Spatio-Temporal Modelling and Mapping of Malaria in Angola

A dissertation submitted to

The University of KwaZulu-Natal

in fulfillment of the requirements for the degree

of

Masters of Science in Statistics

by

Artemisa Aminata Soares Lima

School of Mathematics, Statistics and Computer Science

March 2017

Declaration

I, Artemisa Lima, declare that this dissertation titled, 'Spatio-Temporal Modelling and Mapping of Malaria in Angola' and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- No part of this dissertation has previously been submitted for a degree or any other qualification at this University or any other institution.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this dissertation is entirely my own work.
- All references that were used are duly acknowledged in the Bibliography.

_____ _____

Artemisa Lima (Student) Date

_____ _____

Prof. Henry Mwambi (Supervisor) Date

_____ _____

Dr.Thomas Achia (Co-supervisor) Date

Abstract

About half of the world's population is at risk of contracting malaria. The growing number of malaria cases and deaths due to this disease in Africa has become a major challenge to public health care sector. Malaria is reported to be the primary cause of mortality in Angola, hence major focus needs to be put on intervention and prevention methods that reduce the disease risk and mortality to a level which is no longer a public risk problem. The common risk factors for malaria can be linked to environmental, socio-economic, demographic and climatic factors.

The mortality rate due to malaria in Angola is analyzed using spatial disease mapping models. Such models are widely used to study the disease incidence, spatial distribution of diseases, prediction of the disease outcome and also to inform intervention strategies in various regions across the world. The methodology used is based on the Bayesian hierarchical modelling (BHM) framework. Four models namely the Poisson-gamma model, Poisson log-normal model, conditional autoregressive (CAR) model as well as the convolution model were used to study the relative risk of malaria mortality at provincial level in Angola using the National malaria control programme data from the period of 2003 to 2010. The Deviance Information Criteria selection was applied to compare and select the best fitted model.

A total of $109,320$ deaths due to malaria were observed during the period of 2003-2010 in Angola. The lowest crude death rate was estimated as 124.14 per 100,000 in Lunda Sul province and the highest was 1583.63 per 100,000 in Luanda province. The results revealed that when comparing the four fitted models, the convolution model when we fitted to the data with both spatial structured and unstructured random effects performed better than the other three models. The structured and unstructured random effects were used to capture variation of risk specific to a province and across provinces respectively. The risk maps revealed variation of risk among provinces with very high relative risks in the South-East parts of Angola.

A full Bayesian approach was also applied to perform a spatial and spatio-temporal modeling of malaria prevalence in Angola among children under 5 years using the 2006-2007 and 2011 Angola malaria indicator survey (AMIS) data. The Bayesian logistic model was applied in the spatio-temporal analysis to investigate the relationship between malaria prevalence and some reported socio-economic and demographic factors for data collected over the years 2006-2007 and 2011. The space-time effect of the association between malaria and these factors has practical implications for informing strategies for malaria control. Other than temporal variation, the risks factors were also found to vary spatially.

The study found that there was a significant difference in the effects of socio-economic and demographic variations on malaria between these two time periods. Wealth has a negative relationship with malaria prevalence while age was found to have a positive linear relationship with malaria prevalence which is indicative that this covariate play an important role in contracting malaria. Children living in urban areas and those who had bed nets were less likely to contract malaria as compared to those who lived in rural areas and those who did not have bed nets. The temporal analysis show that the prevalence for malaria was lower in 2011 as compared to the 2006-2007 period.

Acknowledgements

I reserve all glory and honour to the Most High for seeing me through this research, despite all challenges. May your Name be glorified and may I ever be your student.

I wish to express my sincere thanks to my supervisor Professor Henry Mwambi, whose encouragement, diligence, expertise, understanding, patience and professionalism added considerable quality to my masters research project. Despite the challenges encountered along the way, thank you for your inspiration your valuable contributions which were so helpful and words cannot fully explain how grateful I am for your immense intellectual guidance and robust contribution towards the successful completion of this work. I would also like to thank my co–supervisor, Dr Thomas Achia for his unwavering support, motivation, mentorship and encouragement, for contributing positively towards my personal development over the years. It was great working with you, thank you.

Special thanks to my colleagues and friends, Mbali Mlangeni, Elphas Okango, and Irina Saraiva De Sousa. This project would not have been complete without your support, and selfless assistance. The time that you sacrificed to provide me with valuable information that contributed towards the completion of this dissertation is remarkably valued.

I am extremely indebted to my parents, Mr. Paulo Jorge Da Costa, Mrs. Sandra Da Costa and Mr. Luis Cruz Lima and to my siblings Denise Lima, Cleusa Da Costa, Jalane Ribas and my aunt Neusa Faria for their patience, love and unreserved support. I am also eager to express my indebtedness to my fiancé Siphetfo Dlamini and my son Harriston Jorge Dlamini for their love, encouragement, patience and push for tenacity.

Contents

List of Figures

List of Tables

Abbreviations

AMIS	**A**ngola **M**alaria **I**ndicator **S**urvey
BHM	**B**ayesian **H**ierarchical **M**odel
BYM	**B**esag, **Y**ork and **M**ollie
CAR	**C**onditional **A**utoregressive
CI	**C**redible **I**nterval
DHS	**D**emographic **H**ealth **S**urvey
DIC	**D**eviance **I**nformation **C**riteria
INLA	**I**ntegrated **N**ested **L**aplace **A**pproximation
ITN	**I**nsecticide **T**reated **N**et
MCMC	**M**arkov **C**hain **M**onte **C**arlo
NMCP	**N**ational **M**alaria **C**ontrol **P**rogramme
PMI	**P**resident's **M**alaria **I**nitiative
SMR	**S**andardized **M**ortality **R**atio
RR	**R**elative **R**isk
WHO	**W**orld **H**ealth **O**rganization

This piece of work is dedicated to my late Aunt Ana Carla Faria Soares Da Silva 1968-2006.
As with flowers, so with men
They blossom, bloom and wither away
But there are some who always
Leave a fragrance behind
In them you belong. -Japheth
...

Chapter 1

Introduction

1.1 Background to Malaria

The risk of disease is important data for policy makers, health authorities and governments. However, the risks vary across geographic regions, so for effective planning, it is important to analyze these variations in order to identify the possible reasons underlying the differences.

Every year, over 14 million people die globally as a result of communicable diseases (WHO, 2013). There is an apparent disparity in the communicable disease profile between developed and developing countries; with developing countries experiencing higher mortality rates due to such diseases than do developed countries. According to the Center for Disease Control, communicable diseases are illnesses caused by an infectious agent, or its toxins. Generally an infectious disease may

be transmitted to a susceptible animal or human directly or indirectly by the infectious agent, or via a disease vector, or the inanimate environment.

Research shows that with such a variety of communicable diseases, the factors for their transmission may be unpredictable and difficult to understand. For instance the nature of the disease dynamics may be either emerging or re-emerging. This research focuses on malaria, and in particular on developing spatial and spatio-temporal models for malaria in Angola.

The word malaria originates from the old Italian words, *Mala'ria*, which means "bad air". This idea originated about 2000 years ago; from the ancient Romans who believed that the disease was caused by polluted air seeping out of fermenting marshes. Modern knowledge indicates that malaria is an infectious disease caused by a parasite of the genus *Plasmodium*, which is transmitted to humans through the bite of infected female mosquitoes belonging to the genus *Anopheles*. Although over 400 different species of *Anopheles* mosquito exist, there are only four species of it known to cause human malaria: *Plasmodium falciparum*, *Plasmodium vivax*, *Plasmodium ovale*, and *Plasmodium malariae* (WHO, 2011). The symptoms of malaria usually show after 10 to 15 days after a bite by an infected mosquito; which include fever, headache, and vomiting. With adequate treatment, malaria is curable.

As with most communicable diseases, malaria is most prevalent in developing countries; that is in the tropical and subtropical regions of Central and South America, Africa, Asia and Oceania. However, specific species of the malaria parasite are endemic to different geographical regions, depending on climate. In this regard,

according to the Angola Malaria Indicator Survey of 2006-2007 AMIS (2007), *Plasmodium falciparum* is that country's dominant species, being responsible for more than 90% of the malaria infections, whereas *Plasmodium vivax* is responsible for less than 7% of infections.

Globally, malaria has become a serious concern affecting the health and wealth of individuals. The World Health Organization has rated Malaria as the fifth worldwide leading cause of death (WHO, 2010). The mortality rate for children under five in Angola is one of the highest in the world reported by President's Malaria Initiative (PMI, 2010). In Africa, there are about 300 to 500 million new cases of malaria per year, and it is estimated to cause between 1 and 3 million deaths per year (Yeshiwondim et al., 2009). One of the targets of the World Health Organization's millennium development goals for 2010 was that by 2015 child mortality due to malaria should have been reduced by two-thirds (WHO, 2010). Consequently, the public health sector in the World Health Organization has prioritized intervention in the global malaria epidemic, which is particularly important in developing countries, where resources are scarce.

Despite its epidemic status, malaria can be prevented. For instance the use of insecticide treated nets (ITNs) has proved to be effective in reducing the burden of malaria. Over 83 countries globally have adopted the use of ITNs for malaria control; 39 of these countries being in Africa. Furthermore, in 2010 it was estimated that 42% of households in Africa owned or had at least used ITNs, and that 35% of children under five had slept under ITNs (WHO, 2011).

As another preventive measure for malaria, the World Health Organization also supports using indoor residual spraying with insecticides; the most common being dichlorodiphenyltrichloroethane (usually known as DDT). In sub-Saharan Africa, the number of people using indoor residual sprays increased from 13 million in 2005 to 75 million in 2009 (DOH, 2003).

The Angolan government has instituted specific strategies to combat malaria. These include indoor residual spraying of dwellings, promotion of insecticide-treated bed nets as well as intermittent preventive treatment, which is specially designed to prevent women from contracting malaria during pregnancy (PMI, 2010).

With such concerted efforts, malaria still remains a serious public health concern and impacts negatively on the socio-economies of developing countries (WHO, 2010). This is compounded by the endemicity of this disease being highly concentrated among those with poorly developed immunity, particularly young children (Marcus, 2009). Angola is no exception.

1.2 Malaria in Angola

As with other developing countries in the tropics, malaria is a major public health concern in Angola, where it is a primary cause of mortality. The estimate of 3.2 million malaria related cases reported in 2004 had increased to approximately 3.7 million cases by 2010, out of which two-thirds concerned children under five years of age (AMIS, 2007, 2012). The disease accounts for an estimated 35% of deaths in children under five, 60% of hospital admissions in children under five and 10% of hospital admissions among pregnant woman (Ruebush et al., 2005). Nevertheless, in Angola, all the population is at risk of malaria. Angola experienced a three-decade long civil war (1985-2002), which crippled the health infrastructure. Specifically, it is estimated that 80% of the health facilities were destroyed during the war and only about 30% of the population have access to the government health facilities.

To counter the malaria epidemic, in 2005, Angola started an initiative, known as the President's Malaria Initiative (PMI). Part of the initiative was to implement a national survey to provide estimates of malaria, such as prevalence, mortality and relative risk of the disease in the country. In 2008, the PMI team divided the country into three regions based on the levels of malaria endemicity. These regions were classified as hyper-endemic, meso-endemic (stable) and meso-endemic (unstable). These regions are illustrated in Figure 1.1 below, where the darker red shading indicates greater instability of the malaria endemicity (AMIS, 2007).

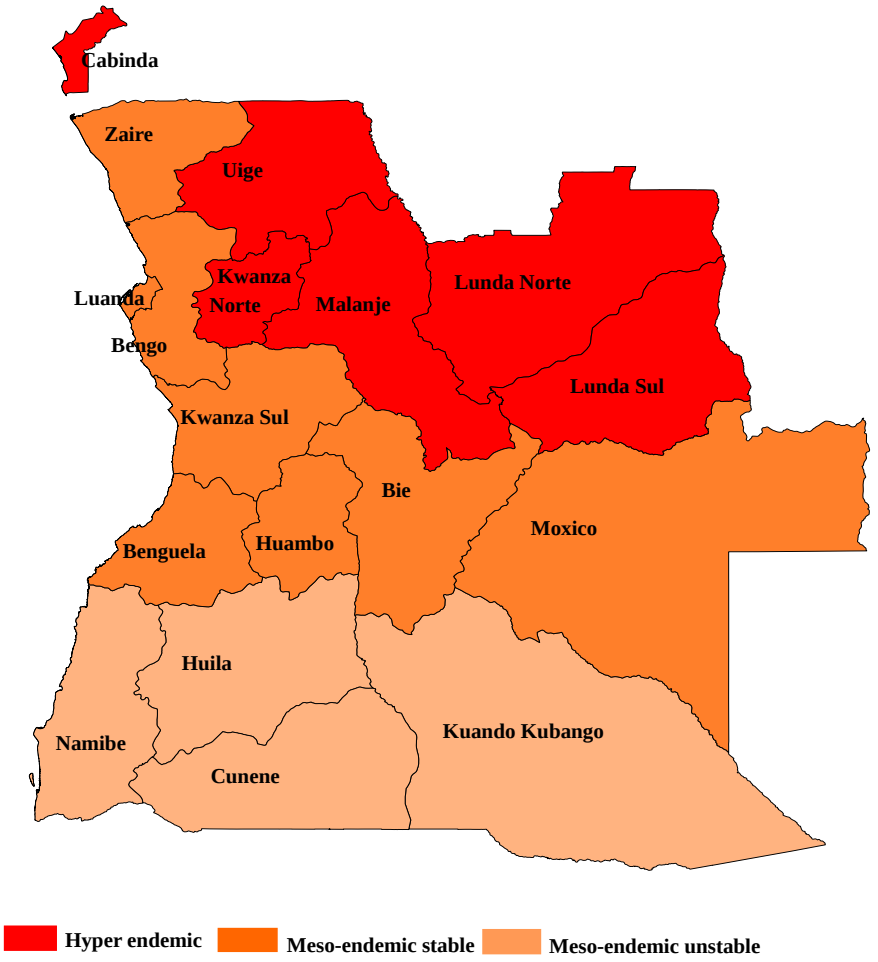

FIGURE 1.1: The geographical representation of malaria endemecity in Angola

.

The hyper-endemic region comprises six provinces in the north-easterly part of the country, where about 28% of the population was at risk. In this region the disease can be transmitted all year round, with highest transmission rates occurring between November and January. The meso-endemic (stable) region covers eight provinces in central Angola. The highest rate of transmission in this region occurs between November and May, where slightly over half of the population (55%) is at risk, while the period of lowest transmission is in July to October. The meso-endemic (unstable) region covers four southerly provinces in Angola, with 17%

of the population being at risk, and where the lowest seasonal transmission rate occurs between May and December.

1.3 Literature Review

As has been shown above, malaria has become a serious global concern. It particularly affects the health and wealth in developing countries. The morbidity and mortality risk associated with malaria are characterized by different factors, leading to spatial and spatio-temporal variation (Yeshiwondim et al., 2009).

1.3.1 Determinants of malaria transmission

Malaria is the most prevalent parasitic disease in humans; transmitted by mosquitoes and influenced by meteorological conditions and the socio-economic status of the population. A number of studies have investigated the risk factors of malaria transmission. Notably, climatic factors, such as temperature, rainfall and altitude have been identified as the most important risk factors for malaria transmission (Montosi et al., 2012; Tanser et al., 2007).

Temperature affects malaria transmission because higher temperatures, that is over 30 °C, allow for faster growth and propogation of Anopheles mosquitos, thereby increasing malaria transmission rates. By contrast, at temperatures below 30 °C the mosquito life cycle is interrupted, therefore reducing the transmission rate. Because lower temperatures usually occur at higher altitudes, malaria transmission becomes less frequent with increasing altitudes.

Rainfall and humidity also play a significant role in influencing malaria transmission. In the immature larval stage, mosquitoes can only survive in water. Rainfall leaves pools of still water, which provide excellent mosquito breeding sites. Nevertheless, heavy rainfall may destroy these ideal breading places and stop mosquito eggs developing into larvae. Thus, mosquitoes are also usually found in areas with annual average rainfall between 1100 mm and 7400 mm and where the humidity is above 60% (Snow et al., 1999).

Temperature and rainfall are thus the most important climatic factors for malaria transmission (Freeman and Bradley, 1996; Tanser et al., 2003). In Zimbabwe, Freeman and Bradley (1996) found that temperature has a strong effect on the severity of malaria incidence and so, for many years, has been a predictor of severe malaria outbreaks. Similarly, a study in Ethiopia by Tulu (1996) shows a strong positive correlation between minimum temperatures and monthly malaria incidence. Mabaso et al. (2006) confirmed that temperature, rainfall and vapour pressure are strong positive predictors for malaria incidence in Zimbabwe.

Besides climatic influences on malaria transmission, there are other risk factors such as human activities, vector control programmes, socio-economic status of the vulnerable population, environmental changes, population movement, urbanization and restricted access to health service that also affects the spread of this disease. These risk factors, some of which are discussed below, may vary across space and time.

1.3.2 Vector control

As has been noted above, malaria is not transmitted directly, but by means of the *Anopheles* mosquito. Vector control has benefited many Southern African countries, and is currently the most effective tool for preventing or controlling malaria transmission. Two primary vector control measures that have been recommended by the World Health Organization are the use of insecticide-treated nets (ITNs) and indoor residual spraying (Mabaso et al., 2004).

Several studies have confirmed that the use of ITNs is the most effective way of preventing and reducing malaria mortality among children under age of five (Alonso et al., 1991; Nevill et al., 1996). In an area of high malaria transmission, Lengeler (2004) has suggested that regular use of ITNs can reduce the malaria mortality among children under 5 years of age by as much as 20%. Yusuf et al. (2010) assert that households with bed nets had a lower chance of infection. More specifically, Haque et al. (2009) conducted a study in the Bangladeshi highlands, which indicated that individuals in households with fewer than three bed nets have twice the risk of malaria infection compared to those in households with three or more bed nets.

Indoor residual spraying has been widely used in areas with high malaria incidence to interrupt its transmission. Angola began indoor residual spraying in 2005, under the National Malaria Control Programme. Spraying was implemented in areas of high epidemic risk in selected provinces to the south of the country (Huila, Namibe, Cunene and Cuando Cubango) and in Luanda. The synthetic pyrethroid sprays

are the preferred insecticides in Angola, because of their low risk and environmental effects unlike other countries in Southern Africa Development Community (SADC) namely; Madagascar, Namibia and South Africa, where malaria has been drastically reduced, or eradicated. They adopted DDT as the insecticide of large choice to prevent malaria (AMIS, 2012).

1.3.3 Socio-economic status of the vulnerable population

Socio-economic factors can have a direct or indirect effect on malaria transmission. Various socio-economic factors are known to promote malaria transmission, as shown by numerous studies conducted in developing countries, which are discussed next. In Indonesia Dale et al. (2005) found that both low to middle income and lower education qualification were significantly associated with malaria incidence. In Tanzania, it was revealed that households in the higher wealth index were infected less often than those in the middle and lower wealth index (Njau et al., 2006). Similarly, a study of three major regions of Ethiopia found that the risk of malaria decreased as household wealth index increased (Graves et al., 2009). Another study by Koram et al. (1995) found that in the Gambia children living in poor quality housing or crowded dwellings were more infected with malaria than those living in the better housing conditions found in peri-urban areas. By way of explaining these differences, Hay et al. (2005) have established that urbanization has an environmental effect on some mosquito species, regarding diversity, numbers, survival rates, infection rates and the frequency with which they bite people.

1.4 Spatial and Spatio-temporal modelling

Theoretically, diseases are mapped in order to identify the geographic distribution of diseases and their risk factors. The use of spatial and spatio-temporal distribution in disease mapping is often understood through the application of statistical techniques to the data to create maps to visually describe spatial and spatio-temporal variation of the disease risk (Currie et al., 2003). There is a vast literature on the development and application of disease mapping techniques; including Besag et al. (1991), Bernardinelli and Montomoli (1992), Best et al. (2005) as some of the early literature on the subject.

Bayesian methods are widely used in disease mapping. The Empirical Bayes (EB) approach was applied by Clayton and Kaldor (1987) for smoothing a map based on the Poisson-gamma and log-normal models. They assumed that the log relative risk were multivariate normal and allowed for spatial correlation via the conditional autoregressive model (CAR). Their model could not be considered to be "fully Bayesian", a quadratic approximation was used for the likelihood which did not account for uncertainty in the estimates of the hyper-parameters.

The fully Bayesian disease mapping was first proposed by Besag et al. (1991) commonly referred to as the BYM model. Besag, York and Mollié were the first to incorporate spatial smoothing into studies of disease mapping. Their proposal was to model the relative risks for each region by means of a combination of fixed and random effects. In CAR, the conditional distribution was modelled via the neighbourhood structure, where the random effects in a region, given all the others,

is simply the weighted average of all the other random effects. Besag et al. (1991) considered the weights average based on neighboring structure on adjacent areas, where regions sharing a common boundary or not; if the regions share the same boundary, the weight is 1, otherwise is 0. Gosoniu et al. (2010) further extended the BYM model by including both spatially unstructured random effects and spatially structured random effects, through the convolution model. According to Gosoniu et al. (2010), the fully Bayesian model produced better estimates than the EB approach developed earlier by Clayton and Kaldor (1987).

In a comparison between the empirical Bayes methods and the fully Bayesian methods for building a disease map, Bernardinelli and Montomoli (1992) found that the fully Bayesian approach offers greater flexibility and better understanding in the statistical analysis of geographical variation in disease rates.

Previous studies on modelling count data have seen considerable development in space-time models. Spatial and spatio-temporal disease mapping models are common tools for the estimation of the risk of disease, identifying regions and periods with high risk, and time trends for disease incidence or prevalence, as are discussed next.

One of the first spatio-temporal models for count data was introduced by Bernardinelli and Montomoli (1992). They assumed a Poisson generalized linear model (GLM) with a predictor containing separate terms for space and time, as well as a space-time interaction effect that allows for different temporal trends in different areas or regions.

Waller et al. (1997) proposed a spatio-temporal model that is an extension of the BYM or the convolution model proposed by Besag et al. (1991). In their study, Waller et al. (1997) assumed that the covariate effects are constant over time and that the disease counts followed a Poisson distribution. When they fitted this model to lung cancer deaths in 88 Ohio counties for the period 1968-1988, they treated each year as a separate time period. Among their important findings were an overall trend of increasing lung cancer deaths and increases in both spatial clustering and uncorrelated heterogeneity. Extending the Waller et al. (1997) model, by assuming that the spatial terms were constant over the research period, Knorr-Held and Besag (1998)) proposed another spatio-temporal model. The latter model consists of a pair of area-specific random effects. It included both unstructured and structured random effects via a convolution prior on space, and also a similar prior for temporal trends to describe spatial and spatio-temporal variation. This model was applied to the same Ohio lung cancer data, but it was not clear whether it revealed additional features or insights into the data (Norton, 2008).

1.4.1 Previous research on malaria using spatial models

Recent research studies in disease incidence and mortality have used spatial and spatio-temporal variation to assess the association and impact of social and environmental factors on transmission of disease including malaria in different parts of the world, including Angola, China, Malawi, South Africa, Thailand and Zimbabwe which are considered next.

To examine the spatio-temporal variation of malaria incidence in the Yunnan province of China, Clements et al. (2009) used Bayesian Poisson regression models. Their study found a strong association between malaria incidence and both temperature and rainfall. Mabaso et al. (2006) investigated the risk factors of malaria incidence at district level in Zimbabwe, by applying Bayesian negative binomial models for spatio-temporal analysis to explore the relationship between climatic variables and malaria incidence. They found inter-annual variability of malaria transmission, wherein the period of highest risk occurred during the wet season and the lowest risk was over dry periods. This seasonal variation was driven mostly by climatic factors.

Kleinschmidt et al. (2002) studied malaria incidence in the period between 1986 and 1990 in two districts of northern KwaZulu-Natal, South Africa, using the conditional autoregressive model (CAR) for mapping incidence rates. Their study found that high malaria incidence occurred in the north-west district and low incidence in the South-East district of KwaZulu Natal.

Lowe et al. (2013) measured the effect of geographic and socio-economic determinants on malaria incidence at district level in Malawi. Their analysis was based on fitting the generalized linear mixed model (GLMM), which included both structured and unstructured spatial and temporal random effects under a hierarchical Bayesian framework that accounted for spatial correlation. Using similar techniques, Lekdee et al. (2014) studied the factors that are related to malaria incidence rates in Thailand. They used both generalized mixed model model (GLMM) and a hierarchical Bayesian approach to produce smooth relative risks maps of malaria

and to model the relationship between climatic factors that are related to malaria transmission. They found that rainfall and temperature were positively related to the malaria morbidity rate.

Ndlovu (2013) in a study of geographical patterns of malaria transmission in KwaZulu-Natal, South Africa produced predicted risk maps using a Bayesian spatio-temporal model. These predictions were based on climatic factors that contributed to or influenced malaria transmission during the study period.

Zhang et al. (2008) applied geographic information system (GIS) to identify spatial distribution and clustering, in order to model populations at high risk for malaria at county level in Anhui province, China. They found that 10 and 24 counties had increased risk of malaria with maximum spatial clusters sizes between 25% and 50% of the total population respectively.

Using GIS mapping techniques and spatio-temporal variation of malaria, Bi (2013) investigated socio-ecological factors relevant to estimating malaria transmission in the Yunnan province of China. The results showed that malaria transmission is driven by a range of socio-ecological factors and varied over both space and time.

The study by Gosoniu et al. (2010) applied a Bayesian geostatistical model to assess effect of interventions after adjusting for environment, climatic and socio-economic factors. Their study revealed that control measures should be concentrated in the southerly, northerly and central parts of the country. The study also confirmed that the highest number of malaria infected children was in the Malanje province which falls in the hyper-endemic region. Using similar methods

(Magalhães et al., 2012) study finding malaria hot-spots in the northen Angola. Their finding identified important individual and modifiable household factors to be associated with malaria risk within meso-endemic area. High risk of malaria endemicity were found in regions covering the northern and eastern areas of Angola (Dando municipality, Bengo province). This two studies appears to be the only such study that has focused on malaria in Angola.

1.5 Problem Statement

Malaria still remains a serious public health issue in Angola. With support from the World Health Organization (WHO), in 1984 Angola started the National Malaria Control Programme (NMCP). To be specific, as stated in the (PMI, 2010) report: "Angola committed itself to the Global Malaria Control Strategy endorsed by the 1992 Ministerial Conference on Malaria in Amsterdam and later signed on to the 1997 African Initiative to Accelerate the Fight Against Malaria". The country also signed other agreements in 1998, 1999 and 2000. Despite these agreements, it was not until 2001 that the government began to allocate a budget for a specific malaria programme. In 2002, a National Commission to combat HIV, malaria and tuberculosis was created by the Board of Ministers, which confirmed that combating the disease was then a priority.

As was mentioned in the previous section, there has been very little statistical work concerning malaria in Angola, especially using disease mapping and other advanced techniques. There is need to consider models that allow the effects of

disease to be observed and how the effects of particular risk factor evolve both over time and in space. In this study we explore these effects by developing a spatio-temporal model and apply into Malaria Indicator Survey (MIS) data.

This study helps to bridge this gap and inform practical strategies in order to help curb child mortality due to malaria in Angola. Children under age of five are most vulnerable to endemic malaria.

1.6 Objectives of the Study

The aim of this dissertation is to adapt the existing statistical models that account for spatial and spatio-temporal modelling and mapping malaria in Angola.

1.6.1 Specific Objectives

The specific objectives of this dissertation are:

- To identity suitable spatial hierarchical models for estimation of the relative risk of malaria mortality at provincial level in Angola, using NMCP data for 2003 to 2010;
- To develop spatial-temporal models for modelling and mapping of malaria among children under the age of 5 years using the 2006-2007 and 2011 Angola Malaria indicator survey (AMIS) data;

- To assess the relationship between malaria transmission and some selected demographic and socio-economic covariates at provincial level using the 2006-2007 and 2011 Angola Malaria indicator survey (AMIS) data.

1.6.2 The Structure of the dissertation

This study is concerned with methods for disease mapping applied to malaria mortality data and spatial-temporal model for modelling and mapping malaria in Angola among children under the age of 5 years. There are five chapters in the dissertaion, which are structured as following:

Chapter 1: This chapter introduces the study, by giving the background, a literature survey, and the aims and objectives of the study.

chapter 2: This chapter presents background information on the study area, a summary of the data set used in the study, and explains the standardization methodology that will be used to compute the expected numbers of deaths and the standardized mortality ratio (SMR).

Chapter 3: In this chapter, we discuss the Bayesian spatial hierarchical model as used in disease mapping. This involves looking at the Poisson-Gamma model, Poisson log-normal model, CAR model and Convolution Priors that are used to model the relative risk of disease in the study, together with their application to malaria mortality using NMCP data.

Chapter 4: Here, we discuss spatio-temporal techniques for modelling and mapping malaria in Angola among children under 5 years, using the MIS data. The

results of the spatio-temporal modelling are outlined, and the resulting malaria prevalence maps are presented. The results of the spatio-temporal modeling are outlined and malaria prevalence maps produced will be presented.

Chapter 5: This chapter gives conclusion and future research from the dissertation.

Chapter 2

Data Description and Baseline Characteristics

2.1 Study area

The Republic of Angola is located within the tropics on the west coast of Southern Africa. It shares bordered with the Republic of Congo-Brazzaville to the north, the Democratic Republic of Congo in the north-east, Zambia on the east and Namibia in the south. Angola comprises 18 provinces and 164 municipalities. The country experiences only two main seasons, summer and winter, as is typical of a tropical climate. Nevertheless, most of Angola lies on a plateau, at an altitude ranging from over 1,200 to 1,800 m which moderates the climate. Furthermore, the coastal region experiences the cooling effect of the Benguela current. The average temperature across the whole country ranges between $20\,^{\circ}\mathrm{C} - 25\,^{\circ}\mathrm{C}$ due to the high

altitudes of the East African Plateau. The seasonal variation in temperature is very small, dropping to 20 °C in winter and exceeding 25 °C in summer as reported in McSweeney (2010). The summer season is rainy and hot and occurs between September and April whilst the dry and cold winter season occurs between between May and August. The tropical conditions in Angola makes it a good breeding ground for mosquitoes which are responsible for causing malaria.

FIGURE 2.1: Map of the Republic of Angola and its neighbors (Source of data: mapsopensource)

.

2.2 Data

In this study we used two different data sets. The first data set was obtained from the National Malaria Control Programme (NMCP) in collaboration with the Ministry of Health in Angola. The aim of NMCP survey was to collect high quality data so as to evaluate the Angolan mortality rate from malaria during the period 2003-2010. As is customary in Angola, every suspected case of malaria diagnosis had been confirmed by blood test, either through microscopy or rapid diagnostics tests. The NMCP is mandated to conduct studies to help establish malaria risk factors, which can then inform suitable intervention strategies to help reduce malaria mortality rates in the general population, and especially among pregnant women and infants.

The second data set for this study was extracted from the Angola Malaria Survey (AMIS) 2006-2007 and 2011 conducted by Demographic and Health Survey (DHS) program. This program aims to collect and analyze reliable demographic indicators and health estimates at regional and family levels to eventually provide for better health interventions. Two questionnaires were used in the survey. The first concerned household information. The second questionnaire collected individual information from women aged between 15 and 64 years, concerning their demographic characteristics and their knowledge of malaria. In particular, this study uses the data on children in the age group 0-5 years. Information on 2,310 children from the first survey had been collected between November 2006 and April 2007 and that for 3,432 children from later survey had been collected between January 2011 and May 2011. The data used represents individuals who

had provided venous blood for malaria testing and also provided full covariate information; obtained with the consent and assistance of the child's parent or guardian. The variables of interest that were subsequently used in this study were residence (urban/rural), province, wealth index, whether the child slept under a mosquito net and the age of the child which was captured as continuous. Readers are directed to the AMIS (2007, 2012) report for more information about the data.

2.2.1 Population data

The calculation of area-specific disease data for mortality or incidence rates requires a good estimate of the population at risk. However, in Angola, the first population census was in 1970, and there was no further census until 40 years later in 2010. Therefore, the estimated population for each province of Angola was obtained from GeoHive-Angola population statistics.

2.3 Baseline characteristics

A total of 109,320 malaria deaths were observed during the period of 2003-2010, from an estimated population size of 18,915,437 in Angola. A summary table of the crude death rates recorded in each province of Angola during the study period 2003-2010 is given below in Table 2.1.

TABLE 2.1: Observed deaths and crude death rates due to malaria per province in Angola

Provinces	Number of deaths	Population	Crude death rate per 100 000 Population
Bengo	951	460508	670.51
Benguela	14142	2110135	670.19
Bie	5065	1794387	282.27
Cabinda	1136	264584	429.35
Cuando Cubango	9372	664415	1410.56
Cuanza Norte	4097	653961	626.49
Cuanza Sul	5725	1666059	343.63
Cunene	6179	687347	898.96
Huambo	10566	2224717	474.94
Huila	8341	1558852	535.07
Luanda	28874	1823282	1583.63
Lunda Norte	1271	700569	181.43
Lunda Sul	650	523596	124.14
Malanje	1995	997626	199.97
Moxico	2034	373667	544.33
Namibe	1401	313667	446.65
Uige	5622	2597682	216.42
Zaire	1899	362852	523.35
Total	109320	18915437	577.94

It can be seen that Luanda province, which includes the capital city of the same name, was observed to have had the highest crude death rate of 1583.63 per 100 000 population. Cuando Cubango was second with 1410.50 per 100 000 population while Cunene was the third with 898.96 per 100 000 population. The lowest crude death rate was observed in Lunda Sul province, with only 124.14 per 100 000 population. In terms of population size, this province is ranked as sixth, one of the smallest in Angola. Lunda Norte had a crude death rate of 181.43 per 100 000, while that of Malanje was 199.97 per 100 000.

A detailed scatter plot of malaria deaths in Angola for the 2003 to 2010 is shown in Figure 2.2, below.

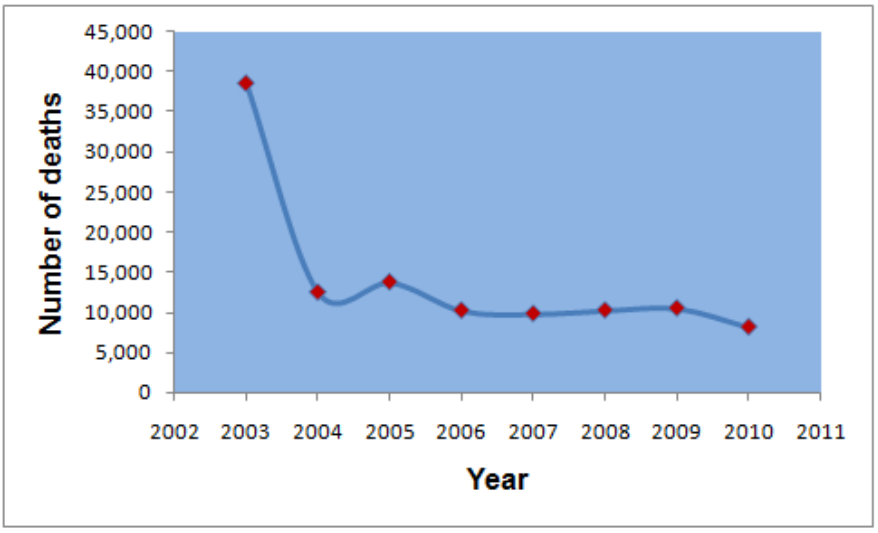

FIGURE 2.2: Annual number of deaths due to malaria. The disease caused the highest mortality recorded in 2003 and the lowest mortality recorded in 2010

.

The figure shows that in 2003 by far the highest number of deaths from malaria recorded was 38 598. The exceptionally high mortality rate in 2003 has several causes, the most notable being that Angola had recently emerged from a protracted civil war. Although the war ended in 2002, Angola's political climate had not effectively stabilized by 2003, and the country was still undergoing post conflict reconstruction. As a result infrastructure, services, facilities and other resources were still grossly inadequate. Therefore, in 2003 public hospitals were simply not equipped to deal with an epidemic of disease such as malaria. In subsequent years the government become increasingly stable and started to put health interventions and control measures in place, with resultant beneficial effect on the malaria mortality rates hence the decline of deaths to malaria as evident in Figure 2.2. However a lot needs to be done since the number of deaths in absolute terms remains high.

2.4 Statistical methods appropriate for of interest in this study

The disease of interest for this study is malaria in Angola. In order to understand the malaria problem this appropriate methods were applied. We used hierarchical spatial modelling techniques namely: the Poisson-gamma model, the Poisson-log-normal model with spatial unstructured random effects (uncorrelated heterogeneity) and the Besag, York and Mollie model Besag et al. (1991) which all incorporate both spatially structured and spatially unstructured random effects. Spatio-temporal models were also used to assess the association between malaria prevalence variation and socio-economic and demographic factors. The analysis in this study were conducted using two software programs; WinBUGS version 1.4 by (Spiegelhalter et al., 2003), and the R statistical software with R-INLA package.

Chapter 3

Spatial Hierarchical Modelling of Malaria in Angola

3.1 Introduction

Mapping of disease involves statistical measures to describe the overall geographical distribution and variation of the disease, so as to help in understanding the demands of the disease. The distributions can be displayed on maps to assess the geographic distribution of a particular disease which is also the approach usual to map the standardized mortality ratio (SMR) across different geographical areas.

Disease mapping generally detects clustering and severity of a disease in a given area, country or region (Clement, 2014). It is aimed at smoothing and predicting some response variables over the geographic area of interest in order to derive

estimate measures of a particular disease, which can then allow policy makers to make better decisions on public health resource allocation.

This chapter introduces the Bayesian hierarchical models (BHMs) used to estimate the relative risk of malaria mortality using NMCP data at provincial level in Angola 2003-2010. These models include the Poisson-gamma model, Poisson log-normal model, conditional autoregressive (CAR) model as well as the convolution model.

In the next section, we calculate the expected number of malaria deaths and crude death rates taking into account the population size.

3.2 Crude Estimation of Malaria Risk

Suppose that the unknown malaria mortality relative risk for a province i is given as θ_i where $i = 1, 2, \ldots, m$ where m is 18 in our case for Angola. Let y_i denote the observed count for malaria deaths and E_i the expected number of deaths due to malaria in the i^{th} province. The expected number of malaria deaths in the i^{th} province can be calculated through a standardization method based on population size P_i by Bivand et al. (2013), which can be written as

$$E_i = rP_i \tag{3.1}$$

where r is the overall risk of death in the entire study population and is given by

$$r = \frac{\sum_{i=1}^{m} y_i}{\sum_{i=1}^{m} P_i}. \tag{3.2}$$

The relative risk of the disease in the study is estimated using the standardized mortality ratio (SMR). This is expressed as:

$$SMR_i = \theta_i = \frac{y_i}{E_i}. \tag{3.3}$$

We assume that the observed number of deaths (y_i) at the i^{th} province occurs independently and is assumed to follow the Poisson distribution for counts expressed as

$$y_i \sim Poisson(E_i\theta_i). \tag{3.4}$$

Under this assumption, the likelihood of the relative risk of θ_i is expressed as

$$l(\theta_i) = \frac{exp(-E_i\theta_i)(E_i\theta_i)^{y_i}}{y_i!} = P(y_i|\theta_i) \tag{3.5}$$

Furthermore, the maximum likelihood estimator of $\hat{\theta_i}$ is the SMR in the i^{th} province, and is given by:

$$\hat{\theta_i} = \frac{y_i}{E_i} \tag{3.6}$$

with variance of the SMR given by

$$var(\hat{\theta_i}) = \frac{\theta_i}{E_i}. \tag{3.7}$$

Hence, the estimated standard error of the SMR is given by:

$$SE(\hat{\theta}_i) = \frac{\sqrt{y_i}}{E_i}. \tag{3.8}$$

Therefore, we can calculate the 95% confidence interval for $\hat{\theta}_i$, which is given by:

$$\hat{\theta}_i \pm z_{\frac{\alpha}{2}} SE(\hat{\theta}_i). \tag{3.9}$$

Although the SMR is a common measure of relative risk, and is often used to estimate and map disease risk, it nevertheless has some limitations. In this regard, according to Lawson et al. (2000) the variance of SMR has a high variability in areas with low population size and small variability in areas with a large population size. This is evident in Equation (3.1) and (3.7). To elaborate, in areas where there are no observed deaths, the SMR is zero, despite the population size. The SMR is more reliable in large highly populated geographical areas such as counties. Hence, to overcome this problem it is beneficial to borrow strength by incorporating information from neighbouring areas, where the disease risks are likely to be more prevalent. This can be accomplished by using Bayesian hierarchical models.

3.3 Bayesian Hierarchical Model

Bayesian hierarchical models (BHMs) are among the statistical techniques most commonly used in analysis of spatial epidemiological data. The Bayesian approach is based on ones prior belief about the parameters of interest. It uses the likelihood

function of the parameters obtained from the observed data, in order to get the posterior distribution, from which the parameters of interest are then estimated. One of the challenges of the Bayesian approach lies in estimating the posterior distribution, which involves using high-dimensional integral functions.

3.3.1 The Likelihood Function

Let y_i, $i = (1, 2, \ldots, n)$ be independent random variables with probability density function $P(y_i|\theta)$, where $\boldsymbol{\theta}$=$(\theta_1, \ldots, \theta_p)$ is a vector parameters. The likelihood function is defined as

$$L(\theta) = P(\boldsymbol{y}|\boldsymbol{\theta}) = \prod_{i=1}^{n} P(y_i|\boldsymbol{\theta}). \tag{3.10}$$

It is the joint p.d.f of $y = (y_1, \ldots, y_n)$ given $\theta = (\theta_1, \ldots, \theta_n)$. This equation is based on the assumption that the sample values of $y = (y_1, \ldots, y_n)^{'}$, given the parameters $\boldsymbol{\theta}$, are independent. This means that the independence assumption makes it possible to express the likelihood function as a product of the individual contributions of $P(y_i|\boldsymbol{\theta})$ in equation (3.10). For this reason, the data is assumed to be conditionally independent. However though independence is an appealing feature, in reality correlated observations do frequently arise, and so sound statistical methods are needed to deal with such data.

3.3.2 Prior Distribution

As mentioned before, Bayesian methods are based on the prior belief about the parameters of interest. This prior belief is described in terms of a density function,

referred to as the prior distribution. Note that, when the sample size is large, the likelihood function will contribute more to the relative risk estimation. Conversely, when the sample has poor data then the prior distribution will dominate the analysis.

All the parameters within Bayesian models are stochastic and can be assigned appropriate prior distributions. The prior distribution is a distribution assigned to the parameter θ before the data y_i are observed. Given a single parameter, θ, the prior distribution can be denoted as $P(\theta)$. However for a parameter vector $\boldsymbol{\theta}$, the same notation can be used for the joint distribution provided $\boldsymbol{\theta}$ is understood to be a vector of parameters. There are different properties that can be associated with the prior distribution. One such property is that, the prior distribution can be improper. An improper prior is defined as the condition where the integration of the prior distribution of a random variable θ over its range Ω is infinite (Lawson, 2008). This can be expressed mathematically as,

$$\int_{\Omega} P(\theta)\, d\theta = \infty. \tag{3.11}$$

Although, the improper property is a limitation of any prior distribution, it is not necessarily the case that an improper prior will lead to the same property in the posterior distribution.

3.3.3 Non-informative Priors

A non-informative prior is a type of prior distribution that is assumed to make a strong preference over the observed values. Thus non-informative prior distributions are sometimes referred to as vague or flat prior distributions. Usually, choosing a non-informative prior distribution for the parameters means that in any posterior analysis, the prior distribution will not impact significantly compared to the likelihood of the data (Lawson, 2008).

According to Lawson (2008), the choice of non-informative prior can be specifically made with some general understanding of the range and behavior of the parameters. Notice that, variance parameters must have prior distributions on the positive real line. Possible prior distributions in this case are often obtained in the gamma, inverse gamma, or uniform families.

3.3.4 Conjugate Prior

In choosing a prior distribution, a particular combination of the prior distribution and likelihood function can lead to the same distribution family in the posterior $P(\theta|y)$ as for the prior distribution $P(\theta)$. Such prior and posterior distributions are then referred to a conjugate distributions, and the prior is called a conjugate prior for that likelihood. For instance, the Poisson likelihood with parameter θ and the gamma prior distribution for θ are conjugate. In this case the posterior distribution of θ is also gamma distributed. Similar results hold for the binomial likelihood, and the beta prior distribution for the probability of success, as well as

the normal data likelihood with a normal prior distribution for the mean (Lawson, 2008). All members of the exponential family have conjugate priors.

The conjugacy prior can be identified by examining the kernel of the prior likelihood product. The prior likelihood product should be in similar form to the prior distribution. In this regard Lawson (2008), stated that conjugacy always guarantees a proper posterior distribution.

3.3.5 Posterior Distribution

The posterior distribution describes the behavior of the parameter of interest after the data is observed and prior assumptions have been made. It is obtained from the product of the likelihood and the prior distribution. The posterior distribution can be expressed as:

$$P(\theta|y) = \frac{P(y|\theta)P(\theta)}{\int P(y|\theta)P(\theta)\, d\theta} \tag{3.12}$$

where $\int P(y|\theta)P(\theta)\, d(\theta)$ is called the normalizing constant, which is equal to the marginal p.d.f of y which has no information about θ therefore just as good as a constant. Because of this, the distribution can be specified in proportional; terms that is it can be expressed mathematically as:

$$P(\theta|y) \propto P(y|\theta)P(\theta) \tag{3.13}$$

where the constant of proportionality is the reciprocal of $\int P(y|\theta)P(\theta)\,d\theta$. Note in Equation (3.13) the right hand side is the product of the likelihood of the data and the prior distribution.

In disease mapping nomenclature where often the data are counts, the obvious assumptions is that the data likelihood is Poisson and that there is a common relative risk parameter with a single gamma prior distribution as explained in Lesaffre and Lawson (2012).

3.4 Markov Chain Monte Carlo (MCMC) Method

In this section, we will briefly describe the basic theory of the Markov chain monte carlo (MCMC) algorithm, which is used to perform posterior inference in the case where the product of the likelihood and the prior are analytically intractable. This method involves a set of steps that use iterative simulation of parameter values from the posterior distribution based on a Markov chain (MC) formulation.

The MCMC techniques enable quantitative researchers to use highly complicated models to estimate the corresponding posterior distribution with high accuracy. This method has contributed a greater deal to the development, advancement and propagation of Bayesian theory. To elaborate, Ntzoufras (2011) stated that the MCMC is a generic algorithm that ensures the Markov chain converges eventually to the target (or stationary) distribution. The target distribution is the posterior distribution $P(\theta|y)$, from which the MCMC algorithm provides samples by direct simulation from it.

It should be emphasized that, as mentioned above, the goal of MCMC techniques is to simulate observation from the target distribution. To achieve this is a number of possible considerations exist, which include importance sampling, rejection sampling, perfect sampling, slice sampling and Laplace approximation.

3.4.1 The Markov Chain Monte Carlo Algorithm

Consider $\theta^{(1)}, \theta^{(2)}, \ldots, \theta^{(T)}$ to be a sample of size T from the posterior distribution $P(\theta|y)$. A Markov chain is a stochastic process in which future states are independent of past states given the present state. Based on the sequence $\theta^{(1)}, \theta^{(2)}, \ldots, \theta^{(T)}$ the process can be specified $P(\theta^{(t+1)}|\theta^{(t)}, \ldots, \theta^{(1)}) = P(\theta^{(t+1)}|\theta^{(t)})$. That is, the distribution of θ at time $t+1$, given all the preceding θ values (for $t, t-1, \ldots, 1$), depends only on the value $\theta^{(t)}$ at the immediate past time t.

As $t \to \infty$, the distribution of $\theta^{(t)}$ converges to its equilibrium, which is independent of the initial value of the chain $\theta^{(0)}$. This condition occurs, when the Markov chain is irreducible, aperiodic and positive-recurrent.

The MCMC method works by constructing the Markov chain in a set of steps that can be outlined as follows:

1. Select an initial value $\theta^{(0)}$

2. Sequentially generate $\theta^{(t+1)}|\theta^{(t)}$ values from $P(\theta|y)$ until the equilibrium distribution is reached.

3. Monitor the convergence of the algorithm using the convergence diagnostics. If convergence diagnostics fails, we generate more samples.

4. Cut off the first B observations. Then B is called the burn-in period; meaning the first B iteration values are eliminated from the sample to avoid the influence of initial values.

5. Consider $\left\{\theta^{(B+1)}, \theta^{(B+2)}, \ldots, \theta^{(T)}\right\}$ as the sample for the posterior analysis.

6. Plot the posterior distribution. Two plots commonly used are the histogram and kernel plots of the distribution.

7. Finally, obtain summary results of the posterior distribution.

All the above mentioned steps are important when assessing the convergence diagnostics. The convergence diagnostics are used to check if convergence has been achieved or not.

3.4.2 Description of the posterior distribution using the MCMC output

Using the techniques described in the previous section, we thus have the MCMC output, which provides us with a random sample $\left\{\theta^{(1)}, \theta^{(2)}, \ldots, \theta^{(T')}\right\}$ after discarding the burn-in observations. From this sample, values or a sample for any desired function $G(\theta)$ of the parameter of interest θ, can be obtained by considering $G(\theta^{(1)}), G(\theta^{(2)}), \ldots, G(\theta^{(T')})$. This is the sample after discarding the burn-in values. The summary of any posterior distribution can also be obtained from

$\theta^{(1)}, \ldots, \theta^{(T')}$ using the traditional sample estimates and sample distribution parameters. For example, we can estimate the posterior mean of $G(\theta)$ by

$$\widehat{E}(G(\theta)|y) = \frac{1}{T'}\sum_{t=1}^{T'} G(\theta^t) \tag{3.14}$$

and the posterior variance by

$$\widehat{V}(G(\theta)|y) = \frac{1}{T'}\sum_{t=1}^{T'-1}\Big[G(\theta^t) - \widehat{E}(G(\theta)|y)\Big]^2 \tag{3.15}$$

and the standard deviation can be obtained as $\sqrt{\widehat{V}(G(\theta)|y)}$. Other measures of interest such as the posterior median or quantiles with 2.5% and 97.5% percentiles provide a 95% credible interval which can be obtained after ranking the posterior draws of $G(\theta)$ then locate the respective quantiles from the rank statistic .

The MCMC output is always referred to as the generated sample after removing the initial interactions produced during the burn-in period. Two of the most popular sampling procedures for the MCMC algorithm are the Metropolis-Hasting algorithm and the Gibbs Sampling. Many variations and extensions of these algorithms have been developed. In the next section, we will briefly discuss these two MCMC sampling algorithms.

3.4.3 The Metropolis-Hasting Algorithm (M-H)

The Metropolis-Hastings (M-H) algorithm is a Markov chain Monte Carlo (MCMC) approach that was introduced by Metropolis et al.(1953) and generalized by Hastings in (1970).

Suppose we have θ as a vector-valued parameter and $P(\theta|y)$ the target posterior distribution from which we wish to generate the sample of size T. In Bayesian inference, the M-H algorithm proceeds through the following iterative steps:

1. Start with initial values $\theta^{(0)}$

2. For $t = 1, \ldots, T$ repeat the following steps

   - Set $\theta = \theta^{(t-1)}$
   - Generate new candidate values θ' from a proposal distribution $q(\theta'|\theta)$.
   - Calculate
   $$\alpha = min\left(1, \frac{p(\theta'|y)q(\theta|\theta')}{p(\theta|y)q(\theta'|\theta)}\right) \tag{3.16}$$
   - Update $\theta^{(t)} = \theta'$ with probability α or $\theta^{(t)} = \theta^{(t-1)}$ with probability $1 - \alpha$.

An important characteristic of this algorithm is that its convergence depends on the proposal distribution. Thus, in practice, it is important to choose the proposal distribution with care, as poor choices considerably delay convergence to the target distribution.

3.4.4 The Gibbs Sampler (GS)

The Gibbs sampler was introduced by German and German (1984). This algorithm is a special case of the Metropolis-Hasting algorithm which uses a proposal density $q(\theta'|\boldsymbol{\theta}^{(t)})$. With the assertion that in this algorithm the desired sample is generated from the full conditional posterior distribution $P(\theta_j|\boldsymbol{\theta}_{-j}, \boldsymbol{y})$, where $\boldsymbol{\theta}_{-j} = (\theta_1, \ldots, \theta_{j-1}, \theta_{j+1}, \ldots, \theta_p)^T$.

One advantage of the Gibbs sampler is that, in each step the random values must be generated from a uni-dimensional distribution, for which a variety of computational tools exists. In most cases, the Gibbs sampler always moves to a new value and does not require specification of a proposal distribution.

In summary the Gibbs Sampler algorithm proceeds as follows:

1. Set initial values $\theta^{(0)}$.

2. For $t = 1, 2, \ldots, T$ repeat the following steps

   - Set $\theta = \theta^{(t-1)}$.
   - For $j = 1, 2, \ldots, p$, update θ_j from $\theta_j \sim P(\theta_j|\boldsymbol{\theta_{-j}}, \boldsymbol{y})$
   - Set $\theta^{(t)} = \theta$ and save it as generated set of values at the $(t+1)$st iteration of the algorithm.

Hence, given a particular state of the chain $\theta^{(t)}$, we generate the new parameter values by

$\theta_1^{(t)}$ from $P(\theta_1|\theta_2^{(t-1)}, \theta_3^{(t-1)}, \ldots, \theta_p^{(t-1)}, \boldsymbol{y})$

$\theta_2^{(t)}$ from $P(\theta_2|\theta_1^{(t)}, \theta_3^{(t-1)}, \ldots, \theta_p^{(t-1)}, \boldsymbol{y})$

$\theta_3^{(t)}$ from $P(\theta_3|\theta_1^{(t)}, \theta_2^{(t)}, \theta_4^{(t-1)}, \ldots, \theta_p^{(t-1)}, \boldsymbol{y})$

$\vdots$

$\theta_j^{(t)}$ from $P(\theta_j|\theta_1^{(t)}, \theta_2^{(t)}, \ldots, \theta_{j-1}^{(t)}, \theta_{j+1}^{(t-1)} \ldots, \theta_p^{(t-1)}, \boldsymbol{y})$

$\vdots$

$\theta_p^{(t)}$ from $P(\theta_i|\theta_1^{(t)}, \theta_2^{(t)}, \ldots, \theta_{p-1}^{(t)}, \boldsymbol{y})$.

Generating values from $P(\theta_j|\theta_1^{(t)}, \theta_2^{(t)}, \ldots, \theta_{j-1}^{(t)}, \theta_{j+1}^{(t-1)} \ldots, \theta_p^{(t-1)}, \boldsymbol{y})$ is relatively easy since it is a univariate distribution and can be written as $P(\theta_j|\boldsymbol{\theta}_{-j}, \boldsymbol{y}) \propto P(\boldsymbol{\theta}|\boldsymbol{y})$, where all the variables except θ_j are kept constant at their given current values.

3.4.5 Metropolis-Hasting versus Gibbs Sampler

There are advantages and disadvantages in using M-H and GS methods. Lawson (2008) reports that the Gibbs Sampler provides a single new value for each θ estimate at each interaction, whereas each M-H step does not require evaluation of a conditional distribution, but it also does not guarantee the acceptance of a new value. Therefore, on the other hand, the Gibbs Sampler can provide faster convergence chains, but are only useful if the computation of the conditional distribution at each iteration are not time consuming. On the other hand the Metropolis-Hasting steps will usually be faster at each iteration, but will not guarantee exploration acceptance of the generated value.

3.5 Assessing and Improving Convergence within the MCMC method

One of the steps required in MCMC methods is the use of diagnostics criteria to assess whether the iterative simulations have reached an equilibrium distribution of the Markov Chain or not. These methods are important as they help to decide how many iterations to use and represent the posterior density to ensure that the Markov Chain converges. Hence, checking convergence is not an issue once the posterior distribution has been determined. In the next sub-section we look at several ways to assess the efficiency and ways of improving the convergence within MCMC.

The efficiency of the MCMC algorithm, in other words its convergence, may be assessed graphically; five such methods are discussed below.

- **Trace plot**: A simple exploration of the trace plots gives a good impression of the convergence. Trace plots are produced for each parameter separately and to evaluate the chain univariately, but can also be used to evaluate and monitor several parameters jointly. Thus, this can be done by monitoring the log of the likelihood or log of the posterior density as discussed in Lesaffre and Lawson (2012).

- **Autocorrelation Function Plot**: When future positions in the Markov chain are highly predictable from the current position, the posterior is slowly explored and one can be able to tell whether the chain has a mixing rate.

A mixing rate is assessed by autocorrelations of different lags. If the mixing rate is low, then autocorrelation decreases slowly with increasing lag.

The autocorrelation also indicates the minimum number of iterations needed for the Markov chain to "forget" its starting position (Lesaffre and Lawson, 2012).

- **Kernel Density Plots**: A more satisfactory density plot for a converged chain would look more for like the bell-shaped density or indicative parameters whose marginal posterior densities are approximately normal (Lawson, 2008).

Formal approaches are also used to assess the MCMC convergence such as:

- **Raftery and Lewis diagnostics** The Raftery and Lewis diagnostic is intended to both detect the convergence to the stationary distribution, and to evaluate the accuracy of the estimated percentiles, by reporting the number of samples to reach the desired accuracy of the percentiles. This diagnostic is designed to test the number of iterations and burn-in needed by first running and taking a shorter pilot chain (Lesaffre and Lawson, 2012).

- **Brooks-Gelman-Rubin diagnostic**

  The Brooks-Gelman-Rubin diagnostic is based on assessing the convergence of multiple chains via comparison of summary measures across chains. This diagnostic compares the within and between chain variances for each variable. When the chains have "mixed"(converged) the variance within each

sequence and the variance between sequences for each variable will be approximately equal. However, this means that the ratio of the two variances converge to 1 if all chains being sampled are identically distributed.

For poorly identified models, the variability of sampled parameter values between chains significantly exceeds the variability within any one chain (Congdon, 2010). The convergence of the Markov chain can be improved by standardizing covariates and the use of unstructured random effects. Although, these methods are the most popular in assessing and improving convergence, we could also assess convergence by examining a time series plot or trace plots of the parameters. There is no guarantee that sampling using an MCMC algorithm will ensure or result to convergence to the posterior distribution.

3.5.1 Model Diagnostics and Selection

Although we consider one of the methods used to select the best fitting model by, there are no clear criteria to compare models in order to select the best fitting model. Nevertheless, the technology to fit complex models by Bayesian hierarchical methods is widely available.

The common problem with using this criteria is how to measure an appropriate penalized complexity of a hierarchical model. The deviance information criteria (DIC) proposed by Spiegelhalter et al. (2002) is the most popular method used for model selection in a Bayesian context. The DIC provides a measure that indicate both model fit and complexity. It is defined as two times the mean of the deviance

minus the deviance of the mean. Thus, the deviance of the posterior expected parameter can be defined as $D(\boldsymbol{\theta}) = 2 \log P(\boldsymbol{y}|\boldsymbol{\theta})$ and it is called the Bayesian deviance. Therefore Spiegelhalter et al. (2002) defined the effective number of parameters p_D in the model, which measure the complexity of the model, and which can be written as

$$p_D = \overline{D(\boldsymbol{\theta})} - D(\bar{\boldsymbol{\theta}}), \tag{3.17}$$

where $\overline{D(\boldsymbol{\theta})}$ is the posterior mean of $D(\boldsymbol{\theta})$, which measures the goodness of the model fit, while $D(\bar{\boldsymbol{\theta}})$ is equal to $D(\boldsymbol{\theta})$ evaluated at the posterior mean of the parameters. Hence, the deviance information criterion (DIC) is defined as

$$DIC = D(\overline{\boldsymbol{\theta}}) + 2p_D \quad \text{or} \quad DIC = \overline{D(\boldsymbol{\theta})} + p_D \tag{3.18}$$

One aspect of both DIC and p_D is that they can be easily calculated from the samples generated by MCMC simulation, as explained in Berg et al. (2004). Therefore to use the DIC small values of $\overline{D(\boldsymbol{\theta})}$ indicate a better fit while small values of p_D indicate model parsimony.

3.6 Bayesian Models for Estimation of the Relative Risk

The Poisson-gamma model and Poisson log-normal model are the two non-spatial models used in this dissertation to estimate the relative mortality risk (RR). In order to control the variability and similarity in estimating the mortality relative

risk of the disease in Angola, the Bayesian model proposed by Besag, York and Mollie Besag et al. (1991) (BYM) model was used.

3.6.1 Poisson-Gamma Model

The Poisson-gamma (PG) model has been widely used in disease mapping to account for extra variability in the data through the use posterior distribution estimation (Lawson et al., 2003). This model, as applied in this study, is explained as follows.

3.6.1.1 Model Description

Let n be the number of provinces under consideration. Also let y_i and E_i be the observed and expected number of malaria deaths in the i^{th} province respectively. Then the numbers of deaths due to malaria in each of the provinces are assumed to be mutually independent and hence follow a Poisson distribution

$$y_i \sim Poisson(\theta_i E_i) \tag{3.19}$$

for i=1,2...,n where in our study n=18 provinces. Here θ_i is the unknown relative risk in the i^{th} province with Poisson mean $\mu_i = \theta_i E_i$. Hence, the likelihood function for y_i is given as

$$P(y|E_i\theta) = \prod_{i=1}^{n} \frac{(\theta_i E_i)^{y_i}}{y_i!} exp(-\theta_i E_i). \tag{3.20}$$

The prior distribution of the relative risk θ_i denoted by $P(\theta_i)$, is assumed to follow a gamma prior distribution with parameters a and $b > 0$ i.e $\theta_i \sim \Gamma(a, b)$ with mean $E(\theta_i) = \frac{a}{b} = \mu$ and variance of $Var(\theta_i) = \frac{a}{b^2} = \sigma^2$. Thus, the prior distribution is given by

$$P(\theta_i|a, b) = \frac{b^a}{\Gamma(a)}\theta_i^{a-1}exp(-b\theta_i). \tag{3.21}$$

Using the equation (3.13), the resulting posterior distribution is then given by

$$P(a, b, \theta|y) \propto \prod_{i=1}^{n} \frac{(\theta_i E_i)^{y_i}}{y_i!}exp(-\theta_i E_i) \prod_{i=1}^{n} \frac{b^a}{\Gamma(a)}\theta_i^{a-1}exp(-b\theta_i) \tag{3.22}$$

3.6.1.2 Parameter Estimation

The parameters a, b and θ are estimated using MCMC via Gibbs sampling. The Poisson-gamma posterior conjugate can be derived, because $P(\theta_i|a, b, y_i) = Gamma(a + y_i, b + E_i)$. Thus, the posterior mean of θ_i is equal to

$$E(\theta_i|y_i, a, b) = \frac{a + y_i}{b + E_i} = w_i SMR_i + (1 - w_i)\frac{y_i}{E_i} \tag{3.23}$$

with $w_i = \frac{b}{(b+E_i)}$. It is a weighted average, which indicates how much the observed SMR_i has been shrunk towards the E_i by taking the posterior mean (Lesaffre and Lawson, 2012).

One disadvantage of the Poisson-gamma model is the inability to cope with spatial correlation. In this regard the existence of spatial correlation in the data violates the assumption of independence inherent in the classical Poisson-gamma model (Lawson et al., 2003). It was therefore important for this study to begin with

the Poisson-gamma model as referred to as a classical model and then proceed with a model that appropriately takes into account spatial autocorrelation, such as Poisson log-normal and the Besag, York and Mollie model.

3.6.2 Poisson Log-normal Model

The Poisson log-normal (PLN) model was proposed in order to account for spatial correlation. Although, the Poisson-gamma model is a simple model offering analytic tractability, it does not allow us to easily incorporate spatial random effects in order to address correlated outcomes. In this regard Clayton and Kaldor (1987) proposed the Poisson log-normal model, in which risk estimation is based on assumption that the logarithm of relative risk follows a multivariate log-normal distribution with normal hyper-parameters μ and σ_v^2 as is outlined next.

3.6.2.1 Model Description

We consider the log-normal model with area-specific random effects or spatially unstructured random effects v_i and without covariates because in our case these were not available, while α_0 is the overall level of the relative risk (intercept). The spatially unstructured random effects were modeled with a zero mean Gaussian prior distribution and variance σ_v^2 so that $v_i \sim N(0, \sigma_v^2)$. It is not easy to specify a prior for σ_v, which denotes the standard deviation of the log residual relative risk and it is a difficult parameter to interpret epidemiologically. Therefore, the choice of the Gamma distribution for the precision, $\tau_v = \frac{1}{\sigma_v^2}$ is suitable because it produces a marginal distribution for the residual relative risk assuming that the hyperprior

distribution for the precision parameter ${\tau_v}^2$ is ${\tau_v}^2 \sim gamma(0.5, 0.0005)$. Using Equation (3.19) the log-normal model for the relative risk is defined as

$$log(\theta_i) = \alpha_0 + v_i \tag{3.24}$$

and the estimated relative risk will be

$$\hat{\theta}_i = exp(\alpha_0 + v_i). \tag{3.25}$$

The random effects are included in the model in order to consider some extra Poisson variation in the log relative risk that might be present in the data. Thus, the log-normal model for the relative risk is more flexible (Lawson, 2008).

3.7 The Besag, York and Mollie Model

Among the models proposed for performing disease risk which have appeared on literature, the Besag, York and Mollie (BYM) model is the most common model used to estimate the spatial pattern of risk in hierarchical Bayesian disease mapping. In this model for relative risk, area-specific random effects are divided into two components. One component is the spatially structured random effects u_i, which take into account effects that vary in space (correlated heterogeneity) and the spatially unstructured random effects v_i that take into account the effects that are inherent within an area (uncorrelated heterogeneity). The BYM model

was introduced firstly by Clayton and Kaldor (1987) and later extended by Besag et al. (1991).

3.7.1 Conditional Autoregressive Model

The conditional autoregressive (CAR) model has been widely used for analysis of spatial data in different fields, such as demography, geography and epidemiology. This model, which assumes spatial correlation between neighbouring regions, was introduced by Besag et al. (1991) as a modern methodology suitable for estimating disease risk. Hence, the structure of the CAR model incorporates the spatial correlation into the model by specifying the distribution of the random effects for all neighbouring regions, based on the idea the that two regions are defined to be neighbours if they share a common border. Thus, the estimation of disease risk at a certain area, depends on the risk at neighbouring areas. This model is fitted by incorporating the spatially structured random effects or the correlated heterogeneity model.

In general, the CAR model is a class of Gaussian Markov random fields characterized by a conditional probability density function. It is intended to model spatial occurrences that are highly related to a specific area, as discussed in Cressie (1993) and Besag (1974). As applied in this dissertation, the CAR model can be described as follows.

Let $S = \{1, 2 \ldots, n\}$ be the provinces under study and $N_i = \{j \in S : \text{j is a neighbour of i}\}$ denote the set of all provinces that are neighbouring province to i, $i \in S$. Let x_i,

$i \in S$ be a random variable. We define the corresponding Gaussian Markov random field $\boldsymbol{u}$ as the vector $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)^{'}$.

The conditional probability density function of a CAR model, according to Mariella and Tarantino (2010), can be written as

$$f(x_i|x_{j\in N_i}) = \sqrt{\frac{1}{2\pi {\sigma_i}^2}} exp\left\{ -\frac{\left[(x_i - \mu_i) - \rho \sum_{j\in N_i} \beta_{ij}(x_j - \mu_j)\right]^2}{2{\sigma_i}^2} \right\}, \tag{3.26}$$

where $\mu_i \in \mathbb{R}$, ${\sigma_i}^2 \in \mathbb{R}^+$, $|\rho| < 1$, $\beta_{(ij)} \in \mathbb{R}$, $\beta_{(ij)} = \beta_{(ji)}$, $\beta_{(ii)} = 0$, and the conditional joint probability density function is

$$f(x) = \frac{1}{(2\pi)^{\frac{n}{2}} det(\boldsymbol{B}^{-1} \sum_D)^{\frac{1}{2}}} exp\left[-\frac{(\boldsymbol{X} - \boldsymbol{\mu})' \sum_D^{-1} (\boldsymbol{X} - \boldsymbol{\mu})}{2} \right], \tag{3.27}$$

where $\boldsymbol{\mu} \in \mathbb{R}^{n\times 1}$ (n-dimensional vector), given by $\boldsymbol{\mu} = (\mu_1, \mu_2, \ldots, \mu_n)^{'}$, B $\in \mathbb{R}^{n\times n}$ invertible matrix defined as

$$\boldsymbol{B} = (\boldsymbol{I} - \rho\boldsymbol{\beta}) \text{ with } B_{ij} = \begin{cases} 1 & \text{if } i = j \\ -\rho\beta_{(ij)} & \text{if } j \in N_i \\ 0 & \text{otherwise} \end{cases}$$

The matrix $\sum_D \in \mathbb{R}^{n\times n}$ is a diagonal matrix of the form $\sum_D = diag({\sigma_1}^2, {\sigma_2}^2, \ldots, {\sigma_n}^2)$ such that ${\sum_D}^{-1}$ is also diagonal, $\beta_{(ij)}{\sigma_j}^2 = \beta_{(ji)}{\sigma_i}^2$, $i, j \in S$. The CAR model for $\boldsymbol{X}$ in equation (3.26) can be expressed by a probability density function defined as

$$X_i|X_{j\neq i} \sim N\left[\mu_i + \rho \sum_{j\in N_i} \beta_{(ij)}(x_j - \mu_j), {\sigma_i}^2 \right], i \in S \tag{3.28}$$

and the joint probability density function in equation (3.27), which is denoted as follows

$$\mathbf{X} \sim N(\boldsymbol{\mu}, \boldsymbol{B^{-1}}\sum\nolimits_{D}) \tag{3.29}$$

For Equation (3.29) to be a valid joint probability density function, the inferred necessary and sufficient condition is that the covariance matrix should not only be symmetric, but also positive definite. Therefore, a common way to do this is to define a symmetric weighted adjacency matrix $\mathbf{W} = (w_{(ij)})$, where

$$\boldsymbol{W} = (w_{(ij)}) \text{ with } w_{(ij)} = \begin{cases} 1 & \text{if } j = i \\ \varphi(ij) & \text{with } j \in N_i\text{: } \forall i, j \in S,\ w_{ij} = w_{ji} \\ 0 & \text{otherwise} \end{cases}$$

where $\varphi(ij)$ is the measure that quantifies the proximity between areas i and j. If $\varphi(ij)$ =1, then i share a common boundary with region j (neighbours). The measure $\varphi(ij)$ could be the distance between the centroids of regions i and j. Thus, also if $\varphi(ij) = 1$, then j is one of the nearest neighbors of i. Let $\mathbf{W}_D$ be the diagonal adjacency matrix of normalization or standardization given by

$$\boldsymbol{W_D} = diag(w_{(1+)}, w_{(2+)}, \ldots, w_{(n+)}). \tag{3.30}$$

Where

$$w_{(i+)} = \sum_{j \in N_i} w_{(ij)}, i, j \in S, \tag{3.31}$$

then we define a matrix of interaction $\boldsymbol{B}$ the normalized adjacency matrix as

$$\boldsymbol{B} = \boldsymbol{W_D}^{-1}\boldsymbol{W} \text{ with } \beta_{(ij)} = \frac{w_{(ij)}}{w_{(i+)}}, i, j \in S. \tag{3.32}$$

Suppose also that the matrix $\boldsymbol{W}_D$ corresponding to a constant diagonal matrix normalized as (3.32), then we will have

$$\boldsymbol{W_D} = \sigma^2 \boldsymbol{W}_D^{-1} \text{ with } {\sigma_i}^2 = \frac{\sigma^2}{w_{(i+)}}, i \in S, \sigma^2 \in \mathbb{R}^+. \tag{3.33}$$

Hence, the Equation (3.29) can be expressed as follows

$$\boldsymbol{X} \sim \left(\boldsymbol{\mu}, \left[\frac{1}{{\sigma_u}^2}(\boldsymbol{W}_D - \rho \boldsymbol{W})\right]^{-1}\right) \tag{3.34}$$

Thus, for the CAR model the variance ${\sigma_u}^2$ represents spatial variability between the areas or locations while ρ is often referred to as spatial dependence parameter or spatial correlation (Mariella and Tarantino, 2010).

Under these methods, we assume that the log relative risk for each province, is modelled through the equation

$$log(\theta_i) = \alpha_0 + u_i \tag{3.35}$$

thus

$$\hat{\theta}_i = exp(\alpha_0 + u_i) \tag{3.36}$$

where α_0 is the overall relative risk (intercept) and u_i is the spatial structure

random effect. For the spatial structured random effects, we used the intrinsic conditional autoregressive prior, similar to that given given by Besag et al. (1991) as follows

$$\left(u_i|u_j, j \neq i, {\sigma_u}^2\right) \sim N\left(\sum_{j \neq i} \frac{w_i u_j}{w_{ij}}, \frac{{\sigma_u}^2}{w_i j}\right) \tag{3.37}$$

where ${\sigma_u}^2$ is the variance parameter that controls the amount of spatial similarity. The hyperparameter ${\sigma_u}^2$ is the variance. The posterior distribution for the CAR model is then implemented using the Gibbs sampler and MCMC algorithm as given by to Besag et al. (1991).

3.7.2 Convolution Model

Besag et al. (1991) further extended their model by assuming that it is possible to include two random effects, as formulated through the following equation.

$$y_i \sim Poisson(\theta_i E_i) \tag{3.38}$$

The log relative risk is modelled through

$$log(\theta_i) = \alpha_0 + u_i + v_i \tag{3.39}$$

thus

$$\hat{\theta}_i = exp(\alpha_0 + u_i + v_i) \tag{3.40}$$

where α_0 is an overall fixed effect on the log relative risk (intercept), u_i is the spatially structured random effect (correlated heterogeneity) and v_i is spatially

unstructured (uncorrelated heterogeneity) effect. Besag et al. (1991) also assumed that the two random effects are independent and require a specification of independent priors. The prior distribution model for the spatially unstructured v_i is assumed to follow a normal distribution and it is given as

$$v_i \sim N(0, \sigma_v^2). \tag{3.41}$$

The prior distribution for the spatially structured random effects is the CAR model as stated in Equation (3.37).

The parameters ${\sigma_u}^2$ and ${\sigma_v}^2$ control the variability of u_i and v_i respectively (Lawson et al., 2003). The main challenge of this model is that these random effects cannot easily be estimated separately, therefore it is only possible to identify the sum $u_i + v_i$ for each area. In this regard, Besag et al. (1991) argue that using the convolution prior distribution allows the data to decide how much of the residuals for disease risk is due to both spatially structured and spatially unstructured random effects.

For all the fitted models, 50000 MCMC iterations were ran, with an initial burn-in period of 10000 and a thinning of every 10^{th} sample value used to assess the convergence by checking trace plots and autocorrelation plots using the Gelman-Rubin convergence diagnostics. The models were compared using the deviance information criterion (DIC), where the best model was chosen to be the one with the smallest DIC value, as suggested by Spiegelhalter et al. (2002). The outcomes/results obtained from these models are described next.

Table 3.1 depicts the results of the parameter estimates and model comparison statistics for the four different mentioned models.

TABLE 3.1: Posterior estimates with corresponding (95% CI) and model comparison of the Bayesian hierarchical model

Parameters	Poisson-Gamma	Log-Normal	CAR	Convolution model
α_0	-	-0.27(-0.65,-0.03)	-0.26(-0.27,-0.25)	-0.26(-0.32,-0.23)
a	2.53(1.30,4.18)	-	-	-
b	2.65(1.22,4.42)	-	-	-
${\sigma_u}^2$	-	-	1.16(0.84,1.65)	1.14(0.77,1.66)
${\sigma_v}^2$	-	0.71(0.51,1.01)	-	0.07(0.01,0.29)
${\tau_u}^2$	-	-	0.81(0.37,1.42)	0.86(0.37,1.69)
${\tau_v}^2$	-	2.17(0.99-3.84)	-	1052(12.02,4771)
$\bar{D}$	198.76	199.08	201.56	198.35
p_D	18.02	18.32	20.77	17.58
DIC	216.76	217.40	222.33	215.92

Based on this analysis, we deduce that the model with both spatially structured and unstructured random effects using the convolution approach offered a better fit (lowest DIC value of 215.92) as can be seen in the table. The overall log relative risk is α_0=-0.26 (95% CI: -0.32,-0.23) which is significantly different from zero and negative. The precision parameter ${\tau_u}^2$=0.86 (95% CI: 0.37,1.69) which indicates a high degree of similarities of malaria mortality risk between the neighbouring

provinces. The model with unstructured random effects captures the variability inherent within a given province. The precision parameter for the unstructured component estimated as $\tau_v{}^2$=1.052 (95% CI: 12.02,4.771). The variability of malaria mortality risk among the provinces could be explained by factors such as demographic, socio-economic and climatic factors. However the data modelled here did not contain these covariates.

All the main convergence diagnostics can be found in Appendix B. Accordingly, we observed that the convergence for the best fitting model (convolution model) was achieved after 50000 iterations. The two chain trace plots in Appendix B Figure B.1 show convergence in the model. The autocorrelation plots showed that there is a slow correlation among some parameters that constitute the chain and for more information see Appendix B Figure B.4. Through the kernel density plot the parameters revealed some stability and more satisfactory kernel density plot for the parameters of interest which were more bell-shaped than others, although it does not need to be symmetric. See also Appendix B Figure B.2 and B.3 for more information.

Figure 3.1 and 3.2 show the box plots of the malaria relative risk in the 18 provinces of Angola for the spatial structured and spatial unstructured random effects from the best fitting model, respectively. The number codes on the box plot represent the provinces (see Table 3.2) for the corresponding provincial names.

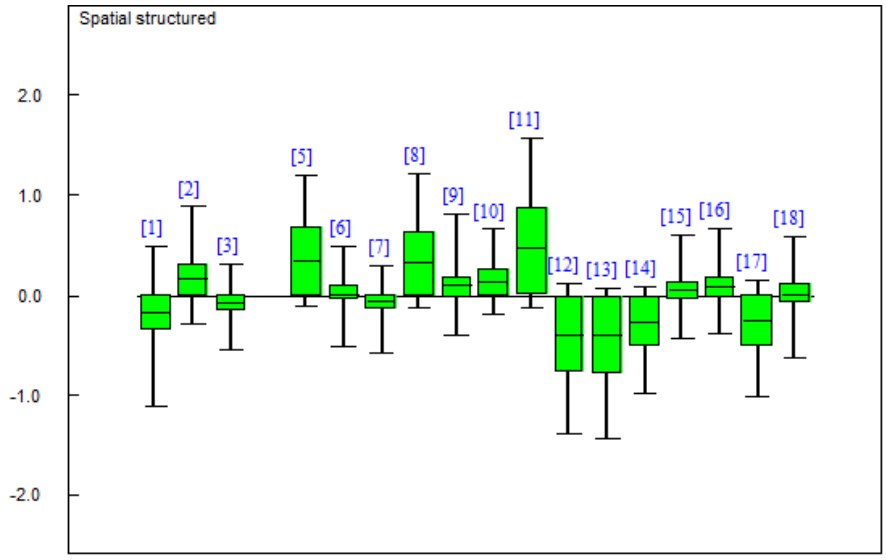

FIGURE 3.1: Spatial structured box plots of malaria mortality risk by provinces based on the best fitting model

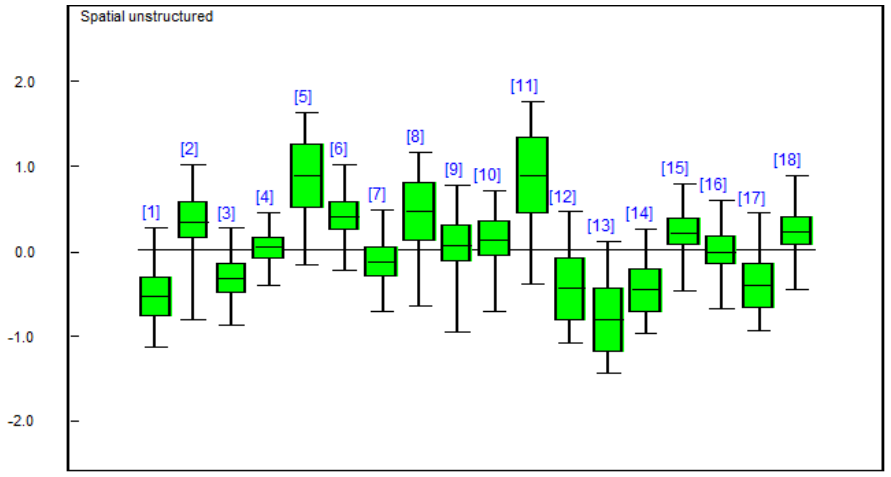

FIGURE 3.2: Spatial unstructured box plots of malaria mortality risk by provinces based on the best fitting model

The spatial unstructured random effects cover covariates that are inherent within locations. The unstructured components are dispersed and mostly a distance from zero although not different from zero. This suggests that effects of the observed covariates inherent within locations vary i.e. they are not homogeneous. Spatially structured random effect on the other hand accounts for any unobserved covariates which vary spatially across the locations under study.

Table 3.2 shows the posterior relative risks and their corresponding 95% credible intervals (CI) for each province in Angola.

TABLE 3.2: Posterior relative risk and corresponding 95% credible interval (CI) based on the best fitting model

Code	Provinces	RR (95% CI)
1	Bengo	0.38(0.35,0.40)
2	Benguela	1.21(1.19,1.23)
3	Bie	0.51(0.50,0.53)
4	Cabinda	0.78(0.74,0.81)
5	Cuando Cubango	2.55(2.50,2.60)
6	Cuanza Norte	1.13(1.10,1.17)
7	Cuanza Sul	0.62(0.61,0.64)
8	Cunene	1.63(1.59,1.67)
9	Huambo	0.86(0.84,0.88)
10	Huila	0.97(0.95,0.99)
11	Luanda	2.87(2.83,2.90)
12	Lunda Norte	0.33(0.31,0.35)
13	Lunda Sul	0.23(0.21,0.24)
14	Malanje	0.37(0.35,0.39)
15	Moxico	0.97(0.93,1.01)
16	Namibe	0.81(0.77,0.85)
17	Uige	0.39(0.38,0.40)
18	Zaire	0.95(0.91,0.99)

From the Table 3.2, we observe that Luanda province has the highest relative risk of individuals dying from malaria with RR: 2.87 (95% CI: 2.83,2.90) and Lunda Sul has the lowest RR of 0.23 (95% CI: 0.21,0.24). At close second is Cuando Cubango with a RR of 2.55 (95%: 2.50,2.60).

Figure 3.3 shows the relative risk map (middle map) of the provinces based on the best fitting model and the associated 95% credible interval (left and right map). These maps show that the country can be divided into four different regions of relative risk which are very low, moderate, slightly high and high risks with respective intervals of RR <0.5, RR:(0.5-1.5), RR:(1.5-2.0) and RR >2.0. Out of the 18 provinces, it can be seen that 2 provinces (Luanda and Cuando Cubango) were

predicted to have the highest malaria relative risk, eleven provinces had a slightly higher to moderate risk (Cunene, Moxico, Bie, Cabinda, Zaire, Namibe, Benguela, Cuanza Norte, Huambo, Cuanza sul and Huila) and five provinces experienced a very low relative risk (Bengo, Uige, Malanje, Lunda Norte and Lunda Sul).

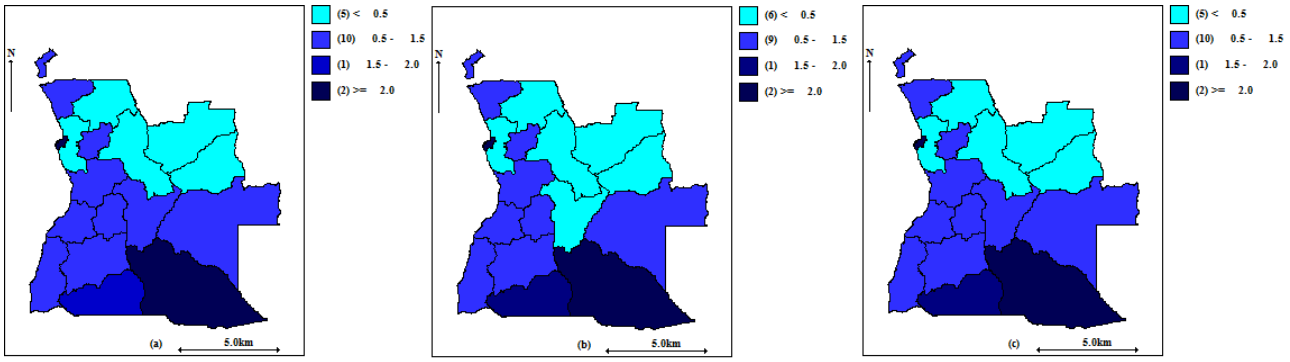

FIGURE 3.3: Maps of mortality risk of malaria in Angola predicted from the best fitting model. (a) mean posterior relative risk, and corresponding lower (b) and upper (c) 95% credible interval

Furthermore, it can be observed that the north, west, south, north-west and central provinces of Angola exhibit a significant slightly high and low disease risk compared to high disease risk in the south-east part of the country.

3.8 Discussion and recommendations

This chapter provided a discussion of the Bayesian hierarchical model, which allow us to estimate the posterior relative risk. We then defined and fitted four different models namely: Poisson-gamma, Poisson log-normal, CAR and convolution model to NMCP data during period of 2003-2010 for the estimation of relative risk at each provincial level in Angola. In order to estimate the model parameters the

MCMC technique was applied. Then the performance for each model was compared through the DIC. However, the DIC results suggest that all the 4 models; Poisson-gamma, Log-normal, CAR and Convolution models had no substantial difference amongst them. The results also revealed that the model with lowest DIC is considered better fit. This implies therefore that the convolution model is considered as the best fitting model for this data. It is important to note that the differences between the values of the DIC for the all models were generally small indicating that the models were fairly good in deriving the relative risk of malaria mortality. We therefore present and discuss the results based on the convolution model which captures both spatially structured and unstructured random effects. The results as depicted on the spatial maps and box plots of relative risk revealed that Luanda and Cuando Cubango experienced the highest relative risk of death due to malaria during the study period. In other provinces, the relative risk was notably lower. Consequently, urgent attention and intervention is needed for decreasing malaria deaths in the two provinces, Luanda and Cuando Cubango.

Chapter 4

Spatio-Temporal Modelling of Malaria among Children in Angola

In this chapter we present the modelling disease risk in space and time. We discuss the spatio-temporal modelling of malaria among children under the age of five years in Angola using the 2006-2007 and 2011 Angola Malaria Indicator Survey data from the DHS program. We discuss and fit suitable models to perform spatio-temporal modelling of malaria among children under the age of five years in Angola. The models that were fitted include the Bayesian logistic model for spatio-temporal analysis in Angola in the years 2006-2007 and 2011. We then extend this method to allow the covariates to vary both in space and time; for the space component we use the conditional autoregressive model, and for time variation we

use the autoregressive model of order 1. The Bayesian logistic model for spatio-temporal analysis is fitted to the data using estimation techniques based on the Integrated Laplace Approximation (INLA) methods. These models for malaria in Angola are then explored, with the results and outcomes being discussed.

4.1 Introduction

Spatio-temporal modelling encompasses a wide range of models that can be used to describe disease distribution in space and its change with time. Such models are a popular tool in disease mapping within the Bayesian hierarchical model framework, because disease mapping seeks to identify similarities and differences across the study area. Spatio-temporal models have been increasingly used in disease mapping in order to identify any simultaneous patterns or trends in space and time. The relevant space and space-time can be either spatially structured, as modelled using CAR model or spatially unstructured using a set of independent random effects.

A number of methods and further studies for modelling spatio-temporal disease risk have been proposed (Assunção et al., 2001; Bernardinelli et al., 1995; Knorr-Held and Besag, 1998; Waller et al., 1997).

MacNab and Dean (2002) proposed a generalised additive mixed model for estimating disease risk, where B-spline smoothing over the temporal dimension provides a flexible means of adjusting the overall time effects, as well as the region-specific

time effects. Martínez-Beneito et al. (2008) developed with an autoregressive approach to spatio-temporal disease mapping method that was based on combining ideas from autoregressive time series in order to link information in time, and by spatial modelling to link information in time and space. Congdon and Southall (2005) also proposed a spatio-temporal modelling method, by means of autoregressive time series models. Schrödle and Held (2011) outlined a spatio-temporal disease mapping model based on the Integrated Nested Laplace Approximation method, as will be used in this dissertation. The particular method we use will be elaborated in the following section.

4.2 Spatio-temporal Methodology

Let y_{ijt} be the malaria status for any given child j at time t:$t = 1, 2$ in province i: $i = 1, 2, \ldots, 18$.

$$y_{ijt} = \begin{cases} 1 & \text{if child } j \text{ in province } i \text{ has malaria at time } t \\ 0 & \text{otherwise} \end{cases}$$

This study assumes the dependent variable y_{ijt} is univariate Bernoulli distributed that is $y_{ijt}|p_{ijt} \sim Bernoulli(p_{ijt})$. The unknown mean response $E(y_{ijt}) = p_{ijt}$ may relate to the independent variable as follows:

$$h(p_{ijt}) = intercept + \underbrace{C + S + T}_{\text{main effects}} + \underbrace{CS + CT + ST + CST}_{\text{interactions terms}}$$

where the intercept terms give the initial amount of risk shared by individuals, in all provinces and time. The function $h(.)$ is the link function in the context of a GLM such as the logit link in our case. The main effects C, S and T represent the covariate, spatial and temporal effects respectively. The second order interaction terms CS, CT and ST represent contribution to the risk due to a combination of main effects that cannot be explained additively by main effects, CST represents the covariate-space-time interaction (López-Quilez and Munoz, 2009).

4.2.1 Spatio-temporal model

In our study we introduce two models namely; a spatial model (model 1) and a spatio-temporal model (model 2). The spatio-temporal model captures covariate space-time interaction (Sun et al., 2000). The models used are;

$$\text{Model 1} \quad : \quad logit(p_{ijt}) = \beta_0 + f(age) + w_{ijt}^T\gamma + v_i \text{ Spatial model}$$

$$\text{Model 2} \quad : \quad logit(p_{ijt}) = \beta_0 + f(age_{jt}) + w_{ijt}^T\gamma_t + v_i \text{ Spatio-temporal model}$$

where

- β_0 is the intercept representing the logit prevalence rate, when all covariates have a zero value
- $f(age)$ represents a function of age. Here age was captured as a continuous and we assume it has a non-linear effect on malaria status

- w_{ijt}^T represents the vector of categorical covariate for child j living in province i at time t, with γ representing the regression coefficients for the spatial model while γ_t represents the time dependent regression coefficients modelled using the autoregressive model of order 1 given by: $X_t = \alpha X_{t-1} + \epsilon_t$ where ϵ_t is a zero mean white noise

- v_i represents the structured random effects in both models.

In Model 1 we do not assume any interaction that is covariate-time, space-time, space-covariate or space-covariate-time interactions. It is a model of continuous covariate age modelled with a random walk model of order 2, as in Lindgren and Rue (2008), which is assumed to have a non-linear effect on malaria prevalence. Categorical covariates are assumed to have a linear effect on malaria status and the structured random effects are modelled using the conditional autoregressive model (CAR) (see Section 3.7.1).

In contrast Model 2, involves space-time and covariate space-time interactions. Here age is modelled as a non-linear function of malaria using the random walk model of order 2, while the other covariates are modelled as functions of time, using the autoregressive process of order 1. The model incorporates the structured random effects, which include any unobserved covariates and vary spatially across the provinces.

4.2.2 Prior parameters estimation

A non-informative normal distribution prior was used for the fixed effects while a random walk model of order 2 was used for the continuous covariate age. The temporal effects were modeled by a first order autoregressive process allowing for correlation between the two time periods and the provinces (Mabaso et al., 2006). The spatial components prior were the CAR model for the structured random effects.

4.2.3 Posterior distribution

The distribution of the parameters after observing the data is the posterior distribution. It is obtained by updating the prior distribution with the observed data. Since our study is fully Bayesian, inference is made by sampling from this posterior distribution. The Markov chain Monte Carlo (MCMC) simulation is the most common approach used to obtain inference for latent Gaussian models. However, this method is slow and performs poorly when applied to such models as discussed in Rue et al. (2009). Instead, the Integrated Nested Laplace Approximation (INLA) criterion is a relatively new technique developed to circumvent these shortfalls (Rue et al., 2009).

4.3 Integrated Nested Laplace Approximation (INLA)

The Integrated Nested Laplace Approximation (INLA) has been proposed by Rue et al. (2009) as a computational tool for Bayesian inference. As an estimation method, INLA is an alternative to the widely used Markov chain Monte Carlo (MCMC) methods. Although, in most cases, similar results can be obtained by either MCMC or INLA inference nonetheless there are fundamental differences in the way that the posterior distribution is estimated. On one hand the MCMC methods can sample directly from a joint posterior distribution, whereas on the other hand INLA uses a close form expression to estimate the marginal posterior distribution. A major advantage of the latter approach is its speed, compared to MCMC simulation; that is INLA estimates the parameters within a short computational time. Moreover, Schrödle et al. (2011) has also investigate the use of the INLA method and showed it produces almost approximately or identical results to the MCMC simulation in a much faster time.

The principle behind INLA lies in estimating the posterior marginal distribution of the parameters in the model, where the model has a latent Gaussian Markov random fields (GMRFs).

Let $\boldsymbol{x}$ be the vector of all Gaussian variables, $\boldsymbol{\theta}$ be the set of hyper parameters and $\boldsymbol{y}$ denote the data. Rue et al. (2009) focus in estimating the posterior distribution for the latent Gaussian model given by

$$\pi(\boldsymbol{x}, \boldsymbol{\theta}|\boldsymbol{y}) \propto \pi(\boldsymbol{\theta})\pi(\boldsymbol{x}|\boldsymbol{\theta}) \prod_{i \in I} \pi(y_i|x_i, \boldsymbol{\theta})$$

$$\propto \pi(\boldsymbol{\theta})|\boldsymbol{Q}(\boldsymbol{\theta})|^{\frac{n}{2}} exp\left(-\frac{1}{2}x^T Q(\boldsymbol{\theta})x + \sum_{i\in I} log\pi(y_i|x_i,\boldsymbol{\theta})\right), \tag{4.1}$$

where Q is the precision matrix of θ. Hence, the posterior marginal distribution of interest can be written as

$$\pi(x_i|\boldsymbol{y}) = \int \pi(x_i|\boldsymbol{\theta},\boldsymbol{y})\pi(\theta|y)\, d\boldsymbol{\theta} \tag{4.2}$$

and

$$\pi(\theta_j|\boldsymbol{y}) = \int \pi(\boldsymbol{\theta}|\boldsymbol{y})\, d\boldsymbol{\theta}_{-j} \tag{4.3}$$

where θ_{-j} is the vector of all parameter components except θ_j. This approach is based on obtaining an approximation to the posterior marginal distribution as it involves the construction of a nested approximation of Equations (4.2) and (4.3),

$$\widetilde{\pi}(x_i|\boldsymbol{y}) = \int \widetilde{\pi}(x_i|\boldsymbol{\theta},\boldsymbol{y})\widetilde{\pi}(\boldsymbol{\theta}|\boldsymbol{y})\, d\boldsymbol{\theta} \tag{4.4}$$

and

$$\widetilde{\pi}(\theta_j|\boldsymbol{y}) = \int \widetilde{\pi}(\boldsymbol{\theta}|\boldsymbol{y})\, d\boldsymbol{\theta}_{-j}. \tag{4.5}$$

Hence, these approximations $\pi(x_i|\boldsymbol{y})$ are computed by approximating $\pi(\boldsymbol{\theta}|\boldsymbol{y})$ and $\pi(x_i|\boldsymbol{\theta},\boldsymbol{y})$ using numerical integration routines to integrate out the hyperparameters $\boldsymbol{\theta}$.

Three steps are involved in the approximation. The first step is to approximate the full posterior $\pi(\boldsymbol{\theta}|\boldsymbol{y})$. To do this a first approximation to $\pi(\boldsymbol{\theta}|\boldsymbol{y})$ is needed

Gaussian distribution can be constructed using the Laplace approximation:

$$\widetilde{\pi}(\boldsymbol{\theta}|\boldsymbol{y}) \propto \frac{\pi(\boldsymbol{x},\boldsymbol{\theta},\boldsymbol{y})}{\widetilde{\pi}_G(\boldsymbol{x}|\boldsymbol{\theta},\boldsymbol{y})}\Big|_{x=x^*(\boldsymbol{\theta})} \tag{4.6}$$

where $\widetilde{\pi}_G(\boldsymbol{x}|\boldsymbol{\theta},\boldsymbol{y})$ denotes the Gaussian approximation to the full conditional distribution of x, and $x^*(\theta)$ is the mode of full conditional distribution of x for a given of θ.

The second step one compute the Laplace approximation of the conditional $\pi(x_i|\theta,y)$ for selected values of θ, given in the following form:

$$\widetilde{\pi}_{LA}(x_i|\boldsymbol{\theta},\boldsymbol{y}) \propto \frac{\pi(\boldsymbol{x},\boldsymbol{\theta},\boldsymbol{y})}{\widetilde{\pi}_G(\boldsymbol{x}_{-i}|\boldsymbol{x}_i,\boldsymbol{\theta},\boldsymbol{y})}\Big|_{x_{-i}=x^*_{-i}(x_i,\theta)} \tag{4.7}$$

where $\widetilde{\pi}_G(\boldsymbol{x}_{-i}|\boldsymbol{x}_i,\boldsymbol{\theta},\boldsymbol{y})$ is a Gaussian approximation to the distribution of $\boldsymbol{x}_{-i}|\boldsymbol{x}_i,\boldsymbol{\theta},\boldsymbol{y}$ and the whole expression is evaluated at $x^*_{-i}(x_i,\theta)$, which is the mode for the given θ.

The third step computes and combines the two full posteriors obtained in the previous steps. Equation (4.4) can therefore be approximated using numerical integration as

$$\widetilde{\pi}(x_i|y) = \sum \widetilde{\pi}(x_i|\theta_k,y)\pi(\theta_i|y)\Delta_k \tag{4.8}$$

where Δ_k represents the weights for each vector of set of values of grid point for θ_k. The INLA approach has been applied in a number of disease mapping approaches, some examples of these implementation can be seen in Ugarte et al. (2014) and Schrödle et al. (2011).

4.4 Application to Angola Malaria Indicator Survey data

The data set for malaria prevalence among children under the age of 5 years using the 2006-2007 and 2011 DHS data was described in Section (2.2). It comprised of socio-economic and demographic factors. The models described in Section (4.1) are the next applied to the data and these analysis were carried out using R statistical software with R-INLA package. In Appendix C we present the R-INLA codes relevant to these models.

Table 4.1 gives the posterior estimates of the odds ratios (OR) and their corresponding 95% credible intervals (CI) for the spatial model.

TABLE 4.1: Odds ratio estimates with corresponding 95% (CI) for spatial the model

Covariates	Children <5 (2006-2007) OR	Children <5 (2011) OR
Residence		
Rural	1	1
Urban	0.21(0.10,0.44)	0.10(0.06,5.67)
Have net		
No	1	1
Yes	0.74(0.57,0.96)	0.65(0.50,0.84)
Wealth index		
Poorest	1	1
Poorer	0.81(0.60,0.93)	0.73(0.96,1.85)
Middle	0.45(0.29,0.69)	0.77(0.90,1.85)
Richier	0.33(0.17,0.59)	0.56(0.39,0.88)
Richest	0.38(0.18,0.73)	0.44(0.27,0.65)
Gender		
Male	1	1
Female	1.02(0.82,1.28)	0.97(0.62,1.03)

The categorical covariates were assumed to have a linear time independent effects on malaria prevalence. All covariates, except gender, showed significant relationship relationships (using CI) with malaria prevalence. For children the odds for those living in urban areas of having malaria in the 2006-2007 and 2011 periods was significantly lower than the corresponding odds for those residing in rural areas, (OR: 0.21, 95% CI: 0.101 to 0.44) and (OR: 0.10 95% CI: 0.06 to 5.67) respectively. Children who had bed nets were also less likely to contract malaria compared to those who did not have nets in 2006/07 and 2011, (OR: 0.74, 95% CI: 0.57 to 0.96) and (OR: 0.65 95% CI: 0.50 to 0.84) respectively. Wealth index had a linear relationship with malaria; the wealthier a child's family, the less likely the child was to contract malaria. In both surveys the ORs decrease as the wealth index rich. With regards to a child's gender there was no significant in malaria prevalence.

4.4.1 The estimated effects of age on malaria

Figure 4.1 show the effects of age on prevalence of malaria. In this study we had assumed that age had a non-linear effect on malaria prevalence for children under the age of five. However, our study revealed that age had a positive linear relationship with malaria prevalence. Thus, the likelihood of malaria infection increases with increasing age. This higher likelihood of increasing malaria prevalence with age was observed among children aged 12 to 59 months (1 to 5 years). It could be due to the cumulative effect of nutrient deficiencies in their diet, which may render the older children more vulnerable to malaria than those below 12 months,

who were more likely to be exclusively breast-fed, and so have greater immunity to infection. In this regard, it has been observed that micro-nutrient deficiencies among infants could impinge on their immune system, increasing their risk for malaria (Nyarko and Cobblah, 2014; Pérez-Escamilla et al., 2009).

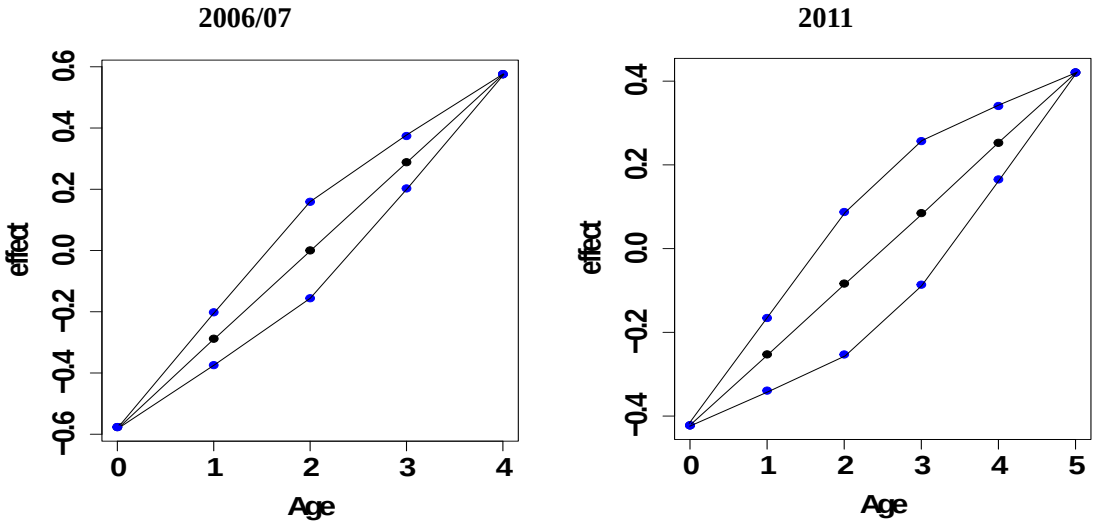

FIGURE 4.1: Effects of age on malaria in 2006-2007 and 2011 in Angola

4.5 Spatio-temporal effects

It is of great importance to study how risk factors and disease prevalence evolve with time. In this section, we discuss briefly the temporal evolution of some risk factors fitted from Model 2 namely place of residence, whether a child slept under a bed net and wealth index.

4.5.1 Malaria prevalence in Angola

The Figure 4.2 displays the spatio-temporal distribution of malaria in years 2006-2007 and 2011 surveys. As can be seen, the prevalence of malaria was high in the hyper-endemic regions in the central parts of the country for both the period 2006-2007 and the year 2011. The highest prevalence at the time of both studies was in the Zaire province. The choropleth maps show that the prevalence of malaria was lower in 2011 than in 2006-2007.

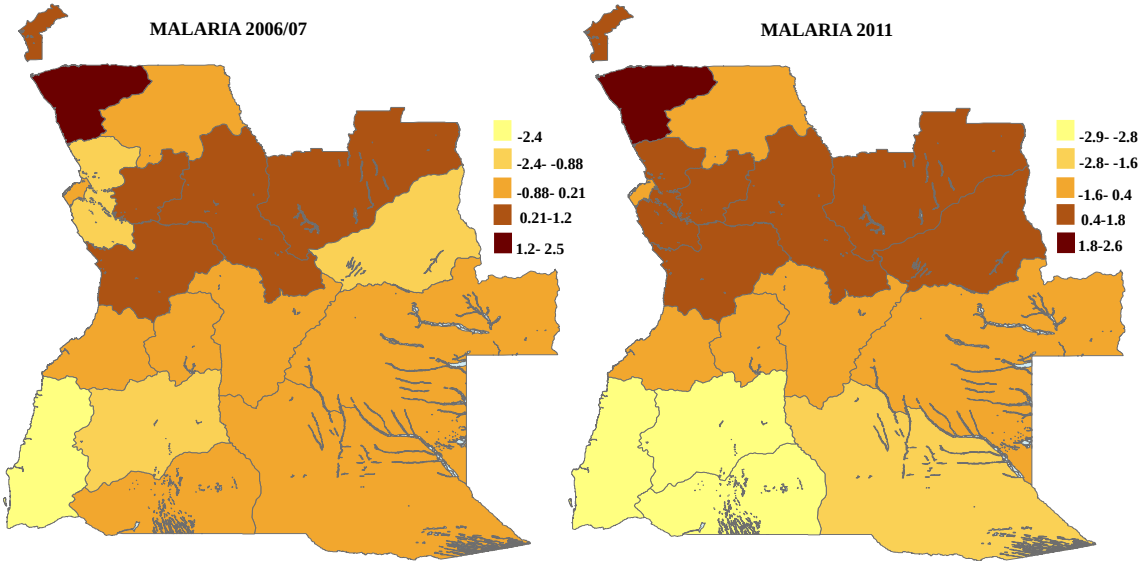

FIGURE 4.2: Spatio-temporal distribution of malaria in Angola

4.5.2 Estimated effects of place of residence on malaria

Figure 4.3 displays the effects of place of residence on malaria prevalence for the two time period.

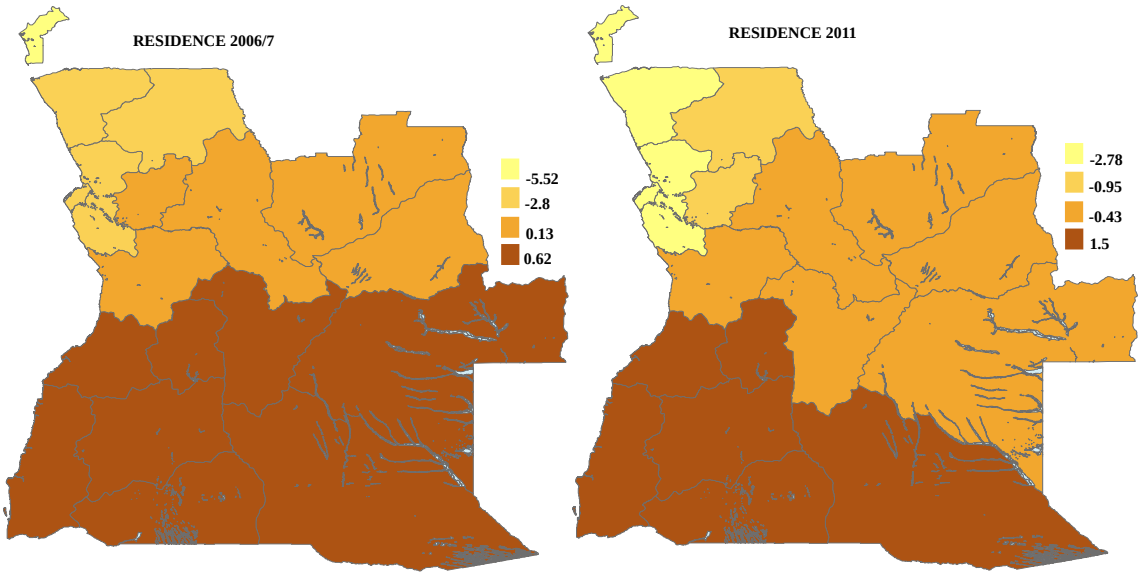

FIGURE 4.3: Effects of place of residence on malaria prevalence for the two time period 2006-2007 (left figure) and 2011 (right figure)

From Figure 4.3 the regions are almost clustered as per malaria endemicity as depicted in Figure 1.1. For the period 2006-2007, the meso-endemic (unstable) and hyper-endemic rural regions had the highest prevalence rates. The 2006-2007 data was gathered between November 2006 and April 2007 when the meso-endemic (stable) region experienced the highest malaria transmission rate (AMIS, 2007). The later 2011 study was conducted between January and May 2011, which is the period of highest transmission for meso-endemic (unstable) regions (AMIS, 2012). The hyper-endemic regions experience a high transmission rate all year, but with the highest transmission rates occurring between November and January (AMIS, 2007). Nevertheless, from the Figure 1.1 it can be seen that some urban regions such as Luanda still recorded high rates of malaria. This indicates that there are other underlying factors such as climate, that drive different malaria transmission among region/provinces. This implies that the fight against malaria is not the same everywhere, and different public health strategies may need to be developed

for each region/province.

4.5.3 Estimated effects of mosquito nets on malaria

Figure 4.4 depicts the effects of using mosquito nets for children in Angola in 2006-2007 and 2011.

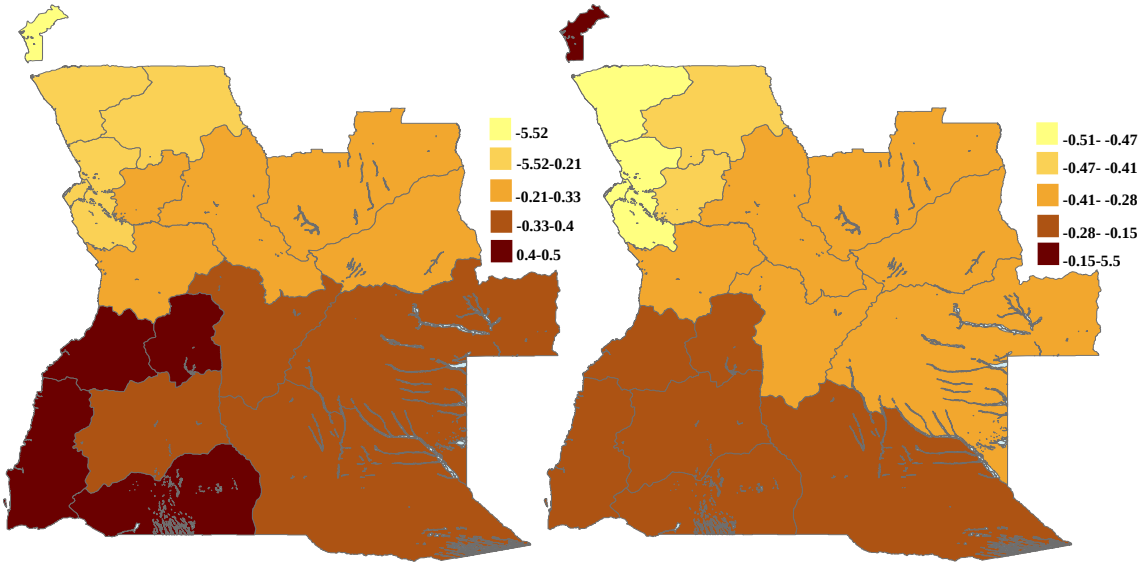

FIGURE 4.4: Effects of mosquito nets on malaria prevalence for the two time period 2006-2007 (left figure) and 2011 (right figure)

The effect of mosquito nets was high in the hyper-endemic and the meso-endemic unstable region. This coincided with the period in time where malaria transmission was highest in these regions; that is between November 2006 and April 2007 and between January and May 2011. This suggests some awareness among the individuals who were vulnerable (or their parents) on the methods of reducing their vulnerability when it was highest. The choropleth maps depict an increased effect of using bed nets on malaria prevalence in 2011 as compared to 2006-2007.

4.5.4 Estimated effects of Wealth Index on malaria

As discussed earlier in Table 4.1, wealth has a negative linear relationship with malaria prevalence. Figure 4.5 shows the effect of wealth on malaria prevalence. The effect of wealth was higher around Luanda, Bengo and Zaire, which are more urbanized and hence have higher standards of living. The effect of wealth was least in the South eastern part of the country. These are some of the poorest region in the country.

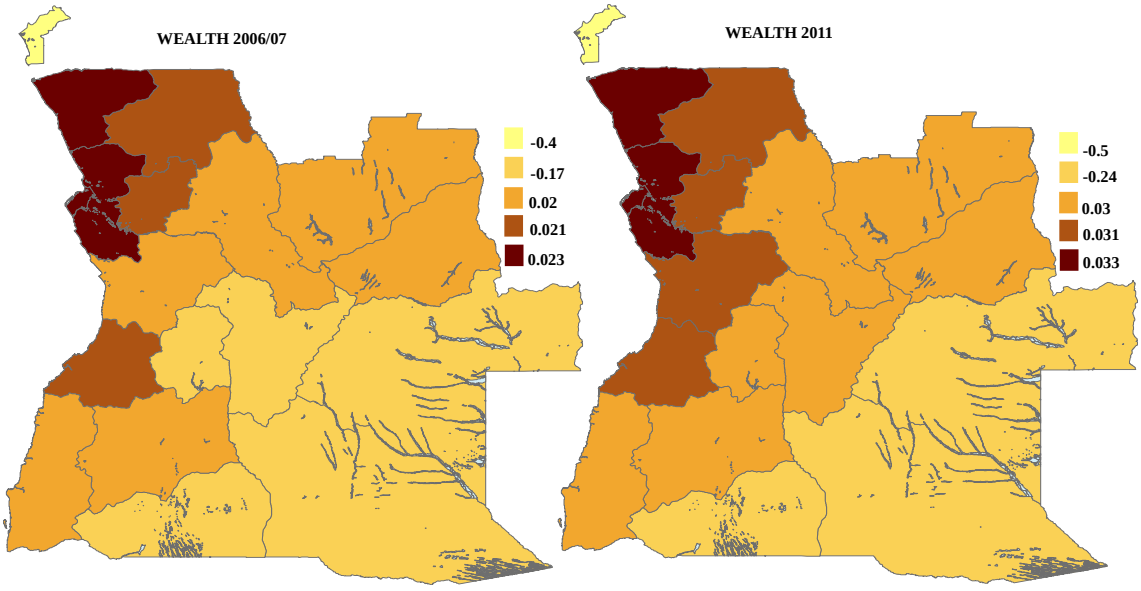

FIGURE 4.5: Effects of wealth index on malaria prevalence for the two time period 2006-2007 (left figure) and 2011 (right figure)

4.6 Discussion

In Chapter 4 we have presented present a full Bayesian approaches to perform a spatial and spatio-temporal modelling of malaria prevalence in Angola among children under the age of five using the 2006-2007 and 2011 Angola malaria indicator survey (AMIS) data. We assumed that all the covariates had a linear effect on malaria prevalence except age which was assumed to have a non-linear relationship and was modeled using random walk 2 model (Lindgren and Rue, 2008). The study nevertheless found that age had a positive linear relationship with malaria prevalence (see Figure 4.1). We speculate that a reason for this could be in children under 1 being more likely to still be breast fed and therefore having stronger immune systems. With increasing age, they are more likely to be weaned as breast feeding is reduced or stopped altogether, and this reduces their immunity, thus making them more vulnerable to malaria. Furthermore, a study by United Nations Children's Emergency Fund (UNICEF) reported a high prevalence of stunted among children under five years old in Angola. This was compounded by the reliance of mothers from rural areas on traditional healers or untrained nursing staff to take care of their children, which could worsen the health of the children's health (UNICEF, 2013).

Wealth was found to have a negative linear relationship with malaria prevalence. In other words, people from a more economically privileged were significantly less likely to be parasitaemic, and significantly more likely to obtain anti-malarial protection drugs when needed, than those in the middle or poor income groups (Njau et al., 2006). It can be interpreted as wealthier individuals being able to

afford better quality medical care, as well as being better able to afford mitigation against malaria, such as treated bed nets or insecticide sprays.

Children living in urban areas were less likely to contract malaria than those living in rural areas. We can suggest three possible reasons for this. One is that individuals living in urban areas are frequently better off economically than their rural counterparts. Another factor is the effect of urbanization on some species of mosquitoes, which affects their diversity, numbers and survival rates, infection rates and the frequency with which they bite people (Hay et al., 2005). All of these can reduce the prevalence of malaria in urban regions. Children who had bed nets were also less likely to get malaria than those who were without nets. This is consistent with other studies, such as (Magalhães et al., 2012; Yusuf et al., 2010). A study conducted by Lengeler (2004) also show that the use of insecticide treated nets (ITNs) are very important tools for controlling malaria in Africa. This is due to the fact that the principal malaria vectors, the Giles *Anopheles gambiae* and *Anopheles funestus* species complexes (White, 1974) primarily feed indoors at night (Pates and Curtis, 2005). Bed nets will therefore reduce exposure and hence transmission.

Spatial effects in the model account for unobserved effects that represent those variables and vary spatially. Identifying high prevalence areas may help in informing intervention strategies for those regions. Figure 4.2 showed the spatial temporal distribution of malaria prevalence in Angola in the years 2006-2007 and 2011. These patterns are important in that they show how malaria prevalence is changing over time and may help unmask new patterns and risk factors. The

prevalence of malaria was high in the hyper-endemic regions in the central parts of the country both in the year 2006-2007 and 2011. The highest prevalence at the time of both studies was in Zaire province. The maps depicts a reduction in malaria prevalence in 2011 as compared to the year 2006-2007. These findings are similar to one reported in (Gosoniu et al., 2010).

The spatio-temporal model in this study extends the existing spatio-temporal models by combining both spatio-temporal and spatially varying coefficients. This helps in observing not only how the disease changes over time but also the effects of the coefficients on the disease change over time.

Chapter 5

Conclusion and Future Research

The aim of the study in this dissertation was understanding and novel application of existing developing statistical methods that account for spatial and spatio-temporal variation in disease mapping with real application to malaria mortality data in Angola (2003-2010) and DHS data (2006-2007 and 2011). To achieve the aims of this study, the specific objectives were set as

- to identity suitable spatial hierarchical models for the estimation of relative risk at provincial level in Angola using malaria mortality from NMCP (2003-2010)
- to develop spatial temporal models for modelling and mapping of malaria among children under of age of 5 years
- to assess the relationship between malaria transmission and some selected demographic socio-economic covariates at provincial level in Angola using the 2006-2007 and 2011 AMIS data.

The findings of this research are chapter specific, and were discussed at length in the relevant chapters. In this chapter, we summarize the results, reflect on our methodology and suggest areas for future studies.

A comprehensive data description and baseline characteristics of the study was introduced in Chapter 2. The crude rates were calculating taking into account the population size of each province. It was established that out of the 18 provinces, Luanda was found to have the highest crude mortality rate while Lunda sul had the lowest crude mortality rates.

In Chapter 3 the Bayesian hierarchical models (BHM) were used to estimate the relative risk (RR) in each province using the malaria mortality data in Angola. Four different models based on the BHM approach were fitted to the data. The DIC model selection criterion was used to compare the models. The criterion indicated that the convolution model that catered for both the spatially structured and spatially unstructured random effects provided the best fitting model. The result and the disease mortality risk maps obtained from the convolution model showed that there are clearly clusters of similar risk that can be defined across the country as depicted in Figure 3.3. A significant spatial variation is also evident in the map with the highest mortality in south eastern part of the country. Our models also revealed that nothing changed between 2003 and 2010 as far as malaria relative risk is concerned.

The results suggest that interventions for reduction, if not eradication of malaria, should focus on such areas with a high mortality risk is the way to go. Other

models may also be considered in estimating relative risk. In particular results from frequentist approach would be particular interesting.

In Chapter 4 we fitted a full Bayesian model to perform a spatial and spatio-temporal modeling of malaria prevalence in Angola among children under the age of five using the 2006-2007 and 2011 Angola malaria indicator survey (AMIS) data. The findings from that model show that some of the risk factors for malaria among children under 5 years change with time, when we compare their effects between the two survey waves of 2006-2007 and 2011. Thus model was also able to reveal the spatial and temporal evolution of the risk factors as well as the malaria prevalence in Angola. This points to the changing dynamics and effects of the socio-economic factors due to reasons including intervention strategies. However, this study did not consider climatic factors which are considered to be major determinants of inter-annual variation in malaria prevalence. This was due to the unavailability of climatic data at the time of the analysis but future studies should incorporate such information.

The analysis presented in this study also shows how spatial and spatio-temporal modeling can be useful to in identifying the key determinant of malaria transmission in Angola among socio-economic and demographic factors. In attempt to determinate the key of malaria, this study found wealth had a negative relationship with malaria prevalence, as increase wealth decrease malaria while age was found to have a positive linear relationship with malaria prevalence which is indicative that this covariate play an important role in contract malaria. In addition, children living in urban areas and those who had bed nets were less likely to

contract malaria than those who lived in rural areas and those who did not have bed nets.

Further studies could be carried out to explore the effects of climatic factors and investigate their effects and association in predicting malaria mortality and prevalence and enhance the findings in Chapter 3 and 4. Further studies also may incorporate weights or missing data analysis techniques to cater for those cases with missing values. This will ensure all available data is used in the analysis and modelling. Since malaria among children depends on their immunity to some extent, studies on relationship between immunity and malaria may help in informing strategies of preventing malaria transmission and hence mortality. It is hoped that the findings of the current work will still be useful to public matter policy makers in their efforts of combat malaria in Angola.

Appendix A

WinBuGS Codes for Chapter 3

Models

```
#################### POISSON-GAMMA MODEL ####################

model

{

for (i in 1:N)

{

# Poisson likelihood for observed counts

O[i]~dpois(mu[i])

mu[i]<-E[i]*theta[i]

# Relative Risk

theta[i]~dgamma(a,b)

}

# Prior distributions for "population" parameters
```

```
a~dexp(0.1)

b~dexp(0.1)

# Population mean and population variance

mean<-a/b

var<-a/pow(b,2)

}

#Data

#Initials

list(a=10,b=10)

list(a=10,b=10)


##################### LOG-NORMAL MODEL ####################

model

{

for (i in 1 : N) {

v[i]~dnorm(0.0,tau.v)

O[i]  ~ dpois(mu[i])

log(mu[i])<-log(E[i])+alpha0+v[i]

RR[i] <- exp(alpha0+v[i]) # Area-specific relative risk (for maps)

}

# Other priors:

alpha0  ~ dflat()

tau.v  ~ dgamma(0.5, 0.0005) # prior on precision

sigma.v <- sqrt(1 / tau.h) # standard deviation
```

```
}

#Data

#Initials

list(tau.v=1, alpha0=0,v=c(0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0))

list(tau.v=1, alpha0=0,v=c(0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0))


#################### CAR MODEL ####################

model {

# Likelihood

for (i in 1 : N) {

O[i]  ~ dpois(mu[i])

log(mu[i]) <- log(E[i]) + alpha0+u[i]

RR[i] <- exp(alpha0 +u[i]) # Area-specific relative risk (for maps)

}

# CAR prior distribution for random effects:

u[1:N] ~ car.normal(adj[], weights[], num[], tau.u)

for(k in 1:sumNumNeigh) {

weights[k] <- 1

}

# Other priors:

# Other priors:

alpha0  ~ dflat()

tau.u  ~ dgamma(0.5, 0.0005) # prior on precision

sigma.u <- sqrt(1 / tau.u)   # standard deviation
```

```
}

#Data

#Initials

list(tau.u = 1, alpha0 = 0, u=c(0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0))

list(tau.u = 1, alpha0 = 0, u=c(0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0))


#################### CONVOLUTION MODEL ####################

model {

# Likelihood

for (i in 1 : N) {

O[i]  ~ dpois(mu[i])

log(mu[i]) <- log(E[i]) + beta +u[i]+v[i]

RR[i] <- exp(beta+u[i]+v[i]) #Area-specific relative risk (for maps)

v[i]~dnorm(0.0,tau.v)

}

# CAR prior distribution for random effects:

u[1:N] ~ car.normal(adj[], weights[], num[], tau.u)

for(k in 1:sumNumNeigh) {

weights[k] <- 1

}

# Other priors:

beta~ dnorm(0.0, 1.0E-5)

tau.u~dgamma(0.5, 0.0005) # prior on precision

sigma.u <- sqrt(1 / tau.u) # standard deviation
```

```
tau.v  ~ dgamma(0.5, 0.0005) # prior on precision

sigma.v <- sqrt(1 / tau.v)   # standard deviation

}

#Data

#Initials

list(tau.u=1, tau.v = 1, beta =0,

u=c(0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0),

v=c(0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0))

list(tau.u = 1, tau.v = 1, beta=0,

u=c(0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0),

v =c(0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0))
```

Appendix B

Markov Chain Monte Carlo Diagnostics

In this section we present the diagnostics tests predicted from Markov Chain Monte Carlo based on the best fitting model. The dynamic trace plots showed a quite dense chains, meaning a good mixing of the parameters. The chains appeared to reached the convergence. Further diagnostics tests Kernel density revealed a clearly smooth and unimodel shapes, and the autocerreletion plots showed that there is no existence of autocorrelation and low correlation was seen among the parameters

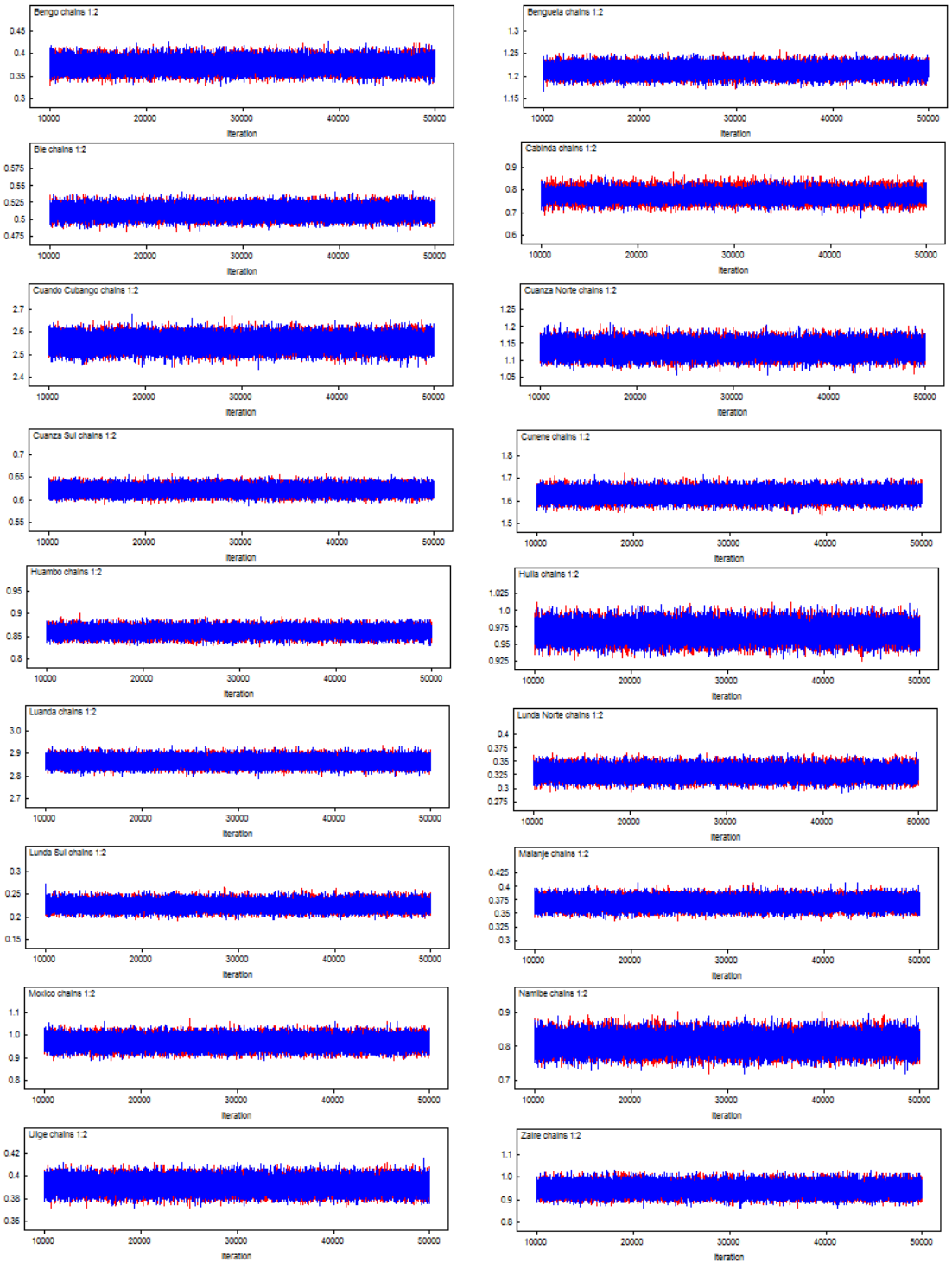

FIGURE B.1: Dynamic trace plots at each provincial level based on the best fitting model

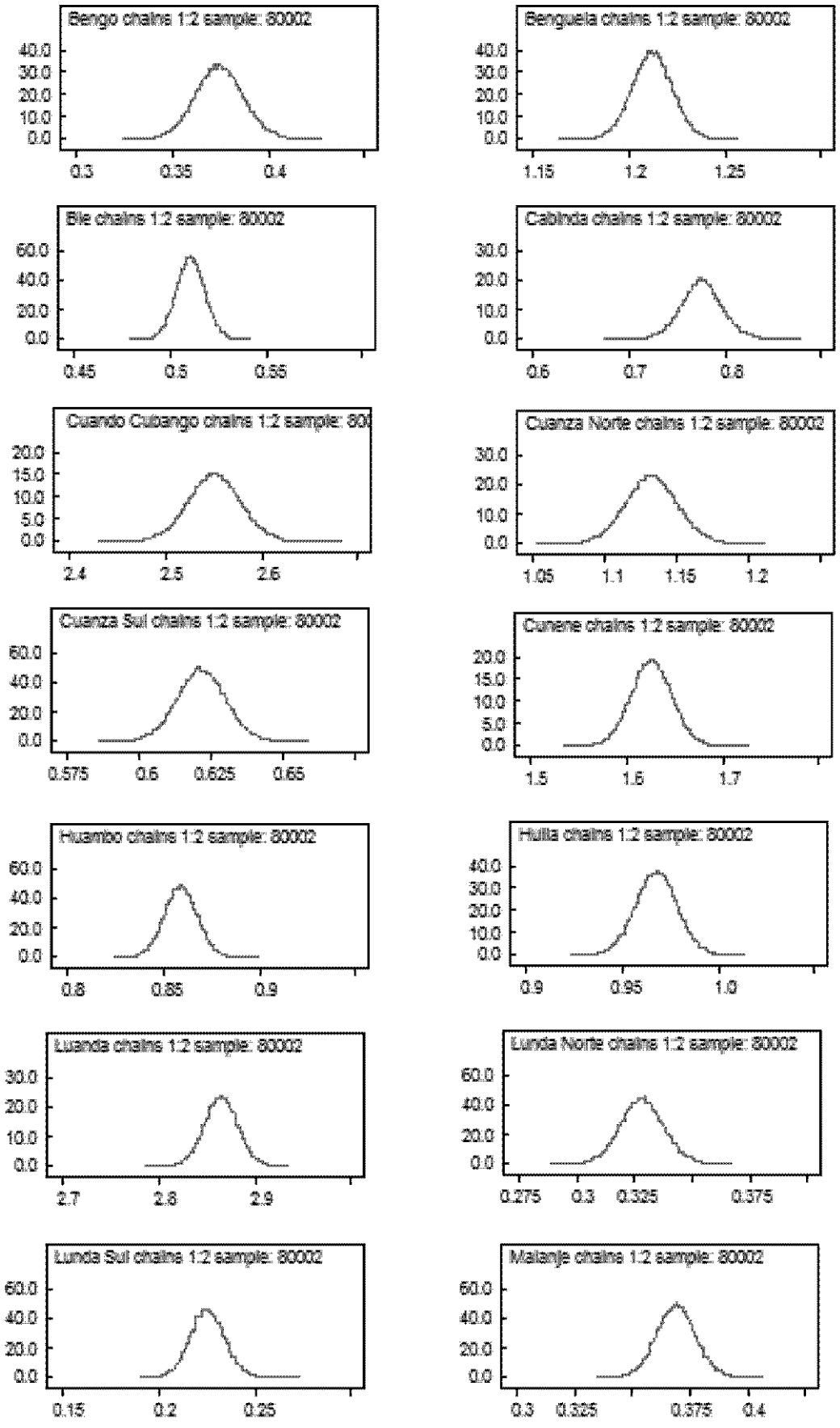

FIGURE B.2: Posterior Kernel density plots at each provincial level based on the best fitting model

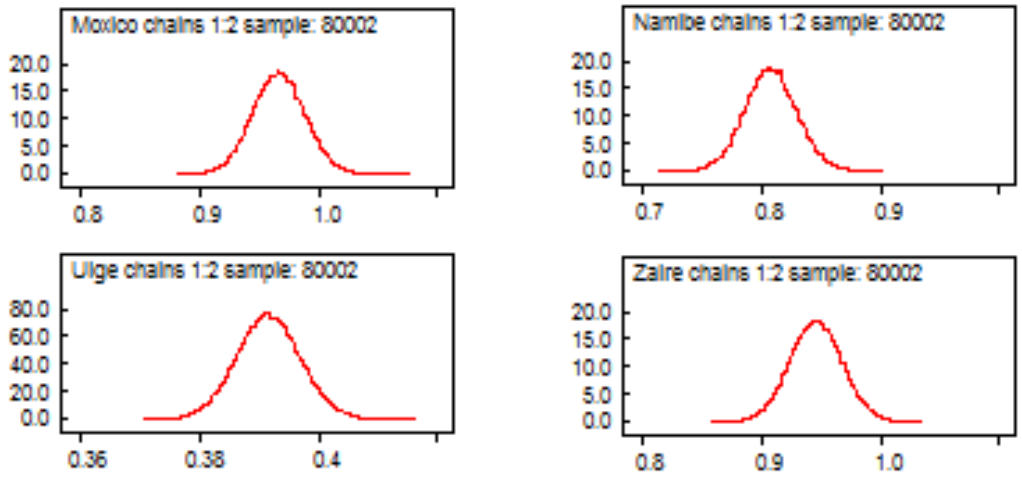

FIGURE B.3: Posterior Kernel density plots at each provincial level based on the best fitting model (continued)

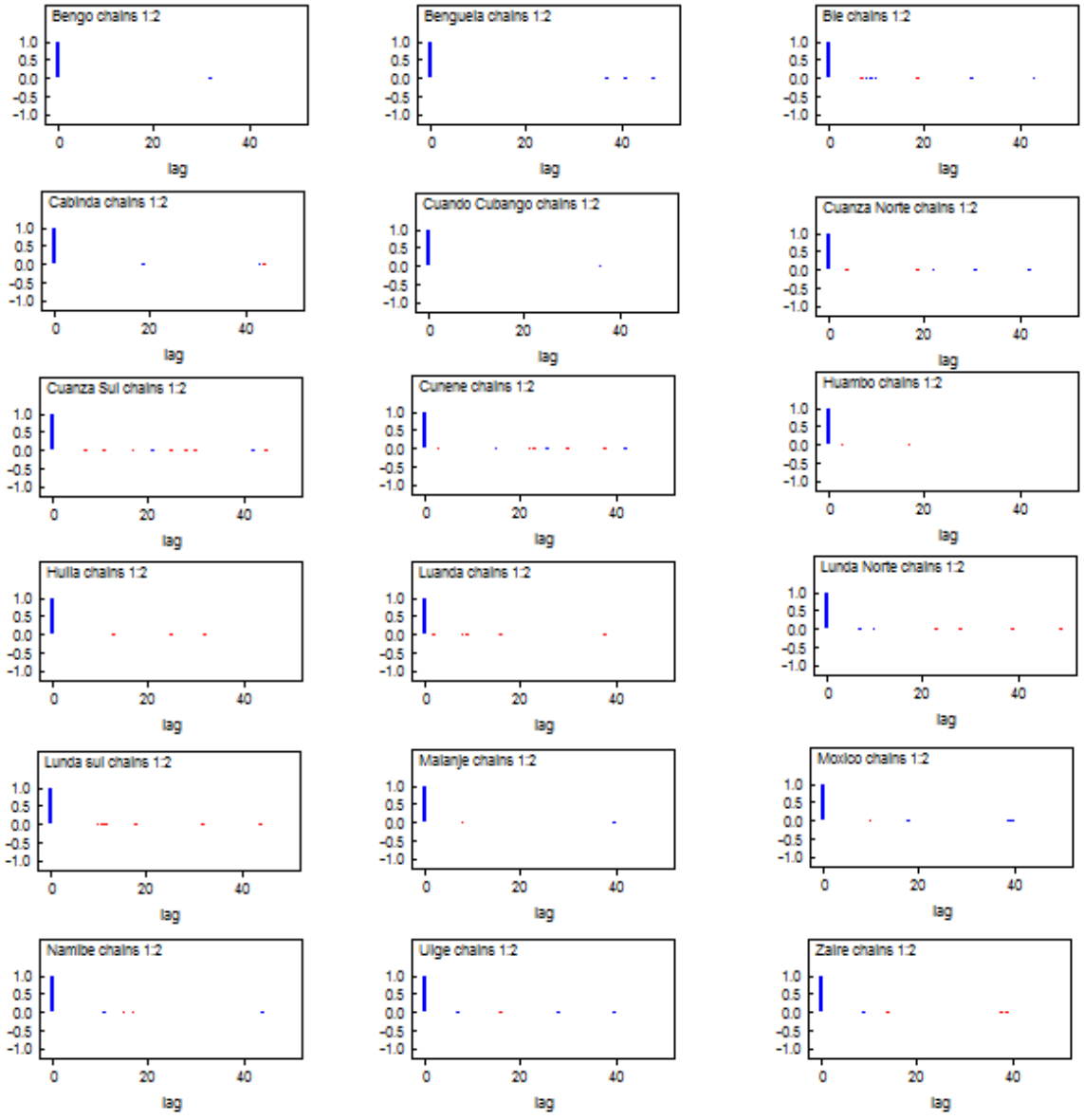

FIGURE B.4: Autoregression plots at provincial level based on the best fitting model

Appendix C

INLA R Code for Chapter 4

```
library("MASS")

library("lattice")

library("ctv")

library("sp")

library(maptools)

library(rgdal)

library(spdep)

require(INLA)

#angola<-read.csv

#angola07<-read.csv("maldata07.csv")

#head(angola07)

#tail(angola07)

#angola07$prov1=angola07$Province

#angola07$prov2=angola07$Province
```

```
#angola07<-read.csv("maldata11.csv")

plot(ang.graph)

#angola11<-read.csv("maldata11.csv")

plot(ang.graph)

adjang<-poly2nb(ang.graph) #Creates adjacency for Angola

formula<-Malaria~1+as.factor(Residence)+

as.factor(Net)+as.factor(Wealth)+as.factor(Gender)+

f(Age,model="rw2")+f(prov1,model="besag",

graph="ang.graph",adjust.for.con.comp = FALSE)+

f(prov2,model="iid",graph="ang.graph",adjust.for.con.comp = FALSE)

result0<-inla(formula,family="binomial",data=angola07,

control.compute=list(dic=TRUE,cpo=TRUE))

summary(result0)


#################### NON LINEAR EFFECT OF AGE ####################

age=result0$summary.random$Age

age

plot(result0$summary.random$Age$ID,result0$summary.random$

Age[,5],xlab="Age",ylab="effect")

par(new=TRUE)

plot(result0$summary.random$Age$ID,result0$summary.random$

Age[,4],col="blue",ann=FALSE, axes=F)

par(new=TRUE)

plot(result0$summary.random$Age$ID,result0$summary.random$
```

```
Age[,6],col="blue",ann=FALSE,axes=F)

#################### SPATIAL EFFECTS ####################

sapt07=result0$summary.random$prov1

sapt07

sapt207=result0$summary.random$prov2

sapt207

#################### MALARIA 2011 ####################

angola11<-read.csv("maldata11.csv")

head(angola11)

tail(angola11)

angola11$prov11=angola11$Province

angola11$prov22=angola11$Province

formula<-Malaria~1+as.factor(Residence)+

as.factor(Net)+as.factor(Wealth)+as.factor(Gender)+

f(Age,model="rw2")+f(prov11,model="besag",

graph="ang.graph",adjust.for.con.comp = FALSE)+

f(prov22,model="iid",graph="ang.graph",adjust.for.con.comp = FALSE)

result1<-inla(formula,family="binomial",data=angola11,

control.compute=list(dic=TRUE,cpo=TRUE))

summary(result1)

#################### NON LINEAR EFFECT OF AGE ####################

age=result1$summary.random$Age
```

```
age

plot(result1$summary.random$Age$ID, result1$summary.random$

Age[,5],xlab="Age",ylab="effect")

par(new=TRUE)

plot(result1$summary.random$Age$ID, result1$summary.random$

Age[,4],col="blue",ann=FALSE,axes=F)

par(new=TRUE)

plot(result1$summary.random$Age$ID, result1$summary.random$

Age[,6],col="blue",ann=FALSE,axes=F)


#################### SPATIAL EFFECTS ####################

sapt11=result1$summary.random$prov11

sapt11

sapt211=result1$summary.random$prov22

sapt211


#################### SPATIO-TEMPORAL ####################

#angola<-read.csv("NEWangola.csv")

angola<-read.csv(ANGOLNEW.csv")

head(angola)

angola$prov3=angola$Province

angola$prov33=angola$Province

angola$prov4=angola$Province

angola$prov5=angola$Province

angola$prov6=angola$Province
```

```
angola$prov7=angola$Province

angola$prov8=angola$Province

#formula<-Malaria~f(Province,Net,model="besag",

graph="ang.graph",

adjust.for.con.comp = FALSE,group=time,

control.group=list(model="ar1"))

#formula<-Malaria~f(Province,Wealth,model="besag",

graph="ang.graph",

adjust.for.con.comp = FALSE,group=time,

control.group=list(model="ar1"))

#mm=table(angola$Net)

#mm


#formula=Malaria~f(Net,model="besag",graph="ang.graph",

adjust.for.con.comp = FALSE,group=time,

control.group=list(model="ar1"))+

f(Wealth,model="besag",graph="ang.graph",

adjust.for.con.comp = FALSE,group=time,

control.group=list(model="ar1"))+

f(Residence,model="besag",graph="ang.graph",

adjust.for.con.comp = FALSE,group=time,

control.group=list(model="ar1"))+

f(Gender,model="besag",graph="ang.graph",

adjust.for.con.comp = FALSE,group=time,
```

```
control.group=list(model="ar1"))+

f(Age,model="rw2")+f(prov3,model="besag",

graph="ang.graph",group=time,

adjust.for.con.comp = FALSE)+f(prov33,model="iid",

graph="ang.graph",group=time,

adjust.for.con.comp = FALSE)


#formula<-Malaria~f(Residence,model="besag",

graph="ang.graph",

adjust.for.con.comp = FALSE,group=time,

control.group=list(model="ar1"))+

f(Net,model="besag",graph="ang.graph",

adjust.for.con.comp = FALSE,group=time,

control.group=list(model="ar1"))+

f(Wealth,model="besag",graph="ang.graph",

adjust.for.con.comp = FALSE,group=time,

control.group=list(model="ar1"))+

f(Gender,model="besag",graph="ang.graph",

adjust.for.con.comp = FALSE,group=time,

control.group=list(model="ar1"))+

f(Age,model="rw2")+f(prov3,model="besag",

graph="ang.graph",

adjust.for.con.comp = FALSE)+f(prov33,model="iid",

graph="ang.graph",
```

```
adjust.for.con.comp = FALSE)

#formula=Malaria~f(Net,model="besag",
graph="ang.graph",
adjust.for.con.comp = FALSE,group=time,
control.group=list(model="ar1"))+
f(Wealth,model="besag",graph="ang.graph",
adjust.for.con.comp = FALSE,group=time,
control.group=list(model="ar1"))+
f(Residence,model="besag",graph="ang.graph",
adjust.for.con.comp = FALSE,group=time,
control.group=list(model="ar1"))+
f(Gender,model="besag",graph="ang.graph",
adjust.for.con.comp = FALSE,group=time,
control.group=list(model="ar1"))+
f(Age,model="rw2")+f(prov3,model="besag",
graph="ang.graph",group=time,
adjust.for.con.comp = FALSE)+f(time,model="ar1")

formula=Malaria~f(Net,model="besag",
graph="ang.graph",
djust.for.con.comp = FALSE,group=time,
control.group=list(model="ar1"))+
f(Wealth,model="besag",graph="ang.graph",
adjust.for.con.comp = FALSE,group=time,
```

```
control.group=list(model="ar1"))+
f(Residence,model="besag",graph="ang.graph",
adjust.for.con.comp = FALSE,group=time,
control.group=list(model="ar1"))+
f(Gender,model="besag",graph="ang.graph",
adjust.for.con.comp = FALSE,group=time,
control.group=list(model="ar1"))+
f(Age,model="rw2")+f(prov3,model="besag",
graph="ang.graph",group=time,
adjust.for.con.comp = FALSE)

formula=Malaria~f(Net,model="besag",
graph="ang.graph",
adjust.for.con.comp = FALSE,group=time,
control.group=list(model="ar1"))+
f(Wealth,model="besag",graph="ang.graph",
adjust.for.con.comp = FALSE,group=time,
control.group=list(model="ar1"))+
f(Residence,model="besag",graph="ang.graph",
adjust.for.con.comp = FALSE,group=time,
control.group=list(model="ar1"))+
f(Age,model="rw2")+f(prov3,model="besag",
graph="ang.graph",group=time,adjust.for.con.comp = FALSE)
result3<-inla(formula,family="binomial",data=angola,
```

```
control.compute=list(dic=TRUE,cpo=TRUE))

summary(result3)


#################### RESULTS ###################

nnet=result3$summary.random$Net

nnet

WWEALTH=result3$summary.random$Wealth

WWEALTH

RESD=result3$summary.random$Residence

RESD

GEN=result3$summary.random$Gender

GEN

Spastruc=result3$summary.random$prov3

Spastruc

AGE=result3$summary.random$Age

AGE

dd<-result3$summary.random$Residence

ddd=dd$"0.5quant"

ddd

mm=ddd[1:18]

mm

zz=ddd[19:36]

ang.graph$welt<-zz

spplot(ang.graph,"welt", col.regions=bpy.colors(20))
```

Bibliography

Alonso, P. L., Lindsay, S. W., Armstrong, J., de Francisco, A., Shenton, F., Greenwood, B., Conteh, M., Cham, K., Hill, A., David, P., et al. (1991). The effect of insecticide-treated bed nets on mortality of Gambian children. *The Lancet*, 337(8756):1499–1502.

AMIS (2007). Angola malaria indicator survey 2006-2007.

AMIS (2012). Angola malaria indicator survey 2011.

Assunção, R. M., Reis, I. A., and Oliveira, C. D. L. (2001). Diffusion and prediction of Leishmaniasis in a large metropolitan area in Brazil with a Bayesian space–time model. *Statistics in Medicine*, 20(15):2319–2335.

Berg, A., Meyer, R., and Yu, J. (2004). Deviance information criterion for comparing stochastic volatility models. *Journal of Business & Economic Statistics*, 22(1):107–120.

Bernardinelli, L., Clayton, D., Pascutto, C., Montomoli, C., Ghislandi, M., and Songini, M. (1995). Bayesian analysis of space-time variation in disease risk. *Statistics in Medicine*, 14(21-22):2433–2443.

Bernardinelli, L. and Montomoli, C. (1992). Empirical Bayes versus fully Bayesian analysis of geographical variation in disease risk. *Statistics in Medicine*, 11(8):983–1007.

Besag, J. (1974). Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society. Series B (Methodological)*, Vol. 36, No. 2:192–236.

Besag, J., York, J., and Mollié, A. (1991). Bayesian image restoration, with two applications in spatial statistics. *Annals of the Institute of Statistical Mathematics*, 43(1):1–20.

Best, N., Richardson, S., and Thomson, A. (2005). A comparison of Bayesian spatial models for disease mapping. *Statistical Methods in Medical Research*, 14(1):35–59.

Bi, Y. (2013). Impact of socio-ecological variability on the transmission of malaria in Yunnan Province, China. *PhD by publication Queensland University of Technology*.

Bivand, R. S., Pebesma, E., and Gomez-Rubio, V. (2013). Spatial Data Import and Export. In *Applied Spatial Data Analysis with R*, pages 83–125. Springer.

Clayton, D. and Kaldor, J. (1987). Empirical Bayes estimates of age-standardized relative risks for use in disease mapping. *Biometrics.*, 43:671–681.

Clement, E. (2014). Small area estimation with application to disease mapping. *International Journal of Probability and Statistics*, 3(1):15–22.

Clements, A., Barnett, A. G., Cheng, Z. W., Snow, R. W., and Zhou, H. N. (2009). Space-time variation of malaria incidence in Yunnan province, China. *Malaria Journal*, 8(180):10–1186.

Congdon, P. and Southall, H. (2005). Trends in inequality in infant mortality in the north of England, 1921–1973, and their association with urban and social structure. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 168(4):679–700.

Congdon, P. D. (2010). *Applied Bayesian hierarchical methods*. CRC Press.

Cressie, N. (1993). Statistics for spatial data: Wiley series in probability and statistics. *Wiley-Interscience New York*, 15:16.

Currie, C. S., Williams, B. G., Cheng, R. C., and Dye, C. (2003). Tuberculosis epidemics driven by HIV: is prevention better than cure? *AIDS*, 17(17):2501–2508.

Dale, P., Sipe, N., Anto, S., Hutajulu, B., et al. (2005). Malaria in Indonesia: a summary of recent research into its environmental relationships. *Southeast Asian Journal of Tropical Medicine and Public Health*, 36(1):1.

DOH (2003). Department of Health. Guidelines for the prevention of malaria in South Africa. *National Department of Health*.

Freeman, T. and Bradley, M. (1996). Temperature is predictive of severe malaria years in Zimbabwe. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 90(3):232.

Gosoniu, L., Veta, A., and Vounatsou, P. (2010). Bayesian geostatistical modeling of malaria indicator survey data in Angola. *PloS ONE*, 5(3):e9322.

Graves, P. M., Richards, F. O., Ngondi, J., Emerson, P. M., Shargie, E. B., Endeshaw, T., Ceccato, P., Ejigsemahu, Y., Mosher, A. W., Hailemariam, A., et al. (2009). Individual, household and environmental risk factors for malaria infection in Amhara, Oromia and SNNP regions of Ethiopia. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 103(12):1211–1220.

Haque, U., Huda, M., Hossain, A., Ahmed, S. M., Moniruzzaman, M., Haque, R., et al. (2009). Spatial malaria epidemiology in Bangladeshi highlands. *Malaria Journal*, 8(10.1186).

Hay, S. I., Guerra, C. A., Tatem, A. J., Atkinson, P. M., and Snow, R. W. (2005). Tropical infectious diseases: Urbanization, malaria transmission and disease burden in Africa. *Nature Reviews Microbiology*, 3(1):81–90.

Kleinschmidt, I., Sharp, B., Mueller, I., and Vounatsou, P. (2002). Rise in malaria incidence rates in South Africa: a small-area spatial analysis of variation in time trends. *American Journal of Epidemiology*, 155(3):257–264.

Knorr-Held, L. and Besag, J. (1998). Modelling risk from a disease in time and space. *Statistics in Medicine*, 17(18):2045–2060.

Koram, K., Bennett, S., Adiamah, J., and Greenwood, B. (1995). Socio-economic risk factors for malaria in a peri-urban area of The Gambia. *Transactions of the royal society of tropical medicine and hygiene*, 89(2):146–150.

Lawson, A. (2008). *Bayesian disease mapping: Hierarchical modeling in spatial epidemiology*, volume 20. Chapman & Hall/CRC.

Lawson, A., Biggeri, A., Boehning, D., Lesaffre, E., Viel, J., Clark, A., Schlattmann, P., and Divino, F. (2000). Disease mapping models: an empirical evaluation. Disease Mapping Collaborative Group. *Statistics in Medicine*, 19(17-18):2217–41.

Lawson, A. B., Browne, W. J., and Rodeiro, C. L. V. (2003). *Disease mapping with WinBUGS and MLwiN*, volume 11. John Wiley & Sons.

Lekdee, K., Sammatat, S., and Boonsit, N. (2014). Spatio-temporal analysis and mapping of malaria in thailand. *World Academy of Science, Engineering and Technology, International Journal of Medical, Health, Biomedical, Bioengineering and Pharmaceutical Engineering*, 8(7):388–392.

Lengeler, C. (2004). Insecticide-treated bed nets and curtains for preventing malaria. *Cochrane Database Syst Rev*, 2(2).

Lesaffre, E. and Lawson, A. B. (2012). *Bayesian biostatistics*. John Wiley & Sons.

Lindgren, F. and Rue, H. (2008). On the second-order random walk model for irregular locations. *Scandinavian Journal of Statistics*, 35(4):691–700.

López-Quilez, A. and Munoz, F. (2009). Review of spatio-temporal models for disease mapping. *University of Valencia*.

Lowe, R., Chirombo, J., and Tompkins, A. M. (2013). Relative importance of climatic, geographic and socio-economic determinants of malaria in Malawi. *Malaria Journal*, 12(1):416.

Mabaso, M. L., Sharp, B., and Lengeler, C. (2004). Historical review of malarial control in southern African with emphasis on the use of indoor residual house-spraying. *Tropical Medicine & International Health*, 9(8):846–856.

Mabaso, M. L., Vounatsou, P., Midzi, S., Da Silva, J., and 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(1):1.

MacNab, Y. C. and Dean, C. (2002). Spatio-temporal modelling of rates for the construction of disease maps. *Statistics in Medicine*, 21(3):347–358.

Magalhães, R. J. S., Langa, A., Sousa-Figueiredo, J. C., Clements, A. C., and Nery, S. V. (2012). Finding malaria hot-spots in northern angola: the role of individual, household and environmental factors within a meso-endemic area. *Malaria journal*, 11(1):1.

Marcus, B. A. (2009). *Deadly Diseases and Epidemics: Malaria*, volume 2nd edition. Chelsea House.

Mariella, L. and Tarantino, M. (2010). Spatial temporal conditional auto-regressive model: A new autoregressive matrix. *Australian Journal Stat*, 39(3):223.

Martínez-Beneito, M., López-Quilez, A., and Botella-Rocamora, P. (2008). An autoregressive approach to spatio-temporal disease mapping. *Statistics in Medicine*, 27(15):2874–2889.

McSweeney, C. M. New, G. L. (2010). The UNDP Climate Change Country Profiles Improving the Acessibility of Observed and Projected Climate Information for Studies of Climate Change in Developing Countries. Bulletin of the American Meteorological Society. Technical Report 91, University of Oxford. School of Geography and the Environment.

Montosi, E., Manzoni, S., Porporato, A., and Montanari, A. (2012). An ecohydrologic model of malaria outbreaks. *Hydrology and Earth System Sciences Discussions*, 9(3):2831–2854.

Ndlovu, N. (2013). Analysis of the Geographical Patterns of Malaria Transmission in KwaZulu-Natal, South Africa using Bayesian Spatio-Temporal Modelling. Master's thesis, School of Agriculture, Earth and Environmental Sciences.

Nevill, C., Some, E., Mung'Ala, V., Muterni, W., New, L., Marsh, K., Lengeler, C., and Snow, R. (1996). Insecticide-treated bednets reduce mortality and severe morbidity from malaria among children on the Kenyan coast. *Tropical Medicine & International Health*, 1(2):139–146.

Njau, J., Goodman, C., Kachur, S., Palmer, N., Khatib, R., Abdulla, S., Mills, A., and Bloland, P. (2006). Fever treatment and household wealth: the challenge posed for rolling out combination therapy for malaria. *Tropical Medicine & International Health*, 11(3):299–313.

Norton, J. D. (2008). *Spatiotemporal Bayesian hierarchical models, with application to birth outcomes*. PhD thesis, The Florida State University.

Ntzoufras, I. (2011). *Bayesian modeling using WinBUGS*, volume 698. John Wiley & Sons.

Nyarko, S. H. and Cobblah, A. (2014). Sociodemographic determinants of Malaria among under-five children in Ghana. *Malaria research and treatment*, 2014.

Pates, H. and Curtis, C. (2005). Mosquito behavior and vector control. *Annu. Rev. Entomol.*, 50:53–70.

Pérez-Escamilla, R., Dessalines, M., Finnigan, M., Pachón, H., Hromi-Fiedler, A., and Gupta, N. (2009). Household food insecurity is associated with childhood malaria in rural Haiti. *The Journal of Nutrition*, 139(11):2132–2138.

PMI (2010). President's Malaria Initiative. Malaria Operational Plan (FY2010). Angola.

Rue, H., Martino, S., and Chopin, N. (2009). Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *Journal of the royal statistical society: Series b (statistical methodology)*, 71(2):319–392.

Ruebush, T., Burkot, T., de Oliveira, A., da Silva, J., Renshaw, M., et al. (2005). Presidents Initiative on Malaria: Needs assessment, Angola. *Washington DC United States Agency for International Development*.

Schrödle, B. and Held, L. (2011). Spatio-temporal disease mapping using INLA. *Environmetrics*, 22(6):725–734.

Schrödle, B., Held, L., Riebler, A., and Danuser, J. (2011). Using integrated nested Laplace approximations for the evaluation of veterinary surveillance data from Switzerland: a case-study. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 60(2):261–279.

Snow, R. W., Craig, M., Deichmann, U., and Marsh, K. (1999). Estimating mortality, morbidity and disability due to malaria among Africa's non-pregnant population. *Bulletin of the World Health Organization*, 77(8):624.

Spiegelhalter, D., Thomas, A., Best, N., and Lunn, D. (2003). Winbugs user manual.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., and Van Der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64(4):583–639.

Sun, D., Tsutakawa, R. K., Kim, H., He, Z., et al. (2000). Spatio-temporal interaction with disease mapping. *Statistics in Medicine*, 19(15):2015–2035.

Tanser, F. C., Pluess, B., Lengeler, C., and Sharp, B. L. (2007). Indoor residual spraying for preventing malaria. *The Cochrane Library*.

Tanser, F. C., Sharp, B., and Le Sueur, D. (2003). Potential effect of climate change on malaria transmission in Africa. *The Lancet*, 362(9398):1792–1798.

Tulu, A. N. (1996). *Determinants of malaria transmission in the highlands of Ethiopia: the impact of global warming on morbidity and mortality ascribed to malaria.* PhD thesis, London School of Hygiene & Tropical Medicine.

Ugarte, M. D., Adin, A., Goicoa, T., and Militino, A. F. (2014). On fitting spatio-temporal disease mapping models using approximate Bayesian inference. *Statistical methods in medical research*, 23(6):507–530.

UNICEF (2013). Massive scale-up targets malnutrition in Angola. Technical report.

Waller, L. A., Carlin, B. P., Xia, H., and Gelfand, A. E. (1997). Hierarchical spatio-temporal mapping of disease rates. *Journal of the American Statistical Association*, 92(438):607–617.

White, G. (1974). Anopheles gambiae complex and disease transmission in Africa. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 68(4):278–298.

WHO (2010). World malaria report. *World Health Organization. Geneva.*

WHO (2011). World malaria report. *World Health Organization. Geneva.*

WHO (2013). World malaria report. *World Health Organization. Geneva.*

Yeshiwondim, A. K., Gopal, S., Hailemariam, A. T., Dengela, D. O., and 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(1):1.

Yusuf, O. B., Adeoye, B. W., Oladepo, O. O., Peters, D. H., Bishai, D., et al. (2010). Poverty and fever vulnerability in Nigeria: a multilevel analysis. *Malaria Journal*, 9(1):235.

Zhang, W., Wang, L., Fang, L., Ma, J., Xu, Y., Jiang, J., Hui, F., Wang, J., Liang, S., Yang, H., et al. (2008). Spatial analysis of malaria in Anhui province, China. *Malaria Journal*, 7(206):10–1186.