

Analysis of Geographical and Temporal Patterns of Malaria Transmission in Limpopo
Province, South Africa Using Bayesian Geo-statistical Modelling

by

Aphelele R. Mgabisa

A dissertation submitted in partial fulfilment of the academic requirements for the degree of
Master of Science in the School of Agriculture, Earth and Environmental Sciences,
University of KwaZulu-Natal

Westville, November 2013

Supervisor

Dr Michael Gebreslasie

As the candidate's supervisor I have/have not approved this thesis/dissertation for
submission.

Signed: _____ Name: _____ Date: _____

Declaration

I declare that the dissertation, which I hereby submit for the Master of Science (MSc) in Environmental Sciences at the University of KwaZulu-Natal, is my own work and has not previously been submitted by me for a degree at this or any other tertiary institution. Any work taken from other authors or organisations is duly acknowledged within the text and references chapter.

.....

Aphelele Ronnie Mgabisa

.....

Dr Michael Gebreslasie

Supervisor

Acknowledgement

As I finalize writing this thesis, I wish to acknowledge a number of people who have contributed to its success either directly or indirectly. To begin, I'm very grateful to my supervisor Dr. Michael Gebreslasie for his valuable and professional guidance and support as well as continuous motivation throughout this project. I appreciate your constructive comments and edits in my work.

I extend special thanks to my Switzerland (Swiss Tropical and Public Health Institute) lecturer Dr Penelope Vounatsou for the mentorship and the patience as I went through the challenges of Bayesian modelling. Thanks for sharing your immense experience with me.

I would also like to thank Federica Giardina for her assistance with statistical analysis and interpretation of the data.

I would like to thank the South African Medical Research Council (MRC) for providing me with the data and funding to complete the thesis.

I am eternally grateful to the National Research Foundation for funding me, which enabled me to get a residence and register during my research.

A special thank you to the Swiss South African Joint Research Programme for funding me during my stay in Switzerland and the funding for the Biostatistics course I attended in Switzerland.

Lastly, I thank my family, friends and colleagues for their patience, continued encouragement and support.

Abstract

South Africa is at the southern fringe of sub-Saharan African countries which persist in experiencing malaria transmission. The purpose of the study is to analyse the geographical and temporal patterns of malaria transmission from 2000 to 2011 using Bayesian geo-statistical modelling in Limpopo Province, South Africa. Hereafter, develop malaria case data-driven spatio-temporal models to assess malaria transmission in Limpopo Province.

Malaria case data was acquired from the South African Medical Research Council (MRC). Population data was acquired from AfriPopo; and Normalised Difference Vegetation Index (NDVI), Land Surface Temperature (LST) and Land Cover data were acquired from MODerate-resolution Imaging Spectro-radiometer (MODIS). Rainfall, Altitude and distance to water bodies' data were acquired from African Data Dissemination Service (ADDS), United States Geological Survey (USGS) and Environmental Systems Research Institute (ESRI), respectively. Bayesian spatio-temporal incidence models were formulated for Gibbs variable selection and models were fitted using the best set of environmental factors. Model-based predictions were obtained over a regular grid of 1 x 1km. spatial resolution covering the entire province and expressed as rates of per 1 000 inhabitants for the year 2010. To assess the performance of the predicted malaria incidence risk maps, the predictions and field observations were compared.

The best set of environmental factors selected by variable selection was Altitude and the night temperature of two months before the case was reported. The environmental factors were then used for model fitting and all of the covariates were important on malaria risk. Predictions were done using all the environmental factors. The predictions showed that Vhembe and Mopani district municipalities have high malaria transmission as compared to other district municipalities in Limpopo Province. Assessment of predictive performance showed scatter plots with the coefficient of determination (R^2). The values representing the statistical correlation represented by the coefficient of determination (R^2) were 0.9798 (January), 0.8736 (February), 0.8152 (March), 0.8861 (April), 0.9949 (May), 0.3838 (June), 0.7794 (July), 0.9235 (September), 0.8966 (October), 0.9834 (November) and 0.8958 (December). August had two values reported and predicted which resulted in R^2 of 1. The numbers of the cases were not enough to outline the correlation for August.

The produced malaria incidence maps can possibly be considered as one of the baselines for future malaria control programmes. The results highlighted the risk factors of malaria in Limpopo Province which are the most important characteristics of malaria transmission.

Dedication

Dedicated to my lovely mother, little sister and brother:

Thoko, Khanyisile and Lungisani Mgabisa.

More especially to my late grandmother:

Nonzingo Constance Mgabisa

Table of Contents

Declaration.....	i
Acknowledgement	ii
Abstract.....	iii
Dedication.....	v
Table of Contents.....	vi
List of Tables	ix
List of Figures.....	x
List of Abbreviations and Acronyms.....	xi
CHAPTER ONE.....	1
INTRODUCTION	1
1.0. Background	1
1.1. Statement of the Problem	3
1.2. Aim and Objectives.....	4
1.3. Structure of the Thesis.....	4
1.4. Conclusion.....	5
CHAPTER TWO	6
LITERATURE REVIEW	6
2.0. Introduction	6
2.1. Malaria	6
2.1.1. Life Cycle of Malaria Parasite	7
2.1.2. Life Cycle of Malaria Vector	8
2.1.3. Symptoms of Malaria.....	10
2.2. Malaria and Poverty	11
2.3. Malaria Transmission.....	12
2.4. Malaria Morbidity and Mortality	14
2.5. Vector Control and Intervention	15
2.6. Modelling Spatial and Temporal Malaria Transmission.....	16
2.6.1. GIS and Remote Sensing for Modelling Malaria Transmission.....	16
2.6.2. Bayesian Geo-statistical Modelling.....	18
2.7. Conclusion.....	20

CHAPTER THREE	21
STUDY AREA	21
3.0. Introduction	21
3.1. Location.....	21
3.2. Climate	22
3.3. Vegetation	23
3.4. Topography	23
3.5. Socio-economic Status	23
3.6. Epidemiological Profile.....	24
3.7. Conclusion.....	25
CHAPTER FOUR.....	26
MATERIAL AND METHODS	26
4.0. Introduction	26
4.1. Data Acquisition and Cleaning	26
4.1.1. Malaria Case Data.....	26
4.1.2. Population Data.....	27
4.1.3. Environmental Factors	28
4.1.3.1. Rainfall	28
4.1.3.2. NDVI, Land Surface Temperature (LST) and Land Cover	28
4.1.3.3. Altitude	28
4.1.3.4. Distance to Water Bodies	28
4.2. Data Processing	29
4.3. Statistical Analysis	29
4.3.1. Data Management	29
4.3.2. NDVI, LST day and LST night and Rainfall.....	30
4.3.3. Land Cover.....	30
4.3.4. Bayesian Geo-statistical Incidence Model.....	31
4.3.4.1. Variable Selection.....	32
4.3.4.2. Model Fitting and Prediction.....	33
4.3.5. Assessment of Predictive Performance.....	33
4.4. Conclusion.....	35

CHAPTER FIVE	36
RESULTS AND DISCUSSION	36
5.0. Introduction	36
5.1. Results	36
5.1.1. Descriptive Analysis	36
5.1.2. Bayesian Geo-statistical Incidence Model Estimates	39
5.1.2.1. Variable Selection.....	39
5.1.2.2. Model Fitting	40
5.1.2.3. Prediction.....	41
5.1.3. Assessment of Predictive Performance.....	44
5.2. Discussion	45
5.3. Limitations of the Study.....	48
5.4. Conclusion.....	49
CHAPTER SIX.....	50
CONCLUSION AND RECOMMENDATIONS	50
6.0. Introduction	50
6.1. Conclusion.....	50
6.2. Recommendations	51
Appendix.....	65

List of Tables

Table 4.1: Environmental data utilised for the analysis.....	29
Table 4.2: Different periods or time intervals used to average data in lags.....	30
Table 4.3: Land cover classification scheme used to categorise land cover data.....	31
Table 5.1: Summary of the results from Gibbs variable selection.....	40
Table 5.2: Summary of the estimates from a fitted Bayesian negative-binomial model.....	41
Table 5.3: Sub-places predicted to have high malaria incidence in 2010 in their local municipalities.....	48

List of Figures

Figure 2.1: Life Cycle of Malaria Parasite.....	7
Figure 2.2: Distinguishing features of Anopheles mosquito and other types of mosquitos....	10
Figure 3.1: Map showing the study area, divided into 5 district municipalities.....	22
Figure 3.2: Map depicting areas affected by malaria in Limpopo Province.....	24
Figure 4.1: Flow diagram of steps followed in the study methods.....	34
Figure 5.1: Graph showing a pattern formed by the total number of cases per year in Limpopo Province from 2001 to 2010. Years compared to the number of reported cases.....	37
Figure 5.2: Graph showing the total number of cases per month in each year in Limpopo Province. Months compared to a total number of reported cases per year (2001 to 2010).....	38
Figure 5.3: Graphs showing the relationship between monthly averages of malaria reported cases and environmental factors in Limpopo Province from 2001 to 2011.....	39
Figure 5.4: Depicts spatio-temporal model based maps of malaria incidence per 1 000 inhabitants in Limpopo Province, from January to June, year 2010.....	42
Figure 5.5: Depicts spatio-temporal model based maps of malaria incidence per 1 000 inhabitants in Limpopo Province, from July to December, year 2010.....	43
Figure 5.6: Graphs showing the relationship between observed and predicted malaria incidence in 2010.....	44

List of Abbreviations and Acronyms

ADDS=African Data Dissemination Service

Amsl=above mean sea level

BIL=Band Interleaved by Line

DEM=Digital Elevation Model

DDT=Dichloro-diphenyl-trichloroethane

EO=Earth Observation

EOS=Earth Observing Systems

ESRI=Environmental System Research Institute

GIS=Geographic Information System

GPS=Global Positioning System

IRS=Indoor Residual Spraying

ITNs=Insecticide-Treated Nets

LLINs=Long-Lasting Insecticide Nets

LST=Land Surface Temperature

McMC=Markov chain Monte Carlo

MODIS=MODerate-resolution Imaging Spectro-radiometer

NB=Negative-Binomial

NDVI=Normalised Difference Vegetation Index

NMMAAPS=National, Morbidity, Mortality and Air Pollution Studies

RBCs=Red Blood Cells

RS=Remote Sensing

SPOT=*Satellite Pour l'Oservation de la Terre*

SSA=Sub-Saharan Africa

USGS=United States Geological Survey

UTM=Universal Transverse Mercator

WHO=World Health Organisation

CHAPTER ONE

INTRODUCTION

1.0. Background

Malaria remains one of the diseases responsible for the greatest cause of morbidity and mortality worldwide (Breman, et al., 2004). In 2011, 3.3 billion people were estimated to be at risk of malarial illnesses. Moreover, an estimation of 80% of malaria cases and 90% of mortality were estimated to occur in the African region (World Health Organisation, 2012). A high mortality and morbidity rate of malaria is basically in the high transmission regions of sub-Saharan Africa. Most of the incidents are amongst children under the age of five years and pregnant women in endemic areas (Amek, et al., 2012; World Health Organisation, 2009). In this instance, malaria remains the world's most important tropical parasitic disease, and one of the major public health challenges in the sub-Saharan Africa (Abellana, et al., 2008; Roca-Felter, et al., 2009; Rowe, et al., 2006).

The study conducted by Snow, et al. (1999) raised an issue that areas such as sub-Saharan Africa are not precise in quantifying the causes of malaria morbidity and mortality. Mostly, due to weak health systems; poor attendance of health facilities; weak civil registration; and death certification systems as a result of malaria deaths which occur at home without any contact to the health system (Mathers, et al., 2005). The highest burden of malarial morbidity and mortality is mostly amongst the poorest countries. Thus, poorer households in poor countries are more vulnerable to deadly consequences of malaria including severe morbidity and mortality due to less access to treatment (Worrall, et al., 2005). There is a strong link between poverty and treatment seeking, with poor countries exposed to self-treatment and less access to private and high level public providers (Worrall, et al., 2005). Therefore, since malaria is mostly the disease dominating poor counties, there is a possibility that the countries can be poorer as they are forced to invest on clinical infections.

The World Health Organisation (WHO) stated that malaria influences the social and economic well-being of societies in affected areas, draining scarce health and human resources, interfering with educational achievement, and causing persistent economic disadvantage. The degree of malaria burden does not only extend beyond immediate threats

to survival. Moreover, the consequences of repeated clinical infection place a burden on households, systems and therefore influence the country's economy (Bloom, et al., 2006). Drug resistance also has an impact on social and economic factors and it increases the burden on the health systems which is one of the main concerns in sub-Saharan Africa (Breman, et al., 2004). The effect of drug resistance can be as high as possible on malaria transmission in low transmission regions due to low immunity in the population (Bjorkman and Bhattarai, 2005). Consequently, there may be high morbidity and mortality rates in all age groups. Not only social and economic factors are influenced by malaria, there are also additional factors of great importance to malaria.

Malaria is an environmental disease, approximately 70 to 90% risk of malaria is considered as a result of environmental factors influencing the abundance and survival of vectors (Amek, et al., 2012; Smith, et al., 1999). There are other additional factors including socio-economic status; population movement; quality and access to health care; control and intervention programmes; human activities (e.g. irrigation) and drug resistance. Basically, environmental factors are the main drivers of malaria transmission since they influence both the life cycle of a vector and the parasite (Bray and Garnham, 1982; Hay, et al., 2002; Jupp, 2005; Mzezewa, et al., 2010). Rainfall provides breeding sites for the mosquito vector which increases the population of mosquitos while temperature and humidity influence both vector and parasite development (Amek, et al., 2012). Therefore, through the relations between the disease and the environment, the burden of malaria can be estimated at places where data on malaria transmission are not available and high risk areas can be identified (Gosoni, et al., 2006). Dealing with the cause of malaria can help in the identification of malaria interventions in high risk areas.

Malaria is caused by protozoan parasites of the genus *Plasmodium*. Mostly, the infections of malaria particularly in sub-Saharan Africa are due to *Plasmodium falciparum* (Midega, et al., 2012) and if untreated, can lead to severe complications and mortality. Malaria parasite is transmitted from human to human via a bite of an infected female *Anopheles* mosquito that often bites at night (Jupp, 2005). Generally, the parasite needs two types of hosts to be transmitted, the human host and the female *Anopheles* mosquito. Human malaria multiplies into two stages in a human body: first in the liver (the hepatic cycle), then in the blood (the erythrocytic cycle) (Bray and Garnham, 1982). When the mosquito bites an infected person, it ingests microscopic malaria parasites found in the persons blood (Aarogya, 2012). The

parasite grows in the mosquito for a week or more before infection can be transmitted to another person. Finally, when the mosquito bites the second human host, the parasite is transmitted from the mosquito's mouth into the person's blood (Midega, et al., 2012).

1.1. Statement of the Problem

A wide range of vector control and intervention programs have been successfully put in place (Blumberg and Frean, 2007; Mabaso, et al., 2004). However, malaria persists on being the global disease burden, more especially in sub-Saharan Africa (Maharaj, 2008). Resistance to insecticides and anti-malaria drugs are the main concern of the global community for malaria elimination efforts. This makes malaria a major threat to public health and economic development of the countries mainly in sub-Saharan Africa (Griffin, et al., 2010).

South Africa is at the southern fringe of sub-Saharan African countries which continue to experience malaria transmission. Approximately 95% of malaria infections found in South Africa result from a parasite *Plasmodium falciparum*, with the mosquito *Anopheles arabiensis* being a major malaria vector (South African National Department of Health, 2007). More or less than 10% of the population resides in malaria endemic areas and is at risk of contracting the disease (Moonasar, et al., 2012). There is no partial immunity in persons living within malaria areas of South Africa due to low intensity seasonal malaria transmission (Department of Health, 2011). In South Africa malaria is considered as seasonal and unstable (South African National Department of Health, 2007), with transmission limited to warm and rainy summer months from September to May (Gerritsen, et al., 2008). According to Blumberg and Frean (2007), malaria is endemic in the low-altitude areas of South Africa along the border shared with Mozambique and Zimbabwe whereby transmission is mainly taking place in three provinces namely: Limpopo, Mpumalanga and KwaZulu-Natal. Over the past years, malaria incidence in South Africa has radically decreased (Maharaj, 2008) verifying that malaria control has been successful in reducing malaria. However, it remains a serious public health challenge in the northern and eastern parts of the country (Jupp, 2005). According to a study by Gerritsen, et al. (2008), over the seasons between 1998 and 2007 there was a dramatic decrease of malaria incidence rate in Limpopo Province. Subsequently, in 2010 Limpopo Province had the highest malaria transmission in South Africa which is greater than or equal to 1 case per 1 000 population, that is 4% of the total population of the country (World Health Organisation, 2011).

This study analysed the geographical and temporal patterns of malaria transmission from 2000 to 2011 using Bayesian modelling in Limpopo Province, South Africa. Hereafter, malaria case data-driven spatio-temporal models will be developed to assess malaria transmission to included environmental influences as these form key aspects of the research/objectives. Finally, predict malaria incidence and produce seasonality maps for 2010 at sub-place level. Reliable maps of malaria transmission can influence intervention strategies and thus improve the use of limited human and financial resources to areas of most need (Gosoni, et al., 2006). Models are to be created for geo-statistical incidence data derived from malaria surveys carried out at a number of fixed locations. Modelling is about the techniques used efficiently to estimate transmission parameters over the earth's surface (Simoonga, et al., 2009). Malaria mapping involves the use of a Geographic Information System (GIS), Remote Sensing and Geo-statistical Modelling (Kleinschmidt, et al., 2000). These data (GIS data and Remote Sensing data) can be utilised with malaria transmission to accurately estimate their association, predict at un-surveyed locations and display spatial patterns of the disease (Hay and Snow, 2006). This may be more important in Limpopo Province since there is no study which has attempted to map malaria at sub-place level.

1.2. Aim and Objectives

The main aim of the study is to analyse the geographical and temporal patterns of malaria transmission for the period starting from May 2000 to June 2011 using Bayesian geo-statistical modelling in Limpopo Province, South Africa.

The specific objectives of the study are to:

- Analyse the influence of environmental factors in malaria transmission;
- Analyse malaria transmission using Bayesian geo-statistical models;
- Predict malaria spatial and temporal based risk maps for 2010; and
- Generate maps to illustrate places that are predicted to have high risk of malaria.

1.3. Structure of the Thesis

The thesis is structured in respect of six chapters as stated below. Chapter one provides the research background. Chapter Two is the literature review based on previous studies that have been conducted related to the study. Chapter Three is the study area describing properties and the location of the place the study was based on. Chapter Four is based on the material and methods utilised for the study to be successful. Chapter Five is on the results

acquired from the study and the discussion of the results. Chapter Six is the conclusion and recommendations of the study.

1.4. Conclusion

This chapter provided the background of the study and the statement of the problem. Subsequently, it provided the proposed methods which were to be used for the study. These are the most important aspects of the research. In order to provide more support and understanding of the study, the literature has to be reviewed. The review of the literature is provided in the next chapter.

CHAPTER TWO

LITERATURE REVIEW

2.0. Introduction

This chapter focuses on providing a review of the literature on malaria disease mapping. It is based on previous studies that have been conducted by researchers. Theoretically, disease mapping has been utilised in such a way that it labels the geographic distribution of different diseases and their risks. This may depend on the main aim and specific objectives of the study. For this study, the reviews were based on malaria. Many studies have been conducted on malaria and some are going to be covered in the review of literature for this study.

2.1. Malaria

The World Health Organisation defined malaria in 2008 as a vector-borne infectious disease that is widespread in tropical and subtropical regions, caused by a parasite that lives part of its life in humans and part in mosquitoes. There are particular factors that need to be considered to identify the presence of the parasite. Anciently, the Romans recognised the relationship which was between stagnant water in the swamps surrounding Rome and the existence of fevers during the summer months. These fevers were said to be caused by ‘bad air’ – ‘*mal aria*’ (Cunha and Cunha, 2008, p196). The fevers were described by Hippocrates in the fifth century B.C. and Galen in the second century B.C. in his detailed treatment of fevers (*de differentiis februm*) (Jarch, 1993). The Romans thought the foul vapours emanating from swamps and stagnant pools were the cause of the disease. This was incorrect, but it made the contribution in presenting the importance of stagnant pools being related to summer or illness among Romans (Van Buren, 1864). Cox (2010, p19) also stated that, ‘for over 2500 years the idea that malaria fevers were caused by miasmas rising from swamps persisted and it is widely held that the word malaria comes from the Italian *mal’aria* meaning spoiled air although this has been disputed’. In 1880 Alphonse Laveran, a French army physician stationed in Algeria, discovered that malaria was caused by a parasite called *Plasmodium* (Tren and Bate, 2004).

2.1.1. Life Cycle of Malaria Parasite

There are only four species of the protozoan *Plasmodium* that cause malaria in humans which are *Plasmodium falciparum*, *Plasmodium vivax*, *Plasmodium ovale* and *Plasmodium malariae* (Griffin, et al., 2010). Amongst the four species of *Plasmodium*, *P. falciparum* is the most prevalent parasite in sub-Saharan Africa (Butler, et al., 2010). The development of *Plasmodium* takes place in two cycles, a sexual cycle within the vector (mosquito) and an asexual cycle in the host (human) as shown in Figure 2.1.

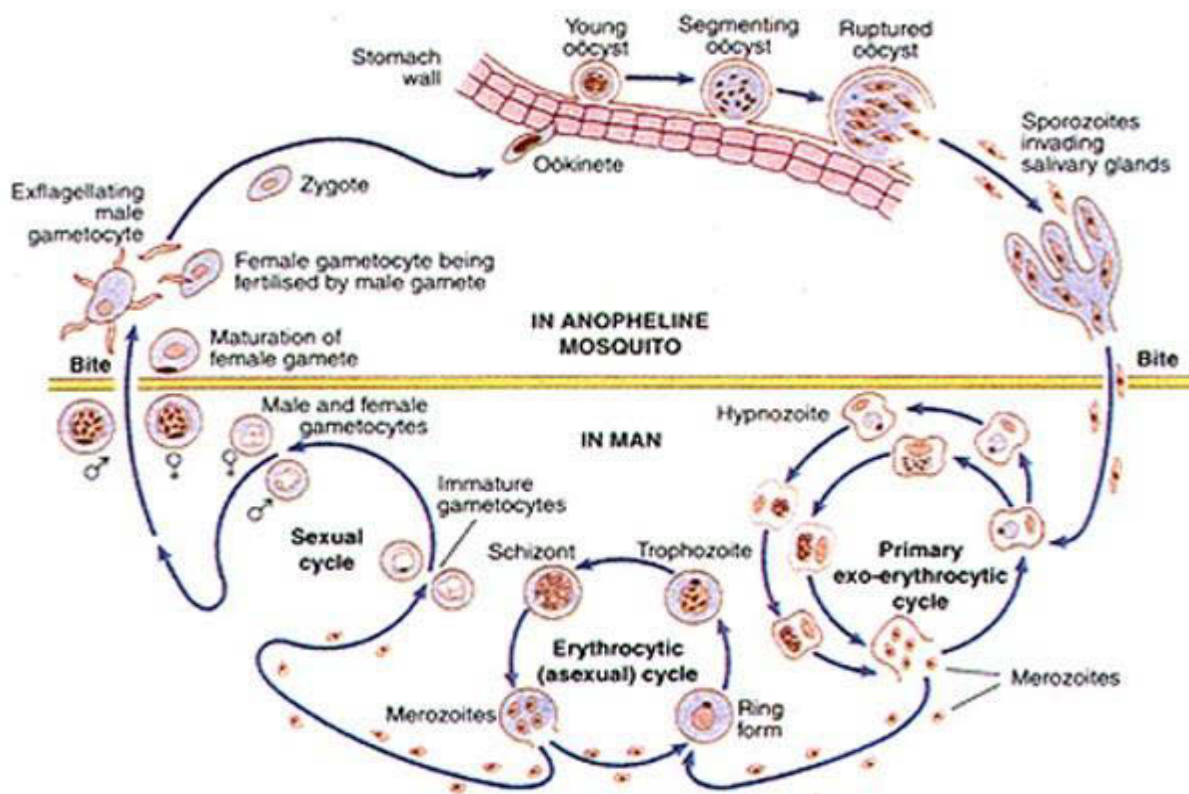


Figure 2.1: Life cycle of malaria parasite. Source: (Aarogya, 2012)

Whenever an infected female *Anopheles* mosquito bites a human host, it takes blood and injects saliva that comprises a parasite at the same time (Bray and Garnham, 1982). The parasite is injected into a person's blood stream. This cycle of blood feeding and egg-laying is known as the *gonotrophic cycle* (Fujioka and Aikawa, 2002). After thirty minutes to four hours, the thread-like sporozoite invades the liver cells (hepatocytes) (Crans, 2004). An asexual replication process, termed *exoerythrocytic schizogony*, follows (see Figure 2.1). The sporozoites grow into schizonts and multiply rapidly producing between 10 000 and 30 000 merozoites per hepatocyte (Medicines for Malaria Venture, 2010). In warm countries it may take 6-16 days for symptoms to be evident depending on the parasite species (Jupp, 2005).

Alternatively, *P. vivax* and *P. ovale* sporozoites turn into hypnozoite and can remain dormant in a liver for months or years. After becoming active, hypnozoites develop into schizonts causing relapses in the person infected (Matteelli and Castelli, 2007).

The merozoites released from the liver upon rupture of schizonts then invade the red blood cells (RBCs). The early RBC (erythrocytic) stage of the parasite is known as the trophozoite (see Figure 2.1). Trophozoites multiply asexually to form schizonts which form up to thirty merozoites after maturity, when they burst they invade the blood circulation (Medical Research Council, 2001). Some of the merozoites develop into macrogametes and microgametes (female and male sexual forms, respectively). The gametocytes circulate the blood stream waiting for a blood seeking female mosquito. Then, the remaining merozoites invade other erythrocytes, multiplying within them and causing their rupture (Whitfield, 1984). The erythrocytes rupture in synchronised 48 hour cycles in *P. falciparum*, *P. vivax* and *P. ovale* infections and in 72 hours cycles for *P. malariae*. Sessions of fever arise after the rupture of erythrocytes in response to the release of toxins into the blood stream (Bray and Garnham, 1982).

Macrogametes and microgametes mate in the mosquito stomach and the resulting zygote develops into a motile ookinete (see Figure 2.1). The ookinetes traverse the stomach membrane to lodge in the stomach outer wall as oöcysts. Over a period of 10-14 days the oöcysts undergo asexual multiplication forming thousands of sporozoites in the mosquito body (Hoffman, 1996). The duration of this process, known as *sporogony*, depends on the *Plasmodium* species and ambient temperature. Optimum temperatures for sporogony for *P. falciparum* are between 25-30°C while *P. vivax* prefers lower temperatures of about 25°C (Jupp, 2005). The sporozoites break out of the oöcyst and migrate to the mosquito's salivary gland. Here they remain infective for the duration of the mosquito's life (Khan and Waters, 2004). Only the older mosquitoes are capable of carrying sporozoites (Amek, et al., 2011). The cycle starts again when the mosquito bites the next human being.

2.1.2. Life Cycle of Malaria Vector

In 1898, Grassi pointed a species of the genus *Anopheles* as vectors which are responsible for malaria transmission (Capanna, 2006). Specifically, mosquitoes enjoy the questionable honour of having been the first insects to be associated with the transmission of a disease (Gillies, 1967). There are 400 different species of *Anopheles* mosquitoes in the world, but only 60 of them are vectors of malaria and only 30 are vectors of major importance (WHO

and CDC, 2005). The *Anopheles* mosquito life cycle and its interaction with malaria transmission need different phases and characteristics to be probable. Johnsen and Rennie (2007) referred mosquitoes as blood-thirsty pests that spread several tropical diseases and ruin people's outdoor activities. All mosquitoes require standing water to complete their cycle. The *Anopheles* mosquito lays eggs on the surface of water at night. The water types used for laying eggs are: irrigation channels, a pool of water in a tree trunk, and a sewage effluent (Rios and Connelly, 2009). In tropical regions, the eggs hatch after two to three days (Coluzzi, 1988). Larvae go through four growth stages below the surface of water (rests parallel to water surface) before becoming pupa. This process happens for five minutes after 7-14 days. The pupa is in a shape of a comma (see Figure 2.2 below) and is the least active stage (transition stage) of the *Anopheles* life cycle. Within 24-48 hours the pupa will metamorphose into an adult, this happens late in the evening (Muir, 1988). In warmer months, it may only take a week for mosquitoes to develop from eggs (Rios and Connelly, 2009).

Male mosquitoes emerge prior to female mosquitoes and will mate usually during flight. Female *Anopheles* mosquitoes generally need to feed on blood (also nectar and damaged fruits) from human to provide proteins to assist with the development of her eggs. They do not feed only on humans but they also feed on mammals (Koutsos, et al., 2007). Male mosquitoes do not bite, but they only feed on nectar and damaged fruits. An adult mosquito generally survives for between one week and one month (Johnsen and Rennie, 2007). *Anopheles* mosquitoes rest with their abdomens sticking up in the air rather (both males and females) than parallel to the surface (see Figure 2.2).

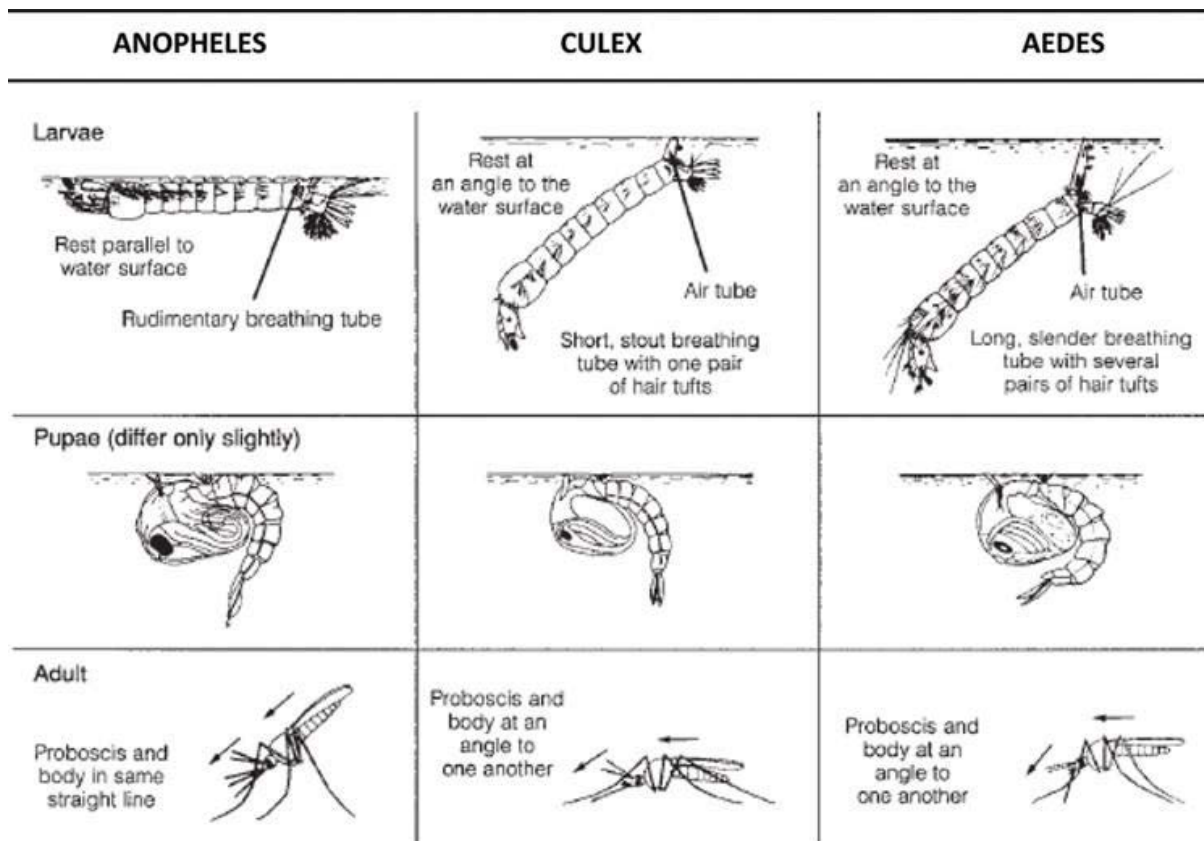


Figure 2.2: Distinguishing features of Anopheles mosquito and other types of mosquitos. Source: (Rozendaal, 1997)

2.1.3. Symptoms of Malaria

People suffering from malaria are generally very sick with high fever, malaise, weakness, gastrointestinal complaints (nausea, vomiting, and diarrhoea), neurologic complaints (dizziness, disorientation, and coma), headache, back pain, myalgia, shaking chills and flu-like illness (CDC, 2011). If an individual has one of the above mentioned illnesses especially fever of unknown origin, irrespective of travel history, diagnosis should be considered. Commonly, malaria is characterized by the underlying malnutrition in children less than 5 years and many children who die also have malaria anaemia (Kiwanuka, 2003). Repeated infections can possibly lead to anaemia, especially in children (Newton, et al., 1998). The symptoms of malaria depend on the age of a patient. In children, hypoglycaemia, convulsions, and severe anaemia are relatively common, while acute renal failure, jaundice, and Acute Respiratory Distress Syndrome (ARDS) are more common in adults. Cerebral malaria, shock and acidosis may occur at any age (Blumberg and Freaun, 2007). During pregnancy, malaria causes low birth weight, which can cause a number of impairments

including cerebral palsy, mental retardation, and cognitive deficits, and it contributes to maternal, fetal and infant mortality (Murphy and Breman, 2001). Pregnant women with severe malaria are at risk of hypoglycaemia, anaemia, and ARDS, as well as fetal loss (Blumberg, 2005). Pregnant women are three times more at risk of severe malaria disease than non-pregnant women obtaining infections from the same area (CDC, 2011). This is in relation with access to treatment in that particular region.

Malaria is a treatable and curable disease, thus the control of malaria is highly dependent on effective diagnosis and efficient drugs (Bjorkman and Bhattarai, 2005). Although malaria can be deadly, illness and death from malaria can regularly be prevented. Previously, chloroquine treated nearly all cases of malaria in many countries worldwide, but recently chloroquine-resistant *P. falciparum* has been observed. Therefore, it is no longer used (NIMR and NVBDCP, 2009). It has been stressed that there is more treatment modalities required due to potentially life-threatening side effects of parenteral quinine utilised for complicated malaria treatment in South Africa. However, the use of artesunate in other malaria-hit countries has been much safer and about 30% more effective (Bateman, 2008). Whenever severe malaria is greatly suspected, blood should be collected for diagnosis testing and parenteral antimalarial drugs may be started as soon as possible (CDC, 2011). Regardless of the fact that malaria is easily preventable and treated cheaply, the disease is still a problem on economic and social costs, especially in poor countries (Tren and Bate, 2004).

2.2. Malaria and Poverty

Beyond the suffering caused by malaria on individuals, families and communities, the disease also extends and strengthens poverty in some of the poorest countries worldwide. According to the World Health Organisation (2008), malaria costs Africa at least US\$ 12 billion in direct losses and much more is lost on economic growth each year. Basically, malaria has a great impact on poorer countries due to physical proximity to water sources and lowered capacity (lack of education and resources) to contact health systems and preventive measures for malaria protection (Koram, et al., 1995). Higher education may give rise to better health awareness, better utilisation of health facilities, higher income and the ability to purchase goods and services that improve infants' health, longer birth intervals, and possibly higher maternal ages (Cleland and van Ginneken, 1989; Gemperli, et al., 2004). Due to the lack of knowledge and access to prevention and treatment faced by poorer countries, some women face illness like fever. They end up believing it is normal during pregnancy, which may blind them from obtaining treatment (Winch, et al., 1996).

In poorer countries, house type, housing and roofing material and house location are factors which influence the availability of a vector. These factors favour the presence of a vector (Koram, et al., 1995). The study by Coleman, et al. (2010) revealed that people living in traditional mud-wall houses had increased risk of acquiring malaria than those living in western-style brick-wall houses. Households that are located nearest to the larval sites are more at risk of malaria and usually these households are in poorer rural locations (Midega, et al., 2012). Antiretroviral therapy is limited in rural areas of South Africa and recurrent and persistent opportunistic infections may be confused with malaria. Householders who open their windows when they sleep are more at risk of contracting malaria than those who close windows (Midega, et al., 2012). It is not easy to manage severe malaria at a rural district hospital due to renal failure and respiratory distress (Mehta, et al., 2007). Therefore, there is a high rate of malaria transmission in poor rural areas as compared to urban areas.

2.3. Malaria Transmission

There are a number of factors influencing malaria transmission such as climatic and environmental conditions (Amek, et al., 2012). The study by Gemperli, et al. (2004) showed malaria prevalence as a measure of malaria transmission. The malaria parasite is typically transmitted from human to human via a bite of an infective female *Anopheles* mosquito (Amek, et al., 2011). Occasionally, a human may be infected through contaminated blood. The transmission may also be from a mother to her foetus before or during delivery (“congenital” malaria). Moreover, since the malaria sporozoite parasite is in the RBCs, malaria can also be transmitted through blood transfusion, organ transplant, or the shared use of needles or syringes contaminated with blood (U.S Department of Health and Human Services, 2007). Usually the transmission of human malaria is through a mosquito (vector) bite.

Anopheles gambiae is considered as the principal vector for human malaria in Africa (Zeng, et al., 1996). Additionally, it (*An. gambiae*) has caused an estimated 200 million clinical cases and more than one million deaths annually. There are two vectors available in South Africa namely *Anopheles arabiensis* and *Anopheles funestus*, belonging to two different complexes of mosquito species (Blumberg and Frean, 2007). The vectors prefer particular types of ecological conditions to transmit human malaria and for breeding. The study by Jupp (2005) confirmed that *An. arabiensis* prefers temporary to semi-permanent ground pools with a moderate amount of vegetation that possibly develop after rain, whereas, *An. funestus* larvae are found mainly in the shady margins of warm perennial streams where the vegetation

is obtainable and the current is slow. Therefore, temperature and rainfall are considered major continuing natural risk factors that affect the life cycle and the breeding of the mosquitoes that transmit malaria (Abellana, et al., 2008). This clarifies that temperature and rainfall are major factors affecting malaria transmission.

Plasmodium parasites depend on temperature, their development slows with the drop or increase in temperature. Some parasites do not develop further in low temperature (e.g. *P. vivax* develop in temperatures >60 degrees Fahrenheit) and others stop developing at higher temperatures (e.g. *P. falciparum*) (Amek, et al., 2012). The transmission of malaria is basically dependent on the presence of environmental conditions linked with the development of both parasites and vector mosquito (Stresman, 2010). For transmission to take place, rainfall; temperature; water bodies; altitude; humidity; vegetation and land use may be considered, as they affect the habitat and vector breeding sites. Mosquitoes cannot be able to survive in low humidity. Rainfall increases breeding grounds and in numerous tropical regions, the growth of malaria cases is during the rainy season. Rainfall has contributed more in the transmission of malaria in Mpumalanga Province because the transmission followed a different rainfall pattern (Ngomane and de Jager, 2012). Mabaso, et al. (2007) revealed that El Niño-Southern Oscillation (ENSO) events cause rainfall patterns to change and this usually affects the mosquito's breeding, which is associated with variation in malaria transmission. Mosquitoes must survive a period longer than the incubation period for the parasite to complete its development within them, then the parasite can be transmitted (Stresman, 2010). Therefore, environmental factors that are relevant to the survival of a mosquito can influence malaria incidence and transmission.

In 2006, Staff at the International Research Institute for Climate and Society (IRI) created an online resource for information about the environmental conditions that are associated with malaria transmission in Africa (Grover-Kopec, et al., 2006). It was found that seasonal characteristics of climate have a significant relationship with the distribution and seasonality of the disease where it is not adequately controlled. The study in the highland province of Burundi by Protopopoff, et al. (2008) revealed that in the untreated zones, children 5-9 years old had the highest prevalence of malaria infection, while individuals over age 50 had the lowest. Malaria prevalence was lower in the hilltops than in the valleys (90% malaria transmission). Children were indeed three times more at risk of malaria infection in the valleys as compared to the hilltops. Hay, et al. (2002) also conducted a study on clinical epidemiology of malaria in the highlands of western Kenya. The study examined

longitudinal, age-structured, clinical data on frequency of admission for severe and complicated *P. falciparum* malaria at the three hospitals located above 1 600m in the highlands of western Kenya. Neither the clinical epidemiology nor estimates of the prevalence of infection in the community substantiated the view that the high altitude areas served by the hospitals in the study support unstable transmission.

In KwaZulu-Natal, South Africa, Kleinschmidt and Sharp (2001) performed a study in the northern part of the province which has experienced low levels of malaria transmission intensity for many years. The study was based on the patterns in age-specific malaria incidence in a population exposed to low levels of malaria transmission intensity. Annual case totals by 5 year age groups and by annual malaria season were extracted from malaria information systems for each enumerator area (EA). It was found that in those areas malaria incidence rose with age until the late teens, and either remains constant or decreased in young adults. Morbidity and mortality are determined by how high or low malaria transmission is, in a given region or population.

2.4. Malaria Morbidity and Mortality

Malaria is a serious and potentially deadly disease caused by a parasite that generally infects a certain type of mosquito, which feeds on humans. People who are most vulnerable to endemic malaria are young children who still have to acquire immunity to the disease, pregnant women with reduced immunity during pregnancy, and non-immune migrants or tourists (Connor, et al., 2006). Mostly, mortality in areas of high malaria transmission is assigned to malaria infection (Marsh and Snow, 1999; Rowe, et al., 2006; Smith, et al., 2001). In sub-Saharan Africa, nearly 4% of all maternal mortality annually is a result of malaria-associated anaemia (World Health Organisation, 2006).

The majority of sub-Saharan African countries have experienced a dramatic decline in mortality especially in infants and children (Rajaratnam, et al., 2010). However, there are concerns that any intervention directed to reduction of malaria transmission might delay the acquisition of immunity, shifting the burden of disease to an older age group (Snow and Marsh, 2002). The more often a person is infected from a young age, the greater the immunity becomes. Areas of high transmission in sub-Saharan Africa have the highest mortality in infants while adults are largely semi-immune (Blumberg and Frean, 2007). This confirms that the immunity has shifted to an older age group due to interventions. Nevertheless, malaria interventions are needed since morbidity and mortality increases if they

are not in place. In this case, it is clear that morbidity and mortality depends on malaria transmission in a given location.

The study by Mehta, et al. (2007) revealed mortality in three of the South African malaria endemic provinces where 248 deaths were notified during the study period, 197 were investigated (Mpumalanga 70; Limpopo 84; KwaZulu-Natal 43). Twenty of them were withdrawn, either because malaria was excluded or another disease was confirmed as the cause of death (HIV-associated cryptococcal meningitis; HIV-associated TB), or malaria was never confirmed. Therefore, mortality was sometimes assumed to be misclassified as being malaria-related in South Africa. This may be because the estimates of the causes of malaria morbidity and mortality are not precise in sub-Saharan Africa. Recently, Ngomane and de Jager (2012) confirmed a gradual decline in malaria morbidity and mortality in Mpumalanga Province over eight malaria seasons (2001-2009), showing an 85% reduction in the annual number of confirmed cases and 74% in the number of deaths-attributed to malaria. The study is consistent with the previous studies which were conducted in the other endemic provinces of South Africa. Due to very low transmission in South Africa, all age groups are vulnerable to severe malaria (Blumberg and Frean, 2007). Beside the fact that there are numerous reasons for the decline in malaria morbidity and mortality, authors have argued that it is due to malaria control and interventions that are put in place (Roca-Felter, et al., 2009).

2.5. Vector Control and Intervention

There are numerous malaria control strategies that have been implemented over the last century, aiming to reduce and eliminate the burden of malaria. The principal tools for vector control in sub-Saharan Africa are insecticide-treated nets (ITNs) and Indoor Residual Spraying (IRS), since the main local vectors are *Anopheles gambiae* and *Anopheles funestus* which feed and rest indoors at night (Maharaj, 2008). Regular ITNs are bed nets that are treated with insecticide about once a year and are less favored in sub-Saharan Africa, some countries contains a net coverage of 60% (Maharaj, 2008). They protect the people underneath against mosquito biting and can kill mosquitoes on the net that try to bite. More recently developed long-lasting insecticide nets (LLINs) do not need re-treatment, because the insecticide lasts as long as the net (World Health Organisation, 2008). IRS is most often carried out with the insecticide DDT (dichloro-diphenyl-trichloroethane) and the spraying of the walls is important as mosquitoes tend to rest there after biting (Bateman, 2008). Instantaneous killing of the mosquitoes is not the essential effect of IRS, but the reduction of the mosquitoes' lifespan. IRS has been highly successful in South Africa but has not been

sustainable on a wide scale elsewhere in sub-Saharan Africa (Blumberg and Freaan, 2007; Moonasar, et al., 2012).

Most of the countries in sub-Saharan Africa describe a recent declining trend in the incidence of malaria, whereas the rest have a little or no change (O'Meara, et al., 2010). Through the recently expanding coverage of malaria prevention and treatment in many countries, malaria has been reduced in most of the sub-Saharan countries. South Africa has succeeded in controlling malaria transmission over the past decade throughout most of the country (Blumberg and Freaan, 2007). IRS had an intense influence on malaria distribution and case numbers in South Africa. The operations of IRS eliminated an extremely efficient malaria vector *Anopheles funestus* and greatly reduced the density and distribution of the major malaria vector *Anopheles arabiensis* (Maharaj, et al., 2012).

The study by Maharaj, et al. (2005) indicated that DDT contributed significantly to the decrease in morbidity as it was the only variable changed at the peak of the epidemic in KwaZulu-Natal. Limpopo Province halved its malaria transmission over the period of ten years (2000-2010) which still leaves a concern due to a slow reduction of the number of deaths (Moonasar, et al., 2012). Recently, Khosa, et al., (2013) conducted a study in Limpopo Province which demonstrated an overall decline of malaria morbidity in Mutale municipality in a five year period (2005-2010), most remarkably after a peak during the 2007/8 season. This is because of the interventions that have been implemented in the area and surroundings over the past years. Due to the rapid increase of malaria control and interventions in sub-Saharan African countries, continued surveillance is required to monitor changes in transmission intensity levels and the burden of malaria disease (Roca-Felter, et al., 2009).

2.6. Modelling Spatial and Temporal Malaria Transmission

2.6.1. GIS and Remote Sensing for Modelling Malaria Transmission

Epidemiology is rapidly changing with the availability of data using new methods of spatial data collection like GIS and Remote Sensing (Saxena, et al., 2009). GIS is a technology designed to capture, store, display, communicate, transform, analyse, and archive geo-referenced information, that is, information tied to a specific location (Patra, 2002). GIS technology has become an essential tool for processing, analysing and visualising spatial data within public health, environmental health and disease ecology (Kistemann, et al., 2001). This makes it an important and reliable component in epidemiological and health projects. GIS has been used with an increasing frequency in epidemiology (Nuckols, et al., 2004).

Remote Sensing is defined as a science (and to some extent, an art) of obtaining information concerning an object, area or phenomenon using a specific instrument without being in physical contact with the object being investigated (Mironga, 2004). Remote Sensing data has been used for more than thirty years to extract information on environmental conditions at high spatial and temporal resolution (Amek, et al., 2011). Simoonga, et al. (2009) referred to Remote Sensing as a technology which is primarily used for the acquisition of satellite environmental data in epidemiology. Mapping and spatial analysis have a historical association with epidemiology (Bhatt and Joshi, 2012). Remote Sensing data in GIS has also been utilised mainly for identification, characterisation, monitoring surveillance of breeding habitats and mapping of malaria risk (Saxena, et al., 2009).

The presence of new Remote Sensing data sources, computerised GIS and geo-statistical methods has brought extraordinary information and capacity for development of malaria risk maps (Hay, et al., 2000). Thus, the applications of GIS and Remote Sensing to vector-borne diseases have progressed far beyond their fruitful pictures which dominated their early use (Kitron, 2000). Spatial technology contributes to systematic and regular monitoring of the earth's environmental conditions supplying large amounts of spatial and temporal data (Saxena, et al., 2009). This has revolutionised the technique in epidemiological research (Bernadinelli and Montomoli, 1992).

GIS and Remote Sensing have previously dominated as important tools in the studies based on mapping malaria. A study by Graham, et al. (2004) demonstrated a range of techniques utilised in GIS, Remote Sensing and spatial analysis that are suitable for epidemiology. It was on "Epidemiology: A Spatial Perspective", the role of GIS, Remote Sensing, spatial analysis, geo-statistics and spatial modelling in epidemiology and public health. Szabo, et al. (2008) performed a study on the mapping of mosquito (culicidae) breeding sites using predictive geographic information methods. Machault, et al. (2010) performed a study on spatial heterogeneity and temporal evolution of malaria transmission risk in Dakar, Senegal, according to remotely sensed environmental data. The objectives of the study were to utilise high spatial resolution SPOT satellite images to identify some urban environmental factors associated with *Anopheles arabiensis* densities in Dakar, to assess the persistence of these associations and to describe spatial changes in at-risk environments at a decadal time scale. Two SPOT images from the 1996 and 2007 rainy seasons in Dakar were processed to extract environmental factors. Supervised image classification of land use and land cover, and a

calculation of Normalised Difference Vegetation Index (NDVI) and distance to vegetation were used for the study.

In 2006, Thomas and Lindsay conducted a study to investigate the local-scale variation in malaria infection amongst rural Gambian children estimated by satellite Remote Sensing. To map these features, 20m grid cell coverage of land cover was derived from a SPOT XS image obtained on 23 September 1988 and was contemporaneous with the first year's field data. Rai, et al. (2011) also conducted a study to show the role of GIS and Global Positioning System (GPS) in vector-borne disease mapping. A database was generated based on a disease incidence report and linked with the vector layer. Digital maps were generated on GIS Platforms and suitable cartographic techniques. Geo-referenced digital maps of villages/tehsils/districts were used for GIS platform. A three tier database was created district-wise, tehsil-wise, and village-wise. Attribute data such as village wise population, schedule caste/ schedule tribe population, medical facilities, primary health centres, were attached to the village maps and were used for the analysis for decision support in formulation of control strategies. Since health phenomena have exposed reliable spatial aspects, maps can show spatial distribution and spatial patterns of diseases (Bhatt and Joshi, 2012).

2.6.2. Bayesian Geo-statistical Modelling

One of the famous definitions of Bayesian approach by Eddy (2004, p1177) states that, "Bayesian approach is to write down exactly the probability we want to infer, in terms only of the data we know, and directly solve the resulting equation which forces us to deal explicitly with all mathematical difficulties, additional assumptions and uncertainties that may arise". Bayesians use the rules of probability to describe any unknown quantity. Bayesian methods, as stated by Bernardo (2003), (1) reduce statistical inference to problems in probability theory, thereby minimise the need to complete new concepts, and (2) serve to discriminate among conventional statistical techniques, by either providing a logical justification to some (and making explicit the conditions under which they are valid), or providing the logical inconsistency of others. Bayes' theorem allows one to revise beliefs regarding a specific parameter, given the data that occurred (Bolstad, 2004).

Previously, Bayesian geo-statistics have been utilised in disease mapping, with the benefits of both environmental covariates and spatial autocorrelation being able to be estimated simultaneously. Additionally, the full posterior distributions which are produced can be used

to quantify uncertainties in parameters of interest (Basanez, et al., 2004). In 2005, modelling approaches were extended and developed as part of the National, Morbidity, Mortality and Air Pollution Studies (NMMAPS). This was done by introducing Bayesian hierarchical distributed lag models with non-normal random effects distributions for estimating associations between short-term variations in ozone and cardio-respiratory mortality (Haug, et al., 2005). Bayesian geo-statistical models implemented via Monte Carlo methods avoid asymptotic inference and computational problems encountered in likelihood-based fitting (Gosoni, et al., 2006). Bayesian inference was used by Kazembe and colleagues in 2007 to estimate three models through Markov chain Monte Carlo (MCMC) (Kazembe, et al., 2007).

Authors have been previously using modelling for mapping malaria in diverse studies. Using Bayesian statistical models, Kleinschmidt and Sharp (2001) investigated spatial and temporal variations in small area malaria incidence rates for the period from mid-1986 to mid-1999 for two districts. A hierarchical fully Bayesian spatial modelling approach was utilised (Bernadinelli and Montomoli, 1992) as cited by (Kleinschmidt and Sharp, 2001). Since this was count data, a Poisson regression model was fitted for the mean count of cases in the i^{th} location and t^{th} year. Zhou, et al. (2004) conducted a study, examining temporal and spatial distribution patterns of *Anopheline* malaria vectors in a highland site and determined the number of houses to be sampled to achieve the targeted precision level. The spatial distribution of the malaria vectors was tested for goodness-of-fit for the Poisson and negative binomial distribution. The Poisson distribution described random patterns and the negative binomial distribution described aggregated patterns.

A study on mapping malaria incidence distribution that accounts for environmental factors was conducted by Zacarias and Andersson (2010). The Poisson model was fitted to analyse the relationship between environmental factors and malaria cases. The study used yearly (2001-2002) aggregated malaria data by district level and additional climate data sources to access the role of climate. If a Poisson model has a mean which is unequal to its variance, then a negative binomial must be used for count data. Bayesian geo-statistical zero inflated binomial models observed with many zeros across locations have better predictive performance than standard binomial analogues (Amek, et al., 2011). Lately, Amek, et al. (2012) used a lag time analysis in STATA (version 9.0) to determine the best combination of lags that favour the mosquito density and infectivity considering seasonality, distance to water bodies and altitude. A geo-statistical zero inflated negative binomial model was fitted to further assess space-time variation with a cycle of 12 months. Fitted in OpenBUGS 3.1.2,

the model also considered the temporal effect and autoregressive term to include temporal correlation (Amek, et al., 2012).

2.7. Conclusion

According to the review of the literature, studies on the geographical and temporal patterns of malaria transmission have been done. Various methods and tools have been used in mapping malaria. This reveals that the fight against malaria is still on and it is approached utilising different methods and tools especially in the sub-Saharan Africa. This study was located in one of the provinces in the sub-Saharan Africa which will be described in the next chapter.

CHAPTER THREE

STUDY AREA

3.0. Introduction

This chapter provides a detailed explanation of the study area. A brief description of geographical information on the Limpopo Province is provided including location, population and extent. Subsequently, a detailed explanation of the environmental and socio-economic factors important for the study is provided. These factors are to be used to achieve the aim and objectives of the study. Therefore, this chapter provide the information on the factors based on the study area.

3.1. Location

The study area is in the Limpopo Province which is located in the uppermost (North) part of South Africa (see Figure 3.1). The province shares a border with three different countries namely, Zimbabwe to the north, Botswana to the west and Mozambique to the east. Limpopo Province covers an area of approximately 125 754 km², which represents 10.3% of the area of South Africa. The geographical location of Limpopo Province is within the extent of latitudes 22° 0 and 25° 33 0 ,and longitudes 2 ° 2 and 32° 2 . It is divided into five district municipalities named Capricorn, Greater Sekhukhune, Mopani, Vhembe and Waterberg. The district municipalities are subdivided into 26 local municipalities which are subdivided into 2 166 sub-places in total. The capital city of Limpopo Province is called Polokwane (DBSA, 1998). Census 2011 reflected a total population of 5 404 868 people, which represents 10.5% of the entire population of South Africa (Statistics South Africa, 2012).

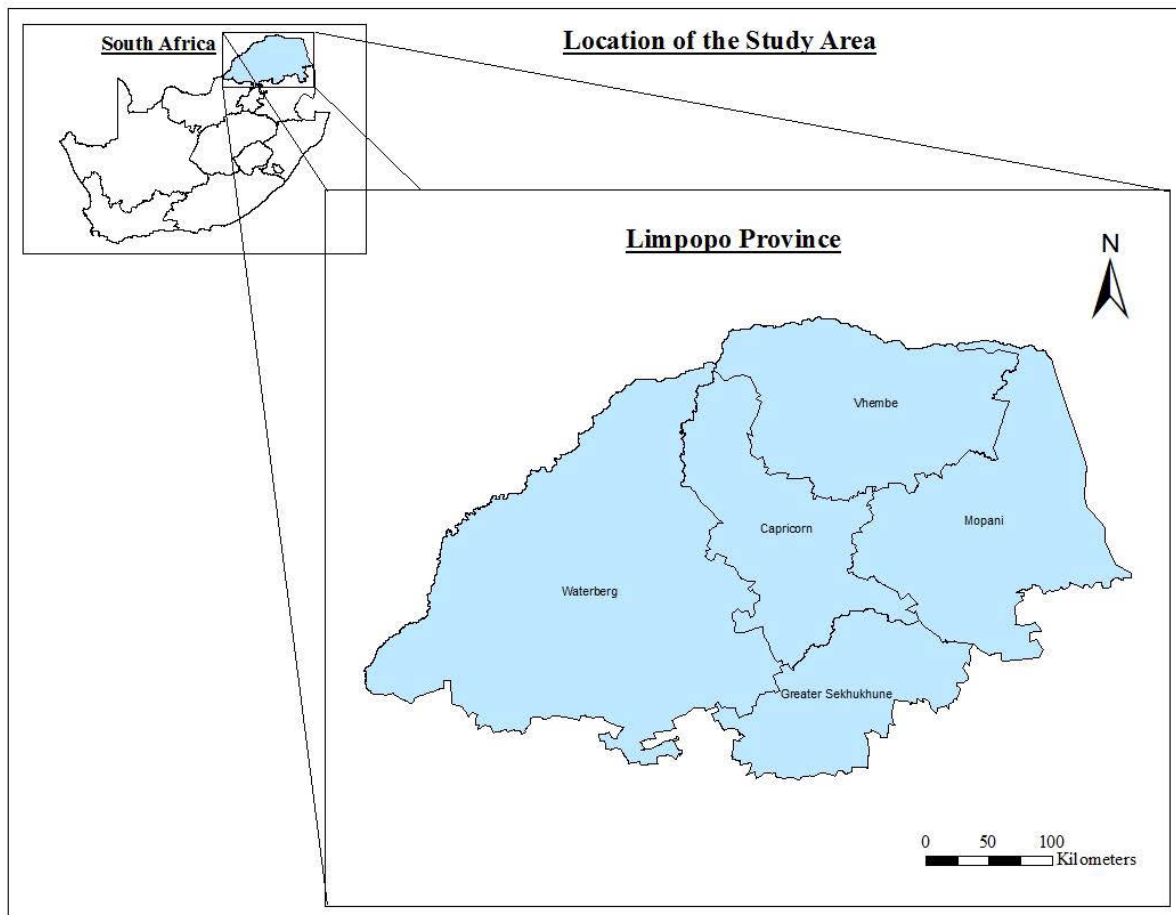


Figure 3.1 Map showing the study area, divided into 5 district municipalities

3.2. Climate

The Limpopo Province contains four climatological regions, namely; the arid far Northern; the arid to semi-arid Northern; the semi-arid Highveld; and the sub-humid low-veld. Generally, the climate is of a sub-tropical type and it falls within a summer-rainfall region (October to March) receiving 90 % of its rainfall annually (Mzezewa, et al., 2010). The average annual rainfall ranges between 300-400 and 600mm. However, the mountain zone has an annual rainfall of 2000mm and the dry low-veld in the Kruger National Park about 400mm (M'Marete, 2006). The eastern and northern parts are sub-tropical, with humid and hot summers. Daily temperatures vary from mid- 20's to mid-30's with an average between 7°C and 27°C in summer and °C to 20°C in winter (Mzezewa, et al., 2010). In winter (May to September), the nights are cold and mostly frost-free, with chilly mornings and dry and sunny days (Tshiala, et al., 2011). The low-veld (eastern part) can get very hot and smouldering. Certain towns can reach up to 5° C, but temperatures in the Kruger National

Park are usually 25-30°C (Tshiala, et al., 2011). Frequent mist occurs in the mountainous areas (Haernertsberg and Magoebaskloof area) (Thomas, 2006).

3.3. Vegetation

The classification of the vegetation found in the study area is inland tropical forest; tropical bush and savannah; pure grassveld; and false grassveld types (DBSA, 1998). There is a combination of bushveld, forest, open grasslands and wetlands classified according to their location and topography (Odhiambo, 2006). The presence of vegetation and wetlands affects the transmission of malaria since they provide breeding sites. The eastern and southern slopes of mountain ranges are subjected to the orographic rainfall blowing from the Indian Ocean (Reyers, et al., 2002). This specifies the relationship between topography and rainfall which may result in the production of vegetation (Cerlini, et al., 2005).

3.4. Topography

The range of elevation varies between 600m to about 900m above mean sea level (amsl) (Odhiambo, 2006). Along the northern edge of the Bushveld, the plains rise to a series of high plateaus and low mountain ranges, forming the southern edge of the Limpopo River Valley (DBSA, 1998). These mountains include the Waterberg and in the far north, the Soutpansberg Mountains. The Soutpansberg range goes to an elevation of 1 700m amsl before dropping off into the Limpopo River Valley and border South Africa and Zimbabwe (M'Marete, 2006). To the east there are high grounds of the Drakensberg. Therefore, the change in topography (altitude and relief) gives rise to varied climatic characteristics.

3.5. Socio-economic Status

The Limpopo Province is predominantly rural with close to 80% of the population falling into this category (Limpopo Provincial Government, 2008). This type of environment prominently forces the population's capacity to be educated which can highly influence the potential for employment in the formal economy. Available information from Census 2011 reveals that the Limpopo Province has the highest proportions of people aged 20 years and older with no formal education. But the percentage has been decreasing from 1996 to 2011. The highest percentage of people in this category (39%) is found in Vhembe District while Capricorn District has the lowest percentage (9%) (Limpopo Provincial Government, 2011).

The number of households has been steadily growing over the past 15 years (1996-2011). With the effects of inflation as well as of increasing access to jobs and a growing economy, the average annual household income increased substantially since 1996 to 2011 (Statistics

3.7. Conclusion

The factors mentioned in this chapter are very useful in mapping malaria. Geographical and temporal patterns of malaria transmission can be analysed using different factors depending on the availability of data and the methods/tools used. The material and methods used to conduct this study are provided in the next chapter.

CHAPTER FOUR

MATERIAL AND METHODS

4.0. Introduction

This chapter briefly explains the material and methods used to conduct the study. A detailed account of the sources and cleaning of the data utilised for the study is explained. Subsequently, the description of the methods applied to achieve Bayesian geo-statistical models and prediction is explained. Finally, the chapter explains the process of the prediction validation of risk maps.

4.1. Data Acquisition and Cleaning

4.1.1. Malaria Case Data

Malaria is one of the notifiable medical conditions in South Africa. Generally in South Africa we have two methods of malaria case data collection namely, active and passive surveillance system. Active surveillance screening measures were done where by teams went to the households. Smears were taken from all household members, but with emphasis on those with fever, a history of fever, a travel history to a malaria risk area or possible migrants from malaria endemic areas (Gerritsen, et al., 2008). Moreover, passive surveillance systems were done where a definitive diagnosis was conducted under a microscope in the laboratories (e.g hospitals). Only cases positively confirmed were notified and entered in the computer systems. The Malaria Control Programme (MCP) manager then identified a staff member to visit health facilities on weekly bases for the collection, verification of records and performing regular checks. Finally, the data was reserved in the MCP and sent to provincial malaria control centres (Gerritsen, et al., 2008).

Confirmed malaria cases were collected from the provincial malaria control centre in the city of Tzaneen, Limpopo Province. They were entered in a Microsoft Access-based national malaria system created for data entry and partial validation. The data was then sent to South African Medical Research Council (MRC) for further cleaning and collation. Archived malaria case data for the analysis was acquired from MRC for the period, from 1998 to 2011. There was a number of 82 fields and 69 348 records in total and not all the fields and records were relevant for this study. Hence, a technique of data mining was developed to filter all relevant epidemiological data for this study. Only the local cases with existing geo-

information were utilised. Malaria cases outside Limpopo Province were excluded. Finally, the spelling and geographical location errors were also corrected to make the data cleaner.

There was a large number of locations as the study area covered the entire province. In total, there were 19 388 locations from May 2000 to June 2011 with 51 681 malaria cases in 1 221 sub-places. The cases were clustered in the north-eastern part of the province and dispersed as you move towards the south-western part. This excludes the Kruger National Park on the eastern part of Limpopo Province. Due to computational and modelling challenges of OpenBUGS/WinBUGS, a subset of 40% locations sample was extracted in STATA in order to accommodate the entire province equitably and fairly distributed (Lasinio, et al., 2013). Consequently, the total number of locations was 7 844 with 21 333 malaria cases in 491 sub-places.

4.1.2. Population Data

The population data was acquired from AfriPop, a website for detailed and free population distribution maps of Africa. The data retrieved had a 100m spatial resolution for the year 2010 as a float file format. The float file format was converted into a raster file format using ArcGIS Version 9.3.1. A population grid of 1 x 1km. was then created to increase the spatial resolution by converting a raster file to a point vector file. This enabled the spatial join of population grid with the sub-place polygon and therefore produced population at a sub-place level. The grid points were allocated into a polygon which they fell within, to get a total population in each sub-place. There was no population data available for the Kruger National Park because it is not inhabited, thus, the park was not included in any of the analysis and modelling in order to avoid bias. Finally, the cases which were greater than the population in each sub-place were dropped.

The population for other years of the study period was calculated using a growth rate. A technique was developed to form a growth rate using the population census of 1996, 2001 and 2007. The population data for those years are available in the Statistics South African databases. The growth rate which was developed by the MRC staff using a geometric regression model was used for the population at each sub-place per year.

4.1.3. Environmental Factors

4.1.3.1. Rainfall

Rainfall data were acquired from the African Data Dissemination Service (ADDS). ADDS uses a technique for estimation of precipitation to augment the rainfall data available from the sparse observational network of rain gauge stations in Africa (Riedel, 2009). The dataset selected during the downloading process had 8 by 8km spatial resolution and 10 days (decadal) temporal resolution (see Table 4.1). The data was downloaded directly from ADDS website as a Band Interleaved by Line (bil) file format which can be opened in ArcGIS for processing.

4.1.3.2. NDVI, Land Surface Temperature (LST) and Land Cover

The data was acquired from MODerate-resolution Imaging Spectro-radiometer (MODIS). MODIS is a scientific instrument launched by NASA in 1999 based on the Terra (EOS AM) and Aqua (EOS PM) Satellite for terrestrial and aquatic data (Tatem, et al., 2004). The MODIS website which was used to download the data is <http://reverb.echo.nasa.gov/reverb/>. The datasets selected during the downloading process had 1 by 1km spatial resolution and 16 days temporal resolution for NDVI; 1 by 1km spatial resolution and 8 days temporal resolution for LST day and night; and 1 by 1km spatial resolution and annual temporal resolution (see Table 4.1). The datasets were retrieved using the 'MODIS Reprojection Tool'. The output files were as a tiff format which can be opened in ArcGIS for processing.

4.1.3.3. Altitude

Altitude was acquired from United States Geological Survey (USGS) a Geological research organisation. The datasets selected during the process of downloading had a 1 x 1km spatial resolution which does not change over time (see Table 4.1). The data was downloaded directly from the USGS website as a Digital Elevation Model (DEM) file format which can be opened in ArcGIS for processing.

4.1.3.4. Distance to Water Bodies

Distance to water bodies was acquired from the Economic and Social Research Institute (ESRI) an international supplier of GIS applications. The files were downloaded directly from ESRI as shapefiles. The latest available files were from 2011 (annual). The shapefiles were spatially joined with the locations of the cases using ArcGIS to get the closest distance of each location to water bodies.

Table 4.1: Environmental data utilised for the analysis

Factors	Spatial Resolution	Temporal Resolution	Period	Source
NDVI	1km	16days	2000-2011	MODIS
LST	1km	8days	2000-2011	MODIS
Rainfall	8km	Decadal (10 days)	2000-2011	ADDS
Land Cover	1km	Annual	2001-2011	MODIS
Water Bodies	Shapefiles	Annual	2011	ESRI
Altitude	1km	-	-	USGS

4.2. Data Processing

Basically, the data was processed using ArcGIS Version 9.3.1. The data was derived with a specific value corresponding to missing values which was zero (0). ArcGIS does not (yet) consider these values as missing and would thus include them when extracting and calculating means (Riedel, 2009). Missing values (0) were reclassified and replaced with “NoData” which is the ArcGIS analogue for missing information. Furthermore, the data was reprojected to contain the same projection and geographic coordinate system. The projected coordinate system and geographic coordinate system used was WGS 1984 UTM Zone 36S.prj and WGS 1984.prj, respectively. The values were extracted at each sub-place level using the unique locations of reported malaria case data. Finally, the points were transferred to STATA using Stat Transfer 7 for the conduction of preliminary analysis.

4.3. Statistical Analysis

4.3.1. Data Management

Preliminary and descriptive statistical analysis was conducted using STATA version 10.1. Primarily, malaria case data was merged with population, environmental and climatic data for statistical analysis. The environmental data had a high temporal resolution and it was categorised to an appropriate temporal resolution for the study. Changes were not made on altitude and distance to water bodies because both were continuous.

4.3.2. NDVI, LST day and LST night and Rainfall

The data was averaged on a monthly basis to avoid temporal inconsistency with malaria case data. Different periods or time intervals prior to the actual reading or month were named lags. Remote sensing data was available from February 2000 and therefore, cases before May 2000 were dropped to accommodate lags calculation. Lags considered include (a) current month (aligned with the case); (b) previous month reading; (c) two months before the case actual reading; (d) average between current month and previous month; and (e) average among the two months before the actual reading and the current month reading (see Table 4.2).

Table 4.2: Different periods or time intervals used to average data in lags

Lag Name	Description
Lag_0	• Current month (aligned with the case).
Lag_1	• Previous month.
Lag_2	• Two months before the case.
Lag_01	• Average between current month and previous month.
Lag_012	• Average among the two months before the case and the current month.

4.3.3. Land Cover

Land cover (categorical variable) was re-grouped into four categories (a) dry non-forest vegetation; (b) forest; (c) wet non-forest vegetation; and (d) non-vegetation. Land cover classification scheme from MODIS was used to categorise land cover (NASA, 2000). Dry non-forest vegetation category contains closed shrublands, open shrublands, woody savannas, savannas and grasslands; forest category contains evergreen broadleaf forest and deciduous broadleaf forest; wet non-forest vegetation category contains croplands and cropland/natural vegetation mosaic; and non-vegetation category contains urban and built-up (see Table 4.3). A category with the highest frequency was made a baseline. Statistically, a baseline is the first category to appear in the frequency table, thus, omitted by default in STATA (Sterne and Tilling, 2002). In this instance, dry non-forest vegetation was the baseline. The category was omitted and made a reference category as it was the most prevalent category. Subsequently, the data was centred and scaled to have zero sample mean and unit sample. This is a binary code which was used to assign a digit for a category from either 1 or 0. The code created was true (1) if the category was selected, otherwise false (0).

Table 4.3: Land cover classification scheme used to categorise land cover data

Category	Land Cover
Dry Non-Forest Vegetation	<ul style="list-style-type: none"> • Closed shrublands, • Open shrublands, • Woody savannas, savannas and • Grasslands.
Forest	<ul style="list-style-type: none"> • Evergreen broadleaf forest and • Deciduous broadleaf forest.
Wet Non-Forest Vegetation	<ul style="list-style-type: none"> • Croplands and • Cropland/natural vegetation mosaic.
Non-Vegetation	<ul style="list-style-type: none"> • Urban and • Built-up.

4.3.4. Bayesian Geo-statistical Incidence Model

The model was formulated in OpenBUGS321 using the above mentioned environmental data as covariates; and population and malaria case data to form malaria incidence data. Malaria incidence is the number of newly diagnosed malaria cases during a specified time period in a specified population (World Health Organisation, 2010). Malaria incidence data was modelled using a negative-binomial (NB) distribution. Unlike Poisson distribution, a negative-binomial can allow the variance to take any value greater or smaller (overdispersed or underdispersed) than the mean (Link and Barker, 2010). Spatial and temporal correlation was taken into account by introducing sub-place and month specific random effects. Spatial random effects were used at sub-place level to take into account spatial correlation present in the data. Temporal random effects were used at monthly intervals to account for temporal correlation. The geo-statistical model assumed that the numbers of observed malaria cases N_{it} at sub-place i and month t follow a negative-binomial (NB) distribution with mean (μ_{it}) which is $N_{it} \sim NB(p_{it}, r)$, where $p_{it} = r / (r + \mu_{it})$ and therefore $N_{it} \sim NB(\mu_{it}, r)$. r is the overdispersion parameter and μ_{it} is an expected average number of cases at sub-place i and month t . Gamma prior distribution was assigned for the parameter r and all the other parameters were assigned a non-informative prior distribution. The general negative-binomial model had the following form of equation:

$$\log(\mu_{it}) = \log(P_{it}) + trend + seasonality + EO + spatial + temporal$$

where P_{it} is the population count at risk in sub-place i and month t , trend which is $f_T(t)$ represents a function of time, seasonality which is $f_S(t)$ identifying the seasonal pattern. Furthermore, $f_T(t)$ and $f_S(t)$ are represented by the following equations: $f_T(t) = \beta * t$ and $f_S(t) = a_1 * \cos(2\pi t / T) + a_2 * \sin(2\pi t / T)$, $t = \dots 2$. EO represents malaria predictors indicated by $\beta_1 * X_{1it} + \beta_2 * X_{2it} + \dots + \beta_k * X_{kit}$. Spatial (ϕ) arises from a spatial Gaussian process $\phi = (\phi_1, \phi_2, \dots, \phi_L)^T$ and $\phi \sim N(0, \Sigma)$, where $\Sigma_{ij} = \sigma^2 \exp(-\rho d_{ij})$ and d_{ij} is the distance between sub-places i and j . Temporal (e) arises from an autoregressive process where $e = (e_1, e_2, \dots, e_T)^T$ with $e_t \sim N(\gamma e_{t-1}, \sigma_e^2)$.

4.3.4.1. Variable Selection

Gibbs variable selection methods (Dellaportas, et al., 2002) were employed using the above mentioned negative-binomial to introduce indicators and select the best set of predictors. For the regression coefficient, it was assumed that $\beta_k \sim N(0, \sigma_k^2)$, σ_k^2 is large (i.e. 1 000). Indicator I_k was introduced for each k^{th} variable. $I_k = 1$, if the variable should be included in the model (otherwise $I_k = 0$), where $I_k \sim \text{Bern}(0.5)$. Then

$$\beta_k \sim I_k * N(0, \sigma_k^2) + (1 - I_k) * N(0, \tau_k^2)$$

where σ_k^2 is a large fixed number and τ_k^2 is chosen large enough to give β_k close to 0 when k^{th} predictor is not included in the model. To consider different lags, a multinomial prior was assigned for Rain (Rainfall), NDVI, TempD (Day Temperature) and TempN (Night Temperature). Different lags formed were all assigned to a categorical distribution.

The total number of candidate covariates was 28 and 7 indicators were introduced as mentioned above. The best set of indicators was indicated by the model after 152 000 iterations. Indicators show whether the covariate is selected or not. The covariates that were included in the spatio-temporal model were those corresponding to indicators with posterior mean greater or equal to 50%. More details on a negative-binomial variable selection method are given in the Appendix.

4.3.4.2. Model Fitting and Prediction

The covariates selected by the variable selection procedures were utilised to fit a negative-binomial model with spatially and temporally structured random effects at sub-place level and monthly temporal autoregressive terms. Bayesian formulation and Markov chain Monte Carlo (McMC) estimation were adopted to enable model fit. A single chain sampler with a burn-in of 20,000 iterations was run for 100,000 iterations with thinning of 1 and the convergence was assessed by graphical inspection of the Markov chain output. The importance of the covariate was assessed by the Bayesian credible interval. If the credible interval did not include 0, the covariate was important. The posterior mean of the regression coefficient was used to quantify the relationship of the covariate with malaria risk. Monthly predictions from the above model were obtained using Bayesian kriging for the year 2010.

Model-based predictions were obtained over a regular grid of 1 by 1km spatial resolution covering the entire province and expressed as rates of per 1 000 inhabitants. Predictions were carried out in R. The maps relied on prediction of risk at locations without observed incidence data. Details on model fitting are provided in the Appendix.

4.3.5. Assessment of Predictive Performance

In order to assess the accuracy of the predicted malaria incidence risk maps, the prediction and observation fields were compared. Comparisons included quantifying the difference between observed and predicted plot-level mean. A linear relationship between observed and predicted mean values was established for each month in year 2010. The relationship was evaluated both graphically and based on the coefficient of determination (R^2). Scatter plots were created to quantify the linear relationship using a trend line. This introduced a calculated value of R^2 along with the linear relationship equation (Tesfamichael, et al., 2010). R^2 determines how closely the plotted data conform to a linear relationship. It considers values from 0 to 1, it becomes larger (1) if there is a perfect fit between data and the trend line, and becomes 0 if there is no statistical correlation (Shtatland, et al., 2000). A linear relationship was defined by the following equation:

$$y = ax + b$$

where a is the slope of the line and b is the intercept (i.e. where the line cuts y-axis). The results were then analysed to get to the end of the study. A general overview of the methods developed in this study is illustrated in Figure 4.1 below.

A General Overview of the Methods Developed In the Study

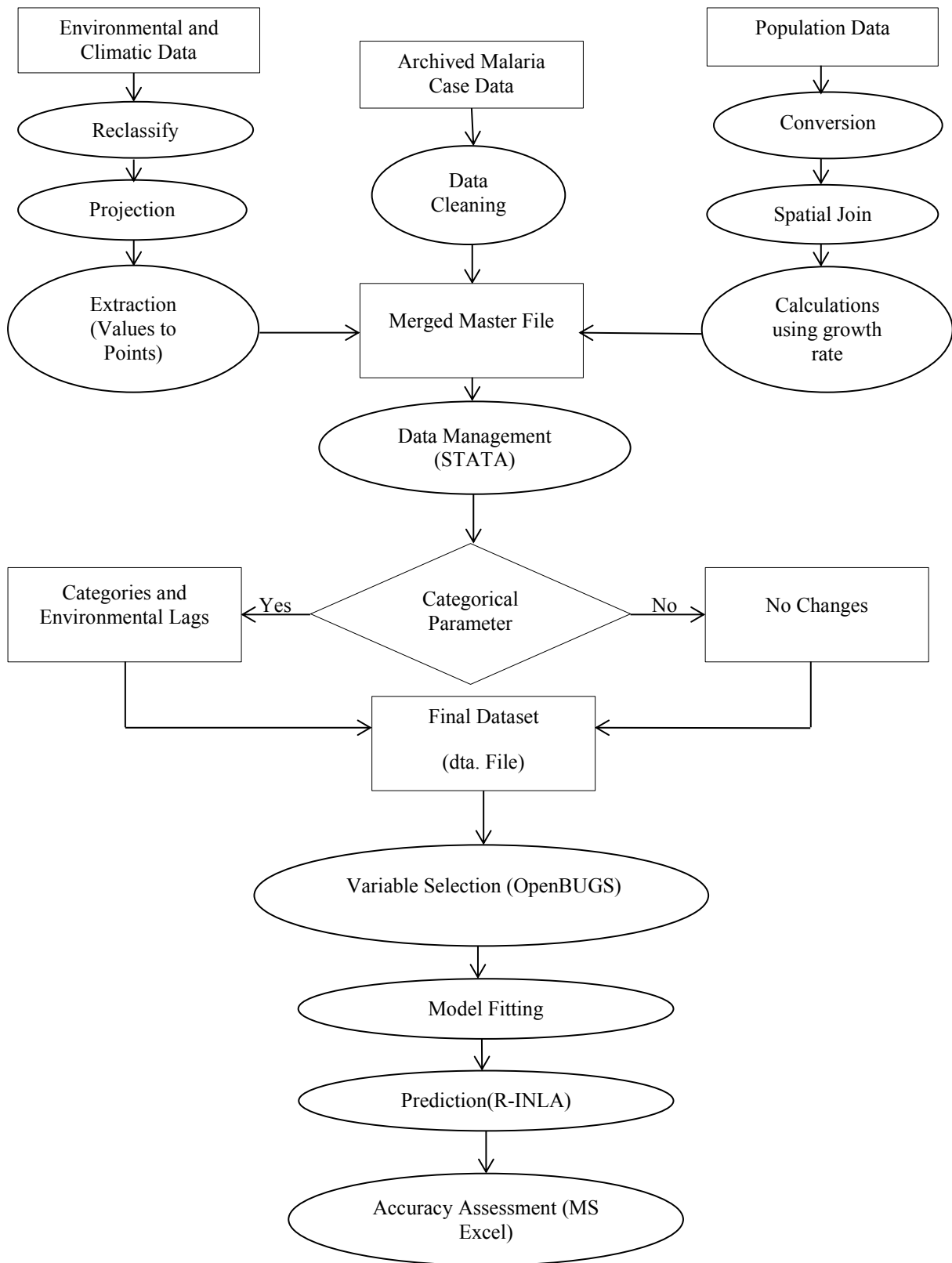


Figure 4.1: Flow diagram showing steps followed for the study methods

4.4. Conclusion

Environmental factors, malaria case data and the population of the Limpopo Province were used to achieve the objectives of this study. The programs mentioned in this chapter were used together to form the results for this study which is the key aspect of the research. The results are presented and discussed in the next chapter.

CHAPTER FIVE

RESULTS AND DISCUSSION

5.0. Introduction

This chapter presents the results and a detailed discussion of the findings of this study based on the aim and objectives of the study. A detailed account of environmental factors which are significant to malaria risk in Limpopo Province is explained. The estimates from a Bayesian negative-binomial incidence model are presented and discussed. Finally, the predicted incidence maps for 12 months in 2010 with accuracy assessment are presented and a detailed discussion is given.

5.1. Results

5.1.1. Descriptive Analysis

During the study period (May 2000 to June 2011), a total number of 51 681 malaria cases from 1 221 unique locations (sub-places) was analysed. Figure 5.1 depicts the total number of malaria cases in Limpopo Province per year over the study period. There were high malaria cases reported in 2001 followed by a decrease in 2002 and an increase again in 2003. Subsequently, it was followed by a dramatic decrease trend between 2003 and 2005 (down to 3000 cases). The recorded number of cases went up again in 2006 showing a marked decrease in 2007 (less than 3 000 cases). In 2008, the reported malaria cases went up (less than 4 500 cases) but not as much as in the previous years. The same pattern was formed for the following years until 2011.

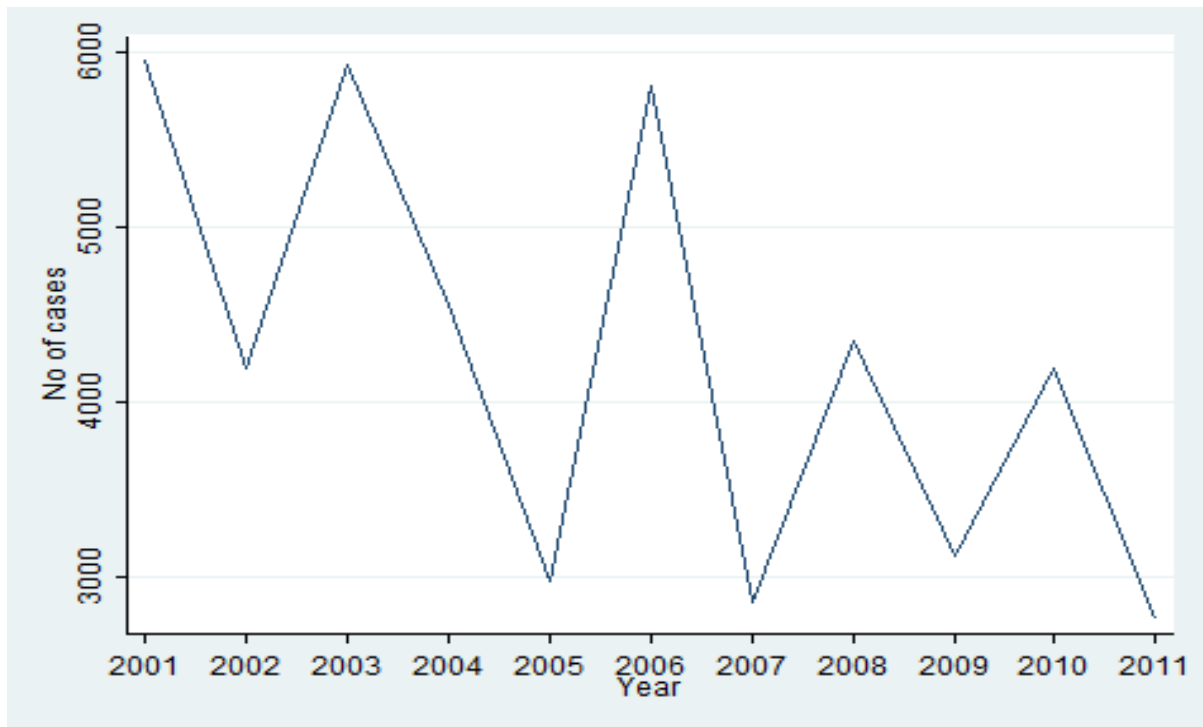


Figure 5.1: Graph showing a pattern formed by the total number of cases per year in Limpopo Province from 2001 to 2011. Years compared to the number of reported cases.

The total number of reported malaria cases by month and year is presented in Figure 5.2 below. Statistically, there was an observed decrease in a number of cases from the months June to August in each year. The cases decrease in these months possibly due to malaria control interventions. The main malaria intervention utilised in Limpopo Province is IRS (using DDT) and contains an effective residual life of 9 months (Maharaj, 2008). The province is sprayed yearly in the warm and rainy summer months (September to May) (Van Dyk, et al., 2010). Malaria interventions affect the peaks of malaria cases in Limpopo Province. After the spraying season, the number of cases decrease and then increase before the next spraying season.

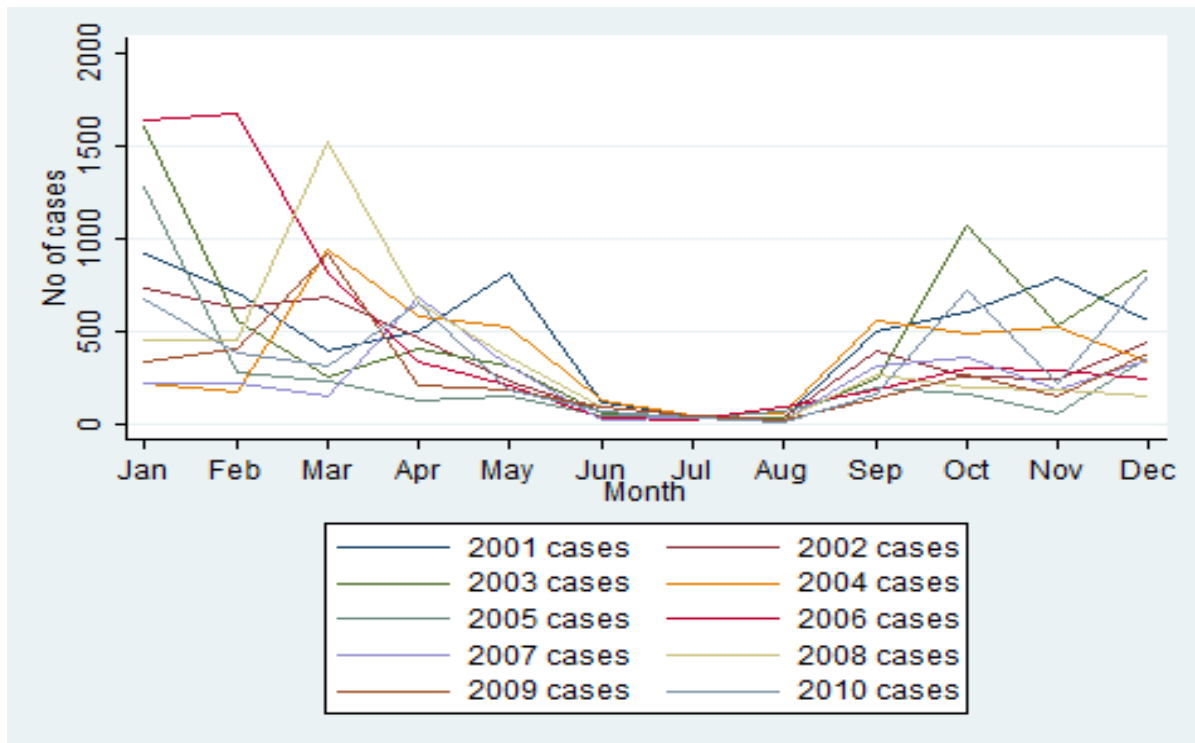


Figure 5.2: Graph showing the total number of cases per month in each year in Limpopo Province. Months compared to a total number of reported cases per year (2001 to 2010).

Monthly averages of malaria cases reported and environmental factors are demonstrated in Figure 5.3 below. The bar graphs portray the number of cases and the line graphs representing each environmental factor (Rain, NDVI, LST_Day and LST_Night). The number of cases decreases with a decrease in rainfall revealing a good relationship between the two. There are fewer cases during the cold and dry months. NDVI is lower from August to October yet cases are fewer from June to August. LST_Day and LST_Night are lower in the cold and dry months with the number of cases also fewer in these months. The lower the environmental factor, the lower the cases reported according to the archived malaria case data for the study.

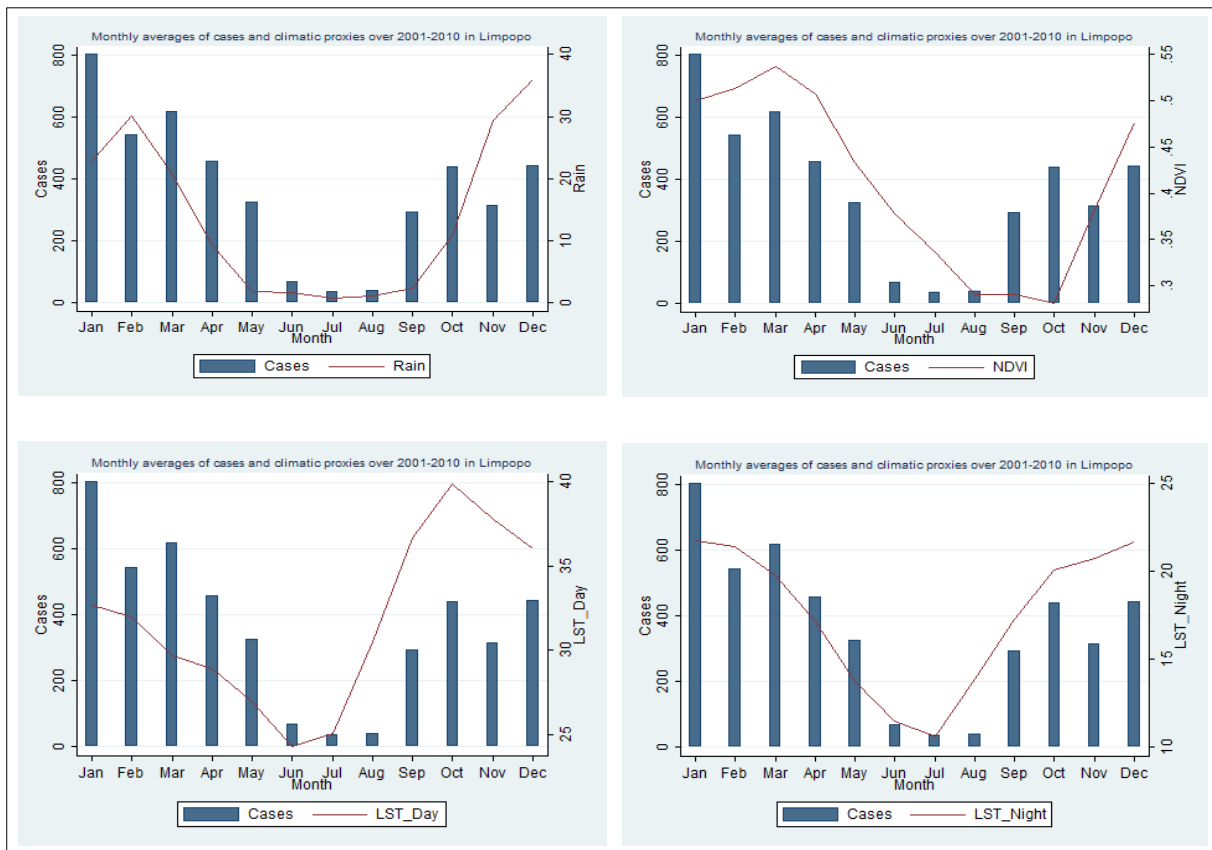


Figure 5.3: Graphs showing the relationship between monthly averages of malaria reported cases and environmental factors in Limpopo Province from 2001 to 2010.

5.1.2. Bayesian Geo-statistical Incidence Model Estimates

5.1.2.1. Variable Selection

Posterior mean, standard deviation and median obtained from MCMC runs of 152 000 iterations using Gibbs variable selection method are presented in Table 5.1 below. Altitude had the highest posterior mean which was 0.6868 (70%) and the median value of the indicator was 1 as shown in Table 5.1. Likewise, night temperature had a median value equal to 3 which is less than 5 denoting that the third lag was selected. The third lag represents the night temperature of two months before the case was reported. Both of the covariates were employed to fit a Bayesian negative-binomial model at sub-place level without indicators.

Table 5.1: Summary of the results from Gibbs variable selection

Indicators	Factors	Mean	Standard deviation	Median
Ind[1]	Altitude	0.6868	0.4638	1
Ind[2]	Distance to Water Bodies	0.1774	0.382	0
Ind1	Land Use/Land Cover	0.001851	0.04299	0
IndTempD	Day Temperature	5.694	0.993	6
IndTempN	Night Temperature	4.378	1.515	3
Indndvi	NDVI	5.604	1.076	6
Indrain	Rainfall	5.673	1.02	6

5.1.2.2. Model Fitting

The estimates of the effects of seasonality, environmental factors (Altitude and Night Temperature) and spatio-temporal variations on malaria risk obtained from a Bayesian negative-binomial model in Limpopo Province are presented in Table 5.2. The negative-binomial analysis indicated that all the covariates were significantly associated with malaria risk statistically. In particular, the posterior mean estimates of seasonality a_1 and a_2 had a positive association with malaria risk. A negative relationship of malaria risk with altitude indicated that the higher the altitude, the lower the risk of malaria. The posterior mean estimate of the night temperature of two months before the case was reported was positively associated with malaria. This suggested that the higher the night temperature of two months before the case was reported, the higher the risk of malaria. The spatial range covariate suggested that spatial correlation was present up to a distance of 0.72 ° which is equivalent to 2 km (0.000 ° = 0.27m) at 95% credible interval. Spatial variability (2.796) was around 8 times higher than the temporal variability (0.336) indicating higher geographical variation. Temporal variation accounted for 11% of the total variance in malaria risk and therefore most of the variability was due to spatial effect.

Table 5.2: Summary of the estimates from a fitted Bayesian negative-binomial model

Covariates	Mean	Standard deviation	Median	95% BCI ^a
Constant	-7.641	0.1374	-7.655	(-7.86, -7.25)
Seasonal a_1	0.3636	0.08478	0.3629	(0.20, 0.53)
Seasonal a_2	0.3196	0.06554	0.3184	(0.19, 0.45)
Altitude	-0.2022	0.07636	-0.1965	(-0.36, -0.07)
Night Temperature	0.1358	0.04142	0.1357	(0.22, 0.05)
Spatial range	0.1172	0.01981	0.1156	(0.083, 0.16)
σ^2 (Spatial variation)	2.796	0.2443	2.778	(2.36, 3.32)
σ_e^2 (Temporal variation)	0.336	0.02391	0.3349	(0.29, 0.39)

^aBayesian Credible Interval

5.1.2.3. Prediction

Smooth maps of monthly predicted malaria incidence rate in 2010 are presented in Figure 5.4 and Figure 5.5 below for the study period. The predicted maps depict the spatial and temporal variation of malaria incidence rate in Limpopo Province. The incidence rate values were classified according to their standard deviation from the mean. The average incidence rate (mean) was 2 cases per 1 000 inhabitants for the predicted maps. The standard deviation was 0.5 with a minimum and maximum number of 0 and 16 cases per 1 000 inhabitants, respectively. However, there were months with a maximum number of 15 cases per 1 000 inhabitants. The distribution of the predicted malaria incidence rate was presented in three different ranges as seen in each map of Figure 5.4 and Figure 5.5.

The province was dominated by incidence values ranging from 0 to 2 cases per 1 000 inhabitants in the majority of the months. The months excluded from the majority were March, June, July and October which were dominated by incidence values ranging from 2 to 4 cases per 1 000 inhabitants. The highest incidence values were greater than 4 to 15 or 16 cases per 1 000 inhabitants in each sub-place. However, high values of malaria incidence were concentrated in the northern, north-eastern and eastern parts of the province for different months.

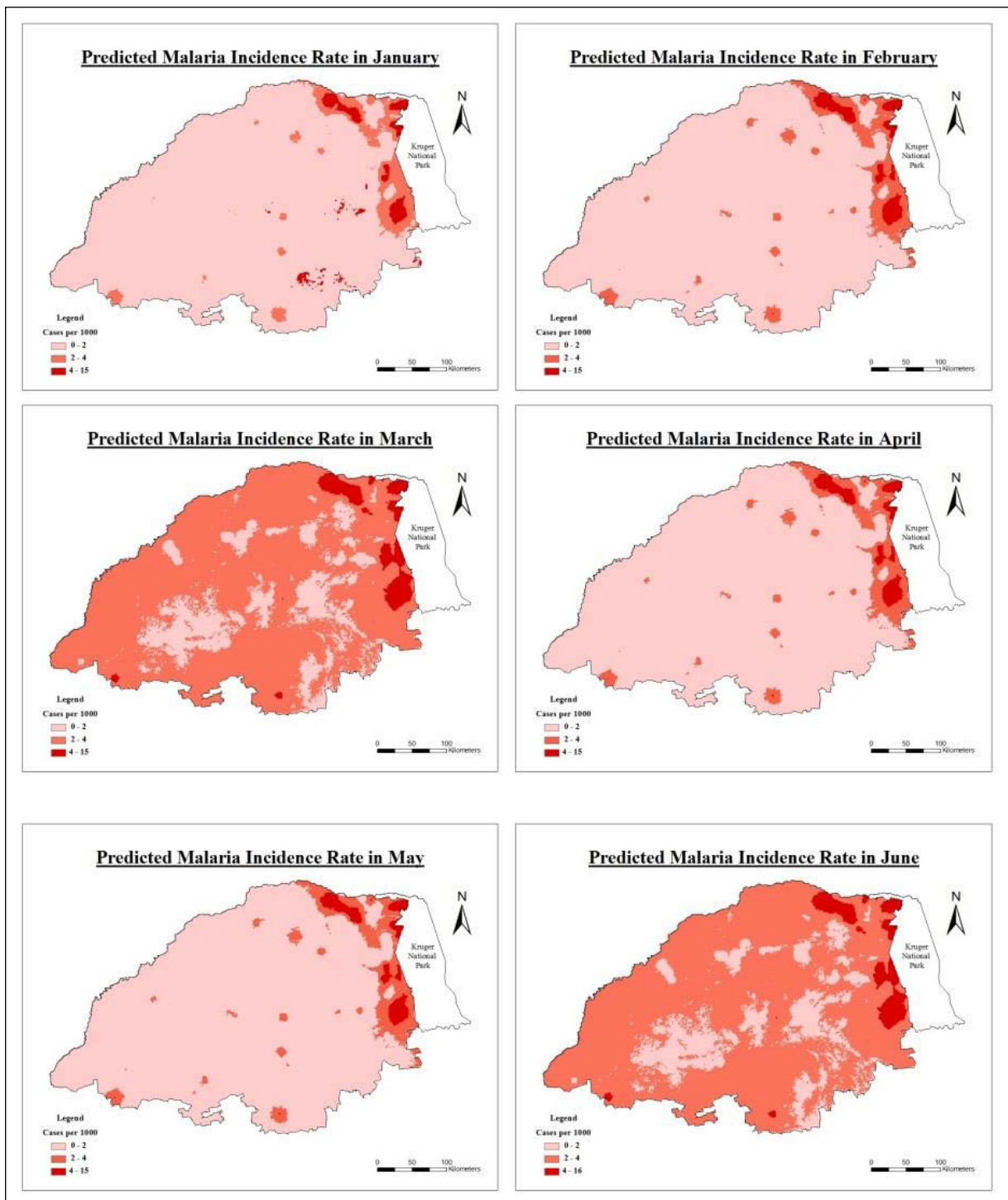


Figure 5.4: Depicts spatio-temporal model based maps of malaria incidence per 1 000 inhabitants in Limpopo Province, from January to June, year 2010

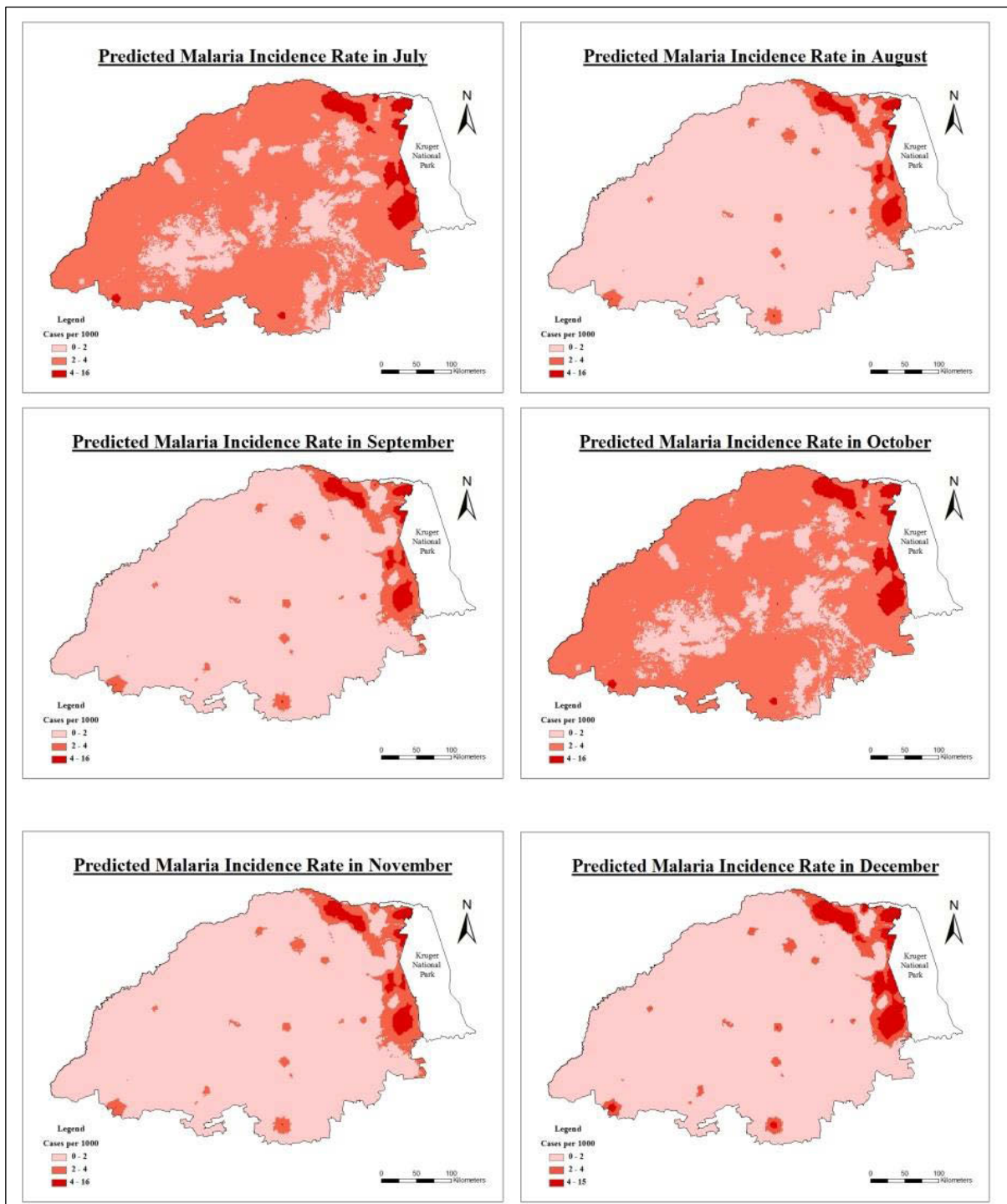


Figure 5.5: Depicts spatio-temporal model based maps of malaria incidence per 1 000 inhabitants in Limpopo Province, from July to December, year 2010

5.1.3. Assessment of Predictive Performance

The scatter plots in Figure 5.5 below present the performance of the predicted malaria incidence rate for the year 2010 in Limpopo Province. The performance was determined by the value of R^2 which is presented in each scatter plot below. The values were 0.9798 (January), 0.8736 (February), 0.8152 (March), 0.8861 (April), 0.9949 (May), 0.3838 (June), 0.7794 (July), 0.9235 (September), 0.8966 (October), 0.9834 (November) and 0.8958 (December). August had cases reported and predicted in only two locations which had a coefficient of determination of 1. Therefore, August had a 100% accuracy of malaria incidence prediction (see Figure 5.5). All the other months had a performance greater than 80% except June which had a 40% performance.

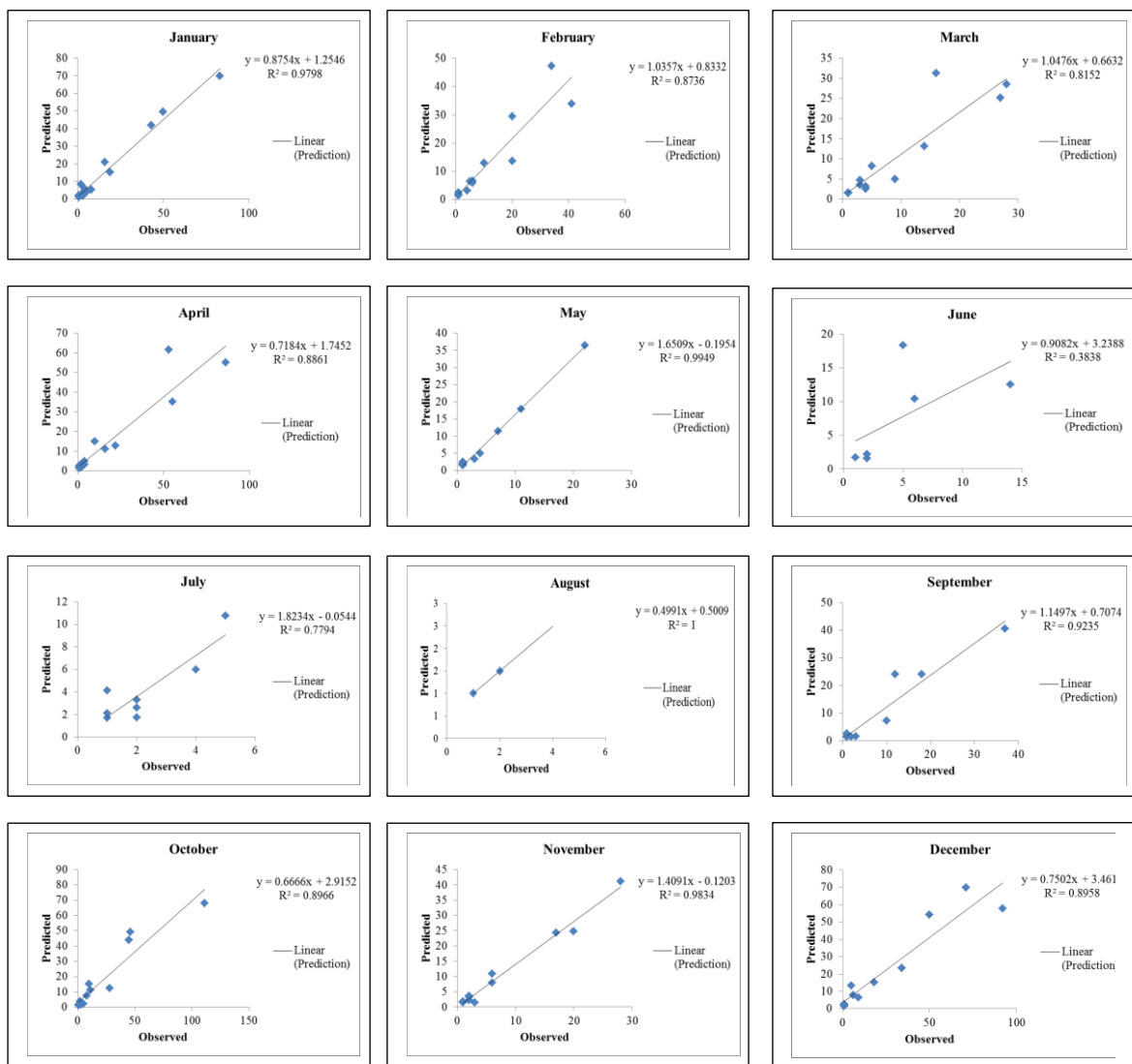


Figure 5.6: Graphs showing the relationship between observed and predicted malaria incidence in 2010

5.2. Discussion

This study estimated geographical and temporal patterns of malaria transmission in Limpopo Province and produced the first malaria seasonality risk maps in the province at sub-place level for the study period (May 2000 to June 2011). Bayesian negative-binomial models were developed for variable selection and model fitting to build a predictive model and assess the effectiveness of environmental factors on malaria transmission.

The results indicated that the negative association between altitude and malaria risk in the results implied that there was a high malaria incidence rate at low altitudes and vice versa. Households built at low altitudes are more at risk of malaria than households built at high altitudes. This may probably be due to the fact that temperatures decrease as one move to higher altitudes. Some parasites do not develop fully at low temperatures (Amek, et al., 2012). An altitude around 1800 to 2000 meters is usually considered the upper limit at which malaria transmission occurs (Lindsay and Martens, 1998). A supporting study revealed that areas in the low lying north-western and south-eastern parts of Zimbabwe have high risk of malaria (Mabaso, et al., 2006). Another study in lower altitudes of Kenya by Saxena, et al. (2009) showed that over 90% of malaria risks are in the low altitude regions. The previous above mentioned studies suggested a linear relationship between altitude and temperature in favour of this study.

In contrast, unusually severe and intense malaria epidemics were reported in several highland areas (1500–2500 m) in four regions on Ethiopia (Yeshiwondim, et al., 2009). Although altitude in the villages of Ethiopia varies from 940 m to 2800 m, there was no relationship between altitude and temperature. Moreover, a study conducted in Kenya (Midega, et al., 2012) in altitudes ranging from 16 to 182m above sea level revealed that altitude did not influence the risk of malaria at lower altitudes. This contradicted with the current study revealing that not all of the previous studies support it. Probably this contrast was due to the differences in altitude range, location or the factors used because not altitude alone affects malaria risk. However, the current study added to the studies that altitude is associated with temperature and malaria transmission.

The positive association between the night temperatures of two months before the case was reported and malaria incidence suggested that recent malaria incidence is affected by bites of two previous months. Studies conducted in China, Yunnan province and Mozambique, Maputo province (Zacarias & Andersson, 2010) referred temperature as the most influential

factor of malaria transmission (Hui, et al., 2009). This attempted to imply that malaria incidence is sensitive to temperature changes. Another study in Burundi (Nkurunziza, et al., 2010) found a significant relationship between the night temperature of one month before the cases were reported and malaria risk. The possible explanations of the results of the current study were as follows. According to the development cycle of the parasite into mosquitoes and the incubation period, people who became ill in a given month were possibly bitten by mosquitoes at night in the two previous months. Usually, mosquitoes are active at night and during the day they hide themselves by resting inside the houses or vegetation. This explains why this factor, which is observed at the same time (at night) when there is mosquito activity, has a great influence on malaria transmission. Furthermore, when the night temperatures are high, people do not cover themselves and sleep with windows open (Coleman, et al., 2010), increasing the risk of being bitten by mosquitoes. Therefore, they are at risk of acquiring malaria.

The spatial correlation from the current study probably arises due to spatial patterns in altitude and the night temperature of two months before the case was reported. This suggested that malaria risk in one sub-place is affected by the risk in the neighbouring sub-place up to 12 km in relation to environmental factors mentioned above. This suggested that there was a strong spatial correlation and therefore malaria risk is more likely to be driven by the two above mentioned factors. In line with this study, the studies conducted in Kenya (Amek, et al., 2012) and northern Ghana (Kasasa, et al., 2013) confirmed a strong spatial correlation of up to 4 km and 2 km. However, the study conducted in Senegal (Giardina, et al., 2012) showed higher spatial correlation as compared to the current study.

Most of the variability in this study was due to spatial effect as it was 8 times greater than temporal effect. The smooth maps generated in this study illustrated that malaria transmission intensity in Limpopo Province varies over space and time, with high transmission occurring in a few hot spots. The coefficient of determination (R^2) for each predicted malaria incidence map revealed that the predicted values were found to be accurate for all the months except June. Loha and Lindtjørn (2010) conducted a study where they also used the coefficient of determination to assess the accuracy of model variations of predictive incidence of *Plasmodium falciparum*. For this study, the coefficient was used to assess the predictive performance of malaria incidence maps.

The predicted maps illustrated that the places which were predicted to have high risk of malaria in 2010 were in two district municipalities which are Vhembe and Mopani. Previous study conducted by Gerritsen, et al. (2008) revealed that Vhembe and Mopani district municipalities have by far the highest incidence rate in Limpopo Province. However, this study went as far as sub-places which fall under five local municipalities of Vhembe and Mopani namely: Musina, Mutale, Thulamela, Great Giyani and Ba-Phalaborwa. The sub-places are presented in the Table 5.3. The possible explanation could be that they (municipalities) are located in areas of low altitudes and high temperatures, as they (altitude and temperature) are significant factors associated with malaria in Limpopo Province. A supporting study showed that high risk of malaria was predicted from south to east of Zambia which is characterized mainly by low altitude (Riedel, et al., 2010). In the same study area, Blumberg and Freaan (2007) revealed that malaria is endemic in the low-altitude areas. Additionally, the areas at risk are sharing the border with Zimbabwe and Mozambique. Both countries fall under the endemic zones in sub-Saharan Africa (World Health Organisation, 2012).

The quantification of relative amounts of spatial risk patterns has helped in highlighting sub-places with low and high proportions of malaria incidence at a given time period. There was less temporal variation revealed in the generated predicted incidence maps. The results showed that malaria is seasonal, with transmission occurring during warm and rainy summer months ranging from September to May. Furthermore, the peaks were observed during October. The same findings were reported by a study conducted in South Africa by Moonasar, et al. (2012). A corresponding study was done in Ethiopia where the authors stated that most malaria occurs from September to December with peaks during the month of October, immediately after the main rainy season (Yeshiwondim, et al., 2009). The seasonal epidemics are probably due to favourable environmental conditions in that region. At very high transmission levels, malaria risk is not very seasonal whereas at low transmission levels the risk is seasonal (Gemperli, et al., 2004).

Table 5.3: Sub-places predicted to have high malaria incidence in 2010 in their local municipalities

Ba-Phalaborwa Municipality	Great Giyani Municipality	Musina Municipality	Mutale Municipality	Thulamela Municipality
Seloane	Homo-North	Nance Field	Nwanedi	Makuleke
Mahale	Nkomo A		Nwanedi Farms	Masetoni
Humulani	Mapayeni		Sanari	
			Mutele A	
			Tshivaloni	
			Tshikondeni Mine	
			Tshikuyu	

5.3. Limitations of the Study

The first limitation was a high number of geographical locations as the study was based on the entire province. Limpopo Province has the highest transmission rate in South Africa which made it challenging for the model to run in WinBugs and/ OpenBugs. These two statistical programmes were not able to run the number of locations because they excluded the maximum limit.

Secondly, due to the large amount of case data, population and environmental data it was not easy for ArcGIS to handle large files during projection, reclassification and extraction of the data. This made the programme to continually crash because common windows were used for batch files.

Finally, the large number of sub-places included in our modelling approach challenged a geo-statistical model fitting, thus resulting in extremely slow Markov chain Monte Carlo (McMC) runs. OpenBUGS took a long period of time to produce results causing the analysis to take a lot more time to complete.

5.4. Conclusion

This chapter presented the environmental factors which have an influence in malaria transmission in the Limpopo Province. The predicted maps showing the regions with the high risk of malaria are presented at sub-place level and the predictive performance for the generated maps which was assessed is provided. The details were discussed in this chapter which was the most important aspect of the study.

CHAPTER SIX

CONCLUSION AND RECOMMENDATIONS

6.0. Introduction

This chapter provides an overall conclusion of all the chapters in this study. Thereafter, provide recommendations for future studies based on the study and other related studies. This may suggest how to improve the study because not all possible factors which may be significant to malaria transmission were evaluated in this study.

6.1. Conclusion

The study analysed the spatial and temporal patterns of malaria transmission for the period starting from May 2000 to June 2011 using Bayesian geo-statistical models in Limpopo Province, South Africa. The models were produced using environmental factors and two of the factors were selected as important factors of malaria transmission. The selected factors were altitude and the night temperature of two months before the case was reported. Subsequently, seasonality maps were predicted for 2010 giving the hot spots of malaria incidence in the province. The hot spots of malaria transmission were found to be Vhembe and Mopani district municipalities. To consider the predicted malaria incidence maps, assessment of the predictive performance was done. The majority of the predicted maps were assumed to be accurate and therefore may be useful in malaria control programmes in Limpopo Province.

Maps produced for malaria risks and transmission intensities are valuable tools to guide, monitor and evaluate effective application of control intervention programs (Griffin, et al., 2010; Gosoniu, et al., 2006). The produced malaria incidence maps can possibly be considered as one of the baselines in future. They can also be useful to tourists for locating places which are endemic to malaria and prevent transferring malaria to their home countries. The results highlighted the risk factors of malaria in Limpopo Province which are the most important characteristics of malaria transmission. Therefore, to reach the goals of malaria elimination, control programmes have to take into consideration the risk factors important for transmission in the hot spots and neighbouring regions. This may help to control malaria in areas which are exposed to transmission without wasting time and money controlling malaria in places which are malaria free.

6.2. Recommendations

Like most of the studies which have been conducted in Limpopo Province, the results of this study suggested that there are malaria hot spots in the study area. It is recommended that the results should be considered in present and future malaria control programmes. Subsequently, for future studies it can be more advantageous to also include factors like IRS, wind speed and direction for mapping malaria in Limpopo Province. Midega (2011) showed that the relationship between distance from a larval site and malaria risk is highly dependent on the wind direction and wind speed. Studies on the effect of wind direction and wind speed can possibly be useful in studying malaria transmission in South Africa. Finally, it is recommended that a different software programme be used for a similar study which can accommodate a large number of locations without crashing.

References

- Aarogya, 2012. *aarogya.com*. [Online], Available at: <https://www.aarogya.com/health-resources/health-programs/2634-national-malaria-control-programme-html.html?>, [Accessed 16 October 2012].
- Abellana, R., Ascaso, C., Aponte, J., Saute, F., Nhalungo, D., Nhacolo, A., Alonso, P., 2008. Spatio-seasonal modelling of the incidence rate of malaria in Mozambique. *Malaria Journal*, 8(228).
- Amek, N., Bayoh, N., Hamel, M., Lindblade, K. A., Gimnig, J., Laserson, K. F., Slutsker, L., Smith, T., Vounatsou, P., 2011. Spatio-temporal modeling of sparse geo-statistical malaria sporozoite rate data using a zero inflated binomial model. *Spatial and Spatio-temporal Epidemiology*, Volume 2, pp. 283-290.
- Amek, N., Bayoh, N., Hamel, M., Lindblade, K. A., Gimnig, J., Odhiambo, F., Laserson, K. F., Slutsker, L., Smith, T., Vounatsou, P., 2012. Spatial and Temporal dynamics of malaria transmission in rural western Kenya. *Parasites & Vectors*, 5(86).
- Basanez, M., Marshall, C., Carabin, H., Gyorkos, T., Joseph, L., 2004. Bayesian statistics for parasitologists. *Trends Parasitol*, Volume 20, pp. 85-91.
- Bateman, C., 2008. Malaria-Now Limpopo and Mpumalanga. *SAMJ*, 98(1).
- Bernadinelli, L. and Montomoli, C., 1992. Empirical Bayes versus fully Bayesian analysis of geographical variation in disease risk. *Stat Med*, Volume 11, pp. 983-1007.
- Bernardo, J. M., 2003. Probability and Statistics. In: R. Viertl, ed. *Bayesian statistics*. Oxford: Encyclopedia of Life Support Systems.
- Bhatt, B. M. and Joshi, J. P., 2012. GIS in Epidemiology: Applications and services. *National Journal of Community Medicine*, 3(2), pp. 259-263.
- Bjorkman, A. and Bhattarai, A., 2005. Public health impact of drug resistance Plasmodium falciparum malaria. *Acta Tropica*, Volume 94, pp. 163-169.
- Bloom, D. E., Bloom, L. R. and Weston, M., 2006. *Business and Malaria: A Neglected Threat?*, Geveva: World Economic Forum.

Blumberg, L. and Frean, J., 2007. Malaria control in South Africa - challenges and success. *SAMJ*, 97(11).

Blumberg, L. H., 2005. Severe Malaria. In: C. Feldman & G. A. Sarosi, eds. *Tropical and Parasitic Infectious in the Intense Care Unit. Perspectives on Critical Care Infectious Disease*. New York: Springer, pp. 1-16.

Bolstad, W. M., 2004. *Introduction to Bayesian Statistics*. Hoboken: John Wiley and Sons.

Bray, R. S. and Garnham, P. C. C., 1982. The Life-Cycle of Primate Malaria Parasites. *British Medical Bulletin*, 38(2), pp. 117-122.

Breman, J. G., Alilio, M. S., Mills, A., 2004. Conquering the intorelable burden of malaria: what's new, what's needed: a summary. *Am J Trop Med Hyg*, 71(2), pp. 1-15.

Butler, A. R., Khan, S., Ferguson, E., 2010. A brief history of malaria chemotherapy. *J R Coll Physicians Edinb*, Volume 40, pp. 172-177.

Capanna, E., 2006. Grassi versus Ross: who solved the riddle of malaria?. *International Microbiology*, Volume 9, pp. 69-74.

CDC, 2011. *Treatment of Malaria (guidelines for clinicians)*, Atlanta: Centers for Disease Control and Prevention.

Cerlini, P. B., Emanuel, K. A., Todini, E., 2005. Orographic effects on convective precipitation and space-time rainfall variability: preliminary results. *Hydrology and Earth System Sciences*, 9(4), pp. 285-299.

Cleland, J. and van Ginneken, J. K., 1989. Maternal education and child survival in developing countries: the search for pathways of influence. *Soc. Sci. Med.*, Volume 27, pp. 1357-1368.

Coleman, M., Coleman, M., Mabaso, M. L. H., Mabuza, A. M., Kok, G., Coetzee, M., Durrheim, David N., 2010. Household and microeconomic factors associated with malaria in Mpumalanga, South Africa. *Trans. R. Soc. Trop. Med. Hyg.*, Volume 104, pp. 143-147.

Coluzzi, M., 1988. Anopheline mosquitoes: genetic methods for species differentiation. In: W. H. Wernsdorfer & S. I. McGregor, eds. *Malaria*. London: Churchill Livingstone, pp. 411-430.

- Connor, S. J., Ceccato, P., Dinku, T., Omumbo, J., Grover-Kopec, E. K., Thomson, M. C., 2006. Using climate information for improved health in Africa: relevance, constraints and opportunities. *Geospatial Health*, Volume 1, pp. 17-31.
- Cox, F. E., 2010. History of the discovery of the parasites and their vectors. *Parasites and Vectors*, 3(5).
- Crans, W. J., 2004. A classification system for mosquito life cycle: life cycle types for mosquitos of the northern United States. *Journal of Vector Ecology*, 29(1).
- Cunha, C. B. and Cunha, B. A., 2008. Brief history of the clinical diagnosis of malaria: from Hippocrates to Olser. *J Vector Borne Dis*, Volume 45, pp. 194-199.
- DBSA, 1998. *Northern Province: Development Profile 1998*, Northern Province: Development Bank of South Africa .
- Dellaportas, P., Forster, J. J. and Ntzoufras, I., 2002. On Bayesian model and variable selection using MCMC. *Statistics and Computing*, Volume 12, pp. 27-36.
- Department of Health, 2011. *Guidelines for Prevention of Malaria in South Africa*, Cape Town: Communicable Disease Control.
- Eddy, S. R., 2004. What is Bayesian statistics. *Nature Biotechnology*, 22(9), pp. 1177-1178.
- Fujioka, H. and Aikawa, M., 2002. Structure and Life Cycle. *Malaria Parasites and Disease*, Volume 80, pp. 1-26.
- Gemperli, A., Vounatsou, P., Kleinschmidt, I., Bagayoko, M., Lengeler, C., Smith, T., 2004. Spatial Patterns of Infant Mortality in Mali: The Effect of Malaria Endemicity. *American Journal of Epidemiology*, 159(1), pp. 64-72.
- Gerritsen, A. M., Kruger, P., Schim van der Loeff, M. F., Grobusch, M. P., 2008. Malaria incidence in Limpopo Province, South Africa, 1998–2007. *Malaria Journal*, 7(162).
- Giardina, F., Gosoniu, L., Konate, L., Diouf, M. B., Perry, R., Gaye, O., Faye, O., Vounatsou, P., 2012. Estimating the Burden of Malaria in Senegal: Bayesian Zero-Inflated Binomial Geo-statistical Modeling of the MIS 2008 Data. *PLoS ONE*, 7(3).
- Gillies, M. T., 1967. Experimental of host selection in the *Anopheles gambiae* complex. *Ann. Trop. Med. Parasitol.*, Volume 61, pp. 68-75.

- Gosoni, L., Vounatsou, P., Sogoba, N., Smith, T., 2006. Bayesian modelling of geostatistical malaria risk data. *Geospatial Health*, Volume 1, pp. 127-139.
- Graham, A. J., Atkinson, P. M., Danson, F. M., 2004. Spatial Analysis for Epidemiology. *Acta Tropica*, Volume 91, pp. 219-225.
- Griffin, J. T., Hollingsworth, T. D., Okell, L. C., Churcher, T. S., White, M., Hinsley, W., Bousema, T., Brakeley, C. J., Ferguson, N. M., Basanez, M. G., Ghani, A. C., 2010. Reducing Plasmodium falciparum Malaria Transmission in Africa: A Model-Based Evaluation of Intervention Strategies. *PLoS Med*, 7(8).
- Grover-Kopec, E. K., Blumenthal, M. B., Ceccato, P., Dinku, T., Omumbo, J. A., Connor, S. J., 2006. Web-based climate information resources for malaria control in Africa. *Malaria Journal*, 5(38).
- Huang, Y., Dominic, F., Bell, M. L., 2005. Bayesian hierarchical distributed lag models for summer ozone exposure and cardio-respiratory mortality. *Environmetrics*, Volume 16, pp. 547-562.
- Hay, S. I., Omumbo, J. A., Craig, M. H., 2002. Clinical Epidemiology of malaria in the highlands of Western Kenya. 8(6), pp. 543-548.
- Hay, S. I., Omumbo, J. A., Craig, M. H., 2000. Earth observation, geographic information systems and Plasmodium falciparum malaria in sub-Saharan Africa. *Advances in Parasitology*, Volume 47, pp. 173-215.
- Hay, S. I. and Snow, R. W., 2006. The Malaria Atlas Project: Developing Global Maps of Malaria Risk. *PLoS Med*, 3(12), p. 473.
- Hoffman, S. L., 1996. *Malaria Vaccine Development*. Washington: ASM Press.
- Hui, F., Xu, B., Chen, Z., Cheng, X., Liang, L., Huang, H., Fang, L., Yang, H., Zhou, H., Yang, H., Zhou, X., Cao, W., Gong, Peng., 2009. Spatio-Temporal Distribution of Malaria in Yunnan Province, China. *American Journal of Tropical Medicine and Hygiene*, 81(3), p. 503-509.
- Hume, A., 2010. *Physiopedia*. [Online], Available at: <http://www.physiopedia.com/File:Cycle2.gif>, [Accessed 24 October 2012].

- Jarch, S., 1993. *Quinine predecessor: Francesco Torti and the early history of Cinchona*. Baltimore: Johns Hopkins University Press.
- Johnsen, M. M. and Renchie, D. L., 2007. Mosquito Life Cycle. *NM state*.
- Jupp, P. G., 2005. Mosquitoes as Vectors of Human Disease in South Africa. *South African Family Practice*, 47(9), pp. 68-72.
- Kasasa, S., Asoala, V., Gosoni, L., Anto, F., Adjuik, M., Tindana, C., Smith, T., Owusu-Agye, S., Vounatsou, P., 2013. Spatio-temporal malaria transmission patterns in Navrongo demographic surveillance site, northern Ghana. *Malaria Journal*, 12(63).
- Kazembe, L. N., Appleton, C. C., Kleinschmidt, I., 2007. Spatial analysis of the relationship between early childhood mortality and endemicity in Malawi. *Geospatial Health*, 2(1), pp. 41-50.
- Khan, S. M. and Waters, A. P., 2004. Malaria parasite transmission stages: an update. *TRENDS in Parasitology*, 20(12), pp. 576-580.
- Khosa, E., Kuonza, L. R., Kruger, P., Maimela, E., 2013. Towards the elimination of malaria in South Africa: a review of surveillance data in Mutale Municipality, Limpopo, 2005 to 2010. *Malaria Journal*, 12(7).
- Kistemann, T., Herbst, S., Dangendorf, F., Exner, M., 2001. GIS-based analysis of drinking water supply structures: a module for microbial risk assessment. *Int. J. Hyg. Environ. Health*, Volume 203, pp. 301-310.
- Kitron, U., 2000. Emerging Infectious Diseases. In: S. I. Hay, S. E. Randolph & D. J. Rogers, eds. *Remote Sensing and Geographic Information Systems in Epidemiology*. London: Academic Press, p. 357.
- Kiwanuka, G. N., 2003. Malaria morbidity and mortality in Uganda. *J Vect Borne Dis*, Volume 40, pp. 16-19.
- Kleinschmidt, I., Bagayoko, M., Clarke, G., Craig, M., Le Sueur, D., 2000. A spatial statistical approach to malaria mapping. *International Journal of Epidemiology*, 29(2), p. 355.

Kleinschmidt, I. and Sharp, B., 2001. Patterns in age-specific malaria incidence in a population exposed to low levels of malaria transmission intensity. *Tropical Medicine and International Health*, 6(12), pp. 986-991.

Kleinschmidt, I., Sharp, B., Mueller, I., Vounatsou, P., 2001. Rises in malaria incidence rates in South Africa: A small area spatial analysis of variation in time trends. *American Journal of Epidemiology*, 155(3).

Koram, K. A., Bennett, S., Adiamah, J. H., Greenwood, B. M., 1995. Socio-economic risk factors for malaria in a peri-urban area of The Gambia. *Trans. R. Soc. Trop. Med. Hyg.*, Volume 89, pp. 146-150.

Koutsos, A. C., Blass, C., Meister, S., Schmidt, S., MacCallum, R. M., Soares, M. B., Collins, F. H., Benes, V., Zdobnov, E., Kafatos, F. C., Christophides, G. K., 2007. Life cycle transcriptome of the malaria mosquito *Anopheles gambiae* and comparison with the fruitfly *Drosophila melanogaster*. *PNAS*.

Lasinio, G. J., Mastrantonio, G., Pollice, A., 2003. Discussing the “big n problem”. *Stat Methods Appl*, Volume 22, p. 97–112.

Limpopo Provincial Government, 2008. *Annual Performance Plan-Vote 7*, s.l.: Limpopo Department of Health and Social Development.

Limpopo Provincial Government, 2011. *Annual Performance Plan 2011/12-2013/14 : Vote 7*, s.l.: Department of Health.

Lindsay, S. W. and Martens, W. J., 1998. Malaria in the African highlands: past, present and future. *Bull World Health Organ*, Volume 76, pp. 33-45.

Link, W. A. and Barker, R. J., 2010. Closed-Population Mark-Recapture Models. In: Elsevier, ed. *Bayes factors and multimodel inference*. California: Elsevier.

Loha, E. and Lindtjorn, B., 2010. Model variations in predicting incidence of *Plasmodium falciparum* malaria using 1998-2007 morbidity and meteorological data from south Ethiopia. *Malaria Journal*, 9(166).

Mabaso, M. L. H., Kleinschmidt, I., Sharp, B., Smith, T., 2007. El Nino Southern Oscillation and annual malaria incidence in Southern Africa. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, Volume 101, pp. 326-330.

- Mabaso, M. L. H., Vounatsou, P., Midzi, S., Da Silva, J., Smith, T., 2006. Spatio-temporal analysis of the role of climate in inter-annual variation of malaria incidence in Zimbabwe. *International Journal of Health Geographics*, 5(20).
- Mabaso, M. L., Sharp, B., Lengeler, C., 2004. Historical review of malarial control in southern African with emphasis on the use of indoor residual house-spraying. *Tropical Medicine and International Health*, 9(8), pp. 846-856.
- Machault, V., Vignolles, C., Pages., 2010. Spatial heterogeneity and temporal evolution of malaria transmission risk in Dakar, Senegal, according to remotely sensed environmental data. *Malaria Journal*, 9(252).
- Maharaj, R., 2008. Controlling Malaria in Southern Africa. *CME*, 26(6), pp. 300-302.
- Maharaj, R., Morris, N., Seocharan, I., Kruger, P., Moonasar, D., Mabuza, A., Raswiswi, E., Raman, J., 2012. The feasibility of malaria elimination in South Africa. *Malaria Journal*, 11(423).
- Maharaj, R., Mthembu, D. J., Sharp, B. L., 2005. Impact of DDT re-introduction on malaria transmission in KwaZulu-Natal. *SAMJ*, Volume 95, pp. 871-874.
- Marsh, K. and Snow, R., 1999. Malaria Transmission and Morbidity. *Parassitologia*, 41(1-3), p. 241.
- Mathers, C. D., Fat, D. M., Inoue, M., Rao, C., Lopez, A. D., 2005. *Counting the dead and what they died from: an assessment of the global status of cause of death dat*, Geneva: World Health Organisation.
- Matteelli, A. and Castelli, F., 2007. *Life Cycle of Parasites*. Brescia: University of Brescia.
- Medical Research Council, 2001. Malaria Advice for Southern Mozambique, Swaziland and South Africa. *Business Fund*.
- Medicines for Malaria Venture, 2010. The Life Cycle of the Malaria Parasite. *Defending Malaria Together*, Volume 07.
- Mehta, U., Durrheim, D. N., Blumberg, L., Donohue, S., Hansford, F., Mabuza, A., Kruger, P., Gumede, J. K., Immelman, E., Sa'nchez Canal, A., Hugo, J. J., Swart, G., Barnes, K. I.,

2007. Malaria deaths as sentinel events to monitor healthcare delivery. *Tropical Medicine and International Health*, 12(5), p. 617–628.

Midega, J. T., Smith, D. L., Olotu, A., Mwangangi, J. M., Nzovu, J. G., Wambua, J., Nyangweso, G., Mbogo, C. M., Christophides, G. K., Marsh, K., Bejon, P., 2012. Wind direction and proximity to larval sites determines malaria risk in Kilifi District in Kenya. *Nature Communications*, 3(674).

Mironga, J. M., 2004. Geographic Information Systems and Remote Sensing in the Management of Shallow Tropical Lakes. *Applied Ecology and Environmental Research*, 2(1), pp. 83-103.

M'Marete, C. K., 2006. Climate and Water Resources In The Limpopo Province. In: D. o. A. a. R. Engineering, ed. *HWR 3541 Hydrologic Measurements*. Limpopo: s.n.

Moonasar, D., Nuthulaganti, T., Kruger, P. S., Mabuza, A., Rasiswi, E. S., Benson, F. G., Maharaj, R., 2012. Malaria Control in South Africa 2000-2010: beyond MDG6. *Malaria Journal*, 11(294).

Muir, D. A., 1988. Anopheline mosquitoes: vector reproduction, life-cycle and biotope. In: W. H. Wernsdorfer & S. I. McGregor, eds. *Malaria*. London: Churchill Livingstone, pp. 431-451.

Murphy, S. C. and Breman, J. G., 2001. Gaps in the childhood malaria burden in Africa: cerebral malaria, neurological sequelae, anaemia, respiratory distress, hypoglycemia, and complications of pregnancy. *Am. J. Trop. Med. Hyg.*, 64(1,2), pp. 57-67.

Mzezewa, J., Misi, T., van Rensburg, L. D., 2010. Characterisation of rainfall at a semi-arid ecotope in the Limpopo Province (South Africa) and its implications for sustainable crop production. *Water SA*, 36(1).

NASA, 2000. *Modis Website*. [Online], Available at: https://lpdaac.usgs.gov/lpdaac/products/modis_products_table, [Accessed 18 July 2013].

NASA, 2004. Terra and Aqua: new data for epidemiology and public health. *International Journal of Applied Earth Observation and Geoinformation*, Volume 6, pp. 33-46.

Newton, C. J., Taylor, T. E., Whitten, R. O., 1998. Pathophysiology of fatal falciparum malaria in African children. *Am J Trop Med Hyg*, Volume 58, pp. 673-683.

- Ngomane, L. and de Jager, C., 2012. Changes in malaria morbidity and mortality in Mpumalanga Province, South Africa (2001-2009): a retrospective study. *Malaria Journal*, 11(19).
- NIMR and NVBDCP, 2009. *Guidelines for Diagnosis and Treatment of Malaria in India*, New Delhi: National Institute of Malaria Research.
- Nkurunziza, H., Gebhardt, A., Pilz, J., 2010. Bayesian modelling of the effect of climate on malaria in Burundi. *Malaria Journal*, 9(114).
- Nuckols, J. R., Ward, M. H., Jarup, L., 2004. Using GIS for Exposure Assessment in Environmental Epidemiology Studies. *Environmental Health Perspectives*, 112(9), pp. 1007-1015.
- Odhiambo, J. J. O., 2006. Physical Features, Soils and Land-Utilization. In: D. o. S. Science, ed. *HWR 3541 Hydrological Measurements*. Limpopo: s.n.
- O'Meara, W. P., Mangeni, J. N., Steketee, R., Greenwood, B., 2010. Changes in the burden of malaria in sub-Saharan Africa. *Lancet Infect Dis*, Volume 10, pp. 545-555.
- Patra, P., 2002. Remote Sensing and Geographic Information System. *The Association for Geographical Studies*, Volume 3.
- Protopopoff, N., Van Bortel, W., Marcotty, T., 2008. Spatial targeted vector control is able to reduce malaria prevalence in the highlands of Burundi. *The American Society of Tropical Medicine and Hygiene*, 78(1), pp. 12-18.
- Rai, P. K., Nathawat, M. S., Mishra, A., 2011. Role of GIS and GPS in Vector Born Disease Mapping: A Case Study. *Journal of GIS Trends*, 2(1), pp. 20-27.
- Rajaratnam, J. K., Marcus, J. R., Flaxman, A. D., Wang, H., Levin-Rector, A., Dwyer, L., Costa, M., Lopez, A. D., Murray, C. J. L., 2010. Neonatal, postneonatal, childhood, and under-5 mortality for 187 countries, 1970-2010: a systematic analysis of progress towards Millennium Development Goal 4. *The Lancet*, 375(9730), pp. 1988-2008.
- Reyers, B., Wessels, K. J., van Jaarsveld, A. S., 2002. An assessment of biodiversity surrogacy options in the Limpopo Province of South Africa. *African Zoology*, 37(2), pp. 185-195.

Riedel, N., 2009. *Manual on how to retrieve, convert and extract Remote Sensing (RS) data*, Basel: s.n.

Riedel, N., Vounatsou, P., Miller, J. M., Gosoni, L., Chizema-Kawesha, E., Mukonka, V., Steketee, R. W., 2010. Geographical patterns and predictors of malaria risk in Zambia: Bayesian geostatistical modelling of the 2006 Zambia national malaria indicator survey (ZMIS). *Malaria Journal*, 9(37).

Rios, L. and Connelly, R., 2009. Common malaria Mosquito *Anopheles quadrimaculatus* Say (Insecta: Diptera: Culicidae). *IFAS Extension*.

Roca-Felter, A., Schellenberg, J. A., Smith, L., Carneiro, I., 2009. A Simple Method of Defining Malaria Seasonality. *Malaria Journal*, 8(276).

Rowe, A. K., Rowe, S. Y., Korenromp, 2006. The burden of malaria mortality among African children in the year 2000. *International Journal of Epidemiology*, 35(3), p. 691-704.

Rozendaal, J. A., 1997. *Vector Control-Methods for use by Individuals and Communities*, Geneva: World Health Organisation.

Saxena, R., Nagpal, B. N., Srivastava, A., Gupta, S. K., Dash, A. P., 2009. Application of spatial technology in malaria research and control: some new sights. *Indian J Med Res*, Volume 130, pp. 125-132.

Shtatland, E. S., Moore, S., Barton, M. B., 2000. Why We Need an R² Measure of Fit (and not only one) In Proc Logistic and Proc Genmod. *Statistics and Data Analysis*, 256(25).

Simoonga, C., Utzinger, J., Brooker, S., Vounatsou, P., Appleton, C. C., Stensgaard, A. S., Olsen, A., Kristensen, T. K., 2009. Remote Sensing, geographic information system and spatial analysis for schistosomiasis epidemiology and ecology in Africa. *Parasitology*, Volume 136, pp. 1683-1693.

Smith, K. R., Corvalan, C. F., Kjellstrom, T., 1999. How much global ill health is attributable to environmental factors. *Epidemiology*, Volume 10, pp. 573-84.

Smith, T. A., Leuenberger, R., Langelier, C., 2001. Child mortality and transmission intensity in Africa. *Trends in Parasitology*, 17(3), pp. 145-149.

- Snow, R. W., Craig, M. H., Deichmann, U., le Sueur, D., 1999. A Preliminary Continental Risk Map for Malaria Mortality among African Children. *Parasitology Today*, 15(3).
- Snow, R. W. and Marsh, K., 2002. The consequences of reducing transmission of Plasmodium falciparum in Africa. *Adv. Parasitol*, Volume 52, pp. 235-264.
- South African National Department of Health, 2007. *Prevalance and Distribution of Malaria in South Africa*, Pretoria: National Malaria Control Policy.
- Statistics South Africa, 2012. *Census 2011*, Pretoria: Statistics South Africa.
- Sterne, J. A. and Tilling, K., 2002. G-estimation of causal effects, allowing for time-varying confounding. *The Stata Journal*, 2(2), pp. 164-182.
- Stresman, G. H., 2010. Beyond temepature and precipitation: Ecological risk factors that modify malaria transmission. *Acta Tropica*, Volume 116, pp. 167-172.
- Szabo, S., Kenyeres, Z., Bauer, N., Gosztonyi, G., 2008. Mapping of Mosquito (Culicidae) Breeding Sites Using Predictive Geographic Information Methods. *Dissertations Commission of Cultural Landscape*, Volume 8.
- Tatem, A. J., Goetz, S. J., Hay, S. I., 2004. Terra and Aqua: new data for epidemiology and public health. *International Journal of Applied Earth Observation and Geoinformation*, Volume 6, pp. 33-46.
- Tesfamichael, S. G., van Aardt, J. A. N., Ahmed, F., 2010. Estimating plot-level tree height and volume of Eucalyptus grandis plantations using small-footprint, discrete return lidar data. *Progress in Physical Geography*, 1(26).
- Thomas, C. J. and Lindsay, S. W., 2000. Local-scale variation in malaria infection amongst rural Gambian children estimated by satellite remote sensing. *TRANSACTIONS OF THE ROYAL SOCIETY OF TROPICAL MEDICINE AND HYGIENE*, Volume 94, pp. 159-163.
- Thomas, R., 2006. Crop Production in The Limpopo Province. In: D. o. P. Production, ed. *Review of Crop Production in The Limpopo Province*. Limpopo: s.n.
- Tren, R. and Bate, R., 2004. *South Africa's War against Malaria-Policy Analysis*, Washington: Cato Institute.

Tshiala, M. F., Olwoch, J. M. and Engelbrecht, F. A., 2011. Analysis of Temperature Trends over Limpopo Province, South Africa. *Journal of Geography and Geology*, 3(1).

U.S Department of Health and Human Services, 2007. *Understanding Malaria: Fighting an Ancient Scourge*. 7 ed. Bethesda: NIAID Science Education.

Van Buren, W. H., 1864. *Quinine as a Prophylastic against Malarious Diseases*. Philadelphia: J.B. Lippincott and Co.

Van Dyk, J. C., Bouwman, H., Barnhoorn, I. E. J., Bornman, M. S., 2010. DDT contamination from indoor residual spraying for malaria control. *Science of the Total Environment*, p. 8.

Whitfield, D., 1984. Two cases of falciparum acquired in Britain. *BMJ*, Volume 289, pp. 1607-1609.

WHO & CDC, 2005. *Malaria control in complex emergencies: agency field handbook*. Geneva: World Health Organisation.

Winch, P. J., Makemba, A. M., Kamazima, S. R., 1996. Local terminology for febril illnesses in Bagamoyo district, Tanzania, and its impact on the design of a community-based malaria control programme. *Social Science and Medicine*, Volume 42, pp. 1057-1067.

World Health Organisation, 1993. World Malaria Situation in 1991. *Weekly Epidemiol. Rec.*, Volume 68, pp. 245-252.

World Health Organisation, 2006. *Malaria kills up to 3 million people each year*, s.l.: Disease Control Priorities Project.

World Health Organisation, 2008. *Key Facts, Figures and Strategies: The Global Malaria Action Plan*, Geneva: The Roll Back Malaria Partnership.

World Health Organisation, 2009. *Malaria Case Management: Operations Manual*. Geneva: WHO Library Cataloguing.

World Health Organisation, 2010. *World Malaria Report*, Geneva: WHO Global Malaria Programme.

World Health Organisation, 2011. *World Malaria Report 2011*, Geneva: WHO Global Malaria Programme.

World Health Organisation, 2012. *World Malaria Report*, Geneva: World Health Organisation.

Worrall, E., Basu, S., Hanson, K., 2005. Is malaria a disease of poverty? A review of the literature. *Tropical Medicine and International Health*, 10(10), pp. 1047-1059.

Yeshiwondim, A. K., Gopal, S., Hailemariam, A. T., Dengela, D. O., Patel, H. P., 2009. Spatial analysis of malaria incidence at the village level in areas with unstable transmission in Ethiopia. *International Journal of Health Geographics*, 8(5).

Zacarias, O. P. and Andersson, M., 2010. Mapping malaria incidence distribution that accounts for environmental factors in Maputo Province-Mozambique. *Malaria Journal*, Volume 9, p. 79.

Zeng, L., Benedict, M. Q., Cornel, A. J., Collins, F. H., Kafatos, F. C., 1996. An Integrated Genetic Map of the African Human Malaria Vector Mosquito, *Anopheles gambiae*. *Genetics Society of America*, Volume 143, pp. 941-952.

Zhou, G., Minakawa, N., Githeko, A., Yan, G., 2004. Spatial Distribution Patterns of Malaria Vectors and Sample Size Determination in Spatially Heterogeneous Environments: A Case Study in the West Kenyan Highland. *Journal of Medical Entomology*, 41(6), pp. 1001-1009.

Zimbabwe, G. O., 2011. *Zimbabwe Malaria Programme Performance Review Aide Memoire*. s.l.: Ministry of Health and Child Welfare.

Appendix

1. Variable Selection

```
model{
for (i in 1:N){
  cases[i] ~ dnegbin(p[i],r)
  p[i] <- r/(r+mu[i])
  log(mu[i]) <- log(population[i])+inprod(b[],X[i,])+e[idtime[i]]+w[idloc[i]]
}

r ~ dgamma(0.01,0.01)

for (i in 1:3){
  b[i]~ dnorm(0.0,0.01)
}

for (i in 1:2){
  Ind[i]~dbern(0.5)
  b[i+3]~ dnorm(0.0,varr[i])
  varr[i]<-pow(1 0000,1-Ind[i])*0.3
}

Ind1~dbern(0.5)
var1<-pow(1 0000,1-Ind1)*0.3
for(i in 1:3){
  b[i+5] ~ dnorm(0.0, var1)
}

Indrain ~ dcat(p1[])
for(j in 1:5) {
```

```

b[j+8]~dnorm(0,tr[j])
tr[j]<-pow(1 0000,1-equals(Indrain,j))*0.3
}

Indndvi ~ dcat(p1[])
for(j in 1:5) {
  b[j+13]~dnorm(0,tn[j])
  tn[j]<-pow(1 0000,1-equals(Indndvi,j))*0.3
}

IndTempD~ dcat(p1[])
for(j in 1:5) {
  b[j+18]~dnorm(0,td[j])
  td[j]<-pow(1 0000,1-equals(IndTempD,j))*0.3
}

IndTempN ~ dcat(p1[])
for(j in 1:5) {
  b[j+23]~dnorm(0,ttn[j])
  ttn[j]<-pow(1 0000,1-equals(IndTempN,j))*0.3
}

# AR(1) prior distribution for temporal random effects:

e[1] ~ dnorm(0.0, tau2)
  for (t in 2:T){
    emean[t-1]<-rho*e[t-1];
    e[t] ~dnorm(emean[t-1], tau.e)
  }

tau2<-(1-pow(rho,2))*tau.e
rho ~ dunif(0, 1)
tau.e ~ dgamma(1,1 )
sigma.e<- 1/sqrt(tau.e )

```

```

# Gaussian process for spatial random effects:
for (i in 1:Nloc) {
  mu1[i]<-0}

w[1:Nloc]~spatial.exp(mu1[], lon[], lat[], tau.sp, phi, 1)
tau.sp~dgamma(1,1)
sigma<-1/tau.sp
phi~dunif(1.56,3256)
rhoinv<-1/phi
Range<-3/phi

}

```

2. Model Fitting

```

model{
  for (i in 1:N){
    cases[i] ~ dnegbin(p[i],r)
    p[i] <- r/(r+mu[i])
    log(mu[i]) <- log(population[i])+inprod(b[],X[i,])+e[idtime[i]]+w[idloc[i]]
  }

  r ~ dgamma(0.01,0.01)

  for (i in 1:5){
    #####cont,sine,cosine, altitude, TempN_lag2#####
    b[i]~ dnorm(0.0,0.01)
  }

  # AR(1) prior distribution for temporal random effects:

  e[1] ~ dnorm(0.0, tau2)

```



```

for (t in 2:T){
    emean[t-1]<-rho*e[t-1];
    e[t] ~dnorm(emean[t-1], tau.e)
}

tau2<-(1-pow(rho,2))*tau.e
rho ~ dunif(0, 1)
tau.e ~ dgamma(1,1 )
sigma.e<- 1/sqrt(tau.e )

# Gaussian process for spatial random effects:
for (i in 1:Nloc) {
    mu1[i]<-0}

w[1:Nloc]~spatial.exp(mu1[], lon[], lat[], tau.sp, phi, 1)
tau.sp~dgamma(1,1)
sigma<-1/tau.sp
phi~dunif(1.56,3256)
rhoinv<-1/phi
Range<-3/phi

}

```