracks precipitation-driven vegetation dynamics in a Kenyan forest-savanna landscape. Movement Ecology.*Doi: 10.1186/2051-3933-2-2.*
7. Graham M.D, Hamilton. D, Adams. M.W and Lee. P (2008).The movement of African elephants in a human-dominated land-use mosaic. *Animal Conservation, 12, 445-455.*
8. Grant M. H, Gareth J. R, Rudi. v. A and Stuart L. P (2008). Rules of habitat use by elephants Loxodonta Africana in southern Africa*: insights for regional management. Oryx, 42, 66–75.*
9. Hoare. R and Toit. J (1999). Coexistence between People and Elephants in African Savannas. *Society for Conservation Biology, 13, 633-639.*
10. James P. D and Pettorelli. N (2012).Exploring the relationship between NDVI and African elephant population density in protected areas. *African Journal of Ecology, 50, 455-463.*
11. Kanisios. M, Gandiwa.E, and Kativu.S (2012).Impact of African elephants on Baikiaea plurijuga woodland around natural and artificial watering points in northern Hwange National Park, Zimbabwe. *International Journal of Environmental Sciences.ISSN:0976-4402.*

12. Merode.E.D, Hillman-Smith.K, Nicholas. A, Ndey. A and Likango. M (2000).The spatial correlates of wildlife distribution around Garamba National Park, Democratic Republic of Congo.*Int. J. remote sensing, 21, 2665–2683.*

13. Mose. V, Huu. T, Western. D, Auger. P and, Nyandwi. C (2013). Modelling the dynamics of migrations for large herbivore populations in the Amboseli National Park, Kenya. *Ecological Modelling, 254, 43-49.*

14. Nduati. E.W, Mundia. C and, Moses M (2013). Effects of Vegetation Change and Land Use/ Land Cover Change on Land Surface Temperature in the Mara Ecosystem. International Journal of Science and Research. *ISSN: 2319-7064.*

15. Omondi. P and Ngene. S (2012). The National Elephant Conservation and Management Strategy (2012–2021) at a Glance. The George Wright Forum, 29, 90-92.

16. Ozinga. T.J, Kleyer. W and Braak. T (2012). Selecting traits that explain species-environment relationships: a Generalized Linear Mixed Model approach. *Journal of Vegetation Science, 24, 988-1000.*

17. Porensky. L, Bucher. S, Veblen. K, Treydte. A, and Young. T (2013). Mega herbivores and cattle alter edge effects around ecosystem hotspots in an African savanna. *Journal of Arid Environment, 96, 55-63.*

18. Rood.E, Abdullah A. and Nijman. V (2010). Using presence-only modelling to predict Asian elephant habitat use in a tropical forest landscape: implications for conservation. *Journal of Conservation Biogeography, 16,975-984.*

19. Vance.A.E, Archie. E.A, Moss. C.J (2008), Social networks in African elephants. *Spring Science+Business Media, 15,273-293.*

20. Viljoen. P (1989).Spatial distribution and movements of elephants (*Loxodonta africana*) in the northern Namib Desert region of the Kaokoveld, South West Africa/Namibia. *Journal of Zoology, 219, 1-19.*

21. Western. D (2006), A half a century of habitat changes in Amboseli National Park, Kenya. *African Journal of Ecology*, ***45***, 302-310.

22. Western.D and Dunne. T (1979). Environmental Aspects of Settlement Site Decisions among Pastoral Maasai. *Human Ecology, 7, 1.*

23. Western.D and Lindsay. W. K (1984), Seasonal herd dynamics of a savannah elephant population. *African Journal of Ecology, 22,229-244.*

24. Western.D and Maitumo. D (2004).Woodland loss and restoration in a savanna park: a 20-year experiment. *African Journal of Ecology, 42,111-121*.
25. Young. H, Douglas J. M, Kristofer M. H, Jacob. R, Goheen, Pringle.R, Truman Young. P and Dirzo.R (2013). Effects of mammalian herbivore declines on plant communities: observations and experiments in an African savanna. *Journal of Ecology, 101*, *1030-1041*.
26. Young. K, Ferreira. S,and Rudolph J. V (2009).The influence of increasing Population size and vegetation productivity on elephant distribution in the Kruger National Park. *Austral Ecology, 34,329-342.*
27. Young .K and Rudolph J. V (2010). Density as an explanatory variable of movements and calf survival in savanna elephants across southern Africa. *Journal of Animal Ecology, 79, 662-673.*

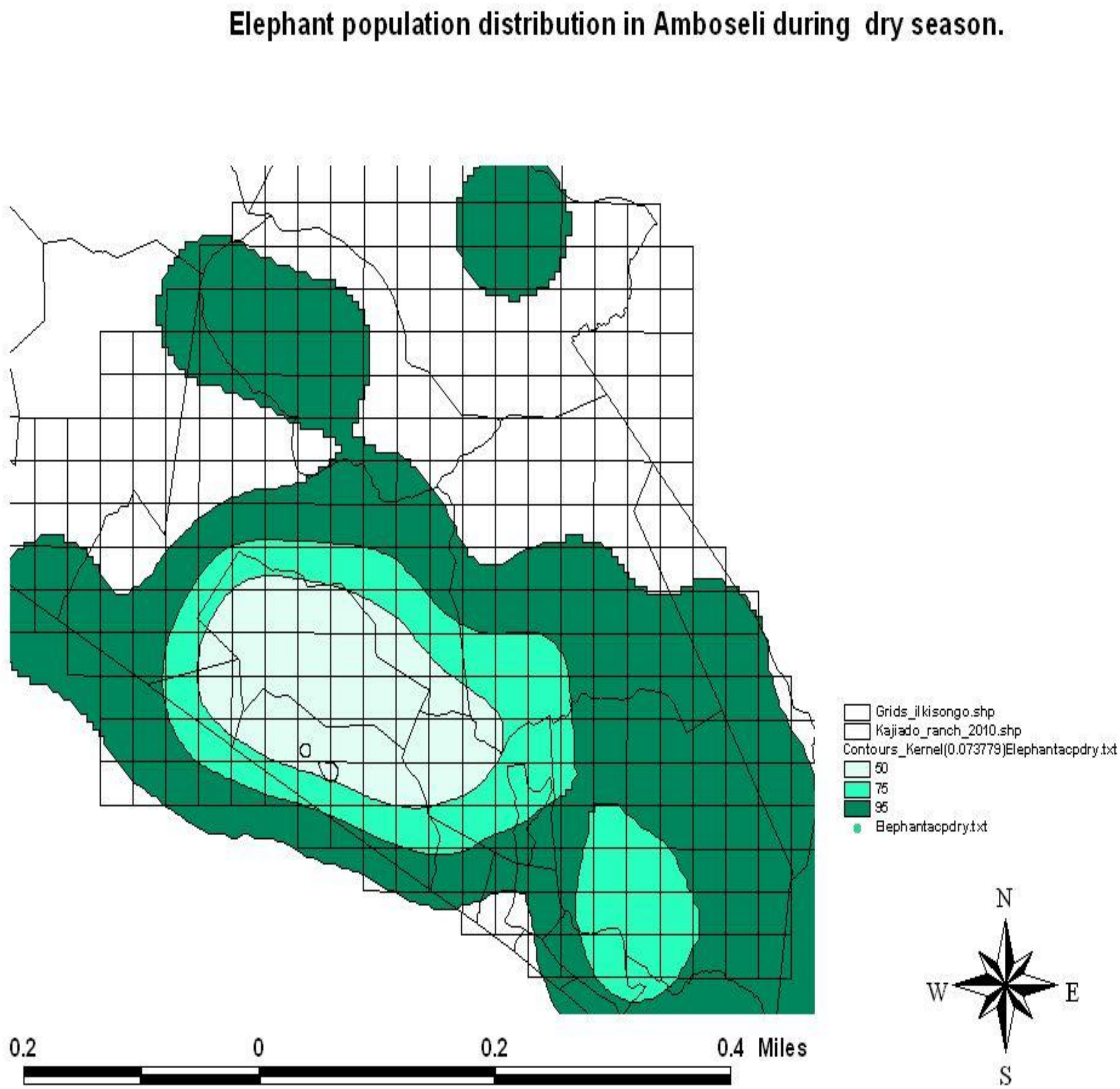
APPENDIX

R codes for generalized least square model and EDA.

```
setwd("C:/ERIC MBOYA_ACP_DATA/2014analysis/eric 2014 analysis/for thesis")

read.csv("Elephantveg1.csv",header=T)

elep<-read.csv("Elephantveg1.csv",header=T)

attach(elep)

par(mfrow=c(4,4))

################qqplots to test normality

qqnorm(Bushland,main="Bushland")

qqline(Bushland)

qqnorm(Grassland,main="Grassland")

qqline(Grassland)

qqnorm(PLAINS,main="Plains")

qqline(PLAINS)

qqnorm(Swamp,main="Swamp")

qqline(Swamp)

qqnorm(Woodland,main="Woodland")

qqline(Woodland)

qqnorm(ElephantBiomass,main="Elephant")

qqline(ElephantBiomass)

qqnorm(Settlements,main="settlements")

qqline(Settlements)

##############ACF plots to test for autocorrelation

acf(Bushland,main="ACF plot of Bushland")
```

```
acf(Grassland,main="ACF plot of Grassland")

acf(PLAINS,main="ACF plot of Plains")

acf(Swamp,main="ACF plot of Swamp")

acf(Woodland,main="ACF plot of Woodland")

acf(ElephantBiomass,main="ACF plot of Elephants")

acf(Settlements,main="ACF plot of settlements")

###############generalized least squares model

read.csv("model.csv",header=T)

mod.gls=read.csv("model.csv",header=T)

attach(mod.gls)

library(nlme)

mod.gls<-gls(Elephant~vegetation.biomass-1+settlement+ ndvi,correlation=corARMA(p=4), method="ML")

summary(mod.gls)

anova(mod.gls)

library(lmtest)

dwtest(Elephant~vegetation.biomass+settlement+ ndvi)

###############test for normality

shapiro.test(Elephant)

shapiro.test(vegetation.biomass)

shapiro.test(settlement)

shapiro.test(ndvi)

###############test for heteroskedasticity

library(lmtest)

bptest(Elephant~vegetation.biomass+settlement+ ndvi)
```

Density Distribution maps for elephants



Elephant population distribution in Amboseli during wet season.