THE APPLICATION OF ECOLOGICAL NICHE MODEL TO MAP OUT THE RIFT VALLEY FEVER RISK AREAS IN KENYA

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF REQUIREMENTS FOR MASTERS OF SCIENCE IN VETERINARY EPIDEMIOLOGY AND ECONOMICS DEGREE OF UNIVERSITY OF NAIROBI

DR. PURITY NKIROTE KIUNGA (B.V.M. UON)

DEPARTMENT OF PUBLIC HEALTH, PHARMACOLOGY AND TOXICOLOGY

2015

DECLARATION

I hereby declare that this thesis is my original work and has not been presented for a degree in any other University.

DR. PURITY NKIROTE KIUNGA (BVM, UON)

Sign……………………………........ Date:……………............................

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

PROF. KITALA P.M. (B.V.M., MSC, PhD)

Sign……………………………........ Date:……………............................

Department of Public Health, Pharmacology and Toxicology, Faculty of Veterinary Medicine, University of Nairobi

DR. BERNARD BETT (B.V.M., MSC, PhD)

Sign……………………………........ Date:……………............................

International Livestock Research Institute (ILRI)

DEDICATION

I dedicate this thesis to my husband Peter Omwenga and our daughters Cheryl Mong'ina and Chanel Moraa, for their endless patience, love and support not forgetting my parents Thomson and Dorcas Kiunga for instilling in me the spirit of hard work and for being proud of me.

ACKNOWLEDGEMENTS

I wish to acknowledge the University of Nairobi, International Livestock Research Institute (ILRI), for availing the opportunity, the funds and resources that were necessary for the completion of the Master’s Degree.

Special thanks go to my supervisors Prof. Kitala P.M. and Dr. Bernard Bett for their encouragement, mentorship and offering excellent guidance and tremendously useful suggestions and corrections.

The people who have inspired my passion to do my studies are too numerous to mention without fear of leaving someone out. I will single out a few who set me on this path and those who offered wisdom, timely insights along the way: My family; Parents Thompson and Dorcas Kiunga, Siblings; Edna, Benjamin, Ruth, Mercy, Mosses, Candy and Chelsea, my friends and classmates for their encouragement and prayers when things got tough.

Finally I sincerely thank my long-time friend and husband Peter Omwenga for financial support and our two daughters Cheryl and Chanel Omwenga.

May God bless you all in your endeavours.

TABLE CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

%:	Percentage
AUC:	Area under Curve
B.V.M:	Bachelor of Veterinary Medicine
DEM:	Digital Elevation Models
Dr:	Doctor
DVS:	Department of Veterinary Services
ECMWF:	European Centre for Medium-Range Weather Forecasts
ELISA:	Enzyme-linked immunoassay
ENSO:	*El Niño*/*Southern Oscillation*
FAO:	Food and Agriculture Organization of the United Nations
FEWS NET:	Famine Early Warning Systems Network
G.o.K:	Government of Kenya
GARP:	Genetic Algorithm for Rule set Production
GDP:	Gross Domestic Product
HWSD:	Harmonized World Soil Database
IIASA:	International Institute for Applied Systems Analysis
ILRI:	International Livestock Research Institute
Km^2:	Kilometre Squared
KNBS:	Kenya National Bureau of Statistics
MBG:	Model Based Geostatistic
MSc:	Masters of Science
NDVI:	Normalized Difference Vegetation Index

PhD:	Doctor of Philosophy
RVF:	Rift Valley Fever
RVFV:	Rift Valley Fever Virus
SDM:	Species Distribution Model
SNS:	Smithburn Neurotropic strain
SRTM:	Shuttle Radar Topographic Mission
SSA:	Sub-Saharan Africa
UoN:	University of Nairobi
WHO:	World Health Organization

ABSTRACT

Rift Valley fever (RVF) is an acute, mosquito-borne zoonotic viral disease of economic importance caused by a virus of the *Phlebovirus* genus, *Bunyaviridae* family that mainly affects ruminants and humans. It causes abortion in gravid animals and high mortality in young animals, characterized by massive hepatic necrosis and pantropic haemorrhage. Rift Valley fever-like disease in livestock was first reported in Kenya in 1912. Numerous studies have shown close relationship between climatic conditions and outbreaks of Rift Valley Fever. *Aedes* and other mosquito species such as *Culex* are the vectors responsible for the disease transmission in both animals and humans.

Various studies carried out to map RVF distribution using a variety of approaches including the use of disease occurrence maps, statistical models which uses presence and absence data e.g. the logistic regression method. However, acquiring correct absence data is not easy and hence maps generated from standard statistical models might not be a true representation of the disease distribution. In this study ecological niche modeling (ENM) was used to model the supporting niche of RVF and determine the distribution of RVF in Kenya using Genetic Algorithm for Rule set Production (GARP) and Random Forest (RFs) which are programs that use presence-only data.

The data were collected at two levels; primary and secondary data collection. For primary data it was acquired by using Global Positioning System (GPS) for georeferencing and also through questionnaire administration to specific farmer affected by RVF in the RVF hotspot areas as per the records obtained from the Director of Veterinary Services (DVS). Secondary data collection included environmental variables which were used as the input data. They included: land use, soil type, elevation, vegetation index (obtained after downloading from

Moderate Resolution Imaging Spectroradiometer (MODIS) satellite spanning from October 2006 to March 2007), rainfall and temperature for the same period of time as the satellite imagery. Of the sampled data ENM was done using Bioclim, GARP and RFs mainly for comparison purposes. In GARP, 70% was used to train the model and 30% to test the model. A parallel analysis that used logistic regression model was done to identify statistical relationships between predictors used in the ENM model and the outcome. This is because ENM are good for prediction but not for analyzing mathematical relationships between variables. The results showed factors that were significant at 95% confidence interval for the outbreak of RVF were; open to closed forests having a crude OR of 1.93, Solonetz soil type having OR of 1.6 and NDVI having OR of 4.66. A one unit increase in temperature decreases the risk of RVF by 10%, and a change in altitude from $\leq$500 to 500 - $\leq$1000 is associated with 94% decrease in outbreak of RVF.

Analysis of the questionnaire data showed that 27.38% of the areas visited had human cases of RVF. The key livelihood activities were: crop farming (contributing 30%) and livestock keeping (35%). The result from ENM mapped the expected distribution of RVF in Kenya. The model was evaluated using the Area Under Cover (AUC) statistic and partial Receiver Operating Characteristic (pROC). The estimates generated from GARP were 0.82 for AUC and 1.77 for pROC respectively indicating that the model predicted the RVF distribution satisfactorily. The results will be used to improve the already existing maps and for better planning of mitigation measures.

CHAPTER ONE

INTRODUCTION

For the 70 percent of the world's poor who live in rural areas, agriculture is the main source of income and employment. But depletion and degradation of land pose serious challenges to producing enough food and agriculture products to sustain livelihoods. In sub-Saharan Africa 63% of the population are in the rural areas (World Bank, 2015). Kenya is one of the countries in sub-Saharan Africa (SSA) where the agricultural sector accounts, on average, for close to 26% of total gross domestic product (GDP) and about 60% of the region's total work force (Food and Agriculture Organisation, 2014; World Bank, 2015).

In Kenya livestock subsector is the core source of livelihood for the majority of the rural population especially in the arid and semi-arid lands (ASALs) and employs about 50% of the Kenya's agricultural labour force (KNBS, 2015), and about 80% of Kenya's land area is arid and semi-arid land and holds over 50% and 58%a of the country's large and small ruminants respectively (KNBS, 2015) which are at a risk of getting Rift Valley Fever (RVF).

These livestock play an important role both at the national and household levels and contributes to 10% (Ksh. 79 billion) of the gross domestic product and depletion and degradation of land due to climate change is a challenge and will affect livestock production. For instance, in 2014, the Agricultural sector in Kenya recorded mixed performance mainly attributable to erratic rains with some regions experiencing depressed rainfall. The lower levels of rainfall resulted in a decrease in pasture regeneration for livestock (KNBS, 2015).

This in connection with disease outbreak can be a very big risk to the livestock sector in the country, thus the importance of mapping the distribution of RVF in Kenya.

Rift Valley Fever is an acute, mosquito-borne viral disease that mainly affects ruminants and humans; it causes abortion and high mortality in young animals. It is also characterized by massive hepatic necrosis and pantropic haemorrhage (Martin, 2008) and thus it is of economic importance in Kenya. As a result it is paramount to know RVF distribution in the country to help in planning and assessment of mitigation measures. In Kenya, RVF-like disease in livestock was first reported in 1912 (Anonymous, 1910; Montgomery *et al*., 1912). They reported an acute and highly fatal disease of lambs on a government farm at the Naivasha area in Rift Valley Province of Kenya. The virus was however isolated and recognized 20 years later in 1931 (Daurbney *et al.*, 1931) confirming the presence of the disease in Kenya.

Numerous studies have shown a close relationship between high and persistent precipitation and outbreaks of RVF. Floodwater *Aedes* spp and other mosquito species such as *Culex* spp are responsible for the transmission of the virus mainly in animals; people often get infected by coming into direct contact with infected animal tissues or fluids. Outbreaks of RVF have also been associated with several risk factors which include: soil types (solonetz, luvisols, vertisols and calcisols), *El Niño*/*Southern Oscillation* (ENSO) leading to extreme increase in precipitation that is above average rainfall resulting to hydrographical modifications/flooding in (‘dambos’, dams, irrigation channels), dense vegetation cover with Normalized Difference Vegetation Index (NDVI) of at least 0.1 units sustained for at least 3 months, altitude of less than 1100 m above sea level (Linthicum *et al*., 1999; Anyamba *et al*., 2009; Hightower *et al.*, 2012; Bett *et al*., 2013). Climate change is therefore likely to influence the risk of the disease by altering the frequency of occurrence of extreme events such as the ENSO weather phenomenon (Martin, *et al*., 2008).

Anthropogenic land use practices alter ecosystems and their ability to control infectious diseases. One mechanism that has been hypothesized is that land use changes cause a reduction in biodiversity and hence a decline in the population of animals that would act as dead-end hosts for infectious pathogens. Affected ecosystems also would lack the capacity to control other disasters/shocks such as floods, soil erosion among others (IPCC, 2007).

This study uses ecological niche model to determine the distribution of RVF in Kenya. The Genetic Algorithm for Rule set Production (GARP) and Random Forest (RFs) algorithms were used because they are suitable for analyzing presence-only data.

1.1 Overall Objective

To modify the existing Rift Valley Fever risk map in Kenya using the ecological niche model.

1.2 Specific Objectives

1. To map out the distribution of Rift Valley Fever risk areas in Kenya using ENM
2. To determine environmental and climatic factors associated with the occurrence of Rift Valley Fever in Kenya

1.3 Justification

Rift Valley Fever is a disease of economic importance in that it causes a lot of losses in terms of mortality and morbidity. It also causes huge economic losses due to quarantine and closure of livestock markets which is the major source of livelihood to a larger population of the country, particularly in the arid and semi-arid areas. Various studies have been carried put to map RVF distribution and predict its future occurrence using standard models, for example logistic regression model. This approach requires both presence and absence data that are not always available because surveillance systems are mostly geared towards identifying outbreaks and not proving the absence of the disease. This study used ecological niche model, which require presence-only data. Such data are available from both the Department of Veterinary Services (DVS) and the Department of Disease Surveillance and Response (Ministry of Health). A refined risk map would be used by decision makers as a tool for targeting interventions, assessing effectiveness of response and for estimating spatially-explicit indices of vulnerability for the disease. It could also be overlaid with the global prediction systems, for example those developed by NASA, to help ground their predictions to real geographical areas in the target area.

CHAPTER TWO

LITERATURE REVIEW

2.1 Rift Valley Fever Disease

2.1.1 Background and causative Agent of the disease

Rift Valley Fever is a mosquito-borne viral zoonotic disease caused by Rift Valley Fever Virus (RVFV) belonging to the family *Bunyaviridae* and genus *Phlebovirus* which primarily affect domestic livestock (Daubney *et al.*, 1931). RVF was first reported among livestock at Lake Naivasha in Kenya in 1912 (Montgomery *et al*., 1912). The outbreak occurred after the introduction of European stock in Africa, but twenty years later the virus was isolated and characterised (Daubney *et al*., 1931) and the disease therefore acquired its name after its endemic location -- the Great Rift Valley -- in Kenya.

The disease has been associated with ENSO (Linthicum *et al.*, 1990) and it is shown to occur in cycles of 5 to 15 years usually following high rainfall resulting to flooding (Davies *et al*., 1980, Linthicum *et al*., 1999). Flooding results from persistent rainfall and accumulation of standing water masses in ‘dambos’ (shallow depressions in arid and semi-arid areas). These dambos get colonized by RVF infected mosquito which hatched from infected eggs. The massive infected mosquitos’ population bites animals that graze or take water from these sites (Davies *et al*., 1980).

2.1.2 Geographical distribution of RVF

RVF outbreaks have occurred in various countries in the sub-Saharan Africa and Madagascar, Saudi Arabia and Yemen. The specific countries that have reported outbreaks in sub-Saharan Africa include Kenya, Somalia, Tanzania, Zimbabwe, South Africa, Egypt, Mauritania, and Senegal (El Akkad, 1978; Saluzzo *et al.*, 1987; Meegan, 1988, Zeller *et al.*, 1997; Abdo Salem *et al.*, 2011; Madani *et al.*, 2003; Gerdes, 2004).

Figures 2.1 and 2.2 show locations where the disease has occurred in Kenya by Province and District while Table 2.1 shows where RVF outbreak has been reported between 1951 and 2006.

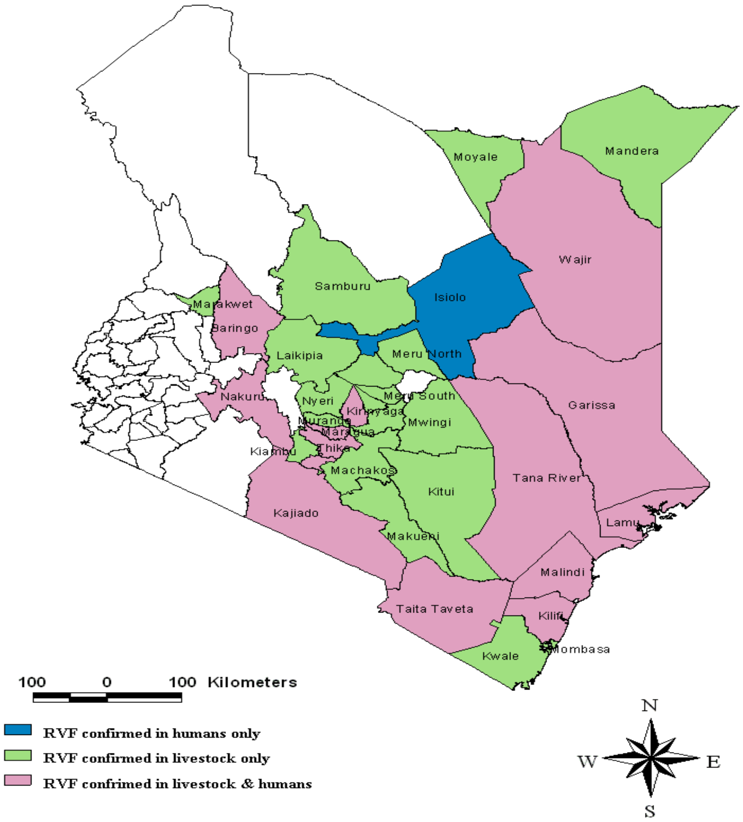

Figure 2.1: Rift Valley Fever Distribution map of Kenya (Source: Department of Veterinary Services Kenya)

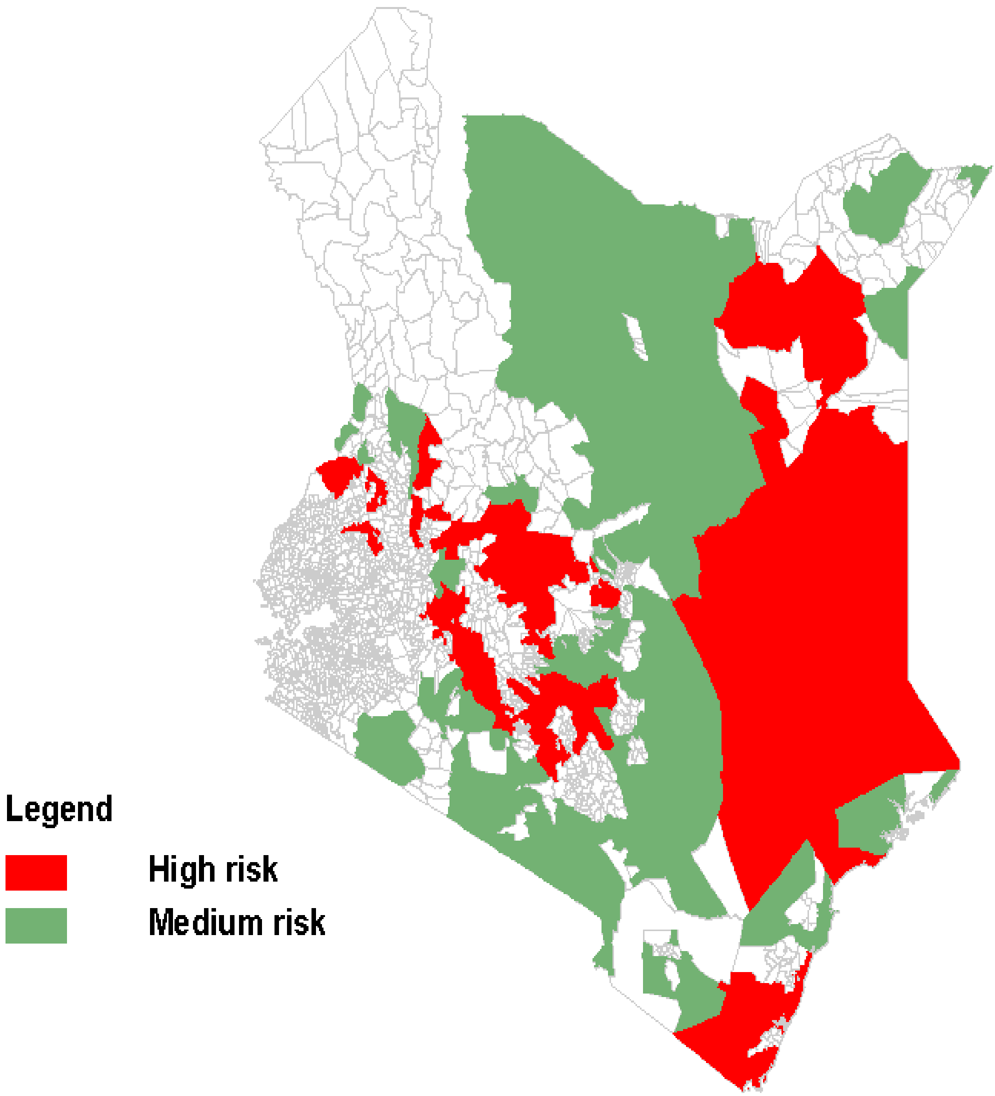

Figure 2.2:Rift Valley Fever risk map of Kenya (Source: Centers for Disease Control).

Table 2.1: Distribution of Rift Valley Fever outbreaks in Kenya by province and district, 1951-2006

District	Province	Year of first reporting RVF disease	Number of outbreak years after RVF introduction	No. of years involved in national outbreaks	Proportion of years involved in national outbreaks after RVF introduction
Nakuru	Rift Valley	1951 (1912)	23	21	91·3
Nairobi	Nairobi	1951	23	20	87·0
Thika	Central	1951	23	17	73·9
Kiambu	Central	1963	19	13	68·4
Maragua	Central	1951	23	15	65·2
Laikipia	Rift Valley	1951	23	15	65·2
Machakos	Eastern	1961	21	13	61·9
Uasin Gishu	Rift valley	1951	23	14	60·9
Nyeri	Central	1990	7	4	57·1
Kilifi	Coast	1961	21	12	57·1
Trans Nzoia	Rift Valley	1951	23	13	56·5
Kwale	Coast	1961	21	10	47·6
Mombasa	Coast	1977	12	5	41·7
Makueni	Eastern	1962	20	6	30·0
Kajiado	Rift Valley	1961	21	6	28·6
Isiolo	Eastern	1961	21	6	28·6
Garissa	Northeastern	1961	21	6	28·6
Wajir	Northeastern	1961	21	6	28·6
Mandera	Northeastern	1961	21	6	28·6
Narok	Rift Valley	1961	21	5	23·8
West Pokot	Rift Valley	1961	21	4	19·0
Tana river	Coast	1961	21	4	19·0
Marsabit	Eastern	1961	21	4	19·0
Kericho	Rift Valley	1968	16	3	18·8
Marakwet	Rift Valley	1981	10	1	10·0
Samburu	Rift Valley	2006	2	2	—
Baringo	Rift Valley	2006	2	2	—
Kirinyaga	Central	2006	2	2	—
Muranga	Central	2006	2	2	—
Malindi	Coast	2006	2	2	—
Kitui	Eastern	2006	2	2	—
Meru South	Eastern	2006	2	2	—
Meru Central	Eastern	2006	2	2	—
Tharaka	Eastern	2006	2	2	—
Mwingi	Eastern	2006	2	2	—
Taita Taveta	Coast	2006	2	2	—
Embu	Eastern	2006	2	2	—
Mbeere	Eastern	2006	2	2	—

(Source: Rift Valley fever in Kenya: history of epizootics and identification of vulnerable districts by Murithi *et al*., 2011)

2.1.3 Transmission of Rift Valley Fever disease

The virus is transmitted by diverse species of mosquitoes in different environments. However, in most of these areas, floodwater *Aedes* mosquito species (*A. mcinthoshi*) is thought to be the principal reservoir of the virus (Linthicum, 1988).

Figure 2.3: Engorged *Aedes mcinthoshi* mosquito

The virus was first isolated from *Aedes caballus sensu lato* and *Culex theileri* in Western Free State of South Africa in 1953 (Gear *et al.,* 1955). Since then, the virus has been isolated from 12 mosquito species in the subcontinent including: five *Aedes*, three *Culex*, three *Anopheleses* and one *Eretomapodites* species (Swanepoel *et al.*, 1974; Mcintosh, 1973). These mosquitoes usually breed in temporary stagnating waters and dambos. The virus can be transmitted to humans by mosquitoes, through the handling of infected animal tissues and fluids during slaughtering or butchering, birthing, conducting veterinary procedures, or from the disposal of carcasses or fetuses (Smithburn *et al.,* 1949; Swanepoel *et al.,* 1979; Mcintosh *et al.,* 1980). There is some evidence that humans may also become infected with RVF by ingesting unpasteurized or uncooked milk from infected animals (Alexander, 1951; Barnard, 1981).

Flooding has been associated with the amplification of mosquito populations. Some of the floodwater mosquitoes that emerge could be infected with the RVFV; these start the infection process especially when they feed on susceptible/amplifying hosts such as sheep, goats and cattle (FAO, 2002). For the infections to lead to a full blown epizootic, floods have to remain for four to six weeks or more to allow the development of large populations of secondary vectors to breed rapidly (FAO, 2002).

2.1.4 Host range of Rift Valley Fever disease

Many species of animals are affected by RVF including the domestic animals cattle, sheep, camels and goats leading to a severe hemorrhagic disease manifested by stormy abortions (Davies *et al.,* 1980). Sheep appear to be more susceptible than cattle or camels. Age is also a significant factor in the animal's susceptibility and development of the severe form of the disease with high mortalities being observed in lambs compared to adult sheep (Davies *et al.*, 1980).

Rift Valley Fever usually produces a febrile influenza-like disease in humans but it may develop into a hemorrhagic fever syndrome (Van Velden *et al.*, 1977; Laughing *et al.*, 1979). The antibodies to the virus have been detected in wildlife species especially ruminants, which include the buffalo, waterbuck, rhino, kudu and impala (Evans *et al.,* 2008).

2.1.5 Clinical signs of Rift Valley Fever disease

In animals, RVF mainly presents with signs of stormy abortions, high fever, bloody diarrhea, jaundice, loss of appetite, dysgalactia, bloody nasal and ocular discharges, severe prostration and finally death especially in sheep. It causes up to 100% mortalities in lambs under five to six days old. It may also present with other symptoms resembling other diseases (Radostits, *et al*, 2000).

The incubation period for RVF varies from 2 to 6 days**.** Those infected either experience no detectable symptoms or develop a mild form of the disease characterized by a feverish syndrome with sudden onset of flu-like fever, muscle pain, joint pain and headache. Some patients develop neck stiffness, sensitivity to light, loss of appetite and vomiting; in these patients the disease, in its early stages, may be mistaken for meningitis. The symptoms of RVF usually last from 4 to 7 days, after which time the immune response becomes detectable with the appearance of antibodies and the virus gradually disappears from the blood (WHO, 2000).

In human, most cases are relatively mild, a small percentage of patients develop a much more severe form of the disease. This usually appears as one or more of three distinct syndromes: ocular **(**eye**)** disease (0.5-2% of patients), meningoencephalitis (less than 1%) or haemorrhagic fever (less than 1%) (WHO, 2000).

2.1.6 Diagnosis of Rift Valley Fever disease

Acute RVF can be diagnosed using several different methods. Serological tests such as enzyme-linked Immunosorbent Assay (ELISA) may confirm the presence of specific antibodies to the virus namely: The IgM in recent infections and IgG antibodies in past infections or vaccinations (Niklasson *et al.*, 1984; Ksiazek *et al.*, 1989). The virus itself may be detected in blood during the early phase of illness or in post-mortem tissue using a variety of techniques including virus propagation in Monkey Derived Kidney cells (MDCK) cultures or inoculation in baby mice, antigen detection tests e.g. RT-PCR and virus neutralization tests (Garcia *et al.*, 2001; Drosten *et al.*, 2002).

2.1.7 Differential diagnosis of Rift Valley Fever disease

Single cases of RVF can be confused with many other diseases, which cause sudden death in sheep and present with similar signs. These include: Nairobi sheep disease, bluetongue, heartwater, ephemeral fever, toxoplasmosis, leptospirosis, brucellosis, Q fever and salmonellosis due to various similar clinical signs (FAO, 2003).

2.1.8 Control of Rift Valley Fever disease

Control measures used in livestock include quarantine, banning slaughter and meat consumption and vaccination. For animals there are two types of vaccines. The first is the attenuated virus vaccine (Smithburn strain) which after inoculation confers immunity lasting 3 years though it has been shown to cause abortions in ewes and is pathogenic to humans (Bernard, 1979; Kark *et al.*, 1982). The other vaccine is a formalin inactivated virus which requires two inoculations and thereafter an annual revaccination. This vaccine induces short lived immunity and is safe to use in pregnant animals (Davies *et al.*, 1992). In humans there is a live attenuated vaccine, MP-12 currently undergoing trials, but it has not yet been approved. A viral glycoprotein vaccine which is still under trial has also been developed (Frank, 2000). Other attenuated vaccine strains have been developed as potential live human vaccines together with formalin-inactivated vaccines and they have been used for a while to protect laboratory workers likely to be exposed to the virus (Eddy *et al.*, 1981; Frank, 2000).

Creating an active animal health surveillance system in order to detect new cases is essential so as to reduce the risk of animal-to-human transmission as a consequence of unsafe animal husbandry and slaughtering and consumption practices. Other useful control measures include: prevention of mosquito bites through the use of: impregnated mosquito nets, personal insect repellent if available, long-sleeved shirts and trousers and by avoiding outdoor activity at peak biting times of the vector species. Also use of larvicides on mosquito breeding sites is effective (Logan *et al.*, 1990; Whittle *et al.*, 1993).

2.2 Surveillance and risk mapping

Current maps indicating the distribution of RVF have been produced either from observation data or statistical models as shown in Figure 2.4, 2.2.and 2.1. However, acquiring correct absence data is not easy and hence maps generated from standard statistical models might be biased.

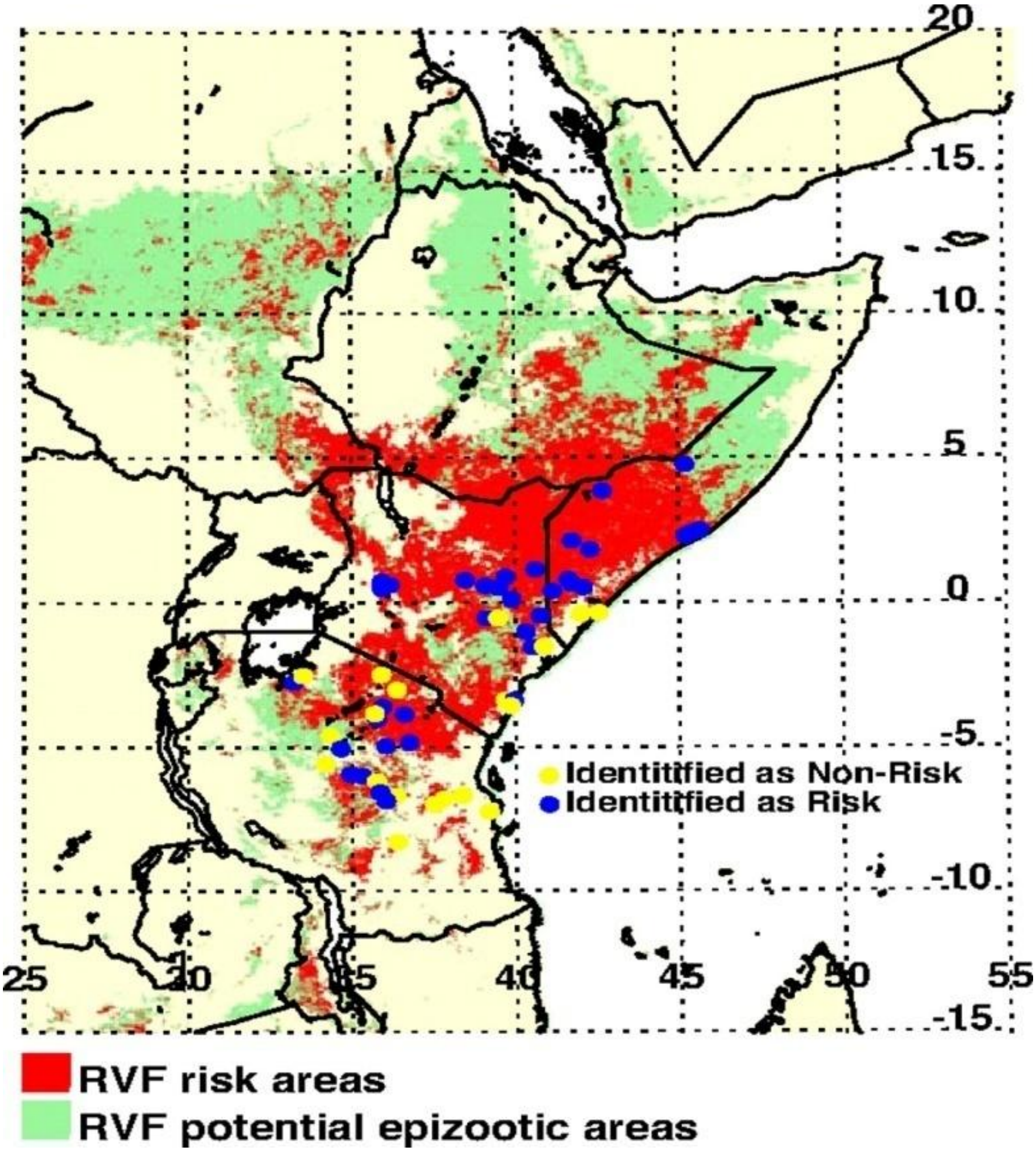

Figure 2.4: Climate-based models predicting RVF in humans and animals (Anyamba *et al.,* PNAS 2009; 106:955-959)

For instance mapping of RVF as shown in Figure 2.1 and 2.2 shows average locations by province and district in Kenya where RVF outbreak has been reported between 1951 and 2006. Thus the need to use ENM modeling which determines the potential distribution of RVF including those that have been identified in these maps, aiding the targeting of mitigation measures.

2.3 Ecological Niche Model

"Ecological niche model (ENM)", "niche-theory model" "species niche model", are terms that are used to describe Species Distribution Model (SDM) which is a strategy used to estimate actual or potential distribution of species as per the environment of the sampled species with specific geographical location eventually enabling identification of habitats having the same environmental characteristic in the entire area of interest (Frankline, 2009).

In this study, the main interest was modeling distribution of RVF in Kenya based on: soil types, precipitation, Normalized Difference Vegetation Index (NDVI), temperature and altitude as predictors of RVF outbreak (Linthicum *et al.*, 1999; Anyamba *et al.*, 2009; Hightower *et al.*, 2012; Bett *et al*., 2013).

A niche is defined as an environment where an organism can survive and grow without the need for an external replenishment (Hutchinson, 1957). A "fundamental niche" is the ecological properties of a species, a conceptual space whose axes include all of the environmental variables affecting that species (Austin *et al*., 1990; Leibold, 1995).

Ecological niche model was generated using Bioclim, Genetic Algorithm for Rule set Production (GARP) in Open Modeler software and Random Forest (RFs) (Stockwell *et al.*, 1991). These algorithms were used to allow for cross-validation of the results. The ENM uses a set of point localities where the species is known to occur and a set of geographic layers representing the environmental parameters that might limit the species' capabilities to survive.

The SDM uses a set of rules of selection, evaluation, testing and incorporation or rejection in modeling such as bioclim rule, logistic regression, range rules, negated range rules to identify environmental conditions under which the species should be able to maintain populations (Peterson *et al.*, 2007). Both GARP and Random Forest algorithms use presence only data and it generates automatically absence data (majorly known as pseudo-absence data) from pixels where presence data are absent. This does not necessarily mean that they are correct absence data like the one collected in the field (Peterson *et al*, 2007). Predictive accuracy of the model is measured by estimating the area under the curve (AUC).

CHAPTER THREE

MATERIALS AND METHOD

3.1 Study area

3.1.1 Location and study area

The study involved the generation of RVF risk map for Kenya using presence data only. The critical decision point that was relevant was whether there exist reliable presence and absence data given that no formal studies have been done to verify absence of disease in areas where outbreaks have not been confirmed. Records available at the DVS that were collected during outbreaks represent presence-only data that is areas the disease were reported and confirmed.

Kenya has a total area of 580,367 km^2 with a land cover of 569,140km^2; the rest is area under water. It lies between latitudes 5°N and 5°S, and longitudes 34°E and 42°E and lies on the equator with the Indian Ocean to the south-east, Tanzania to the South, Uganda to the West, South Sudan to the north-west, Ethiopia to the North and Somalia to the North-East with 47 administrative regions known as counties (Figure 3.1).

Figure 3.1: Map of Kenya showing the 47 administrative counties

3.1.2 Climatic Condition of Kenya

As shown in Figure 3.2, Kenya has various eco-climatic zones varying from tropical eco-zone along the coast, arid zone in the North and North-East. Less than 15 percent of the country receives somewhat reliable rainfall of 760 millimeters or more per year, mainly the southwestern highlands near Lake Victoria and the coastal area, which is tempered by monsoon winds. Most of the country experiences two wet and two dry seasons. Kenya has two rain seasons: short rains (October to December) and long rains (March to June). The hottest period is from January to March.

The driest month is August, with an average of 24 millimeters average rainfall, and the wettest is April, the period of “long rains,” with an average of 266 millimeters. The hottest month is February, with temperatures of 13°C to 28°C, and the coolest is July, with temperatures of 11°C to 23°C. The highlands feature a bracing temperate climate.

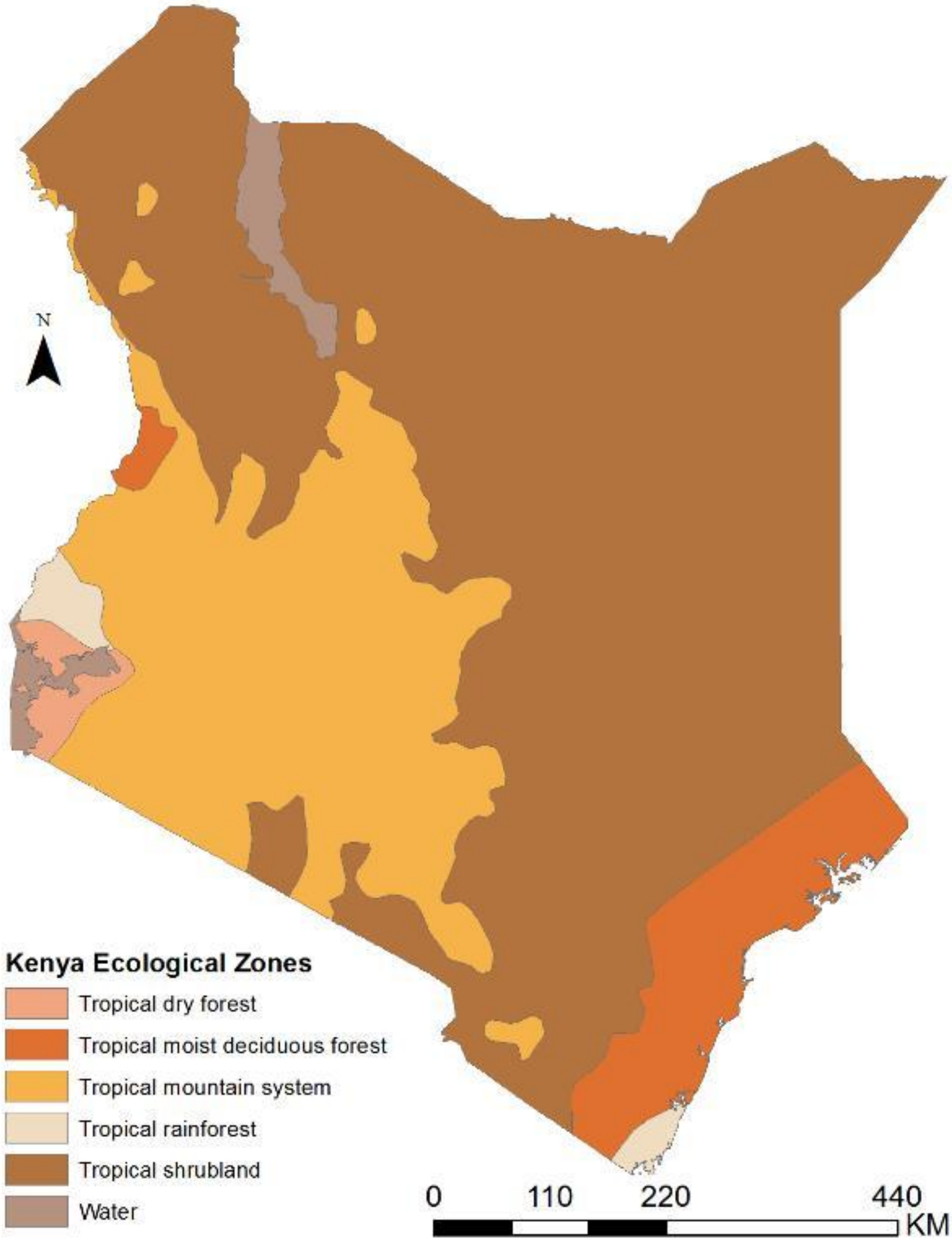

Figure 3.2: Map showing Eco-climatic Zones of Kenya

3.1.3 Human Population in Kenya

Kenya is a multi-ethnic state and is primarily inhabited by Bantu and Nilotic population with some Cushitic ethnic minority in the north. Its total human population is estimated to be 44,037,656. Kenya has no single prominent culture; instead it has various cultures practiced by different communities.

3.2 Study Design

These study uses disease classification framework adopted by Hay *et al*., (2013) as shown in Figure 3.4, which outlines the framework that was used to support the choice of ecological niche model to map RVF distribution in Kenya.

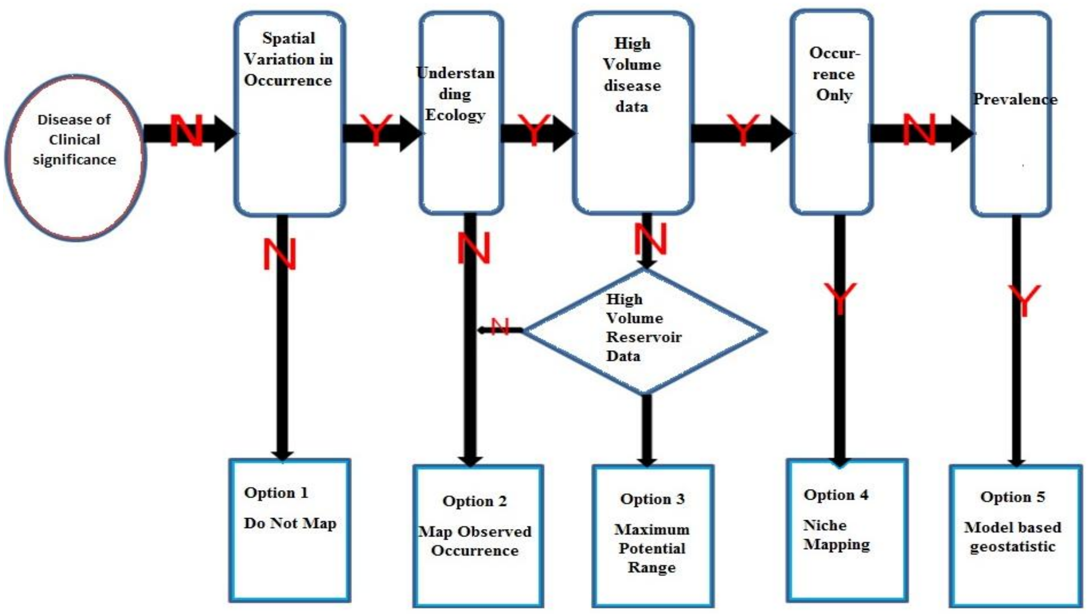

Figure 3.4: A schematic process of disease classification (N=No and Y=Yes).

The framework involved classification of whether the distribution of the disease in question has a spatial variation as expected, it will not be possible to develop a risk map for a disease that is homogenously distributed in space because there won't be any difference in risk of disease distribution the disease will be expected to occur everywhere, the ecology of the disease need to be understood: in RVF case, previous studies have confirmed that RVF occurs in defined climatic conditions. They have also defined environmental factors associated with the disease outbreaks which include: Precipitation, temperature, NDVI, soil types, implying that the ecology of the disease is fairly known.

This indicates that these surveillance data can be best analyzed using ENM (option 4) than model-based geo-statistics (option 5). The latter requires both presence and absence data in order to estimate odds of disease presence

3.3 Data sources

3.3.1 Primary Data

The areas affected in the 2006 to 2007 outbreak were obtained from the DVS. These areas were visited and geo-referenced using Garmin® Global Positioning System (GPS) (Garmin International, Inc., USA) hand receiver to obtain the GPS readings (Easting, Northing and Altitude) in Universal Transverse Mercator (UTM) units.

Later, questionnaires(Appendix I) were administered to specific farmers affected by RVF in the areas identified from DVS from the 2006 to 2007 RVF outbreak to identify their livelihood activities, type of livestock keeping and their response in case of RVF outbreak.

3.3.2 Secondary data

The secondary data mainly included satellite data which were obtained from on-line databases. The data were: land cover data assembled by FAO from the Global Land Cover analysis, precipitation data and temperature estimates was downloaded from European Centre for Medium-Range Weather Forecasts (ECMWF), Normalized Difference Vegetation Index (NDVI) data was obtained from SPOT VEGETATION (http://free.vgt.vito.be/), elevation data was generated by NASA Shuttle Radar Topographic Mission (SRTM) based on Digital Elevation Models (DEM) (http://www.cgiar-csi.org/data/srtm-90m-digital-elevation-database-v4-1), data on soil types was extracted from the Harmonized World Soil Database (HWSD) developed by FAO and the International Institute for Applied Systems Analysis (IIASA) (FAO/IIASA/ISRIC/ISS-CAS/JRC, 2009).

The land cover was global data, Gridded ERA-Interim reanalysis precipitation data and minimum and maximum temperature estimates were gridded ERA-Interim, optimized (global best estimates) to fit both short-range forecasts (from a model) and observed data. Normalized Difference Vegetation Index (NDVI) data, which is defined as a measure of amount and vigor of vegetation on land surface, was derived from radiometric sensor measures of reflectance for both red and near infrared bands on two separate channels or images. Usually NDVI estimates are derived by subtracting red band measures from the near-infrared and dividing the difference by the sum of the two measures. These values range between -0.1 and 1.0; negative values indicate clouds and water, positive values near zero indicate bare soil and higher values indicate dense vegetation. Extracts of NDVI are available on 10 day-intervals at a spatial resolution of 1km. For this study, minimum, maximum and average values for each division were extracted.

The elevation data was digital and data on soil types had a resolution of 1km and over 15000 different soil mapping units were recognized in the database. The database contained information of the soil units, soil properties and other parameters such as organic carbon, pH, water storage capacity, soil depth, etc.

3.4 Data Analysis

3.4.1 Descriptive Analysis

3.4.1.1 Analysis of data from the questionnaire

The data collected from the questionnaire surveys which included socio-economic activity, production systems, livestock species and community intervention for the future outbreak of RVF were coded and entered into database designed using Microsoft Excel software (Microsoft Corporation, USA). The above data were summarized through descriptive analysis such as proportions.

3.4.1.2 Spatial data sets

Spatial characterization of relative distribution of soil type and division with RVF in Kenya, elevation and land cover were done using spatial data through maps as shown in Figures 3.5, 3.6, 3.7 and 3.8. From the spatial maps, RVF outbreak is shown to be generally associated with soil types (solonetz, luvisols, planosols), and an altitude of less than 1,100 m above sea level, which is in agreement with various studies that have been done (Linthicum *et al*., 1999; Anyamba *et al*., 2009; Hightower *et al*., 2012; Bett *et al*., 2013).

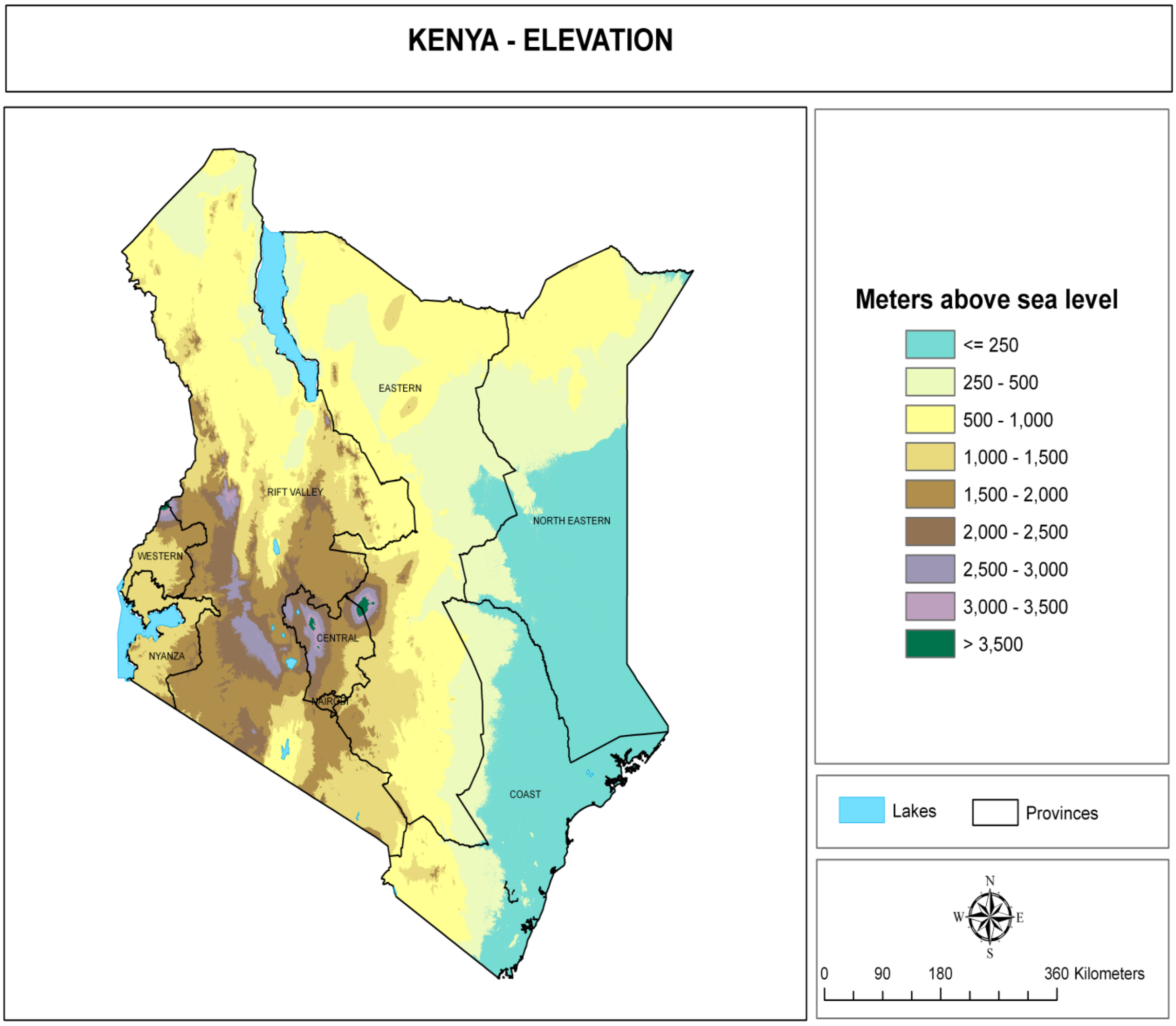

Figure 3.5:Map showing Kenya elevation (Source: ILRI GIS unit, 2013)

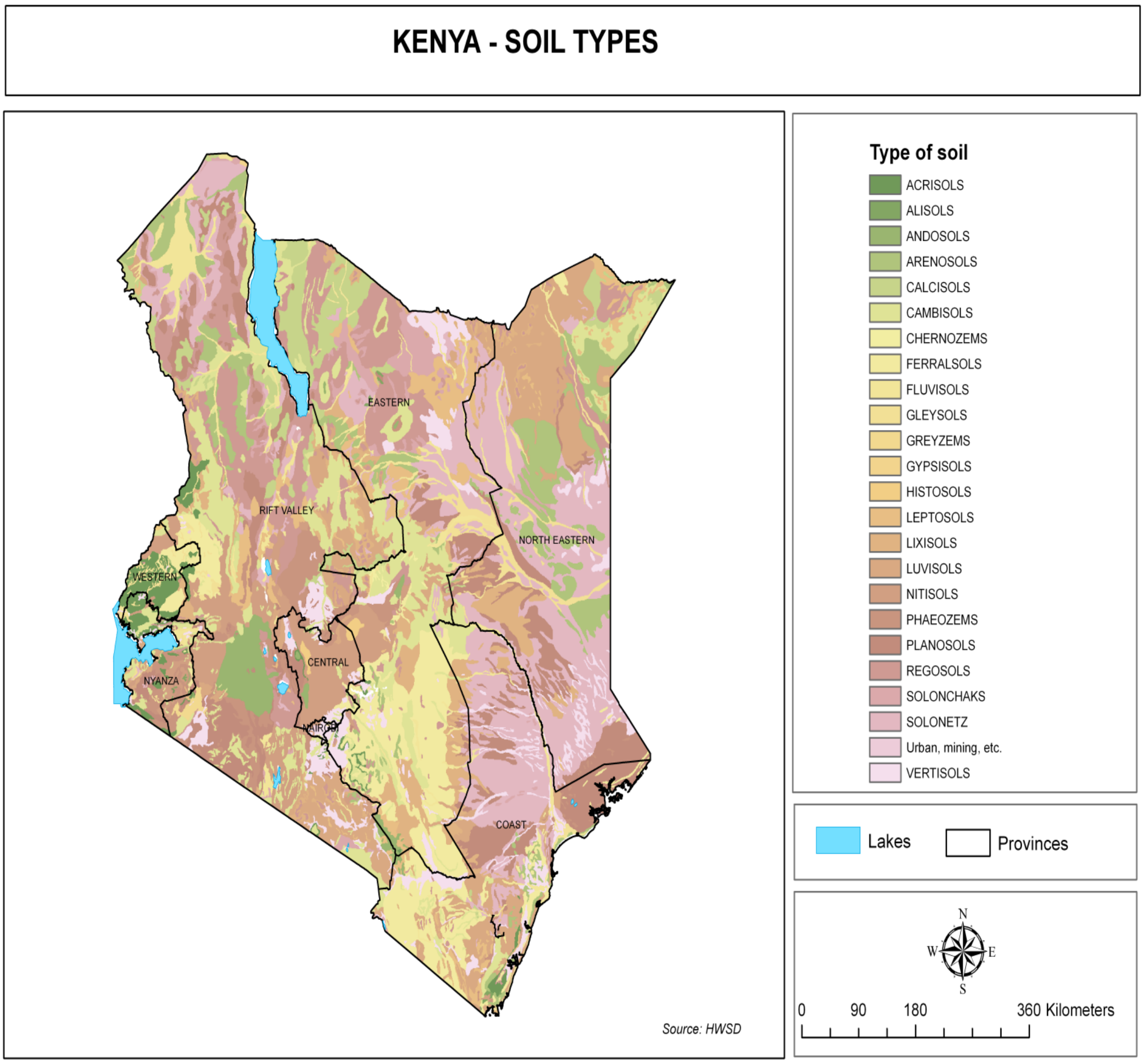

Figure 3.6 : Map showing soil type of Kenya(Source: ILRI GIS unit, 2013).

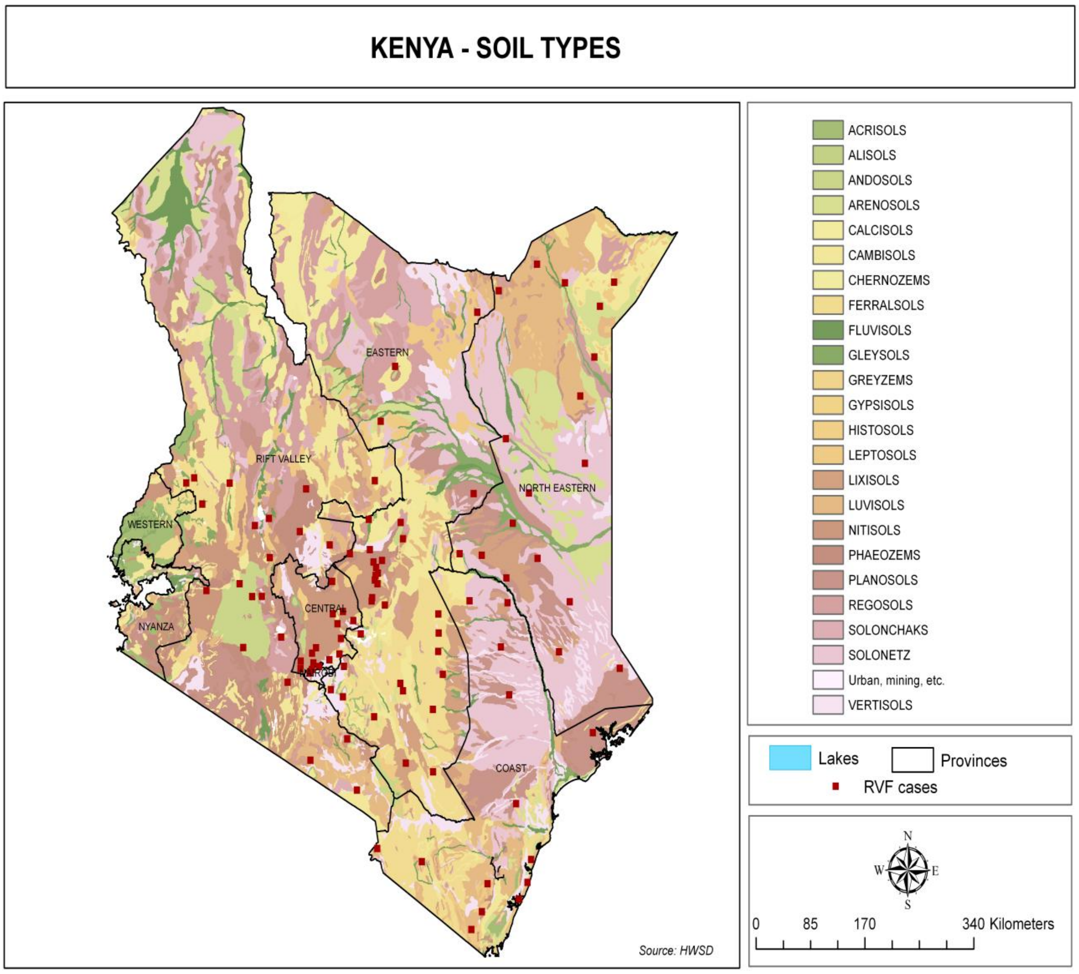

Figure 3.7: Map showing Relative Distribution of Soil type and Divisions with RVF in Kenya(Source: ILRI GIS unit, 2013).

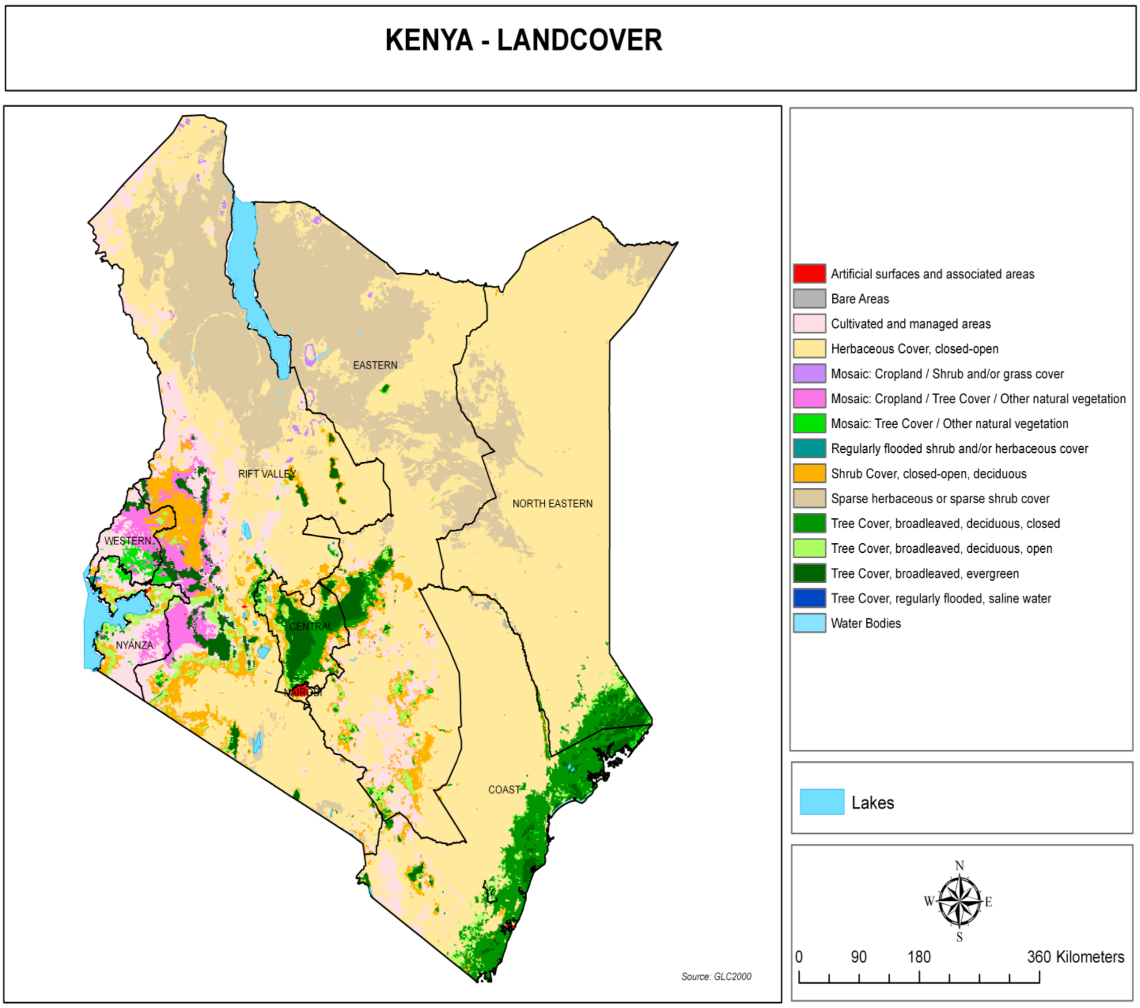

Figure 3.8: Map showing land cover of Kenya(Source: ILRI GIS unit, 2013).

3.4.2. Ecological Niche Analysis

Ecological niche model was generated using; Bioclim, Genetic Algorithm for Rule set Production (GARP) in Open Modeler software and Random Forest (RFs) (Stockwell *et al.*, 1991). Ecological niches and associated potential geographic distributions can be approximated via correlative approaches that relate known point-occurrence data to digital GIS data layers summarizing spatial variation in relevant environmental dimensions .The ENM uses a set of rules of selection, evaluation, testing and incorporation or rejection in modeling. Predictive accuracy of the model is measured by estimating the area under the curve (AUC).

3.4.2.1 Genetic Algorithm for Rule set Production Analysis

Genetic Algorithm for Rule set Production (GARP) is an evolutionary-computing method that builds models based on non-random associations between known occurrence points for species and sets of GIS coverage describing the ecological landscape. Occurrence data are used by GARP as follows: 50% of occurrence data points are set aside for an independent test of model quality (extrinsic testing data); 25% are used for developing models (training data); and 25% are used for tests of model quality internal to GARP (intrinsic testing data). Distributional data are converted to raster layers and by random sampling from areas of known presence (training and intrinsic test data) and areas of 'pseudoabsence' (areas lacking known presences).

The genetic algorithm produces a logic model, rather than a strictly derived mathematical model. An initial condition (first rule applied) is created in GARP by application of a single inferential tool randomly selected from a defined set. This set includes 4 basic rule types (bioclimatic rules, atomic rules, range rules and logistic regression), each of which implements a different method for building prediction models. Subsequent combinations of rules with specially defined operators (e.g. crossover, mutation) are then used to modify the initial rules, and through iteration and optimization, models are "evolved". After each modification, the quality of the rule is tested (to maximize both significance and predictive accuracy) and a size-limited set of the best rules is retained. Because rules are tested based on independent data (intrinsic test data), performance values reflect the expected (general) performance of the rule, an independent verification that gives a more reliable estimate of true rule performance. The final result is a set of rules that can be projected onto a map to produce a potential geographic distribution for the species under investigation.

To produce a final prediction model (map), 10 individual GARP models were created, each with 100,000 maximum iterations and a convergence criterion of 0.0001 from 159 point localities that had been sampled as shown in Figure 3.9 together with environmental parameters were used with replacement. Fifty (50) GARP runs were run and a rule set to pick only 20 runs that had hard omission error of 10%, commission error of 50% and 50% of the 20 models was picked for further analysis. The best subset procedure as defined by Anderson *et al*., (2003) was used to filter model by model. The final prediction maps were produced by summing these 10 high-quality models. Color gradations are used to indicate the proportion of times out of 10 that specific areas (pixels) were included in the predicted distribution of RVF in Kenya.

Model quality and accuracy evaluation was done using Area Under Cover (AUC); if a model has AUC of 0.5-0.7 it is considered as having a poor predictive ability while that with AUC of 0.7-0.9 and >0.9 are considered as having a moderate and high predictive abilities, respectively (Swets, 1988; Manel *et al*., 2012). The model was also evaluated using partial Receiver Operating Characteristic (pROC) which plots sensitivity against 1-proportion of area predicted. It shows relationship between the proportion of observed presence correctly predicted and 1-proportion of area predicted because this study is dealing with presence data only (Townsend, 2012).

To assess and determine the relative importance of the individual ecological parameters and its influence on the model, a jackknife procedure was performed, involving construction of a series of ENMs, each systematically omitting one of the *n* layers, following procedures outlined (Peterson et al., 1999).

This manipulation resulted in *n* - 1 maps, each representing the predicted distribution of the disease without consideration of the information in a particular parameter; effects of these manipulations were summarized by a calculation of percent difference (across all pixels in the map) from the map produced using all variables.

The empirical contribution of the information contained in each layer toward creation of the comprehensive ENM (i.e., the statistical significance of each parameter within the overall model) was assumed using a single sample Student's *t*-test ($H_0 = 0$) to evaluate differences in the mean number of pixel matches between the comprehensive ENM (based on *n* variables) and each derived ENM (based on *n*-1 variables). To accomplish this test, each pixel in the map was assigned a value between 0 and 10 corresponding to the frequency of positive prediction in the 10 summed models (see above). The mean difference in predicted level for matched pixels across the population of pixels in the comprehensive versus derived ENMs was then compared to a hypothesized value of zero (signifying that the derived and comprehensive ENMs were identical). Kappa statistics were also used to assess levels of agreement between the comprehensive and derived ENMs.

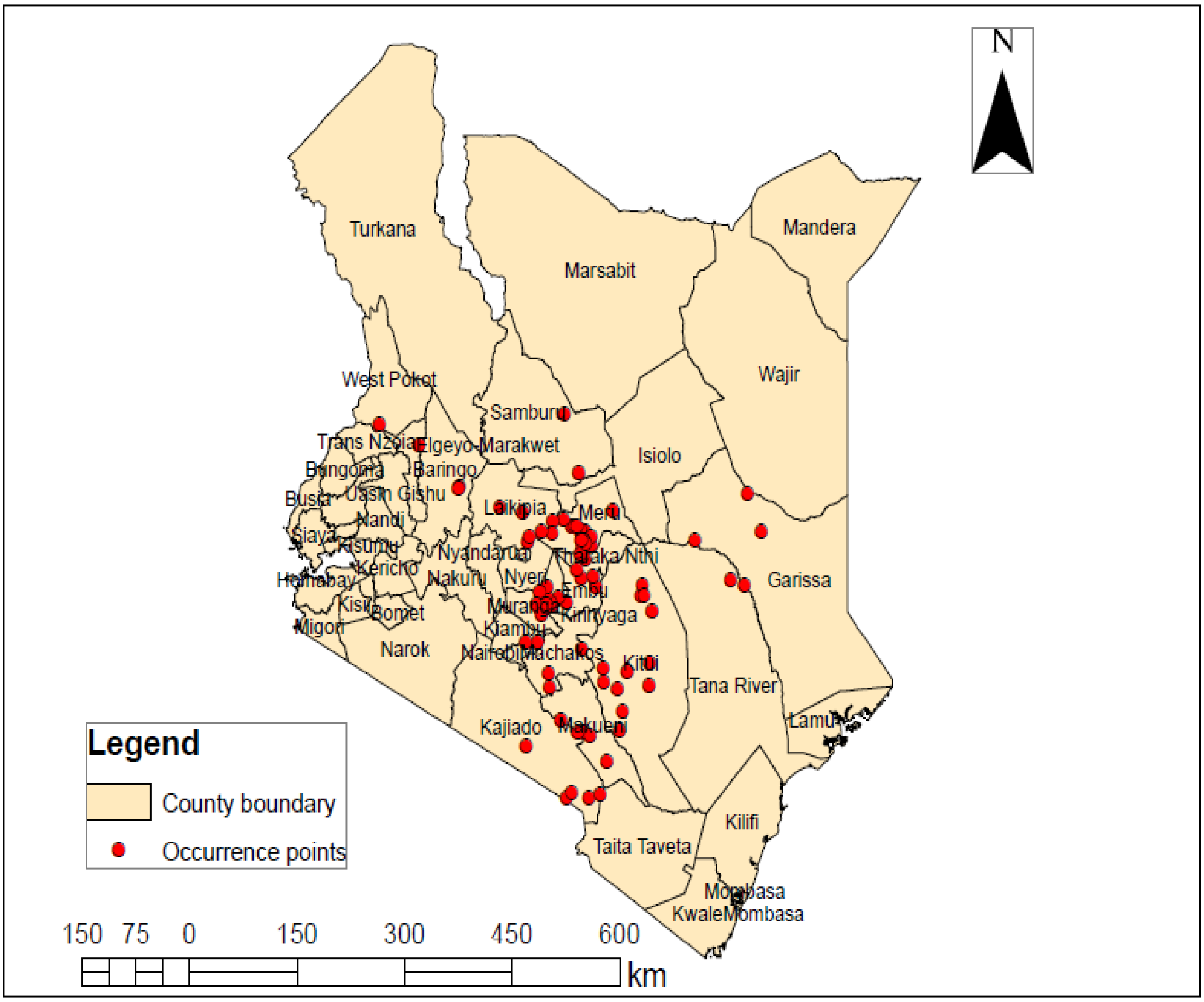

Figure 3.9: Map of Kenya showing Rift Valley Fever georeferenced areas

3.4.2.2 Random Forest Analysis

Random forests are an ensemble learning method for classification, regression and other tasks, that operate by constructing a multitude of decision trees at training time and outputting the class that is the mode of the classes (classification) or mean prediction (regression) of the individual trees. Specifically, it is an ensemble of trees constructed from a training data set and internally validated to yield a prediction of the response given the predictors for future observations. There are several variants of RF which are characterized by: the way each individual tree is constructed, the procedure used to generate the modified data sets on which each individual tree is constructed and the way the predictions of each individual tree are aggregated to produce a unique consensus prediction.

In this model, RFs algorithm with the help of R package, built a random forest classifier model (where the response variable, presence, was regressed against the explanatory variables, environmental variables) using 1000 trees from the 159 georeferenced spatial extent and resolution of the environmental data layers which were harmonized and the occurrence data sub-sampled to take care of sampling bias. Pseudo-absence data were generated and both occurrence and pseudo-absence data were used to extract the predictor values at respective presence/pseudo-absence points. The data were separated into training and testing data. The training data (75%) were then used to calibrate the model while the rest (25%) of the data were used to test the model. The model was used to generate a prediction map of the possible distribution of RVF in Kenya and result compared with those of GARP.

Model quality evaluation was assessed using Area under Curve (AUC) to show model accuracy.

3.4.3 Logistic regression

Data used for logistic regression model were obtained by overlaying a grid of 25 x 25 km on the entire country. A total of 1093 grids were obtained in this process. Grids that fell in the areas geo-referenced and identified as hotspots were assumed to be infected, and so coded as RVF positive while the rest were coded as RVF negative. The period considered for the analysis was 2006 to 2007. Grids that were assumed to be infected were considered as having been positive during the months when outbreaks occurred, i.e., October to December 2006 and January to February 2007. The logistic regression model used a case-control design whereby the grids that were positive represented cases while the other grids were used as controls. In this case, 221 of 1093 grids were positive, representing a prevalence of 20.22%. Predictors used in the analysis – which were also extracted using the grid included soil type, rain, NDVI, altitude, temperature, land-cover and livestock population. The strength of association between predictors and the outcome (RVF infection) was estimated by odds ratios (OR) which were directly derived from estimates of logistic regression.

The odd ratio is a relative measure of risk that describes how much more likely it is that RVF will occur if risk factor is present compared to if there is no risk factor. If odds ratio is close to 1, the risk factor is unlikely to be associated with RVF disease. For an odds ratio greater or smaller than 1, the likelihood that the risk factor is associated with risk of disease increases, and the stronger the association. Further, if the 95% CI of the odds ratios includes the value 1, this implies that the odds ratio obtained in the study is statistically consistent with a true odds ratio of 1, "not statistically significantly different. Odds ratios from logistic regression are interpreted as a multiplicative factor of risk of disease when the risk factor is present.

The logistic model for the probability of the *i*th risk factor to contribute to RVF outbreak with only one predictor was computed as : $\Pr\{Yi = yi\} = \pi_i^{yi} (1 - \pi i)^{1-yi}$. The significance level was set at $p \leq 0.05$. A multivariable logistic regression model was then built using variables that were found significant during the univariate analysis. Variables were added to the model as follows: $\text{logit} \Pr\{Yi = yi\} = \pi_i^{yi} (1 - \pi i)^{1-yi}$. Model building used backwards elimination method to identify factors to include in the model based the likelihood ratio test ($p \leq 0.05$).

CHAPTER FOUR

RESULTS

4.1 Descriptive Analysis

Descriptive statistics generated from the analysis of the questionnaire data were from eighty four (84) specific farmers. The farmers details were obtained from confirmed 2006 and 2007 RVF outbreak cases from the DVS. The descriptive were as summarized below.

Figure 4.1 shows the proportion of male and females interviewed in visited hotspots areas while Figure 4.2 shows the proportion of human RVF cases in RVF hotspots sites in 2006/2007.

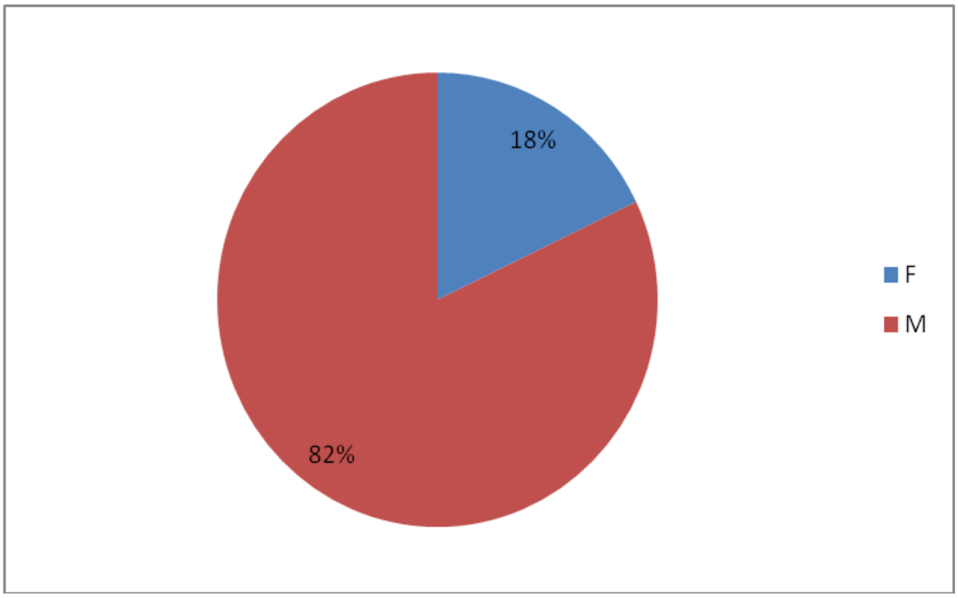

Figure 4.1 Proportions of males and females interviewed in the field survey

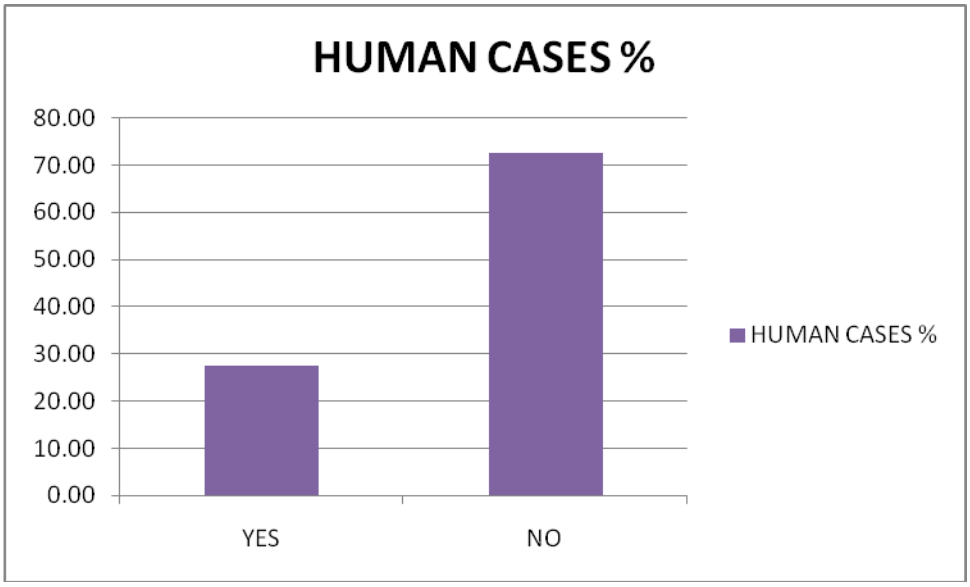

Figure 4.2: Proportion of human RVF cases in 2006/2007 RVF outbreak in visited hotspot areas

Figure 4.3 shows various combination of livelihood activities identified in the RVF hotspots which were over ten livelihood combinations. Livestock production and crop farming had the highest proportion of 43%.

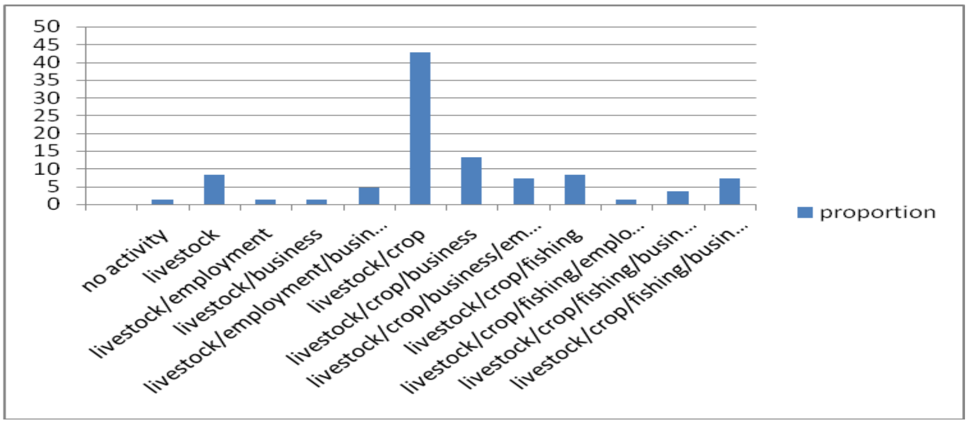

Figure 4.3 Proportion of the various combination of livelihood activities identified in the RVF hotspot area

Figure 4.4 shows relative proportions of livestock production system carried out in the areas visited. Of the three livestock production systems, extensive system comprised the highest proportion (48%) followed by semi-intensive production system (32%) and finally intensive system (20%).

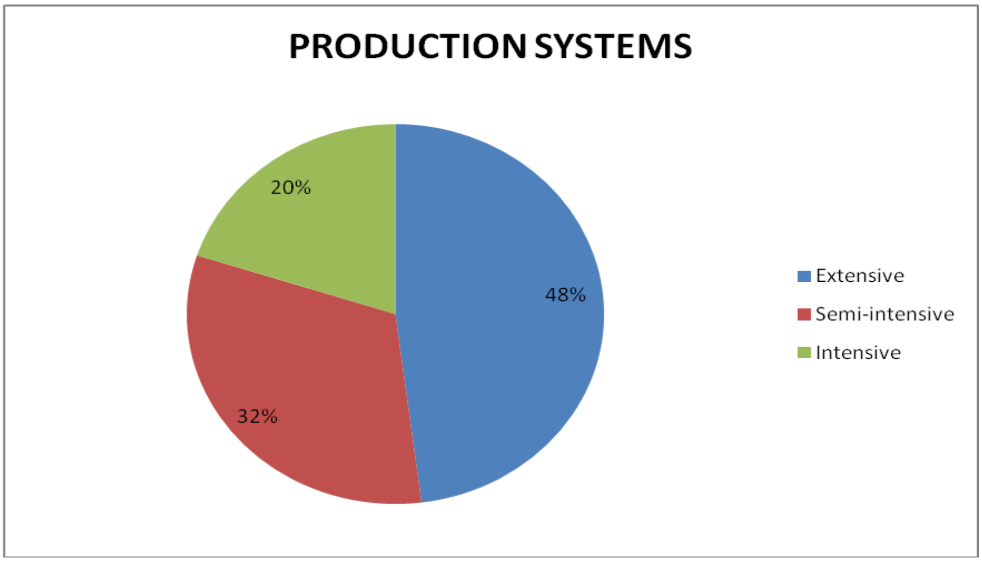

Figure 4.4 Production systems and their relative proportion in RVF hotspots areas visited

Figure 4.5 shows the proportion of various livestock combination kept in RVF hotspot areas visited. Cattle sheep and goat combination had the highest proportion of 76%.

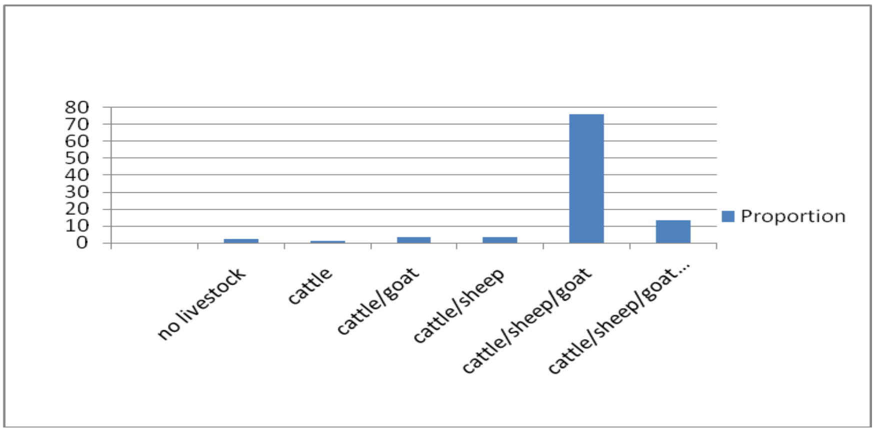

Figure 4.5 Proportion of livestock species combination kept in RVF hotspot areas visited

Various RVF outbreak intervention measures were identified by the respondents as shown in Table 4.1. Most of the farmers used vaccines to prevent the disease, while others either did nothing or remained vigilant and alerted veterinarians in the area whenever suspicious cases were noted.

Table 4.1 Community interventions used during the 2006/2007 RVF outbreak

Intervention	**proportion**
Vaccination	45
Others	5
Alert to report	27.5
Nothing	22.5
Total	**100**

4.2 Ecological Niche Model outputs

Three sets of maps of Rift Valley Fever (RVF) distribution were generated; one used Bioclimatic variables and the others used environmental variables customized for the outbreak period (October 2006 to February 2007). The GARP and Random Forest algorithms that used customized variables had better predictions and were able to show all the regions that had reported RVF before compared to models that used bioclimatic variables. The bioclimatic variables exaggerated the distribution of the disease. The GARP algorithm with customized climate variables produced a map (Figure 4.7) with Area Under Cover (AUC) of 0.82 compared to similar outputs from Random forest (Figure 4.8) which had an AUC of 0.99.

Output from the Bioclim algorithm (Figure 4.6) had an AUC of 0.69. A Partial ROC analysis for GARP also indicated that the customized variables with a value of 1.77 gave a better prediction than bioclimatic variables which had a value of 1.10.

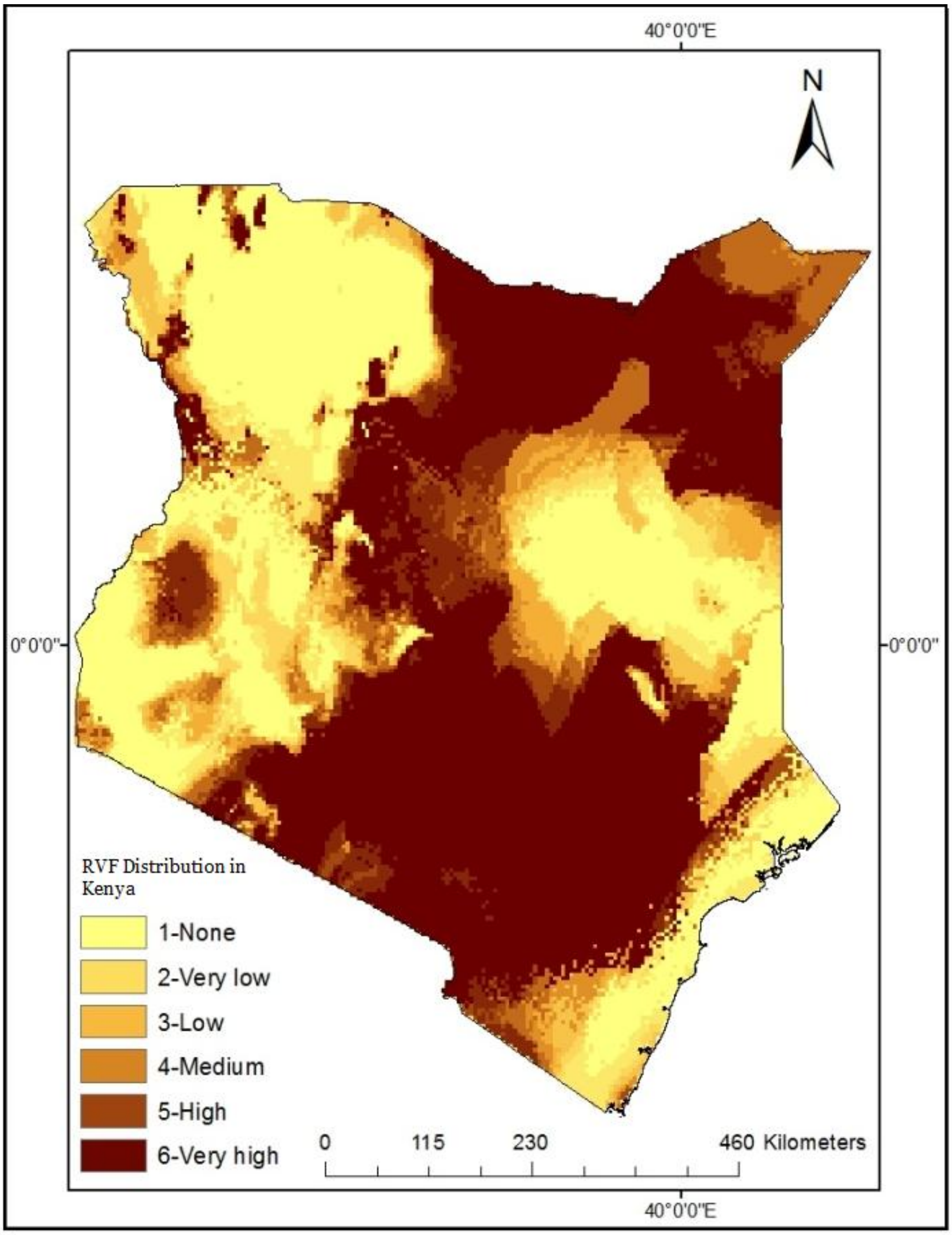

Figure 4.6: Map showing RVF distribution generated from Bioclimatic Variables.

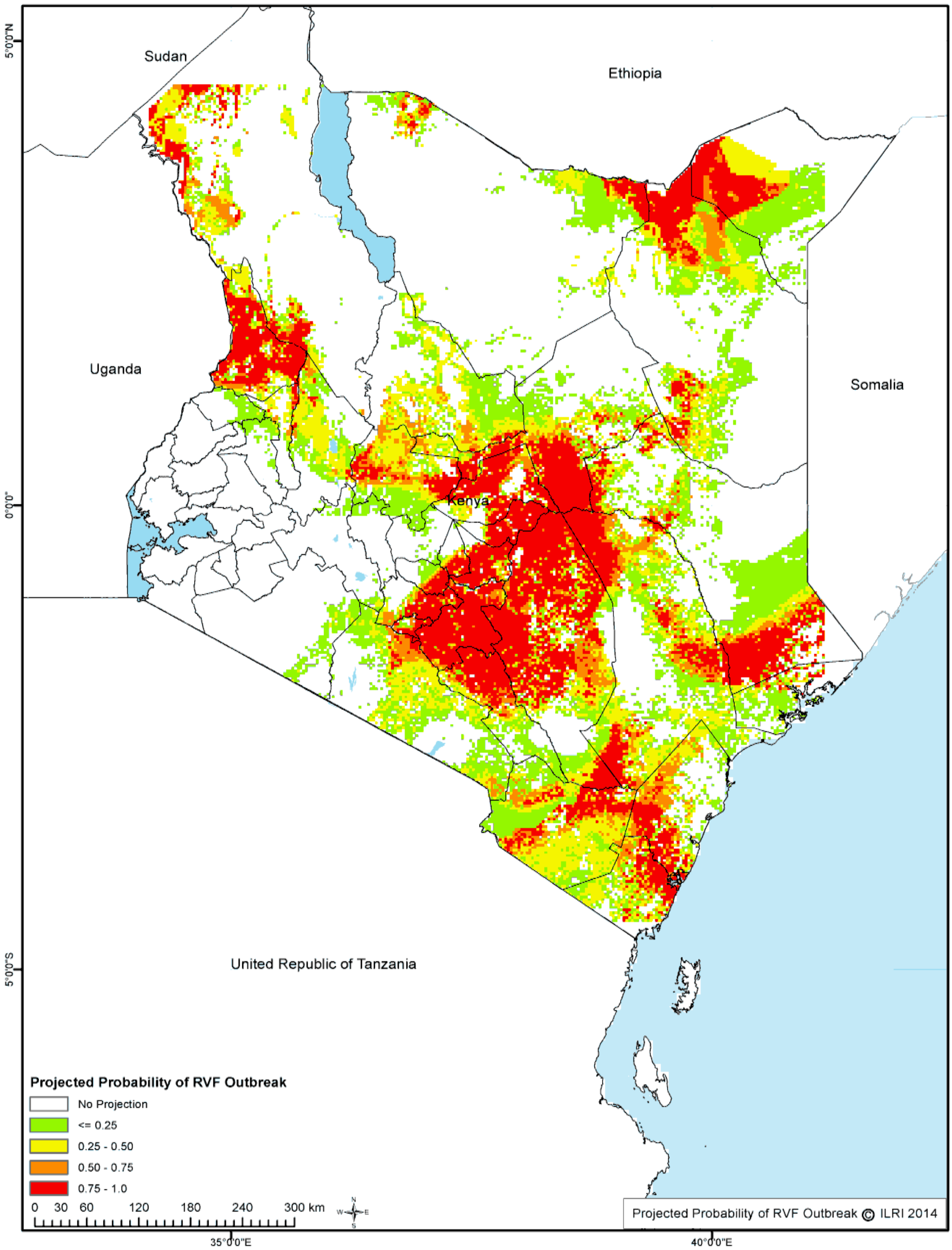

Figure 4.7: Map showing RVF distribution generated from GARP algorithm.

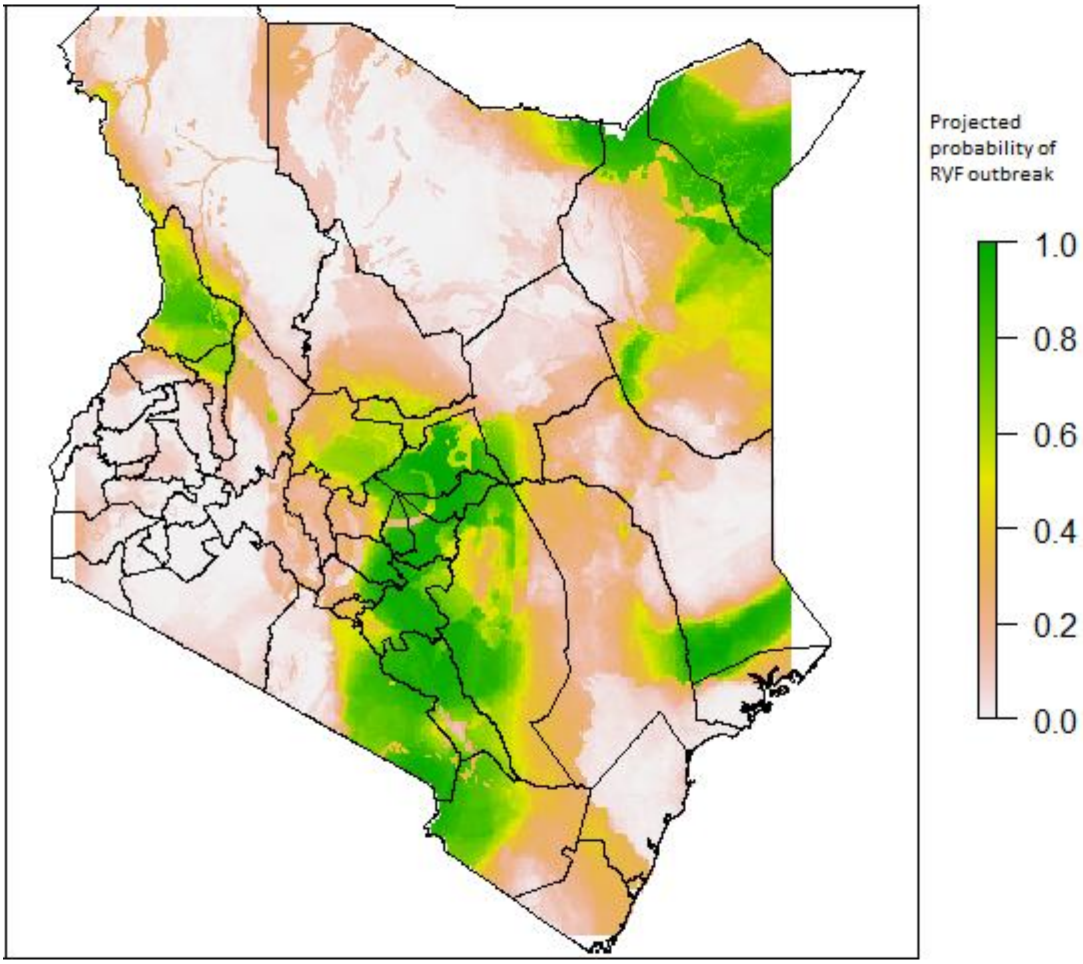

Figure 4.8: Map showing RVF Distribution generated from Random Forest algorithm.

Results from jackknife analysis identified importance of individual environmental variables to RVF outbreak. The results showed NDVI for March, 2007 had the highest influence on the model while the least influence of NDVI was for December, 2006 (Figure 4.9).

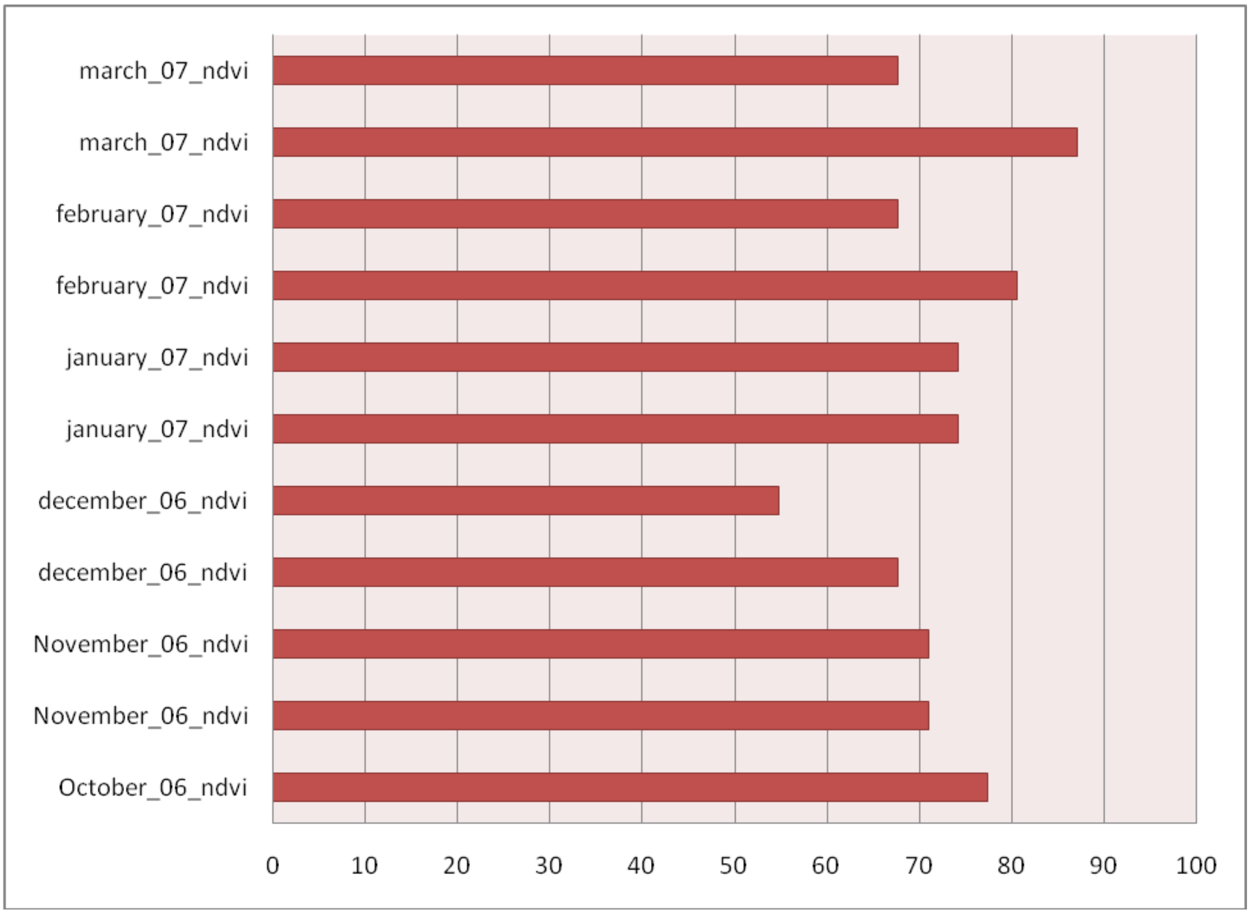

Figure 4.9: Jackknife Analysis result for the NDVI Variables .

Temperature and rainfall data had relatively equal influence on the model for all the months. January, 2007 rainfall and temperature had the highest influence on the model and the December temperature as well (Figure 4.10).

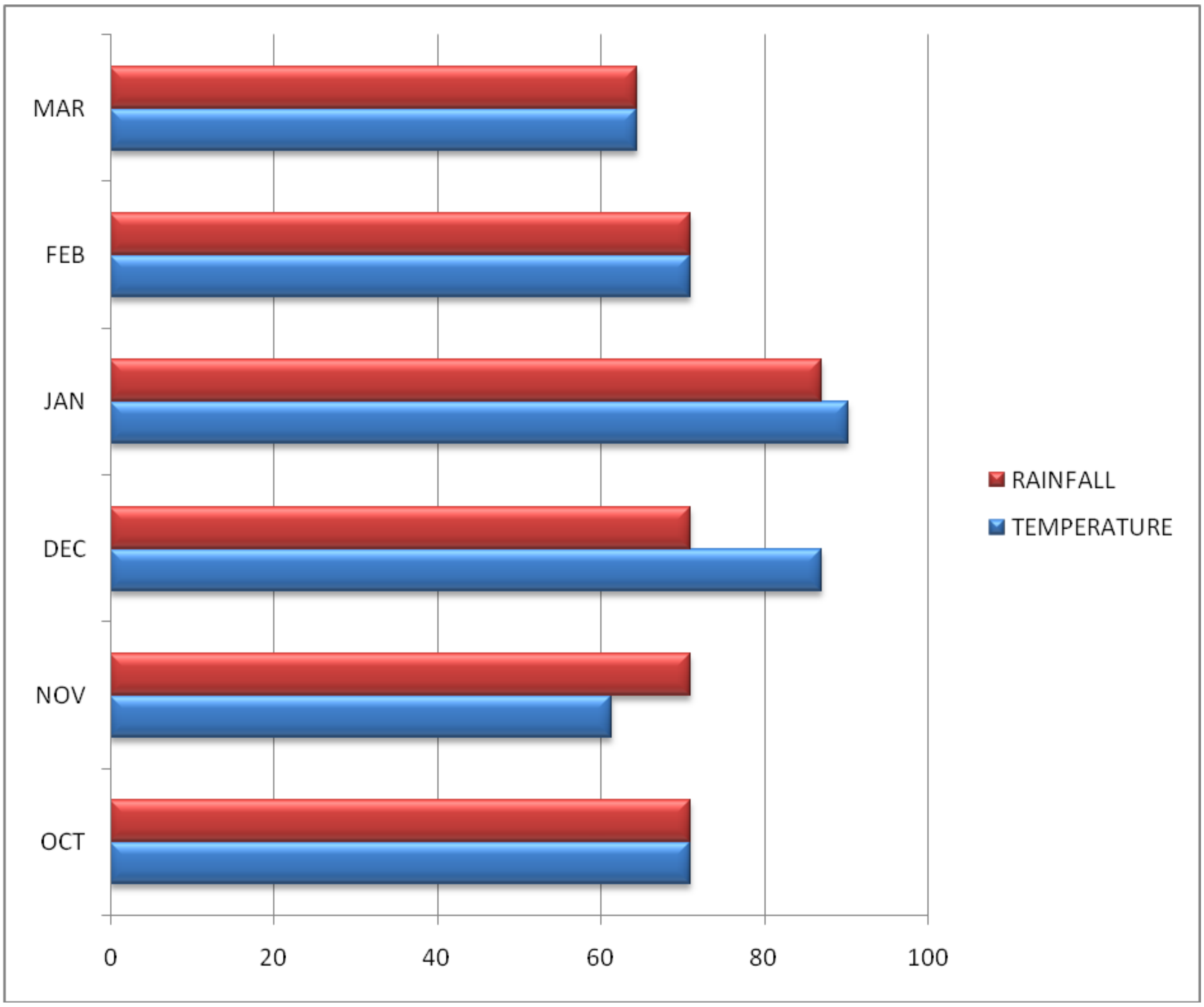

Figure 4.10: Jackknife Analysis result for Rainfall and Temperature Variables.

4.3 Factors associated with Rift Valley Fever from the logistic regression model

The risk of RVF occurrence was determined using the logistic regression model; odds ratios generated by the model indicated the risk of RVF outbreak.

Table 4.2 summarizes individual variable contribution to the outbreak of RVF. Factors that were shown to be significant at 95% confidence interval for the outbreak of RVF were; open to closed forests having a crude OR of 1.93, Solonetz soil type having OR of 1.6 and NDVI having OR of 4.66. A one unit increase in temperature decreases the risk of RVF by 10%, and a change in altitude from $\leq$500 to 500 - $\leq$1000 is associated with 94% decrease in outbreak of RVF.

Table 4.2: Descriptive Analysis from Univariate analysis

Variable		n	Crude OR	95% CI
Land cover	Artificial/bare areas	120	0.16	0.06 – 0.40
	Open to closed forests	67	1.93	1.12 – 3.31
	Grassland/shrub land	610	1.00	
	Mosaic croplands/vegetation	296	0.99	0.22 – 0.33
Cattle			1.00	0.99 – 1.00
Goats			1.00	1.00 – 1.01
Camels			1.02	0.97 – 1.08
Soil type	Others	803	1.00	-
	Luvisols	82	1.45	0.85 – 2.48
	Solonetz	156	1.60	1.08 – 2.39
	Vertisols	52	1.66	0.88 – 3.14
Altitude	≤500	454	1.00	-
	500 - ≤1000	315	0.06	-0.32 – 0.44
	1000 - ≤1500	149	0.75	0.32 – 1.18
	>1500	175	0.47	0.04 – 0.89
Rain	Last 2 months Cumulative		1.09	1.08 – 1.11
Temperature			0.90	0.89 – 0.92
NDVI			4.66	3.20 – 6.81

Regression model was then built using significant variables resulting to Table 4.3

Table 4.3 shows odds ratios (ORs) from logistic regression models and p-values postulated to be associated with RVF at 95% confidence interval as indicated below.

	RVF	**Odds Ratio**	**95% Confidence Interval**		P>|Z|
Soil types	Luvisols	1.61	1.03	2.52	0.038
	Solonetz	2.19	1.53	3.12	0.000
	Vertisols	1.41	0.78	2.59	0.253
Rain	Last 2 months cumulative (cum2)	1.09	1.07	1.10	0.881
Soil and Rain Interaction	Luvisols*cum2	1.05	1.00	1.11	0.041
	Solonetz*cum2	1.11	1.07	1.14	0.000
	Vertisols*cum2	1.06	1.00	1.12	0.056
	NDVI	8.08	2.68	24.37	0.000
	NDVIsq	1.57	0.46	5.33	0.470
Altitude	>500 - ≤1000	0.41	0.25	0.68	0.001
	1000 - ≤1500	0.19	0.11	0.33	0.000
	Temperature	0.87	0.84	0.90	0.000
	cons	$3.97e^{+}16$	$1.51e^{+}12$	$1.04e^{+}21$	0.000

Table 4.3: Regression Analysis of Variables for RVF Outcome

The interaction of soil type and rain is well elaborated in Figure 4.11 where X-axis shows change in level of rainfall while Y-axis shows log odds of RVF (predicted probability). The interaction term indicates that the effect of rain differs depending on soil type; the log odds of the outbreak increases much faster in areas with vertisols soil type than those with luvisols, solonetz or the other soil types.

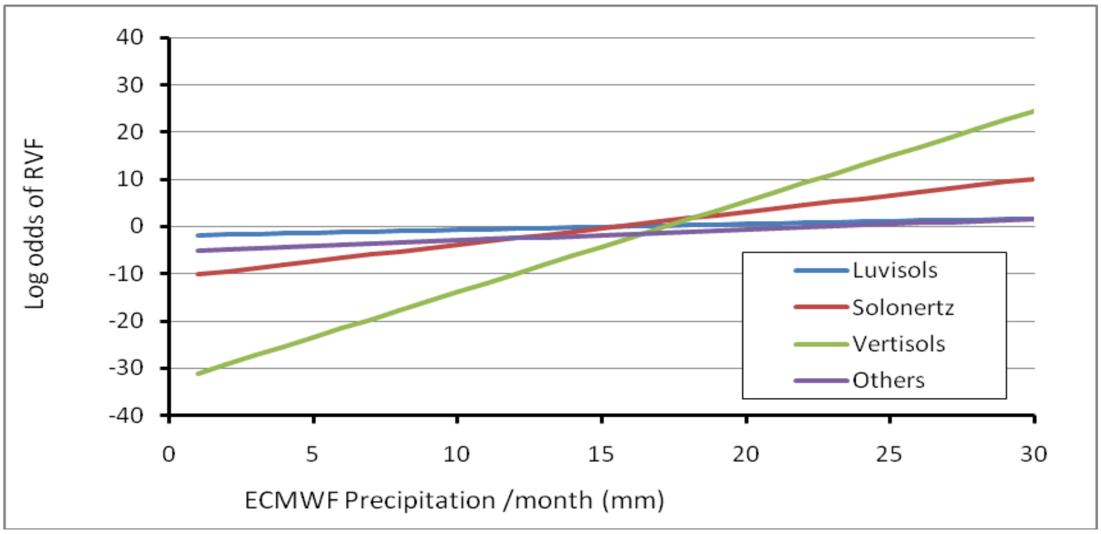

Figure 4.11: Interaction between Soil type and Rainfall

4.4 Comparison of GARP, Random forest and Logit models

From these results, both GARP and RFs showed consistency in distribution of RVF in Kenya. Random forest had the highest AUC of 0.98 thus an excellent map compared to GARP map. The GARP results showed temperature and rainfall variables influenced RVF outbreak almost equally and NDVI showed highest influence on March of 2007.These variables shows relationship with each other but does not show the significance association of individual variable and their combined effect. This was well elaborated by logit regression where OR at 95% C.I. was used to show how each input increases the odds of RVF outbreak

CHAPTER FIVE

DISCUSSION

The results of this study showed that by using ENM to model the distribution of RVF, the map had the best resolution thus giving strength of this analysis. This is because outbreak sites were geo-referenced to gain a best resolution of occurrence data (instead of using a district as the unit of analysis) and multiple analytical/mapping techniques (GARP, Random Forest and logistic regression) were implemented and their outputs compared. Compared to other studies, this study further complemented the work that has been done on RVF risk mapping in Kenya where initial maps were generated by mapping occurrence/observed data using a district as the unit of analysis. Since then new maps have been developed using serological data from wild ungulates and camels (Britch *et al*., 2013).

This study was also able to characterize land cover and livelihood activities used in the RVF "hotspots" such as: land use, topography, vegetation type which has been shown to be very important (Pearson *et al*., 2003). In this case using remote sensing variables such as NDVI, land use, soil map to map species were used as distribution increases accuracy (Woodward *et al.*, 1997).

This study used three methods to identify RVF niches in Kenya. The first analysis which used Bioclim data did not yield a good map as shown in Figure 4.6. This is because the map had AUC of 0.82 and this is explained from the fact that bioclim data show annual patterns of temperature and rainfall and so they are not sensitive to sudden changes in rainfall which is usually important for RVF occurrence. These bioclim data do not also include other environmental factors that influence the distribution of RVF such as soil type, land-cover type and the distribution of potential hosts. From the map, bioclim model over-predicted the RVF distribution and thus they are not the best predictors to use in this study. Compared to other studies for instance, Petersen *et al* (2005), there is an observation that bioclim data are averaged over 50 years making them less suitable for analyses implemented at the global levels and less so for those implemented at the local /country scale and this concurred to the result obtained from the Bioblim analysis and a lot of information is lost thus the need to use machine learning method the second with time-specific variables (customised for the outbreak period) which gives the best prediction.

In this study, machine learning method; Ecological Niche Model (ENM) used GARP and Random Forests because they use presence only data and they generate automatically absence data, majorly known as pseudo-absence data from pixels where presence data are absent. This does not necessarily mean that they are correct absence data like the one collected in the field (Peterson *et al*, 2007). Though a lot of research on ecological niche modeling using these variables have been done (Berry *et al.*, 2002; Peterson *et al* 2002; Thuiller *et al* 2005; Araujo *et al* 2006), the validity of the approach have been questioned (Araujo *et al*, 2007) but with the approach of both GARP and comparing the output with those of Random Forest showed consistency to the distribution of RVF.

The RVF distribution from this study was predicted better with Random Forests algorithm with area under curve of 0.99 than GARP which had AUC of 0.82 though AUC cannot be used to compare one software to the other (Peterson *et al*., 2007). When the GARP model was evaluated using partial receiver operating characteristic (pROC) value of 1.77 was obtained as opposed to the bioclimatic prediction which had a lower AUC of 0.69 and partial ROC value of 1.10. Thus the map generated from GARP was satisfying as per the AUC results and also the RFs RVF distribution map having the highest AUC and showing consistency compared with GARP RVF distribution map.

From this study, the RVF maps generated were more refined for the outbreak sites were geo-referenced to gain a best resolution of occurrence data used in that it was able to show the distribution of RVF in the country and in a particular province where the disease has never been reported. The map showed potential niches of disease occurrence, for instance Turkana county and the western part of Kenya were shown as RVF free zones though had a potential. Comparing with other studies done, the maps generated from those studies were not refined and were more generalized in that the results mapped RVF using a district as a unit of analysis.

From GARP analysis, jackknife analysis done was to show importance of individual environmental parameter on model. The soil map influence the model by 74.19% with NDVI, which is a measure of the vegetation cover, influenced the model most by 87.09% in March 2007. The rainfall and temperature for December and January influenced the model most. This is because these regions are ASAL areas and normally experience highest temperature and no rainfall at all during this period. However during El Nino, they receive a lot of rainfall that causes flooding therefore influencing the model.

Based on the results of logistic regression analysis, solonetz soil type, solonetz soil type interaction with the last two months cumulative rain, NDVI and temperature were significant factors contributing to RVF occurrence on the variables agreed with other studies that showed RVF outbreak are associated with soil types (solonetz, vertisols, planosols) and increase in precipitation leading to flooding and increase in vegetation cover (Linthicum *et al.,* 1999; Anyamba *et al*., 2009; Hightower *et al*., 2012; Bett *et al*., 2013).

The results of the study agreed with those from the past spatial data analysis which showed altitude as a risk factor which contributes to RVF occurrence at less than 1100m above sea level. However, in this study RVF cases were observed up to 2,300 m above sea level that is around Mount Kenya regions and is in agreement with similar observations were made in Madagascar where RVF was reported to occur in a mountainous region of >1,500m (Chevalier *et al*., 2011). The results of this study further confirm that altitude was significant up to altitude above 1500m above sea level.

Soil is another factor that supports persistence of RVF outbreak. From logistic regression results in this study, solonetz soil type having OR of 1.6 had significance thus agreeing with other studies (Linthicum *et al.,* 1999; Anyamba *et al.*, 2009; Hightower *et al.,* 2012; Bett *et al*., 2013) of their significant association with RVF outbreak. Thus the variables used in this study were actually associated with the outbreak of RVF. The maps generated can be deducted to be the actual distribution of RVF in Kenya and the areas identified have a potential risk of RVF occurrence.

The results from the affected farmers after 2006/2007 RVF outbreak through the questionnaire showed that livestock production combined with crop farming as a source of livelihood activity had the highest proportion of 43%. Extensive livestock production system was practiced in most of the areas where RVF outbreaks were reported. Ruminants (cattle, sheep and goats) contributed to 76% of the livestock species kept. The livestock were important in that they are the hosts for RVF. From the affected farmer's data, community outbreak interventions showed 22.5% of the communities simply wait for outbreaks to occur. Apart from policy makers, researchers will be able to use this information on surveillance of RVF.

From the interviews carried out in the study areas, human cases of RVF were reported and thus confirming the disease is also risky to humans.

CHAPTER SIX

CONCLUSIONS AND RECOMMENDATIONS

6.1 Conclusions

1. This study showed that ecological niche modeling is better placed in generating better maps that show true distribution of the species. This is because the results generated in this study were used to improve the already existing maps. For better planning of mitigation measures, environmental and climatic factors associated with the occurrence of RVF were identified. Correlation was established between the factors and disease outbreak. A comparison of outputs with those of a standard regression model also showed interaction.

2. The study showed that elevation, solonetz soil type, open to closed forest land cover and livestock keeping as part of livelihood activity together with crop production are factors that causes RVF outbreak. Other past studies agree that the factors causing RVF outbreak and includes soil types (solonetz, luvisols, planosols), an altitude of less than 1100 m above sea level and closed to open landcover which is in agreement with various studies that have been done (Linthicum *et al*., 1999; Anyamba *et al*., 2009; Hightower *et al*., 2012; Bett *et al*., 2013).

6.2 Recommendations

1. Human cases were present however, this study was unable to relate its outbreak to the variables that caused RVF outbreak in animals thus a gap that can be researched on to enable achievement of one health concept
2. Simulating future risks of RVF based on climate and land use changes is a gap that needs to be studied in future.
3. Specific training on use of ENM should be carried out as it has not been widely used in RVF studies.
4. When policy makers are implementing the prevention and control programs they should concentrate in areas where the disease shows the potential risk of occurrence.
5. Policy makers should educate the community on importance of vaccination of livestock before perceived outbreak and control of mosquitoes so as to prevent spread of the disease.
6. Further studies needs to be done to confirm if the variables that cause outbreak of RVF in humans are the same as those in livestock to enable mapping of risk of human RVF distribution so as to have a one health approach

REFERENCES

Abdo-Salem S, Tran A, Grosbois V, Gerbier G, Al-Qadasi M, Saeed K, Etter E, Thiry E, Roger F, Chevalier V. (2011) Can Environmental and Socioeconomic Factors Explain the Recent Emergence of Rift Valley Fever in Yemen2000-2001? *Vector Borne Zoonotic Disease 2011 Feb 1.*

Alexander, R. A. (1951). Rift Valley Fever in the Union. *Journal of the South African Veterinary Medical Association*, **22**:105-109.

Anyamba, A., Chretien, J. P., Small, J., Tucker, C. J., Formentry, P. B., Richardson, J. H., Britche, S.C., Schnabelf, D. C., Ericksonb, R. L., Linthicume K. J. (2009). *Prediction of a Rift Valley Fever outbreak.* Proceedings of the National Academy of Sciences, 106(3): 955-959.

Anonymous (1910). Kabete Veterinary Laboratory Report

Bernard, B. J. (1979). Rift Valley Fever vaccine- antibody and immune response in cattle to live and inactivated vaccine. *Journal of the South African Veterinary Associations*, **50**:155-157.

Barnard, B. J. (1981). Rift Valley Fever in South Africa. *Proceedings of the 49th General Session of the office International Des Epizooties, Paris* , 25-30.

Besselaar, T. G., & Blackburn, N. K. (1991). Topological mapping of the antigenic sites on the Rift Valley Fever virus envelope glycoproteins using monoclonal antibodies. *Archives of Virology* , **121**:11-124.

Bett, B., Kiplimo, J., Notenbaert, A., & Kamp, S. (2013). Quantifying Weather and Climate Impacts on Health in Developing Countries(QWeCI): Mapping the distribution of potential Rift Valley Fever hotspots in East Africa. *4th Annual East*

African Community Health and Scientific Conference. Kigali.

Daurbney R, H. J. (1931). Enzootic hepatitis or Rift Valley Fever: an undescribed virus disease of sheep cattle and man from East Africa. *Pathology and Bacteriology,* **34**:545-576.

Davies, F. G., & Highton, R. B. (1980). Possible Vectors of Rift Valley Fever in Kenya. *Transactions of the Royal Society of Tropical Medicine and Hygiene journal,* **74**:815-816.

Davies, F. G., Kilelu, E., Linthicum, K. J., & Pegram, R. G. (1992). Patterns of Rift Valley Fever activity in Zambia. *Epidemiology and Infection* , **108**:185-191.

Donald, D. R., Mary, E. P., Sonja, R. G., & Janet, L. S. (2010). Structure of theRift Valley Fever Virus nucleocpsid protein reveals another architecture for RNAencapsidation. *PNAS 2010* , **107(26):** 11769-11774.

Drosten C., Gottig S., Schilling S., Asper M., Panning M., Schmitz H., (2002). Rapid Detection of Quantification of RNA of Ebola and Marburg Viruses, Lassa Virus, Crimean-Congo Hemorrhagic Fever, Rift Valley Fever, Dengue Virus and Yellow Fever Virus by Real Time Reverse Transcription PCR. *Journal of Clinical Microbiology* **40**, 2323 – 2330.

Eddy, G. A., Peters, C. J., Meadors, G., & Cole, F. E. (1981). Rift Valley Fever Vaccine in humans. *Contributions to Epidemiology and Biostatistics, 3. Rift Valley Fever* , 124-141.

El Akkad, A. M. (1978). Rift Valley Fever in Egypt. *Egypt Public Health Associations***,** **53**:137-146.

Evans , A., Gakuya, F., Paweska, J. T., Rostal, M., Akoolo, L., Van Vuren, P. J.,

Manyibe T., Macharia J.M., Ksiazek T.G., Feikin D.R., NjengaM.K., (Hay, et al., 2013)**(2008).** Prevalence of antibodies against Rift Valley Fever Virus in Kenyan Wildlife. *Epidemiology and Infection* , **136:**1261-1269.

Food and Agricultural Organization of the United Nations (FAO) (2002 and 2003) Preparation of Rift Valley fever contingency plans, 2002; Recognizing Rift Valley Fever, 2003

Food and Agricultural Organization of the United Nations (FAO) (2015) FAOSTAT, Website; http://faostat3.fao.org/

Frank, P. N. (2000). Response of laboratory staff to vaccinatiob with an inactivated Rift Valley Fever Vaccine-TSI-GSD 200. *African Journal of Medicine and Medical Sciences* , **29**:930-932.

Frankline, J. (2009).*Mapping species distributin: spatial inference and prediction.* UK: Cambridge University Press, Cambridge.

Garcia S., Crance J.M., Brillecocq A., Peinnequin A., Jouan A., Bouloy M., Garin D. (2001). Quantitative real time PCR detection of Rift Valley Fever virus and its application to evaluation of antiviral compounds. *Journal of Clinical Microbiology*, **39**: 4456 - 4461

Gear J, H. S., De Meillon, B., Le Roux, A. F., Kofski, R., Rose, I. R., & Steyn, J. J. (1955). Rift Valley Fever in South Africa: study of the 1953 outbreaks in the Orange Free nState, with special reference to the vectors and possible reservoir hosts. *South African Medical Journal* , **29**:514-518.

Gerdes, G. H. (2004). Rift Valley Fever. *Revenue Scientific et Technique* , **23**:613-623.

Kenya National Bureau of Statistics (KNBS) (2015) Website, http://www.knbs.or.ke/

Hay, S. I., Battle, K. E., Pigott, D. M., Smith, D. L., Moyes, C. L., Bhatt, S., et al. (2013). *Global mapping of infectious disease.* uk: Royal society.

Hightower, A., Kinkade, C., Nguku, P. M., Anyangu, A., Mutonga, D., Omolo, J., Njenga M. K., Feikin D. R., Schnabel D., Ombok M., Breiman R. F. (2012). Relationship of Climate, Geography, and Geology to the Incidence of Rift Valley in Kenya during the 2006-2007 Outbreak. *American Society of Tropical Meddicine and Hygiene***, 86(2):**373-380.

Intergovernmental Panel on Climate Change (IPCC) (2007). *Contribution of working group II to the Fourth assesment Report on the Intergovermental Panel on Climate Change.* Cambridge, UK, 976 pp: Cambridge

University Press.

J, F. (2009).*Mapping species distribution: spatial inference and prediction.* UK: Cambridge University Press, Cambridge.

Kark, J. D., Aynor, y., & Peters, C. J. (1982). A Rift Valley Fever Vaccine Trial, I. Side Effects and Serological Response over a Six Month Follow-Up. *African Journal of Epidemiology* , **116**:808-820.

Ksiazek T.G, Jouan A., Meegan J.M., Leguenno B., Wilson M.I. and Peters C.J., (1989). Rift Valley Fever Among Domestic Animals in the Recent West African Outbreak. *Research in Virology,* **140**: 67-77.

Laughing , L. W., Meegan, J. M., Strasbaugh, L. J., Morens, D. M., & Watten, H.

(1979). Epidemic of Rift Valley Fever in Egypt. Observation of the Spectrum of Human Illness. *Transactions of the Royal Society of Tropical Medicine and Hygiene***,** **73:**630-633.

Linthicum KJ, B. C. (1988). The horizontal distribution of Aedes pupae and their subsequent adult within a flooded dambo in Kenya: implications for rift valley fever control. *j. Am . Mosq. Contol Assoc.* **4**, 551-554.

Linthicum, K. J., Bailey , C. L., Tucker , C. J., Mitchell, K. D., Logan, T. M., & Davies, F. G. (1990). Application of Polar Orbiting Meteorological Satellite Data to Detect Flooding of Rift Valley Fever Virus Vector Mosquito Habitats in Kenya. *Medical and Veterinary Entomology***, 4**:433-438.

Linthicum , K. J., Anyamba, A., Tucker, C. J., Kelley, P. W., Myers, M. F., & Peters, C. J. (1999). Climate and Satellite indicators to Forecast Rift Valley Fever Epidemics in Kenya. *Science* , **285**:397-400.

Linthicum , K. J., Anyamba, A., & Tucker, C. J. (2001). Climate-disease Connections Rift Valley Fever in Kenya. *Cadernos de Saude Publica 2001* , **17**:133-140.

Logan , T. M., Linthicum, K. J., Wagateh, J. M., Thande, P. C., Kamau, C. W., & Roberts, C. R. (1990). Pretreatment of Floodwater Aedes Habitats (Dambos) in Kenya with a Sustainable Release Formulation of Methoprene. *Journal of the American Control Associations***, 6:**736-738.

Madani , T. A., Mazrou, Y. Y., Jeffri, M. H., Mishakhas, A. A., Rabeah, A. M., Turkistani, A. M., et al. (2003). Rift Valley Fever Epidemic in Saudi Arabia: Epidemiological, Clinical and Laboratory Characteristics. *Clinical Infectious Diseases* , **37:**1084-1092.

Martin, V., Chevalier, V., Ceccato, p., Anyamba , A., De Simone , L., Lubroth J, et al. (2008). The impact of climate change on the epidemiology and control of Rift Valley Fever. In climate change: impact on the epidemiology and control of animal diseases (S. de La Rocque, S. Morand, G. Hendrickx). *Rev. sci. tech. Off. int. Epiz* , **27(2)** 413-426.

Mcintosh, B. M. (1972). Rift Valley Fever: I. Vector Studies in the field. *Journal of the South Africa Veterinary Association* , **58**:127-132.

Mcintosh, B. M. (1973). A taxonomic re-assesment of Aedes (Ochlerotatus) caballus (Theobald) (Diptera: Culicidae) including a description of a new of Ochleratus. *Journal of Entemological Society of South Africa*.

Mcintosh, B. M., Jupp, P. G., Dos Santos, I., & Barnard, B. J. (1980). Vector Studies on Rift Valley Fever Virus in South Africa. *South African Medical Journal***,** **58**:127-132.

Meegan, J. M., & Bailey, C. H. (1988). Rift Valley Fever In Monath T.P (ed) The Arboviruses: Epidemiology and Ecology VOL IV. *Boca Raton: CRC Press Inc* , 51-76.

Montgomery R.E and Stordy R. J. (1912).*Nairobi Sheep Disease.* Annual report Department of Agriculture British East Afrrica.

Munyua, P., Murithi, M., Ithondeka, P., Hightower, A., Anyangu, S. A., Bett, B., et al. (n.d.). Risk Mapping for Rift Valley Fever Epidemics and Identification of Predictive Factors of High Risk Regions, Kenya. *In Press*.

Murithi, R.M., Munyua, P., Ithondeka, P.M., Macharia, J.M., Hightower,A., Luman.E.T., Breiman, R.F., Njenga, M.K., (2011). Rift Valley fever in Kenya: history of epizootics and identifications of vulnerable districts. . *Epidemiol Infect* **139(3):**372-80.

Niklasson B., Peters C.J, Grandien M.A., and Wood O., (1984). Detection of human immunoglobulin G and M antibodies to Rift Valley Fever Virus by enzyme linked immunosorbent assay. *Journal of Clinical Microbiology,* **19**: 225-229.

Radostits, O.M., Clive C.G., Douglas C.B., and Kenneth W. H (2000). Veterinary Medicine: A Textbook of the Diseases of Cattle, Sheep, Goats, Pigs and Horses, ninth ed. W.B. Saunders Company Ltd, London, pg,1042-1043.

Ropelewski, C. F., & Halpert, M. S. (1989). Global and Regional Scale Precipitation Patterns Associated with El Nino/Southern Oscillation. *Monthly weather review* , **115**:1606-1627.

Saluzzo, J. F., Digoutte, J. P., Chartier, C., Martinez, D., & Bada, R. (1987). Focus of Rift Valley Fever Transmission in Southern Mauritania . *Lancet* , **1:**504.

Smithburn, K. C., Nagaffy, A. F., Haddow, A. J., Kitchen, S. F., & Smith, J. F. (1949). Rift Valley Fever:mAccidental infections among Laboratory workers. *Journal of Immunology* , **62**:213-227.

Swanapoel, R., Manning, B., & Watt, J. A. (1979). Fatal Rift Valley Fever Affecting Humans in South Africa: A clinic- Pathological Study. *Central African Journal of Medicine* , **25**;1-8.

Swanepoel, R., & Cruikshank, J. G. (1974). Arthropode borne viruses of medical importance in Rhodesia. *Central African journal of medicine* , **20**:71-79.

Van Velden, D. J., Meyer, J. D., Oliver , J., Gear , J. H., & Mcintosh, B. (1977). Rift Valley Fever Affecting Human in South Africa. *South African Medical Journal ,* **51**:867-871.

Whittler, K., Linthicum, K. J., Thande, P. C., Wagati, J. N., Kamau, C. M., & Roberts, C. R. (1993). Effects of controlled burning of survival of floodwater Aedes eggs in Kenya. *Journal of the American mosquito control associations***, 9**:72-77.

World Bank, (2015) *World Development Indicators* 2015, World Bank, Washington, DC.

World Health Organization (WHO) (2000) Fact sheet no. 207

Zeller , H. G., Fontenille, D., Traore-Lamizana, M., Thiongane, Y., & Digoutte, J. P. (1997). Enzootic Activity of Rift Valley Fever Virus in Senegal. *Journal of Medical Hygiene* , **56(3)**: 265-272.

APPENDICES

Appendix I: Questionnaire

Date of the interview	
Name of the researcher	
Name of the respondent	
Mobile no. of the respondent	
Village:	Sub-location:
Location:	district:
GPS coordinates (decimal degrees)	Latitude:______________________________________ Longitude: ______________________________________ Altitude:______________________________________
What is the dominant livelihood activity of the village (pursued by a majority of the people in the village)? Estimate % of households depending on each activity	Livestock keeping [__] __________% Crop farming [__] __________% Fishing [__] __________% Business [__]__________% Employment [__]__________% Others (specify)________________________________
If livestock farming is the key livelihood	Extensive [__]

activity, indicate production system used	Semi-intensive [__] Intensive [__] Other________________________________
Which livestock species are kept in the village (can tick more than one)?	Cattle [_] Sheep [_] Goats [_] Camels [_] Others:______________________________
What is the dominant land cover/vegetation type in the village (tick more than one)?	
Savannah grassland [__] _________% Artificial surfaces (towns) [__]_________% Cultivated area [__] _________% Forest [__]_________% Shrub land [__]_________% Water bodies [__] _________% Mosaic (cropland/tree cover/grassland) [__]_________% Other____________________________	
What are the common wildlife species found in the area (give a list)?	
Have there been outbreaks of RVF in the area? If yes, describe how the disease appeared?	Yes [_] No [_]

(clinical signs and post mortem lesions)	
Year(s) when the area was affected by RVF (the most recent outbreak(s))	
What do you associate the outbreak with?	
Which livestock species were affected by the outbreak (can tick more than one)?	Cattle [_] Sheep [_] Goats [_] Camels [_] Others:______________________________
Were there any human cases in the village?	Yes [_] No [_]
What measures were taken to manage the outbreak	
What is the community doing to enable them manage any future outbreaks	
Any other information	

Appendix II: Summary of Jackknife Analysis for Bioclim Variables

Layer	Accuracy without layer
October_06_ndvi	77.4194
November_06_ndvi	70.9677
November_06_ndvi	70.9677
december_06_ndvi	67.7419
december_06_ndvi	54.8387
january_07_ndvi	74.1935
january_07_ndvi	74.1935
february_07_ndvi	80.6452
february_07_ndvi	67.7419
march_07_ndvi	87.0968
march_07_ndvi	67.7419
december_06_rainfall	70.9677
november_06_rainfall	70.9677
october_06_rainfall	70.9677
february_07_rainfall	70.9677
january_07_rainfall	87.0968
march_07_rainfall	64.5161
Altitude	61.2903
Landcover	67.7419
Soil	74.1935
december_06_temperature	87.0968
february_07_temperature	70.9677
january_07_temperature	90.3226
march_07_temperature	64.5161
november_06_temperature	61.2903
october_06_temperature	70.9677
Accuracy	74.1935
Bias	-49.6278

SUMMARY OF JACKKNIFE ANALYSIS FOR BIOCLIM VARIABLES

Layer	**Accuracy without layer**
bio_1.asc	90
bio_10.asc	83.3
bio_11.asc	83.3
bio_12.asc	86.7
bio_13.asc	83.3
bio_14.asc	90
bio_15.asc	70
bio_16.asc	76.7
bio_17.asc	80
bio_18.asc	76.7
bio_19.asc	70
bio_2.asc	80
bio_3.asc	83.3
bio_4.asc	86.7
bio_5.asc	76.7
bio_6.asc	73.3
bio_7.asc	76.7
bio_8.asc	83.3
bio_9.asc	90
Accuracy	76.6667
Bias	78.9474

Appendix III: Data Set Summary

Variable Total	Total N= 26232
NDVI	
median(IQR)	0.3 (0.2,0.5)
Lag rain	
median(IQR)	0.2 (0,1.5)
Temp	
median(IQR)	298.6 (295.3,300.2)
Logcattle	
median(IQR)	2 (0.8,2.9)
Goats	
median(IQR)	10.1 (4.8,21.5)
Camels	
median(IQR)	1.1 (0,2.9)
height_m	
median(IQR)	595 (309.8,1174.2)
Symbol	
	504 (1.9)
Acrisols	480 (1.8)
Alisols	48 (0.2)
Andosols	480 (1.8)
Arenosols	1584 (6)
Calcisols	792 (3)
Cambisols	2640 (10.1)
Chernozems	24 (0.1)
Ferralsols	2136 (8.1)
Fluvisols	840 (3.2)
Gleysols	456 (1.7)
Greyzems	48 (0.2)
Histosols	72 (0.3)
Leptosols	768 (2.9)
Lixisols	1104 (4.2)
Luvisols	1968 (7.5)
Nitisols	1248 (4.8)
Phaeozems	840 (3.2)
Planosols	2160 (8.2)
Regosols	2544 (9.7)

Solonchaks	504 (1.9)
Solonetz	3744 (14.3)
Vertisols	1248 (4.8)
Texture	
	504 (1.9)
Coarse	2664 (10.2)
Fine	11160 (42.5)
Medium	11904 (45.4)
Type	
	504 (1.9)
clay (light)	4704 (17.9)
clay loam	3864 (14.7)
clay(heavy)	3840 (14.6)
Loam	1680 (6.4)
loamy sand	744 (2.8)
Sand	1128 (4.3)
sandy clay	648 (2.5)
sandy clay loam	6168 (23.5)
sandy loam	2424 (9.2)
silt loam	264 (1)
silty clay	168 (0.6)
silty clay loam	96 (0.4)
Landcover	
Artificial areas	24 (0.1)
Bare areas	2352 (9)
Closed broadleaved deciduous forest	648 (2.5)
Closed broOpen broadleaved deciduous forestadleaved deciduous forest	480 (1.8)
Closed to open broadleaved evergreen or semi-deciduous forest	408 (1.6)
Closed to open broadleaved forest regularly flooded (fresh-brackish water)	48 (0.2)
Closed to open grassland	6360 (24.2)
Closed to open shrubland	1032 (3.9)
Mosaic Croplands/Vegetation	2352 (9)
Mosaic Forest-Shrubland/Grassland	5112 (19.5)
Mosaic Grassland/Forest-Shrubland	192 (0.7)
Mosaic Vegetation/Croplands	4152 (15.8)
Open needleleaved deciduous or evergreen forest	24 (0.1)
Rainfed croplands	600 (2.3)

Sparse vegetation	1944 (7.4)
Water bodies	504 (1.9)
Soiltype	
Acrisols	480 (1.9)
Alisols	48 (0.2)
Andosols	480 (1.9)
Arenosols	1584 (6.2)
Calcisols	792 (3.1)
Cambisols	2640 (10.3)
Chernozems	24 (0.1)
Ferralsols	2136 (8.3)
Fluvisols	840 (3.3)
Gleysols	456 (1.8)
Greyzems	48 (0.2)
Histosols	72 (0.3)
Leptosols	768 (3)
Lixisols	1104 (4.3)
Luvisols	1968 (7.6)
Nitisols	1248 (4.9)
Phaeozems	840 (3.3)
Planosols	2160 (8.4)
Regosols	2544 (9.9)
Solonchaks	504 (2)
Solonetz	3744 (14.6)
Vertisols	1248 (4.9)

Appendix IV: Summary tables of interviewed farmers

ID	interview date	respondent name	village	sub-location	location	livestock keeping	crop farming	fishing	business	employment	extensive production	semi-intensive production	intensive production	cattle kept	sheep kept	goat kept	camel kept	others species kept	human cases
1	15-Aug-13	Julia Naitole	Runywene	Gaitu	Gaitu West	TRUE	TRUE	TRUE	TRUE	FALSE	FALSE	TRUE	FALSE	TRUE	TRUE	TRUE	FALSE	TRUE	FALSE
2	14-Aug-13	Garishon Kaae	Kaborene	Njuki Njiru	Miriga	TRUE	TRUE	FALSE	FALSE	FALSE	FALSE	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE	TRUE	FALSE
3	14-Aug-13	David kamathi	karimaiga	Bugui	muranthakari	TRUE	TRUE	FALSE	FALSE	FALSE	FALSE	TRUE	FALSE	TRUE	TRUE	TRUE	FALSE	TRUE	FALSE
4	14-Aug-13	Erick Gitonga	Gantukune	Gakoromone	Kooje Municipality	TRUE	TRUE	FALSE	TRUE	FALSE	FALSE	FALSE	TRUE	TRUE	TRUE	TRUE	FALSE	FALSE	FALSE
5	14-Aug-13	Erick Ngaruni	Kinani	Nkoune	Kaaga	TRUE	TRUE	FALSE	FALSE	FALSE	FALSE	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE	FALSE	FALSE
6	14-Aug-13	Lydia Kaburu	Angirine	Kemuitari	Thuurta	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE	TRUE	FALSE	TRUE	TRUE	TRUE	FALSE	TRUE	FALSE
7	14-Aug-13	Martin Mutethia	Ntima	Kagaa	Kambakia	TRUE	TRUE	TRUE	FALSE	TRUE	FALSE	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE	TRUE	TRUE
8	13-Aug-13	Mary Wambui	Karuku	Wachoro	Karaba	TRUE	TRUE	TRUE	TRUE	FALSE	FALSE	TRUE	FALSE	TRUE	TRUE	TRUE	FALSE	TRUE	FALSE
9	13-Aug-13	Peter Muthii	Gakendu	Gategi	Karaba	TRUE	TRUE	FALSE	FALSE	FALSE	FALSE	TRUE	FALSE	TRUE	TRUE	TRUE	FALSE	FALSE	TRUE
10	12-Aug-13	Charles Njiru	Ngoce	Ndurumori	Ndurumori	TRUE	TRUE	FALSE	FALSE	FALSE	FALSE	TRUE	FALSE	TRUE	TRUE	TRUE	FALSE	TRUE	FALSE
11	12-Aug-13	Margaret Wanjovi	Karurumo	Karurumo	Karurumo	TRUE	TRUE	TRUE	FALSE	FALSE	TRUE	TRUE	FALSE	TRUE	TRUE	TRUE	FALSE	TRUE	TRUE
12	12-Aug-13	Fredrick Maringa	Ciamugu	Evurore	Ishiara	TRUE	TRUE	FALSE	FALSE	FALSE	FALSE	TRUE	FALSE	TRUE	TRUE	TRUE	FALSE	TRUE	FALSE
13	09-Aug-13	Joseph Machira	Research	Tebere	Tebere	TRUE	TRUE	FALSE	FALSE	FALSE	FALSE	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE	FALSE	FALSE
14	09-Aug-13	Stephen Wamwea	Kiyuyu	Rukanga	Rukanga	TRUE	TRUE	TRUE	FALSE	FALSE	TRUE	TRUE	FALSE	TRUE	TRUE	TRUE	FALSE	TRUE	FALSE
15	09-Aug-13	Eunice Wanjau	Maganjo	Kariti	Sagana	TRUE	TRUE	TRUE	FALSE	FALSE	FALSE	TRUE	FALSE	TRUE	TRUE	TRUE	FALSE	FALSE	TRUE
16	09-Aug-13	Zacharia Njeru	Kiajang'a	Kiajang'a	Mwerua	TRUE	TRUE	FALSE	FALSE	FALSE	FALSE	TRUE	FALSE	TRUE	TRUE	TRUE	FALSE	TRUE	TRUE
17	08-Aug-13	Newton Maina	Burguret	Gathiuru	Gakawa	TRUE	TRUE	FALSE	FALSE	FALSE	TRUE	FALSE	FALSE	TRUE	TRUE	TRUE	FALSE	TRUE	TRUE
18	08-Aug-13	Lucy Wamaitha	Ngamwa	Ngamwa	Rutune	TRUE	TRUE	FALSE	TRUE	FALSE	FALSE	TRUE	FALSE	TRUE	TRUE	TRUE	FALSE	TRUE	TRUE
19	07-Aug-13	Olivia Mungai	Gikindu	Gakoigo	Ngenda	TRUE	TRUE	FALSE	FALSE	FALSE	FALSE	TRUE	FALSE	TRUE	FALSE	FALSE	FALSE	TRUE	TRUE
20	07-Aug-13	Martha Njambi	Ngaru	Kiria	Kiria	TRUE	TRUE	FALSE	TRUE	FALSE	FALSE	TRUE	FALSE	TRUE	TRUE	TRUE	FALSE	TRUE	FALSE
2	07-Aug-13	Pius Irungu	Wathiani	Wathiani	Sabasaba	TRUE	TRUE	TRU	FALSE	FALSE	FALSE	TRUE	FALSE	TRUE	TRUE	TRUE	FALSE	TRUE	FALSE

1								E											
22	07-Aug-13	Joseph Ndolo	Upendo Rurii	Mirira	Gikindu	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE	TRUE	FALSE	TRUE	TRUE	TRUE	FALSE	TRUE	FALSE
23	20-Aug-13	Stanley King'ori	Juja Farm	Kalimoni	Juja	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE	TRUE	TRUE	TRUE	TRUE	FALSE	TRUE	FALSE
24	19-Aug-13	Symon Kariuki	Gatong'ora	Gatong'ora	Gikumari	TRUE	TRUE	FALSE	TRUE	FALSE	TRUE	FALSE	TRUE	TRUE	TRUE	TRUE	FALSE	TRUE	TRUE
25	18-Aug-13	Josphat Kathurima	Kamuramba	Muriinya	Ntugi	TRUE	TRUE	FALSE	TRUE	TRUE	FALSE	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE	TRUE	FALSE
26	18-Aug-13	Elvis Koome	Maitei	Naari	Naari	TRUE	TRUE	FALSE	FALSE	FALSE	TRUE	FALSE	TRUE	TRUE	TRUE	TRUE	FALSE	TRUE	FALSE
27	18-Aug-13	Elius Riungu	Kanondone	Maitei	Maitei	TRUE	TRUE	FALSE	FALSE	FALSE	TRUE	FALSE	TRUE	TRUE	TRUE	TRUE	FALSE	TRUE	FALSE
28	17-Aug-13	Damaris Karimi	Baibariu	Baibariu	Kawiru	TRUE	TRUE	FALSE	FALSE	FALSE	TRUE	FALSE	FALSE	TRUE	TRUE	TRUE	FALSE	TRUE	TRUE
29	17-Aug-13	Rose Mwikali	Kiamuriuki	Nthambo	Mugumuni	TRUE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	TRUE	TRUE	TRUE	FALSE	FALSE	TRUE	FALSE
30	17-Aug-13	Mweandi Mbuna	Kathurine	Mugumango	Mikui	TRUE	TRUE	FALSE	FALSE	FALSE	FALSE	TRUE	TRUE	TRUE	TRUE	FALSE	FALSE	TRUE	FALSE
31	17-Aug-13	Abednego Gitonga	Kithima	Kiraro	Chogoria	TRUE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	TRUE	TRUE	TRUE	TRUE	FALSE	TRUE	FALSE
32	16-Aug-13	Charles Kimani	Manyangaro	Ngarendare	Ngarendare	TRUE	TRUE	FALSE	TRUE	FALSE	TRUE	FALSE	FALSE	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE
33	16-Aug-13	Jane Wanjiru	Ethi	Ethi	Mugogondo	TRUE	FALSE	FALSE	TRUE	FALSE	TRUE	FALSE	FALSE	TRUE	TRUE	TRUE	FALSE	TRUE	FALSE
34	16-Aug-13	Symon Gekunda	Rugindaru	Kirimara	Timau	TRUE	TRUE	FALSE	FALSE	FALSE	TRUE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
35	15-Aug-13	Frankline Mwiti	Ntharene	Ntharene	Kithangari	TRUE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	TRUE	TRUE	TRUE	TRUE	FALSE	FALSE	FALSE
36	15-Aug-13	Stella Mutweri	Ntonyero	Kiria	Kiria	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE	FALSE	TRUE	TRUE	TRUE	TRUE	FALSE	TRUE	TRUE
37	15-Aug-13	Ireri Gerald	Karia	Karia	Igoji West	TRUE	TRUE	TRUE	TRUE	FALSE	FALSE	TRUE	FALSE	TRUE	TRUE	TRUE	FALSE	TRUE	FALSE
38	15-Aug-13	Wilson Mathiu	Tune	Kilendene	Mitunguni	TRUE	TRUE	FALSE	TRUE	FALSE	FALSE	TRUE	FALSE	TRUE	TRUE	TRUE	FALSE	FALSE	FALSE
39	15-Aug-13	Charles Mutuiri	Nguchia	Mbajone	chaaria	TRUE	TRUE	FALSE	FALSE	FALSE	FALSE	TRUE	FALSE	TRUE	TRUE	FALSE	FALSE	FALSE	FALSE
40	15-Aug-13	Peter Kimathi	Kirirwa	Kiria	Kiria	TRUE	TRUE	FALSE	TRUE	FALSE	FALSE	FALSE	TRUE	TRUE	TRUE	TRUE	FALSE	FALSE	FALSE
41	15-Aug-13	Festus Gitonga	Gitie	Gitie	Mujwa	TRUE	TRUE	TRUE	FALSE	FALSE	FALSE	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE	TRUE	FALSE
42	15-Aug-13	Francis Mungakia	Rugongo	Kiria	Kiria	TRUE	TRUE	TRUE	FALSE	FALSE	FALSE	FALSE	TRUE	TRUE	TRUE	TRUE	FALSE	TRUE	TRUE
43	05-Aug-13	Geoffrey Ruto	Tabar Kasige	Kimaus	Koibirir	TRUE	TRUE	FALSE	TRUE	TRUE	TRUE	FALSE	FALSE	TRUE	TRUE	TRUE	FALSE	TRUE	FALSE
44	06-Aug-13	Peter chebii	Resim	Cheptembererwo	Chesumen	TRUE	TRUE	FALSE	FALSE	FALSE	TRUE	FALSE	FALSE	TRUE	TRUE	TRUE	FALSE	TRUE	FALSE
4	07-Aug-13	Dr. Merisya		West Pokot		TRUE	TRUE	FALS	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE

5		James						E											
46	25-Jul-13	Francis Lanaiba	Ndonyo Wasin	Ndonyo Wasin	Ndonyo Wasin	TRUE	FALSE	FALSE	TRUE	TRUE	TRUE	FALSE	FALSE	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE
47	25-Jul-13	Lepine Lerurini	Laresoro	Laresoro	Loseria	TRUE	FALSE	FALSE	TRUE	TRUE	TRUE	FALSE	FALSE	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE
48	25-Jul-13	Joseph Lesubeer	Lerata	Lerata	Waso East	TRUE	FALSE	FALSE	TRUE	TRUE	TRUE	FALSE	FALSE	TRUE	TRUE	TRUE	FALSE	TRUE	FALSE
49	25-Jul-13	Peter Paraine	Sere Olipi	Sere Olipi	Sere Olipi	TRUE	FALSE	FALSE	TRUE	TRUE	TRUE	FALSE	FALSE	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE
50	18-Jul-13	Patrick Muthee	Sweet waters	Marura	Marura	TRUE	TRUE	FALSE	TRUE	TRUE	FALSE	TRUE	FALSE	TRUE	TRUE	TRUE	FALSE	TRUE	FALSE
51	24-Jul-13	Lelewai Lerungus	Chongoti	Thome	Mutara	TRUE	TRUE	FALSE	TRUE	TRUE	TRUE	FALSE	FALSE	TRUE	TRUE	TRUE	FALSE	TRUE	TRUE
52	15-Jul-13	Wesley Lelerima	Kiserian B	Kiserian	Kiserian	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE	FALSE	TRUE	TRUE	TRUE	FALSE	TRUE	TRUE
53	15-Jul-13	Lekumbe Ngiruchi	Sokotei	Kiserian	Kiserian	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE	FALSE	TRUE	TRUE	TRUE	FALSE	TRUE	TRUE
54	16-Jul-13	Paul Kariithi	Cinder Wood Farm	Timau	Timau	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
55	19-Jul-13	Peter Jessel	Jessel Ranching	Impala	Segera	TRUE	FALSE	FALSE	FALSE	FALSE	TRUE	FALSE	FALSE	TRUE	TRUE	TRUE	FALSE	FALSE	FALSE
56	11-Sep-13	Siad Abdulahi	Shantabaq	Ahamedtukale	Guteli	TRUE	TRUE	FALSE	FALSE	FALSE	TRUE	FALSE	FALSE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
57	09-Sep-13	Benson Muthu	Mang'uu	Kavuti	Ngomeni	TRUE	TRUE	FALSE	FALSE	FALSE	FALSE	TRUE	FALSE	TRUE	FALSE	TRUE	FALSE	TRUE	FALSE
58	09-Sep-13	Paul Muthui	Malawa	Mitamisyi	Mitamisyi	TRUE	TRUE	FALSE	FALSE	FALSE	TRUE	FALSE	FALSE	TRUE	TRUE	TRUE	FALSE	TRUE	FALSE
59	09-Sep-13	Mwanzia Masya	Ikime	Kavaani	Kavaani	TRUE	TRUE	FALSE	FALSE	FALSE	TRUE	FALSE	FALSE	TRUE	TRUE	TRUE	FALSE	TRUE	TRUE
60	06-Sep-13	Erick Mwema	Maliku	Maliku	Maliku	TRUE	TRUE	TRUE	FALSE	FALSE	TRUE	TRUE	FALSE	TRUE	TRUE	TRUE	FALSE	TRUE	FALSE
61	06-Sep-13	Hellena Kasemba	Kalikuvu	Kakuuni	Itoteka	TRUE	TRUE	FALSE	FALSE	FALSE	FALSE	TRUE	FALSE	TRUE	FALSE	TRUE	FALSE	TRUE	FALSE
62	05-Sep-13	Dina Wakula	Ithumula	Ngungi	Sombe	TRUE	TRUE	FALSE	FALSE	FALSE	TRUE	FALSE	FALSE	TRUE	TRUE	TRUE	FALSE	TRUE	FALSE
63	05-Sep-13	Mbaluka Kitheka	Kinanie	Ndetani	Endau	TRUE	TRUE	FALSE	FALSE	FALSE	TRUE	FALSE	FALSE	TRUE	TRUE	TRUE	FALSE	TRUE	FALSE
64	04-Sep-13	Julius Nziga	Mutulu	Kyoani	Ikutha	TRUE	TRUE	FALSE	FALSE	FALSE	TRUE	FALSE	FALSE	TRUE	TRUE	TRUE	FALSE	TRUE	FALSE
65	04-Sep-13	Ephantus Mwangangi	Ngozi	Kathungu	Ikanga	TRUE	TRUE	FALSE	TRUE	FALSE	TRUE	FALSE	FALSE	TRUE	TRUE	TRUE	FALSE	TRUE	FALSE
66	04-Sep-13	Nzeeni Mbithuka	Ndileu	Kituti	Athi	TRUE	TRUE	FALSE	FALSE	FALSE	TRUE	FALSE	FALSE	TRUE	FALSE	TRUE	FALSE	TRUE	FALSE
67	04-Sep-13	Wambua Kimwele	Yakilindi	Kalambani	Muthao	TRUE	TRUE	FALSE	FALSE	FALSE	TRUE	FALSE	FALSE	TRUE	TRUE	TRUE	FALSE	TRUE	FALSE
68	10-Sep-13	Pauline David	Ngauluka	Ukasi	Ukasi	TRUE	FALSE	FALSE	FALSE	TRUE	TRUE	FALSE	FALSE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE

69	11-Sep-13	Bare Osman	Orahei	Urgaad	Danyiri	TRUE	TRUE	FALSE	TRUE	FALSE	TRUE	FALSE	FALSE	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE
70	12-Sep-13	Siradhu Hussein	Bula Argi	Bula Argi	Bula Argi	TRUE	TRUE	FALSE	FALSE	FALSE	TRUE	TRUE	FALSE	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE
71	12-Sep-13	Magow Kassim	Didkalkash	Madogo	Madogo	TRUE	FALSE	FALSE	FALSE	FALSE	TRUE	FALSE	FALSE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
72	11-Sep-13	Abdi Abdulahi	Buratensa	Balambala	Balambala	TRUE	TRUE	FALSE	FALSE	FALSE	TRUE	FALSE	FALSE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
73	08-Oct-13	Willy Maingi	Konza	Mumandu	Lumbwa	TRUE	FALSE	FALSE	FALSE	FALSE	TRUE	FALSE	FALSE	TRUE	TRUE	TRUE	FALSE	FALSE	TRUE
74	08-Oct-13	David Muli Mutiso	Miwani	Mjini	Township	TRUE	TRUE	FALSE	TRUE	TRUE	TRUE	FALSE	TRUE	TRUE	TRUE	TRUE	FALSE	FALSE	FALSE
75	09-Oct-13	Paul Maithia	Kinyaua	Masimba	Kiboko	TRUE	FALSE	FALSE	FALSE	FALSE	TRUE	FALSE	FALSE	TRUE	TRUE	TRUE	FALSE	TRUE	FALSE
76	09-Oct-13	Benson Mwengi	Sekeleni	Kasuvi	Kiboko	TRUE	TRUE	FALSE	FALSE	FALSE	TRUE	FALSE	FALSE	TRUE	TRUE	TRUE	FALSE	TRUE	TRUE
77	09-Oct-13	Julius Mwema	Maiku	Thange	Utithi	TRUE	TRUE	FALSE	FALSE	FALSE	TRUE	FALSE	FALSE	TRUE	TRUE	TRUE	FALSE	TRUE	FALSE
78	10-Oct-13	Natoi Kereto	Olchoro	Lugulului	Lugulului	TRUE	TRUE	FALSE	TRUE	FALSE	TRUE	FALSE	FALSE	TRUE	TRUE	TRUE	FALSE	TRUE	FALSE
79	10-Oct-13	Maria Saning'o	Mbironi	Kimana	Kimana	TRUE	TRUE	FALSE	TRUE	FALSE	TRUE	FALSE	FALSE	TRUE	TRUE	TRUE	FALSE	TRUE	FALSE
80	10-Oct-13	Kordillo Philip	Kuku center	Kuku	Kuku	TRUE	FALSE	FALSE	FALSE	FALSE	TRUE	FALSE	FALSE	TRUE	TRUE	TRUE	FALSE	TRUE	FALSE
81	10-Oct-13	Joshua Saruni	Iltila	Iltila	Kuku	TRUE	TRUE	FALSE	FALSE	FALSE	TRUE	FALSE	FALSE	TRUE	TRUE	TRUE	FALSE	TRUE	FALSE
82	11-Oct-13	Kideri Sokonoi	Mailua	Mialua	Mailua	TRUE	FALSE	FALSE	FALSE	FALSE	TRUE	FALSE	FALSE	TRUE	TRUE	TRUE	FALSE	FALSE	TRUE
83	11-Oct-13	Japeth Kakuo	Empiunoto	Simba	Kenyawa	TRUE	FALSE	FALSE	FALSE	FALSE	TRUE	FALSE	FALSE	TRUE	TRUE	TRUE	FALSE	TRUE	FALSE
84	11-Oct-13	Daniel Kaata	Enooretet	Sultan-Hamud	Nkaama	TRUE	TRUE	FALSE	FALSE	FALSE	TRUE	FALSE	FALSE	TRUE	TRUE	TRUE	FALSE	TRUE	FALSE