Modelling the seasonal and spatial variation of malaria transmission in relation to mortality in Africa

Inauguraldissertation
zur
Erlangung der Würde eines Doktors der Philosophie

vorgelegt der
Philosophisch-Naturwissenschaftlichen Fakultät der
Universität Basel

von
Susan Fred Rumisha
aus Tanga, Tanzania

Basel, 2013

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät auf Antrag von
Prof. Dr. M. Tanner, Dr. P. Vounatsou, Dr. T. Smith und Prof. Dr. Heiko Becher

Basel, den 13. December 2011

Prof. Dr. Martin Spiess
Dekan

… to my beloved parents, sisters and brothers.

Time,
The most essential professional tool,
The necessary connection between events,
The device that prevents everything from happening at once,
Time.
~Penelope Lively

Table of Contents

List of abbreviation i
List of Tables ii
List of Figures ii
Summary iii
Zusammenfassung vii
Acknowledgement xii
Chapter 1 Introduction 1
1.1. Background 2
1.1.1. Malaria burden 2
1.1.2. Species, vector and behavioral characteristics 3
1.1.3. Clinical features and malaria diagnosis 5
1.2. Measuring malaria transmission 5
1.3. Malaria control and interventions 7
1.4. Malaria transmission and mortality 10
1.5. Demographic surveillance systems, malaria transmission intensity and mortality 11
1.6. Mapping of malaria transmission 12
1.7. Geostatistical modeling 15
1.7.1. Bayesian modeling of large geostastical data 15
1.7.2. Approximation of spatial processes using a subset of data 16
1.7.3. Modeling large sparse zero-inflated entomological data 17
1.7.4. Modelling seasonality and temporal effect in malaria data 17
1.8. Aims and objectives 19
1.8.1. Specific methodological objectives 19
1.8.2. Epidemiological questions addressed by the developed methods 19
Chapter 2 Assessing seasonal variations and age patterns in mortality during the first year of life in Tanzania 21
Abstract 22
2.1. Introduction 23
2.2. Data 25
2.2.1 Infant Mortality 25
2.2.2 Seasonality 26
2.3. Methodology 26
2.3.1 Seasonality Index 26
2.3.2 Goodness-of-fit test 27
2.3.3 Modeling 27
2.3.4 Description of harmonic cycle model 28
2.3.5 Bayesian model specification 29
2.3.6 Model assessment and selection 30
2.4. Results 31
2.4.1 Descriptive analysis 31
2.4.2 Seasonality index 33
2.4.3 Goodness-of-fit test 34
2.4.4 Modeling 34
2.5. Discussion 39

Chapter 3 Bayesian modeling of large geostatistical data to estimate seasonal and spatial variation of sporozoite rate 43
Abstract 44
3.1. Introduction 45
3.2. Data and Study site 48
3.3. Model formulation 49
3.3.1 Geostatistical modeling via Gaussian spatial processes 50
3.3.2 Geostatistical modeling via Gaussian spatial processes approximations 50
3.3.3 Selection of knots 51
3.3.4 Bayesian kriging 52
3.3.5 Model validation 52
3.3.6 Practical implementation 53
3.4. Results 54
3.4.1. Descriptive analysis 54
3.4.2. Selection of knots 55
3.4.3. Model validation 56
3.4.4. Model-based estimates 57
3.5. Discussion 62
Chapter 4 Modelling heterogeneity in malaria transmission using large sparse spatio-temporal entomological data 69
Abstract 70
4.1. Introduction 71
4.2. Methodology 74
4.2.1. Study Site 74
4.2.2. Mosquito Data 74
4.2.3. Environmental and Climatic Data 75
4.2.4. Statistical analysis 75
4.3. Results 81
4.3.1. Data Description 81
4.3.2. Geostatistical model results 82
4.3.3. Mapping EIR 83
4.3.4. Population adjusted EIR 87
4.4. Discussion 89
Chapter 5 Assessing the relation between child survival and malaria transmission: an analysis of the MTIMBA data in Rufiji DSS, Tanzania 93
Abstract 94
5.1. Introduction 95
5.2. Methods 97
5.2.1. Study Area 97
5.2.2. Data 97
5.2.3. Linking mortality with other databases 98
5.2.4. Statistical analysis 99
5.3. Results 100
5.3.1. Mortality data 100
5.3.2. Exploratory analysis 102
5.3.3. Model-based results 104
5.4. Discussion 105

Chapter 6 Malaria transmission intensity and mortality in older children and adults in Rufiji DSS, Tanzania 111
Abstract 112
6.1. Introduction 113
6.2. Methods 114
6.2.1. Study site and the data 114
6.2.2. Statistical analysis 115
6.3. Results 116
6.3.1. Descriptive statistics 116
6.3.2. Model based results 118
6.3.3. Effect of EIR on mortality from birth to adulthood 119
6.3.4. Excess mortality attributed to malaria transmission 120
6.4. Discussion 121
Chapter 7 General discussion and conclusion 127
7.1. A preamble 128
7.2. Justification for the research and goals 129
7.3. Structure 130
7.4. Statistical contribution 131
7.5. Epidemiological contribution 133
7.6. Limitations and challenges 136
7.6.1 Awareness and capacity in statistical data analysis 136
7.6.2 Epidemiological 137
7.6.3 Statistical 138
7.7. Future research and extension 141
7.7.1 Non-stationary spatial process and space-time interactions 141
7.7.2 Jointly modeling of malaria transmission and mortality accounting for spatial misalignment and species heterogeneity 142
7.7.3 Improving EIR estimates and assessing other measures of transmission 142
7.7.4 Converting malaria transmission risk on mortality to disease burden 143
7.7.5 From DSS to vital registration system 143
7.8. Implication and concluding remark 144
References 145
Curriculum Vitae 165

List of abbreviation

ACT	Artemisinin-based Combination Therapy
ADDS	Africa Data Dissemination Service
AIC	Akaike Information Criteria
AIDS	Acquired Immunodeficiency Syndrome
BCI	Bayesian Credible Interval
CDC	Centers for Disease Control and Prevention
CSP	Circumsporozoite Protein
DALYs	Disability-Adjusted Life Years
DDT	dichlorodiphenyltrichloroethane
DHS	Demographic and Health Surveys
DSS	Demographic Surveillance Systems
EIR	Entomological Inoculation Rate
ELISA	Enzyme Linked Immuno-Sorbent Assay
GIS	Geographical Information Systems
GMAP	Global Malaria Action Plan
GMEP	Global Malaria Eradication Program
GPS	Global Positioning System
HIV	Human Immunodeficiency Virus
INDEPTH Network	The International Network for the Demographic Evaluation of Populations and Their Health in developing countries
ITNs	Insecticide-Treated Nets
IRS	Indoor Residual Spraying
IPTi	Intermittent Preventive Treatment for infants
IPTp	Intermittent Preventive Treatment in pregnancy
LLITNs	Long Lasting Insecticide-Treated Nets
MAP	Malaria Atlas Project
MARA	Mapping Malaria Risk In Africa
MCMC	Markov Chain Monte Carlo
MDGs	Millennium Development Goals
MICS	Multiple Indicator Cluster Survey
MIM	Multilateral Initiative on Malaria
MIS	Malaria Indicators Survey
MMV	Medicines for Malaria Venture
MTIMBA	Malaria Transmission Intensity and Mortality Burden across Africa
MODIS	Moderate Resolution Imaging Spectroradiometer
MVI	Malaria Vaccine Initiative
MVN	Multivariate Normal
NB	Negative Binomial
NDVI	Normalized Distance Vegetation Index
NIMR	National Institute for Medical Research
PATH	Program for Appropriate Technology In Health
RBM	Roll Back Malaria
RDSS	Rufiji DSS
RS	Remote Sensing
SAPV	Spatially Averaged Predicted Variance
SAVVY	Sample Vital Statistics with Verbal Autopsy
SARIMA	Seasonal Auto-Regression Integrated Moving Average
SES	Socio-Economic Status
SR	Sporozoite Rate
SRS	Sample registration systems
SSA	sub-Saharan Africa
SVR	Sample vital registration
TDHS	Tanzania Demographic and Health Survey
TEHIP	Tanzania Essential Health Interventions Project
UNDP	United Nations Development Programme
UNICEF	United Nations Children's Fund
VC	Vectorial Capacity
WHO	World Health Organization

List of Tables

Table 2.1: Age-specific point estimate and model based (with 95% Bayesian Credible Interval (BCI)) seasonality indices by year and on pooled data 33
Table 2.2: Model assessment: the error sum of squares, number of parameters, and variance and autocorrelation parameters from models fitted with data for infants of 10-months of age 35
Table 2.3: Results of M12 for infants age 2, 5, 8 and 11 months and for pooled data 37
Table 3.1: Environment and Climate data 48
Table 3.2: Space-time posterior estimates for predictors, space and time parameters 57
Table 4.1: Results of association of environment/climate variables on sporozoite rate and mosquito density and spatio-temporal parameters 82
Table 4.2: Overall predicted EIR with the percent (%) attribute of each species 87
Table 4.3: Distribution of predicted EIR over the RDSS area by Year, N*(%): 87
Table 5.1: Number of individuals, deaths and locations after merging mortality database with entomological, socio-economic and malaria interventions databases 98
Table 5.2: Descriptive statistics on mortality at different age groups of child health, Oct 2001- Sept 2004, Rufiji DSS 101
Table 5.3: Mortality rate according to ITN possession, SES and EIR levels (natural scale) in the Rufiji DSS 103
Table 5.4: Parameter estimates obtained from Bayesian spatial-temporal models on neonates, infants and older children survival in the Rufiji DSS 104
Table 6.1: Model estimates from Bayesian discrete time Bernoulli survival geostatistical models on school age children, adults and older individuals' survival in the Rufiji DSS, Oct 2001-Sept 2004 118
Table 7.1: ITN ownership (%) by age in the Rufiji DSS, Oct 2001- Sept 2004 135

List of Figures

Figure 1.1: Risk of malaria transmission worldwide (Source: WHO, 2010) 2
Figure 1.2: Malaria transmission cycle (Source: CDC) 4
Figure 1.3: Countries with DSS sites within the INDEPTH Network (Source: SIDA Review, 2010) 11
Figure 2.1: Death counts and age-specific mortality rates (MR) of infants in Rufiji DSS, October 2001- September 2004 31
Figure 2.2: Age-specific mortality rates by calendar months 32
Figure 2.3: Mortality rates (MR) for the Infants during the period of October 01 – September 04, Rufiji DSS 32
Figure 2.4: Error Sum of Squares for Models 1-12 for all age groups* 36
Figure 2.5: Monthly time series of mortality rates aligned with the fitted values, RDSS 38
Figure 3.1: (a) Crude monthly sporozoite rate for *An. funestus* and *An. gambiae* (b) Monthly rain and temperature 54
Figure 3.2: Box plots for the distance measure D for *An. funestus* (left) and *An. gambiae* (right) between the spatial parameters estimated by the empirical variables of the sub-sample 55
Figure 3.3: Grid showing original and sub-locations selected using a balance sampling for *An. funestus* (left) and *An. gambiae* (right) 56
Figure 3.4: Proportion of points included in the CI of the posterior predictive distribution 56
Figure 3.5: Predicted sporozoite rates *An. funestus* for Jan-Dec in Rufiji DSS, Tanzania 59
Figure 3.6: Predicted sporozoite rates *An. gambiae* for Jan-Dec in Rufiji DSS, Tanzania 60
Figure 3.7: Mean predicted sporozoite rates for *An. funestus* and *An. gambiae* 61
Figure 4.1: Seasonal variations of a) rainfall, temperature and b) mosquitoes densities of *An. gambiae* and *An. funestus* in the Rufiji DSS October 2001- September 2004 81
Figure 4.2: Selected EIR maps showing the spatial distribution and the seasonal pattern, for the period of Oct 2001- Sept 2004. A) Dry months followed by the period of short rains, B) Months immediately after the onset of heavy rains during the 1st year (very wet), C) Months immediately after the onset of heavy rains during the 2nd year (dry) and D) Months immediately after the onset of heavy rain season during the 3rd year (normal rains) 84
Figure 4.3: Predicted monthly EIR median and attribute of each species in Rufiji DSS 85
Figure 4.4: Spatial temporal distribution of annual EIR with prediction error maps. 86
Figure 4.5: Distribution of households in the Rufiji DSS area (N=14,516), Source: TEHIP, 2002 88
Figure 5.1: Child death rates by age categories and month of death, RDSS 101
Figure 6.1: Gender and age specific estimates of mortality rates in children and adult population, RDSS, Oct 2001- Sept 2004 117
Figure 6.2: Estimates of HR of EIR on mortality (with 95% BCI) from a geostatistical model for different age categories. The under five years were added to allow clear presentation of the trend 119
Figure 6.3: Age-specific patterns of excess mortality by transmission intensity 120
Figure 7.1: Age-specific patterns of excess mortality by transmission intensity (Source: WHO/Health Metrics Network) 144

Summary

Based on the recent world malaria report, about three billion people, which is almost half of the world population, are estimated to be at risk of malaria transmission. In developing countries, malaria is believed to be a major cause of morbidity and mortality. The disease affects all age groups with children under five years carrying the highest morbidity and mortality burden. It is among the indirect causes of maternal mortality. In endemic areas many infants' deaths are attributed to low-birth-weight resulted from malaria infection during pregnancy. Malaria brings huge economic burden due to number of days lost during sickness and deaths, sustaining a vicious cycle of disease and poverty in sub Saharan Africa (SSA). In World Health Organization Regional Office for Africa (WHO/AFRO), nine percent of deaths and more than ten percent of disability-adjusted life years (DALYs) are attributable to malaria alone.

A number of malaria control interventions to reduce intensity of transmission have been successfully implemented in endemic areas in the regions of SSA. In the past, elimination of malaria succeeded in many developed countries but it is still a dream in many developing countries today. Failures in global eradication are related to resistance in insecticides and anti-malarial drugs and health systems related factors. In the last decade, the Roll Back Malaria (RBM) partnership reinforced new strategies to combat malaria with long-term goal of eradicating the disease globally. This facilitated about tenfold rise in global funding for malaria research and development of multi disciplinary initiatives to combat malaria. Malaria is now in the main agenda of all international health and development forums and part of the Millennium Development Goals (MDGs). Recent global health statistics report declines in mortality especially in children. The reduction in mortality is associated with achievements in intervention strategies and also improvements in malaria diagnosis and treatment. The coverage of the most efficient control intervention, the use of Long Lasting Insecticide-Treated Nets (LLITN)/ITN has increased in many countries resulting in substantial lessening of malaria transmission and infection rates. However, poor natural acquisition of malaria immunity in children as a consequence of weak or no exposure is a major epidemiological concern and brings a fear of higher mortality rates or shifting of age of death to older children. Understanding possible links between transmission, intervention, immunity and mortality is important for sustainable progress towards the targets set by RBM or MDGs, is crucial.

Comprehensive analysis of information on malaria transmission, vital events, drivers of transmission and mortality-related risk factors is required to assess and quantify how intervention, transmission and mortality are related. The data should have high spatial and temporal resolutions to assess micro scale variability of transmission and seasonality. Lack of vital registration systems in developing countries hinders availability of such data. Establishment of Demographic Surveillance Systems (DSS) in many developing countries aims to fill these information gaps. Within a DSS, a defined population is routinely

monitored on vital events and other demographic indicators which create a platform for evaluating public health indicators. One of the initiatives integrated within DSSs is the Malaria Transmission Intensity and Mortality Burden across Africa (MTIMBA) project. The project compiled a database of mosquito collections at selected sites in Africa over a large number of locations, using standardized methodologies for a period of three years. The entomological parameters were linked with routinely monitored vital events within the DSS. The MTIMBA database is the most comprehensive entomological database ever collected in Africa which allows studying small area spatial-temporal variation in malaria transmission in relation to mortality.

Malaria is an environmental disease hence transmission varies with climate as it modifies population, survival, distribution and infectivity of malaria vectors. Quantification of association between climate and transmission is important to allow prediction of risk even in areas that field data cannot be easily obtained. Development in geographical information systems (GIS) and availability of remote sensing (RS) data facilitates availability of environment and climate data at high space and time resolutions allowing accurate estimation of outcome-factor relationship. However, DSS data are collected longitudinally at fixed locations (i.e. geostatistical data) and are characterized by seasonal patterns, spatial and temporal correlations due to similar exposures shared within small proximities and close time profiles.

Additionally, DSS data are large, sparse and zero-inflated (e.g. mosquito data). Standard models assume independence between observations, an assumption which do not hold for correlated data, hence utilizing these models might result into biased estimates. Geostatistical modeling of large, sparse and zero inflated space-time data is computational challenging specifically in the estimation of the spatial processes. The spatial correlation is accounted by introducing location-specific random effect parameters which are assumed to arise from a spatial process quantified by a multivariate normal distribution. The models are highly parameterized and their fit is computationally intensive. Bayesian computational algorithms such as Markov Chain Monte Carlo (MCMC) can be used to fit these models. Estimation of the spatial process requires inversion of the covariance matrix at each simulation point. The dimension of the matrix increases exponentially with number of locations and the inversion becomes infeasible when the size is too large. Recent techniques overcome this problem by approximating the spatial process from a subset of locations. These methods have been applied on Gaussian outcomes observed over a grid. However, the MTIMBA data are neither Gaussian nor regular in space. Extension of methodology is required to address MTIMBA data characteristics. Another important aspect is the modeling and estimation of malaria transmission using Entomology Inoculation Rate (EIR). EIR is typically treated as a continuous outcome and modeled using linear regression models after performing a logarithmic transformation to meet the assumption of normality. However, it is difficult to achieve normality from sparse data with large number of zeros. Formulation of rigorous

methods to efficient model these data are needed to allow precise prediction of malaria transmission at locations with mortality data to enhance studying the association. Lastly, seasonality in climatic conditions introduces seasonal patterns in transmission and mortality data. Although in statistical literature modeling formulations of seasonality patterns are well developed, most epidemiological applications ignore or do not fully explore this aspect of the data.

The objectives of this thesis were to i) develop Bayesian geostatistical models to analyze very large and sparse geostatistical and temporal non-Gaussian data with seasonal patterns and ii) apply these models to (a) estimate space-time heterogeneity in malaria transmission (b) assess mortality variations between different ages during the first year of life while adjusting for seasonality and (c) determine the relation between transmission intensity and risk of mortality in children and adult population after taking into account control interventions. This work used an extract of MTIMBA data from the Rufiji DSS (RDSS) collected between October 2001 and September 2004.

Evaluation of different approaches to capture seasonal pattern using infant mortality data is discussed in Chapter 2. The aim was to estimate mortality peaks and assess whether they differ at different stages of infant life. The peaks of mortality were observed to be seasonal and aligned with climate conditions. However, no difference in the timing of peaks was observed among age groups. Among the statistical approaches assessed, cosine terms were able to best capture seasonality with mixture of cycles.

In Chapter 3, models approximating the spatial process from a subset of locations were developed to assess seasonal and spatial pattern of sporozoite rate (SR) of *An. funestus* and *An. gambiae* and produce smooth monthly maps of SR for RDSS. Effect of climate was also determined. The SR data were selected due to moderate number of locations involved so we could easily validate our model performance. Balanced sampling was employed to draw 5000 samples of the original data with size varying from 50 to 350. A distance measure was used to compare the variogram of the full data and the samples. The location coordinates were used as auxiliary variables to ensure a balance sample. Results showed that it is not the size of the sample that determines how good is the approximation to the spatial process but rather the sample spatial configuration. The results emphasized the importance of understanding vector behaviors and how they respond towards changes in different attributes.

A rigorous approach to analyze EIR data is discussed in Chapter 4. EIR is the product of mosquito density and SR which are count and binomial data, respectively. Separate models for SR (binomial) and density data (negative binomial) were fitted by species and the EIR was calculated using model-based predictions of SR and mosquito density. Zero-inflated analogue of negative binomial was used to account for over-dispersion and zero-inflation in the mosquito data. The large spatial process was approximated by a sample of locations obtained via minimax space-filling methods. The model took into

account seasonality, temporal trends and correlations and adjusted for climatic predictors. High resolution EIR maps were produced for the RDSS. Temporal, spatial and seasonal patterns of EIR were highly influenced by environmental predictors.

Using the EIR estimated in Chapter 4, we aligned all locations (households) with mortality data within the RDSS and predicted malaria transmission on monthly intervals. Bernoulli discrete-time regression models were used to assess the relationship between all-cause mortality and malaria transmission (i.e. EIR) in different age categories. Separate analyses were performed for neonates (0-28days), infants (1-11months), children (1-4yrs), school-children (5-14yrs), young adults (15-60yrs) and older individuals (>60yrs). The results of these analyses are presented in Chapters 5 and 6. Models include spatial and non-spatial random effects at village level and monthly temporal correlation and were adjusted for age and ITN possession. The EIR was incorporated in the model as a covariate with measurement error to account for the prediction uncertainty. The results indicated that i) the effect of malaria transmission on mortality differ by age with school-children having higher impact as compared to younger children and adults possibly due to weak immunity development during childhood which emphasize on the importance of preventive interventions in this age group ii) ITNs had a protective effect but rather not significant on neonate mortality, however, interactions between ITN ownership and family socio-economic status (SES) are likely, iii) it is importance to incorporate information on exposure history and malaria immunity when studying the transmission-mortality relation to assess how that modifies the association.

This work is a building block on the insight and understanding of association between malaria transmission and all-cause mortality. The major strength of results of this work relies on EIR estimates predicted at high spatial (household level) and temporal resolution by employing rigorous geostatistical models fitted on large entomological data. The better exposure estimates obtained (which include measure of uncertainty) are able to more accurately estimate the relation between malaria transmission and mortality which is important for determining the consequences of malaria intervention on transmission and mortality trends. However, comparison of analyses conducted within and between different sites in SSA is critical to better understand the reproducibility of most of conclusions. That might facilitate developing a refined model to relate different measures of malaria transmission with mortality which is essential for the targets set by RBM initiatives.

Zusammenfassung

Einer weltweiten Schätzung zufolge leben etwa drei Milliarden Menschen in Gefahr sich mit Malaria zu infizieren, was etwa der Hälfte der Weltbevölkerung entspricht. In den Entwicklungsländern ist Malaria wahrscheinlich die Hauptursache von Morbidität und Mortalität. Die Krankheit befällt alle Altersgruppen, allerdings tragen Kindern unter fünf Jahren die höchste Last an Morbidität und Mortalität. Malaria ist ausserdem eine der Hauptursachen der Müttersterblichkeit und viele Todesfälle in Kleinkindern gehen auf eine Infektion der Mutter während der Schwangerschaft zurück. Durch die Krankheit kommt es zu einer enormen wirtschaftlichen Belastung aufgrund der hohen Anzahl an Fehltagen und wegen der Todesfälle, was einen durch Krankheit und Armut bedingten Teufelskreis in Subsahara-Afrika (SSA) aufrecht erhält. Im Gebiet des Regionalbüros der Weltgesundheitsorganisation in Afrika (WHO/AFRO) sind allein neun Prozent der Todesfälle und mehr als zehn Prozent der sogenannten disability-adjusted life years (DALY) auf Malaria zurückzuführen.

Eine Reihe von Malaria-Kontroll-Interventionen um die Intensität der Übertragung in den endemischen Gebieten von SSA zu verringern wurde erfolgreich umgesetzt. In der Vergangenheit ist die Ausrottung der Malaria in vielen Industrienationen gelungen, allerdings bleibt sie auch heute noch immer ein Traum in vielen Entwicklungsländern. Misserfolge in der weltweiten Ausrottung der Krankheit sind auf Resistenzen gegenüber Insektiziden und Anti-Malaria-Medikamenten zurückzuführen, als auch auf bestimmte Faktoren in Bezug auf die Gesundheitssysteme. In den letzten zehn Jahren verstärkte die Roll Back Malaria (RBM) Partnerschaft neue Strategien zur Bekämpfung der Malaria mit dem langfristigen Ziel der weltweiten Ausrottung der Krankheit. Diese verzehnfachte die globalen Mittel zur Erforschung der Malaria und zur Entwicklung von multidisziplinären Initiativen zu deren Bekämpfung. Inzwischen ist Malaria ein Teil der Agenda aller internationalen Gesundheits- und Entwicklungsforen und gehört zu den sogenannten Millennium Development Goals (MDGs). Aktuelle globale Statistiken berichten von einem Rückgang der Sterblichkeit, vor allem bei Kindern. Diese Verringerung der Mortalität ist eng mit Errungenschaften der Interventionsstrategien und auch der Verbesserungen der Diagnostik und Behandlung verbunden. Die Flächendeckung der effizientesten Intervention, der Einsatz von langlebigen und mit Insektizid imprägnierten Moskitonetze (LLITN)/ITN, hat sich in vielen Ländern erhöht, was zu einer wesentlichen Verminderung der Malaria-Übertragung und Infektionsraten führte. Allerdings steigert die verringerte Immunität gegen Malaria bei Kindern als Folge der schwachen Exposition die epidemiologische Sorge einer erhöhten Sterblichkeit und der Verschiebung des Todesalters hin zu älteren Kindern. Daher ist das Verständnis möglicher Verbindungen zwischen Übertragung, Intervention, Immunität und Mortalität von entscheidender Bedeutung um nachhaltige Fortschritte bei der Verwirklichung der Zielsetzung von RBM oder den MDGs zu erreichen.

Es wird eine umfassende Datenbank benötigt, welche Informationen über die Malaria-Übertragung, Geburts- und Sterbefälle, ausschlaggebende Faktoren zur Übertragung und Mortalität beinhaltet, um die Zusammenhänge von Intervention, Übertragung und Mortalität zu untersuchen und zu quantifizieren. Die enthaltenen Daten sollten eine hohe räumliche und zeitliche Auflösung besitzen, damit Unterschiede in der Übertragung auf Mikroebene und Saisonalität betrachtet werden können. Der Mangel an staatlich erfassten Geburts- und Sterbefällen in den Entwicklungsländern behindert allerdings die vollständige Verfügbarkeit dieser Daten. Die Gründung der demographischen Überwachungssysteme (DSS) in vielen Entwicklungsländern soll diese Lücke beheben. Innerhalb des Einzugsgebietes eines DSS wird eine definierte Bevölkerung routinemäßig hinsichtlich der Geburts- und Sterbefälle und anderer demographischer Indikatoren überwacht, um eine Plattform zur Bewertung gesundheitlicher Indikatoren zu schaffen. Eine der Initiativen die in die DSS integriert sind ist das sogenannte Malaria Transmission Intensity and Mortality Burden across Africa (MTIMBA) Projekt. Dieses Projekt erstellt eine Datenbank mit Hilfe standardisierter Methoden zur zweiwöchentlichen Erfassung von Mücken an einer Vielzahl von ausgewählten Standorten in Afrika über einen Zeitraum von drei Jahren. Die entomologischen Parameter werden mit den routinemäßig überwachten demographischen Daten innerhalb des DSS verbunden. Die MTIMBA Datenbank ist die umfassendste entomologische Datenbank die jemals in Afrika erstellt wurde und die sich dem Studium der mikroräumlichen und zeitlichen Variation der Malaria-Übertragung in Bezug auf die Mortalität widmet.

Malaria ist eine umweltbezogene Krankheit, da die Übertragung stark mit dem Klima schwankt, welches die Bevölkerung, das Überleben, die Verteilung und die Infektiösität des Vektors beeinflusst. Die Quantifizierung des Zusammenhangs zwischen Klima und Übertragung ist daher wichtig, um eine Vorhersage des Risikos in Gebieten ohne zuverlässige Studiendaten zu ermöglichen. Die Entwicklung der geografischen Informationssysteme (GIS) und die Verfügbarkeit von Satellitendaten erleichterte die Verfügbarkeit von geographischen und klimatischen Daten mit hoher räumlicher und zeitlicher Auflösung, was eine genaue Abschätzung der Faktor-Wirkungsbeziehung ermöglicht. Die DSS-Daten werden zu verschiedenen Zeitpunkten an festen Standorten (d.h. es handelt sich um geo-statistische Daten) gesammelt und sind durch saisonale Muster und räumliche und zeitliche Korrelationen gekennzeichnet, die aufgrund ähnlicher beeinflussender Faktoren in Nachbarschaften und zeitlicher Nähe entstehen. Darüber hinaus handelt es sich um sehr große, räumlich karge und zu Null tendierende (z.B. Mückenanzahl) Daten. Die Standard-Modelle gehen allerdings von der Unabhängigkeit der Beobachtungen aus, eine Annahme die nicht auf korrelierte Daten zutrifft. Daher kann die Nutzung dieser Modelle zu verzerrten Schätzungen führen. Wohingegen die geo-statistische Modellierung dieser Art von Daten eine rechenintensive Herausforderung, speziell bei der Schätzung der räumlichen Prozesse, darstellt. Die räumliche Korrelation wird durch die Einführung von Orts-spezifischen

Zufallsparametern ermöglicht, deren Verteilung durch eine multivariate Normalverteilung des räumlichen Prozess angenommen wird. Die Modelle enthalten dadurch eine Vielzahl von Parametern und sind sehr rechenintensiv. Berechnungsalgorithmen nach Bayes, wie Markov Chain Monte Carlo (MCMC), können verwendet werden, um diese Modellierungen durchzuführen. Allerdings erfordert die Schätzung der räumlichen Prozesse die Inversion der Kovarianzmatrix bei jedem Schritt der Simulation. Die Dimension dieser Matrix steigt exponentiell mit der Anzahl der Standorte, was deren Inversion bei sehr vielen Studienorten undurchführbar macht. Aktuelle Techniken zur Überwindung dieses Problem nähern sich dem tatsächlichen räumlichen Prozess mittels einer Teilmenge der Standorten an. Diese Methoden wurden bereits an Gauß-verteilten Ereignissen an Rasterpunkten verwendet, allerdings sind die MTIMBA Daten werden weder nach Gauß noch regelmäßig im Raum verteilt. Daher ist eine Erweiterung dieser Methodik erforderlich, um die MTIMBA Daten auszuwerten. Ein weiterer wichtiger Aspekt ist die Modellierung und Schätzung der Malaria-Übertragung mittels der entomologischen Inokulationsrate (EIR). Die EIR wird in der Regel als kontinuierliches Ereignis angesehen und mittels linearer Regression, nach der Durchführung einer logarithmischen Transformation um normalverteilte Daten zu erzielen, modelliert. Allerdings ist es schwierig, Normalität aus spärlichen Daten mit einer großen Anzahl von Nullen zu erreichen. Die Formulierung von rigorosen Methoden zur effizienten Modellierung dieser Daten wird benötigt, um genaue Vorhersagen der Malaria-Übertragung zu ermöglichen. Außerdem führt die Saisonalität der klimatischen Bedingungen zu saisonalen Mustern der Übertragungs- und Mortalitätsraten, und obwohl Saisonalität in der statistischen Literatur ausreichend betrachtet wurde, wird sie in den meisten epidemiologischen Betrachtungen ignoriert oder nicht vollständig zu untersucht.

Die Ziele dieser Arbeit waren, i) die Entwicklung geostatistischer Modelle nach Bayes um sehr große zeitlich verteilte aber räumlich spärliche geo-statistische und Nicht-Gauß-verteilte Daten mit saisonalen Mustern zu analysieren, ii) diese Modelle anzuwenden auf (a) die Abschätzung der räumlich-zeitlichen Heterogenität der Malaria-Übertragung, (b) die Beurteilung der höchsten Sterblichkeitsrate in den verschiedenen Stadien der Kindheit unter Einbeziehung der Saisonalität, und (c) das Bestimmen des Verhältnisses zwischen der Intensität der Übertragung und des Mortalitätsrisikos bei Kindern und Erwachsenen mit Berücksichtigung von Interventionen. Diese Arbeit verwendet einen Teil der MTIMBA Daten aus dem Rufiji DSS (RDSS), welche zwischen Oktober 2001 und September 2004 gesammelt wurden.

Die Auswertung der verschiedenen Ansätze um saisonale Muster in der Säuglingssterblichkeit Daten zu erkennen wird in Kapitel 2 diskutiert. Ziel war es, die Sterblichkeitsrate abzuschätzen und zu beurteilen, ob sie sich in den verschiedenen Stadien eines Säuglings unterscheidet. Die höchste Sterblichkeit ist saisonal bedingt und wird durch klimatische Bedingungen beeinflusst. Es wurde jedoch kein Unterschied

zwischen den Altersgruppen beobachtet. Unter den verschiedenen statistischen Ansätzen stellten sich Kosinus-Terme mit einer Mischung aus Zyklen als am besten heraus um die Saisonalität wiederzugeben.

In Kapitel 3 wurden Modelle zur Abschätzung des räumlichen Prozesses aus einer Teilmenge der Standorte entwickelt, um zeitliche und räumliche Muster der Sporozoiten-Rate (SR) von *An. funestus* und *An. gambiae* zu bewerten und um Karten der SR für jeden Monat im Gebiet des RDSS zu erstellen. Der Einfluss des Klimas wurde ebenfalls bestimmt. Die SR-Daten wurden an einer moderaten Anzahl von Standorten erfasst, so dass die Modelle leicht gegeneinander getestet werden konnten. Balance-Sampling wurde hierbei eingesetzt, um 5000 verschiedene Stichproben mit zwischen 50 und 350 Elementen aus den ursprünglichen Daten zu erhalten. Zusätzlich wurden Variogramme der einzelnen Stichproben und der vollständigen Daten erstellt und deren Abstände ermittelt. Die Ortskoordinaten wurden als Ausgleichsvariablen verwendet. Die Ergebnisse zeigten, dass nicht die Größe der Stichprobe die beste Annäherung an den tatsächlichen räumlichen Prozess bestimmt, sondern die räumlichen Konfiguration der Stichprobe. Die Analysen zeigten außerdem Unterschiede in der Übertragung zwischen den beiden Arten, wobei *An. funestus* verantwortlich ist für die Trockenzeit-Übertragung und *An. gambiae* für die Regenzeit-Übertragung. Dies führt zu einer ganzjährlichen Übertragung.

Ein umfassender Ansatz um EIR Daten zu analysieren wird in Kapitel 4 behandelt. EIR ist das Produkt der Mücken-Dichte und der SR. Separate Modelle für die SR (binomialverteilt) und Dichte-Daten (negativ binomial-verteilt) wurden für die verschiedenen Mückenarten durchgeführt und die EIR wurde abschließend unter Verwendung modellbasierter Vorhersagen der SR und der Moskito-Dichte berechnet. Eine zero-inflated negative Binomialverteilung wurde unter Berücksichtigung der Überdispersion und der Null-Inflation der Mücke-Daten verwendet. Der räumlichen Prozess wurde durch eine Stichprobe, welche durch das Minimax space-filling Verfahren ermittelt wurde, angenähert. Das Modell berücksichtigte außerdem die Saisonalität, zeitliche Trends und Zusammenhänge, und klimatische Prädiktoren. Hochauflösende EIR-Karten wurden für den RDSS produziert. Zeitliche, räumliche und saisonale Muster des EIR wurden stark von umweltbezogenen Faktoren beeinflusst.

Mit dem in Kapitel 4 berechneten EIR wurden alle Standorte (Haushalte) mit den Sterblichkeitsdaten innerhalb des RDSS verbunden und die monatlichen Malaria-Übertragungsrate vorhergesagt. Zeitdiskrete Bernoulli-Regressionsmodelle wurden verwendet, um die Beziehung zwischen der allgemeinen Mortalität und der Malaria-Übertragung (d.h. EIR) in den verschiedenen Alterskategorien zu beurteilen. Separate Analysen wurden für Neugeborene (0-28 Tage), Säuglinge (1-11 Monate), Kinder (1-4 Jahre), Schüler (5-14 Jahre), Jugendliche (15-60 Jahre) und älteren Personen (>60 Jahre) durchgeführt. Die Resultate dieser Analysen werden in den Kapiteln 5 und 6 dargestellt. Die Modelle

beinhalteten sowohl räumliche und nicht-räumliche zufällige Effekte auf Dorfebene als auch monatliche zeitliche Korrelationen und bezogen das Alter und den Besitz von ITNs mit ein. Die EIR wurde in das Modell als Kovariate samt Messfehler in die Rechnung mit aufgenommen. Die Ergebnisse zeigten, dass i) die Auswirkung der Malaria-Übertragung auf die Sterblichkeit je nach Alter unterschiedlich ist, wobei schulpflichtige Kindern die größeren Auswirkungen im Vergleich als jüngere Kinder und Erwachsene erleiden, möglicherweise auf Grund der geringen Immunität, was von Bedeutung für die präventiven Maßnahmen ist, ii) der fehlende Zusammenhang zwischen Malaria-Übertragung und Mortalität bei jungen Erwachsenen untermauert die Behauptung, dass die Interaktion zwischen HIV und Malaria nicht wesentlich zur erhöhten Erwachsenensterblichkeit beiträgt, iii) ITNs eine schützende Wirkung haben, allerdings nicht signifikant auf die Neugeborensterblichkeit, wobei Wechselwirkungen zwischen ITN-Besitz und dem sozio-ökonomischen Status (SES) der Familie wahrscheinlich sind, iv) es wichtig ist, Informationen über die Expositionsvergangenheit und Malaria-Immunität einbeziehen, wenn die Zusammenhänge zwischen der Übertragung und der Mortalität zueinander untersucht werden.

Diese Arbeit ist ein Baustein zum Verständnis der Zusammenhänge zwischen Malaria-Übertragung und Mortalität. Die große Stärke unserer Ergebnisse stützt sich auf die räumlich (Haushaltsebene) und zeitlich hoch aufgelösten EIR-Vorhersagen erstellt mittels umfassender geo-statistischer Modelle für die große Anzahl an entomologischen Daten. Wir glauben, dass die bessere Abschätzung der Einflüsse eine genauere Bestimmung des Verhältnis zwischen Malaria-Übertragung und Mortalität ermöglicht, was wichtig für die Ermittlung der Folgen von Malaria-Interventionen auf die Übertragungsrate der Krankheit und die Entwicklung der Sterblichkeit ist. Allerdings brauchen wir noch weitere Ergebnisse von anderen Standorten um die Reproduzierbarkeit der meisten unserer Ergebnisse zu untersuchen. Die Entwicklung eines stochastischen Modells um verschiedene Maßnahmen der Malaria-Übertragung mit der Sterblichkeit in Zusammenhang zu setzen könnte dabei das Ziel nachfolgender Forschungen sein.

Acknowledgement

Looking back at hours, days, months and finally years that I spent on this work... I see so many that I am indebted to. Even without mentioning one name after the other, I would like you to know that, I am grateful to all the little-little things each one contributed to make my 'journey' smooth and successful.

First of all, I would like to thank Penelope Vounatsou for a great guidance and supervision, it was a pleasure working with you. The roles you played from the initial stages of my work to where it stands today. The continuously tireless support you provide, off hours I consumed which you should have spent differently but where instead dedicated to me. Thanks for the clarifications and expertism whenever I sounded naïve. With you, I gained self-confidence, focus and intellectual competence.

Tom Smith, it would not be kind of me if I don't specific appreciate your assistance. Thanks for opening your doors and willingness to respond to all of my questions about mosquitoes, not a single question that was considered irrelevant. Given the amount of other responsibilities you had, that was well appreciated. Thank you for critical epidemiological inputs and comments you provided in my work.

I would like to thank members of the Swiss TPH. I appreciate the fantastic atmosphere at the Institute, open-minded people, loving and supportive. The international *mix* was a great opportunity to scrutinize my personality and character to enable working with people with various cultures. Considering space limitation, I will just mention a few. I thank Don de Savigny who beside a number of academic and professional advices, I great value the moral support you always provided, thanks for the nice accessories (from Tanzania and other places). It was kind knowing that someone realize your "home sickness" and found a way to cheer you up. To Christian Lengeler, thanks for the explanation on lots of epidemiological concepts and assisting me in shipping a big part of my *cargo* to Tanzania; to Christine Mensch, Margrit Slaoui, Zsuzsanna Györffy, Christine Walliser, Beatrice Wäckerlin, for all the administrative support, just to make sure I lived inline with the *Swiss*. For that I say *asanteni sana*. Few more include Mitchell Weiss, Nino Künzli, Jürg Utzinger, Jacob Zinsstag, Brigit Obrist, Conny and Stefan Dongus, the IT crew, and without forgetting Konstantina Boutsika (oh I will miss your smile). Lastly, a big thank to the Director of the Swiss TPH, Marcel Tanner, for a warm welcome, letting me feel a part of the Swiss TPH family and create loveable working environment. Thanks for the wonderful and cheerful moments, the excellent Swahili phrases and greetings that I was always curious to listen to. You always reminded me that *Umoja ni Nguvu*….

I thank the whole Biostatistics Unit including the DLD family, Dominic, Laura and the little Daria who were colleagues and friends that I could knock at their door for anything and at anytime, thanks for being there; Nadine Riedel (oopss Nadine Schur) a lovely kid sister and a friend. I come from an extended family which I missed a lot. It was great to have a baby sister whom I felt responsible for (didn't work very well though as she is pretty matured☺), thanks for the times we spent together, coffee breaks, girly outings, Tandem Language sessions, the list continues without forgetting the German translation for my work, *Asante mamii*.... (*I would be indebt not mentioning Nadine's Mother, Cornelia, who was always thinking of me, asking how I'm doing work wise but also making sure I was 'keeping the looks' despite the PhD hassles around (wink), Danke vielmal Mum*); Ronaldo Scholte; Federica Giardina; Verena Jürgens; Frédérique Chammartin; Eveline Hürlimann; Alex Karagiannis; Anna-Sofie Stensgaard, Amina Msengwa (whom we shared a small hostel bed on my first visit to Basel) and last my friends-brothers-co-MTIMBA crews, Amek Nyaguara Ombek, Simon Kasasa and Eric Diboulo.

I appreciate the teamwork and cooperation I received from fellow students. I thank Mercy Ackumey (for being a sister and a spiritual booster); Sandra; Karin; Bianca; Gonzalo; Ricarda; Carol; Thomas; Steffi, Valerie; Ellen….was just nice having you at that time. The Tanzanian crew; Henry Mwanyika, Ndeniria Swai (my lovely sister), Robert Tillya, Angelina Mageni, Bernadette Huho, Pax Masimba, Angel Dillip, Amani Shao, Ester

Stöcklin, Fatuma Manzi, Mulokozi Kabanywanyi, Boniphace Idindili, Mwifadhi Mrisho, Rashidi Khatibu, Nassor Laizer, Felista Mwingira, Paschal Mugasa, Simon Vendeline, Marianna Balampama, Judith Kahama-Maro, Ritha Njau, Leah Mtui and Peta Mhoma… without forgetting the families of Beatrice Jensen, Happiness Minja, Eliza Mpare, Rashid Mhehe and Jembe Mkenya. It was nice knowing you, thanks for keeping me up to date on various things happening in my *own* country. You often took me by surprise. I am honored to give special thanks to my dear friend Maxmillian Mpina. Thank you for all the prayers, support, encouragement, chats and… without forgetting the morning "*alarms*". All through, I could count on you and there was always a shoulder to lean on.

I thank the administration of the National Institute for Medical Research in Tanzania for giving me time and opportunity to pursue my career. I appreciate the patience of my colleagues in all those tasks that I couldn't perform due to the pressure of my PhD. Mentioning a few, I give thanks to Leonard Mboera, Elizabeth Shayo, Kesheni Senkoro, Benjamin Mayala, Andrew Kajeguka and Debora Bulemo. It is my hope that the knowledge gained will produce fruitfully.

Special heartfelt thanks should go the Christian Family at the Basel Christian Fellowship. I cannot put in words the appreciation I have by being a part of this family of believers in Christ. The love, the encouragements and open arms that were provided to me all the time will never be forgotten. You are a testimony to many changes in my life and I will always remember the moment I walked in and the transformation I have at the moment, knowledgeable, stronger, and blessed much more. I personally thank the Pastor, David Manduka, for the leading, preaching, the Bible Studies, the *Pasta's* lunch (which I initially confused with *Pastor's* lunch). Manduka was (still is) a silent father to me and a guardian at the same time. Everyone was nice in her or his own way but I cannot list them all, to mention the few I thank my grands, Anita and Jim Cooper; and Auntie Mary and Uncle Niels Bloms. God bless you.

Is a great opportunity to express my deepest gratitude to my family and family-friends. First my Parents, Fred Rumisha and Beatrice Elingisanja Mlay who where ALWAYS there, I have no idea how to thank you enough in writing, let me just say *Aika Mbee, Aika Mae, Ruwa nakuringe nakurikire na iyoo*. Second my sisters, Eileen-Elizabeth *(+ family: Husband Zawadieli Maruchu and their lovely kids Ivan, Joan and my God Daughter Laura*), Frida-Ndeta, Ikunda-Geraldine, Grace-Mkunde and my brothers David-Kiwera and Leonard *(+ family: Wife Felista and lovely kids Sarah, Peter and Aminieli)*. And many among you who I couldn't catch up for a long time, specifically Bibi Maua Mongi. I apologies for all important events I missed. It is my wish that God will bless me to make up to some of you, somehow. Third, I thank the family of Annemie Janssens (Belgium), and my baby sister Sofie *(+Husband Isidro and son Son-Gautama*), Oma and Opa, Uncle Jan *(+family: Wife Dorah, Son Jan and lovely daughters Carolien and Salome)* in Bern. Thanks to other family-friends in Belgium, Switzerland and UK.

Finally I thank God for choosing me to receive this opportunity. I glorify him for the wisdom, humbleness, and possibility to love and care for those around me even at the times I was overwhelmed with many responsibilities, for the good health and lastly step-by-step guidance untill where I am today. For those who will come after me, desiring to pursue this journey, I cluster the whole of it in a nutshell… and call it "*a road to perseverance full of blessings"*.

From the fullness of his Grace I have received one blessing after another.

Come, Thou Fount of Every Blessing.

Bless be your name…

This work received a financial support from the Swiss National Foundation grant Nr. 325200-118379

Chapter 1 Introduction

1.1. Background

1.1.1. Malaria burden

Malaria is among the oldest infectious disease in human populations, and currently still endemic in over 90 countries worldwide, mainly in the tropics (Garcia 2010; Hay et al. 2009) (Figure 1.1). In 2009, 169–294 million cases of malaria with about 780 thousands deaths were estimated worldwide (WHO 2010b). Around 78% of these cases and over 90% of the deaths occurred in the African region. Most deaths (85%) are observed in children under five years of age (Snow et al. 1999; WHO 2007, 2010a). Reduction and resurgences in malaria cases and deaths have been reported in different regions, nevertheless, the situation in SSA is still intolerable (Bryce et al. 2005; WHO 2006, 2010a). A number of vector control and intervention programs have been put in place to reduce transmission levels and burden of the disease. However, for decades, due to lack of longitudinal surveillance data, no clear evidence on the association between effectiveness of interventions and changes in transmission intensities has been confirmed (Smith et al. 2001; Snow et al. 1997). Evidence-based and effective vector control and elimination/eradication programs are needed to clearly understand dynamics in malaria burden, transmission and their association with other health outcomes (The malERA Consultative Group on Vector Control 2011). This creates an urge to generate accurate and comprehensive data on these parameters to allow rigorous and precise analysis of all potential linkages (de Savigny and Binka 2004; Greenwood et al. 2005).

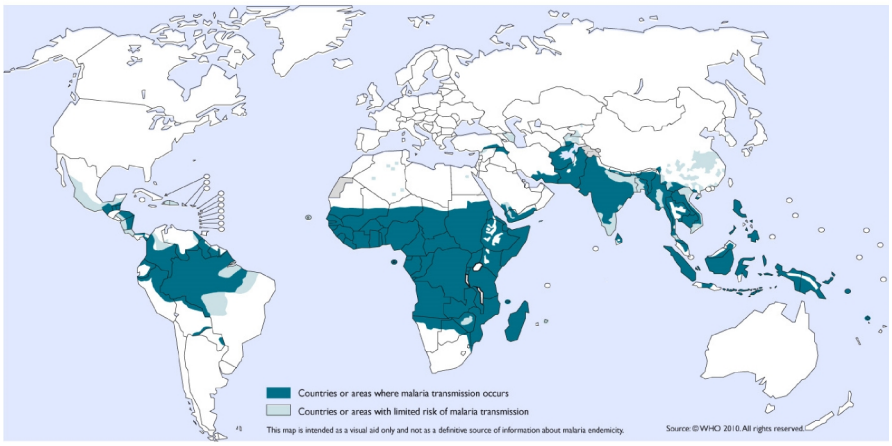


Figure 1.1: Risk of malaria transmission worldwide (Source: WHO, 2010)

1.1.2. Species, vector and behavioral characteristics

Malaria is an infectious disease caused by parasitic protozoa of the genus Plasmodium transmitted by a bite from an infected mosquito. The four species of *Plasmodium* that cause human malaria are *P. falciparum*, *P. vivax*, *P. ovale*, and *P. malariae* (Warrell & Gilles, 2002). The predominant species is the *P. falciparum* which is confined in the tropics and is known to cause the most dangerous form of the disease. *P. vivax* is present in the tropics, subtropics and in temperate zones while no specific areas are defined for *P. malariae* (Garcia 2010). *P. ovale* is most prevalent in the region of West Africa. A fifth species, *P. knowlesi*, primarily infecting monkeys, has been identified in Southeast Asia (Greenwood et al. 2008; Jongwutiwes et al. 2004; Singh et al. 2004).

The vector of the malaria parasites are the female anopheline mosquitoes. The mosquito is the primary host while humans act as intermediate host and reservoir of the parasite. When an infected mosquito bites an individual, the sporozoites, which are contained in the salivary glands, are injected into the human bloodstream and migrate to the liver (Beier et al. 1999). The liver cells become infected and form schizonts which rupture and release merozoites. This stage is referred to as the exo-erythrocytic cycle. The merozoites attack the red blood cells and feed on the haemoglobin. The immune system fights most of the merozoites, however, some invade the cells to initiate the so called erythrocytic cycle. Within this cycle, the infected cells form into a ring-shaped immature trophozoite which later develops new schizonts. With time and several generations of this cycle, some merozoites develop into male and female forms, namely gametocytes. When a mosquito takes a blood meal from an infectious person and picks gametocytes, the sporogonic cycle is initiated. Within the mosquito gut, the male and female gametocytes mate and mature via different stages into oocysts. The oocysts grow and rupture to release new sporozoites which migrate to the salivary glands of the mosquito. In case the mosquito is taking another blood meal, the sporozoites are injected to an individual and the cycle starts again (Figure 1.2).

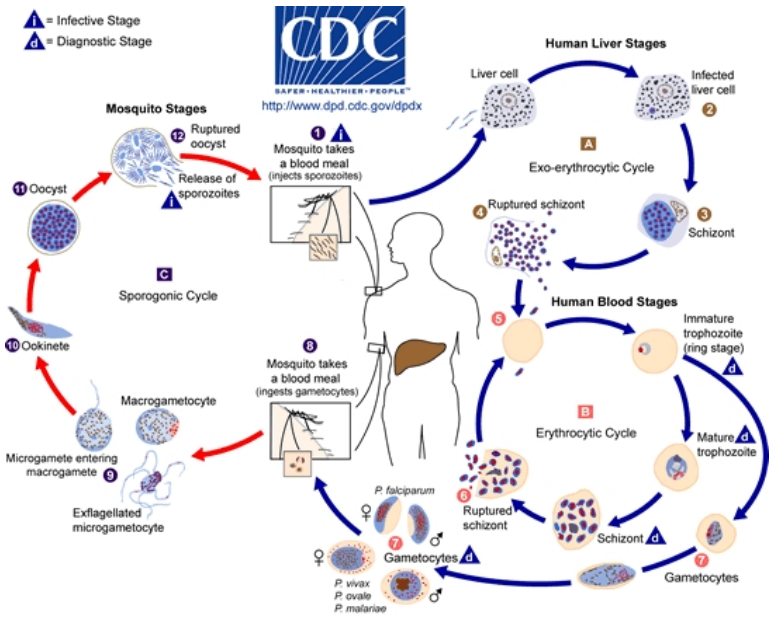


Figure 1.2: Malaria transmission cycle (Source: CDC)

Approximately 430 *Anopheles* species are known, of which only 30–50 are able to transmit human malaria. In SSA, the primary vectors for malaria transmission are *An. gambiae* complex and *An. funestus* (Coetzee and Fontenille 2004; Gillies and De Meillon 1968; Hay et al. 2005). Behavioral characteristics of these vectors, such as feeding and habitation preferences, distinguish their potential to transmit malaria. Environmental and climatic conditions, mainly temperature, rainfall and humidity, are the key drivers of the vector's population density, distribution and survival. *An. funestus* is responsible for dry season malaria while *An. gambiae* is a dominant vector in the rainy seasons (Fontenille et al. 1997a). High temperature (between 25 and 30 °C) catalyzes development of parasites within mosquitoes and shortens the gonotrophic cycle (simply the time between two consecutive blood meals) which increases mosquito population and biting frequency (Afrane et al. 2005; Charlwood and Graves 1987; Lardeux et al. 2008; Quiñones et al. 1997). Parasite development ceases at 16°C while mosquito thermal death is expected at temperatures above 40 °C (Haddow 1943; Craig et al. 1999). Rainfall creates breeding sites and favorable humid conditions for mosquito growth. However, heavy rain flushes away mosquito larva preventing development of adult vectors. In tropical regions, transmission of malaria follows a distinctly seasonal pattern influenced by seasonality in climatic factors (Hay et al. 1998; Thomson et al. 1997).

1.1.3. Clinical features and malaria diagnosis

The symptoms of malaria, specifically *P. falciparum* infections, include headache, joint aches, fatigue or malaise, fever, and sometimes diarrhea and vomiting. Initial symptoms typically appear about seven days after receiving an infectious bite. In most cases, untreated or poorly managed malaria case progress to severe form of the disease which might even cause death, sometimes within several hours after the first symptoms. Therefore, treatment within 24 hours of onset is highly recommended. Severe case of the disease might result to cerebral malaria, severe anaemia (especially in children), other serious complications such as brain damage (due to high fever and seizures) and kidney failure might (mainly in adults) (WHO 2010a). In remote areas where no laboratory facilities are available, malaria is often diagnosed presumptively. However, proper diagnosis requires microscopic examination of blood films and other specialized tests, such as serology and polymerase chain reaction (CDC 2010). Recently, rapid diagnostic tests have been introduced to improve diagnosis and quality of management of malaria patients (D'Acremont et al. 2009; WHO 2009).

1.2. Measuring malaria transmission

The malaria transmission cycle involves both the host and the vector (Figure 1.2), hence the intensity of transmission is assessed using parasitological indicators (such as parasite density, prevalence or incidence rates describing the strength in human) and/or entomological parameters (such as mosquito density, sporozoite rates, survivorship or vectorial capacity from which human exposure to infectious mosquitoes can be measured) (MacDonald 1957). These measures quantify the risk and intensity malaria transmission and can be used to evaluate effectiveness of intervention strategies (Molineaux 1988; Smith and McKenzie 2004; Shaukat et al. 2010).

The most direct measure of malaria transmission intensity is the entomological inoculation rate (EIR). The EIR estimates the number of effective bites per person at a certain unit of time and quantifies the level of exposure to infected mosquitoes (Burkot and Graves 1995; Beier et al. 1999; Billingsley et al. 2005; Killeen et al. 2006; Kelly-Hope and McKenzie 2009). It is derived by the product of man-biting mosquito density, SR (proportion of infected mosquitoes) and

human blood index (proportion of blood meals obtained from humans) (Garrett-Jones 1964; Birley and Charlewood 1987; Drakeley et al. 2003). EIR varies between (small) regions due to differences in environmental conditions, type and effectiveness of interventions (Kelly-Hope and McKenzie 2009; Shaukat et al. 2010). As it measures human risk of exposure to infection it is preferable and widely used. Additionally, EIR can be estimated under (manageable) field conditions and directly associated with prevalence and incidence rates (Trape and Rogier 1996; Beier et al. 1999; Smith et al. 2004) and other public health indicators such mortality (Smith et al. 2005; Ross et al. 2006).

The length of the gonotrophic cycle is another measure to estimate the strength of transmission and survival of anopheline mosquitoes (Rodriguez et al. 1992; Fernandez-Salas et al. 1994; Quiñones et al. 1997). It is governed mainly by temperature (MacDonald 1957; Gu et al. 2006) and comprises of three stages: i) search for a host and taking of a blood meal, ii) blood digestion and development of the ovaries, and iii) search for a suitable breeding place and oviposition. Additionally to temperature, distance between breeding sites and hosts, and use of insecticide treated nets alter the length of this cycle (Detinova 1962; Hii et al. 1995; Gu et al. 2006). These factors need consideration when assessing malaria transmission.

Alternative transmission measures include infectious reservoir and vectorial capacity (VC). The infectious reservoir estimates the probability of a mosquito to become infected when taking a feed from an infected human (Charlewood 1997; Killeen et al. 2000; Pethleart et al. 2004; Smith and McKenzie 2004) and it is useful in assessing changes in infection rates after introducing control measures (Jeffery and Eyles 1955; Pethleart et al. 2004). The VC expresses the potential of malaria transmission. Assuming a perfect transmission (i.e. a human host always contract an infection when bitten by infective mosquitoes), it represents the expected number of infectious bites from a single vector after feeding on an infectious host. VC is the best measure to quantify receptivity to malaria infections (Adlaoui et al. 2011). However, its estimation requires knowledge of several parameters such as the emergence and survival rates of mosquitoes. These parameters vary significantly in place, time and species and are difficult to obtain in practical scenarios (Dye 1986). Mathematical models using theoretical data and assumptions on transmission competence and vector survival are usually

used to estimate VC but suffer from bias and limited practical applicability (Molineaux 1988). Estimation of infectious reservoir of *Plasmodium* requires direct feeding of mosquitoes through human volunteers or artificial membranes. Seeking ethical consents laboratory work and resources are among few limitations for routine application of this measure (Githeko et al. 1992; Drakeley et al. 2000; Shaukat et al. 2010).

Simplicity in estimation and direct interpretation made EIR a commonly used measure of malaria transmission as compared to other measures such as infectious reservoir and VC. However, in addition to the spatio-temporal heterogeneity in mosquito density and SR which result into a large variability of EIR, in some settings malaria transmission might be concentrated in a small fraction of population making it difficult to select a representative sample of human population to where EIR should be estimated (Smith et al. 1995; Woolhouse et al. 1997; Billingsley et al. 2005). In most cases, information on important factors that alter EIR such control interventions (e.g. ITNs, IRS), demographic (e.g. urbanization, migration, high population density) and climatic (e.g. rain, temperature) are not readily available to be accounted for, resulting into inaccurate estimation of EIR, poor association with other disease indicators and lack of comparability between and within regions (Service 1977; Hay et al. 2000b; Robert et al. 2003; Mboera 2006; Kelly-Hope and McKenzie 2009).

1.3. Malaria control and interventions

Campaigns to fight malaria started before the World War II with main focus in larval control and source reduction managed to eradicate malaria in regions including Italy, United States and Israel (Kitron and Spielman 1989). In the year 1939, a Swiss chemist, Paul Muller discovered that the chemical dichlorodiphenyltrichloroethane (DDT) could eradicate many insects without harming human beings. Availability of DDT shifted malaria control focus to strategies that reduce vector populations through indoor residual spraying (IRS) which were complimented with case management using effective drugs (quinine and chloroquine) for prophylactic and treatment (Harrison 1978). This resulted to a substantial reduction of malaria in many states, for example in India. The shift provided great hope on the fight against malaria resulting to initiation of WHO Global Malaria Eradication Program in 1955 which ran until 1970s (WHO

1957). The initiative aimed at global eradication of the disease using vector control strategies. Great success was achieved in Europe, some African countries and the Middle East. Unfortunately, the program excluded most of the SSA countries due to difficulties in logistics, operational and financial issues (Wernsdorfer and McGregor 1988; Nájera 2001; Greenwood and Mutabingwa 2002). In 1978, following the Alma Ata Declaration, most African states adopted the primary health care strategy as an alternative approach by integrating malaria control activities within the health systems (Muturi et al. 2008; Christopher et al. 2011). This strategy was not very effective as it required funds and well functioning health systems which were not in place in most of these countries at the time. Other reasons contributed to poor success were increasing anti-malarial drug resistance and failure in vector control programs due to insecticide resistance (WHO 1984; Sachs and Malaney 2002; Alilio et al. 2004). With time, the burden of malaria disease was increasing significantly (WHO 1984).

In 1988, the RBM partnership was launched with the idea of redesigning strategies and approaches previously used with additional components in the health systems targeting case management, improve use of ITNs and preventive treatment (RBM, WHO 1999). Partners of RBM include WHO, United Nations Children's Fund (UNICEF), the United Nations Development Programme (UNDP) and the World Bank. Other programs and international initiatives supporting the RBM vision include the Multilateral Initiative on Malaria (founded in 1997) which focuses on research and training; Medicines for Malaria Venture (founded in 1999) dealing with development and delivery of new antimalarial drugs; and the Malaria Vaccine Initiative (founded in 1999) which works on the development of vaccine (Davies 1999; Wheeler and Berkley 2001; Rabinovich 2002; PATH, Program for Appropriate Technology in Health 2007). Via improved treatment and prevention strategies targeting the most vulnerable groups of malaria infections, i.e. children under five and pregnant mothers, RBM aimed to halve malaria mortality in SSA by 2010 and to reduce cases and malaria deaths close to zero by 2015 as compared to the 2000 statistics (WHO 2000; Yamey 2000; Sachs and McArthur 2005; RBM:GMAP 2008). The long term vision of RBM is a world-wide eradication of malaria through progressive country by country elimination (RBM:GMAP 2008).

Funding in malaria research has increased widely in the past decade with a number of funding agents involved in development of new tools and interventions to achieve eradication (Roberts and Enserink 2007; WHO/UNICEF/PATH 2010). Malaria control intervention strategies principally aim at i) reducing the vector population by implementing IRS and source reduction, ii) reducing human-vector contact by promoting the use of LLINs/ITNs (Maxwell et al. 2002, 2003; Flaxman et al. 2010), iii) preventing the vulnerable groups from infections through interventions such as intermittent preventive treatment in pregnancy (IPTp), intermittent preventive treatment for infants (IPTi) and development of vaccines (Bejon et al. 2008), and iv) ensuring prompt and accurate case management through monitoring treatment policies and drug resistance. Among those available today, the use of LLINs/ITNs is the most effective (Lengeler 2004; Fegan et al. 2007; WHO 2008a). According to the World Malaria Report, more than 80 million ITNs were distributed to SSA by 2009 which have contributed to a significant reduction in mortality, low birth weight and mosquitoes population (Gamble et al. 2007; Killeen et al. 2007a; Noor et al. 2009; WHO 2010b). Introduction of IPTp and IPTi have also shown positive effects on the burden of malaria in a number of areas (Schellenberg et al. 2001a; Eijk et al. 2004, 2005; Hommerich et al. 2007; Mockenhaupt et al. 2007; Manzi et al. 2009). Replacing single antimalarial therapy by artemisinin-based combination therapy (ACT) and use of rapid diagnostic tests have had great improvement in the treatment and accurate diagnosis of malaria in endemic countries (D'Acremont et al. 2009; MMV 2009; Snow and Marsh 2010; WHO 2010b).

The multi -innovative and -sectorial efforts are expected to sustainably lessen malaria burden and to significantly reduce or even interrupt transmission. However, weak health systems, lack of tools and understanding of the heterogeneity in malaria transmission create doubts in the achievement of the RBM vision (Tanner and de Savigny 2008).

1.4. Malaria transmission and mortality

In areas with high levels of malaria transmission, most deaths are associated with malaria infection (Marsh and Snow 1999; Smith et al. 2001; Korenromp et al. 2003, 2004; Idro et al. 2006; Rowe et al. 2006; Okiro et al. 2009). Recently, most of SSA countries have experienced a downshift trend in mortality especially in infants and children (Rajaratnam et al. 2010a; WHO 2010b). The advances in implementation, access and success in malaria control interventions strategies on transmission, and efficient malaria treatment are among the factors driving the observed mortality pattern (Ahmad et al. 2000; Kleinschmidt et al. 2009a; WHO 2010a). However, from a clinical point of view, reducing malaria transmission may affect acquisition of malaria immunity in children and shift the mean age of child mortality to older ages (Snow and Marsh 2002; Lindblade et al. 2004; Omeara et al. 2008). Todate arguments regarding the way malaria transmission affects mortality are contradictory. However, to sustain progress in malaria control, adequate knowledge is required on the effect of interventions on transmission and consequences of altering transmission on malaria morbidity and mortality.

Several efforts aimed to understand the relationship between malaria transmission and mortality in Africa (Lengeler et al. 1995; Snow and Marsh 1995a; Trape and Rogier 1996; Smith et al. 2001; Phillips-Howard et al. 2003; Lindblade et al. 2004; Lim et al. 2011). Yet, opposing results and conclusions have been reported. For example, Smith et al. (2001) showed a positive association between infant mortality rate and EIR in Africa. A study in Western Kenya, (Lindblade et al. 2004) observed no difference in mortality rates between areas with and without ITNs interventions. Other studies which relate ITN efficacy and child mortality include those in Tanzania(Schellenberg et al. 2001b), Kenya (Fegan et al. 2007) and the Gambia (D'Alessandro et al. 1995). In another attempt, Demographic and Health Surveys (DHS) and Mapping Malaria Risk in Africa (MARA) databases were linked to assess the transmission-mortality relation, but no clear relationship was observed (Gemperli 2003). Based on the results of these studies, it is difficult to generalize and foresee the effect of interventions on malaria burden. However, most of these studies are either small scale trials or they are based on aggregated data collected at different time points (seasons) and places, and probably not aimed to investigate the effect of malaria transmission on mortality. Analysis of a comprehensive database linking vital events, information on transmission, interventions and other risk factors related to poverty and health systems, is required to study the malaria-mortality relationship.

1.5. Demographic surveillance systems, malaria transmission intensity and mortality

Most African countries do not have reliable vital registration systems, resulting into invalid estimates of all-cause or malaria specific deaths. Long term and standardized monitoring of entomological parameters and evaluation of prevention and control interventions is also lacking. To address data needs, demographic surveillance systems (DSS) were established in a number of developing countries (INDEPTH Network 2002; Jamison and Bank 2006). DSS monitor birth, deaths, cause of death, in- and out- migration and other demographic information such as social economic status, education and access to health services of a defined population. DSS serve as the best platform of accurate and reliable data to understand population diversities (in regions of similar conditions/settings) and to study all-cause and cause-specific mortality (UNICEF/MICS 1995; Korenromp et al. 2004).

The International Network for the Demographic Evaluation of Populations and Their Health in developing countries (INDEPTH) Network was formed in 1998 to strengthen capacity, to monitor and to facilitate analysis of DSS data, and to conduct multi-site comparative demographic research (Sankoh and Binka 2005a). Figure 1.3 shows a map with countries involved in this network.

Figure 1.3: Countries with DSS sites within the INDEPTH Network (Source: SIDA Review, 2010)

In 2002, the INDEPTH Network launched the Malaria Transmission Intensity and Mortality Burden across Africa (MTIMBA) initiative aiming to collect information that will guide evidence-based malaria control policies in Africa (Sankoh and Binka 2005b; INDEPTH Network- Homepage 2011). Particularly, the project was designed to elicit understanding of the relationship between malaria transmission intensity, mortality and the effect of control interventions. Countries involved in MTIMBA project were Ghana, Kenya, Mozambique, Burkina Faso, Senegal and Tanzania. Integrated within the DSS, the MTIMBA project collected entomological data at a large number of survey locations (households) for a period of three years. The tools and methodology were standardized across all sites. The project covered over 700,000 people living in about 10,000 km^2 and collected hundreds of thousands of mosquitoes (Kasasa et al. in preparation). The entomological parameters were linked to time, place and vital events in a population which is rigorously monitored through the DSS. To date, the MTIMBA database is the only comprehensive geo-referenced entomological database with information which allows studying space-time heterogeneity and seasonality of malaria transmission in relation to mortality.

1.6. Mapping of malaria transmission

Climate and environment are the main factors of malaria transmission. Precise mapping of malaria transmission requires reliable data on the outcome and risk factors data, as well as appropriate techniques in estimating existing relationship. Malaria mapping involved use of geographical information systems (GIS) techniques, remote sensing data and geostatistical modeling (Beier et al. 1999; Kleinschmidt et al. 2000, 2001a; Thomson and Connor 2000; Diggle et al. 2002; Gemperli 2003). GIS can manage, manipulate, analyze and display geo-referenced data from different sources. Remote sensing (RS) is a technique to derive information about ground objects on the earth surface via sensors attached on satellites. RS data has been used for more than thirty years to extract information on climate conditions at high spatial and temporal resolution. These data can be linked with malaria transmission and endemicity to accurately estimate their association, identify potential risk factors, predict at un-surveyed locations and display spatial patterns of the disease (Hay et al. 2000a, 2009; Tanser et al. 2003; Hay and Snow 2006; Usher 2010).

Malaria transmission can be mapped using parasitological (prevalence) data (compilation of historical or contemporary national surveys), clinical incidence information and data from entomological surveys. However, historical data are heterogeneous in age, period, methodology, diagnostic methods and might include sparse survey locations. These heterogeneities and lack of important information such as intervention and other demographic factors at the time of collection complicates spatial mapping (Gemperli et al. 2006a; Gosoniu 2008). National parasitological surveys are routinely carried out as part of Malaria Indicators Survey (MIS). These provide the best source of data for estimating the geographical distribution of the disease risk at national scale but cannot be used to estimate seasonal and temporal variations. Incidence data especially in infants and children might reflect a direct impact of malaria transmission, however, should be used carefully as they require precise estimates of population at risk. All these data indirectly measure intensity of malaria transmission. Entomological data, if collected longitudinally using standardized techniques, would be most appropriate to directly assess malaria transmission since entomological parameters, such as EIR can be estimated.

Attempts to produce global and regional maps of malaria transmission initially involved mainly models based on climatic suitability for vectors supported by expert opinion. Lysenko and Semashko (1968) produced the first global map of malaria risk based on relationship between the duration of sporogony and temperature (global isotherms). In considering temperature limits of transmission, elevation and rainfall conditions ideal for parasites and vector survival, Dutta & Dutt (1978) mapped a possibility of malaria occurrence worldwide. Combining GIS techniques, spatial interpolation and a developed fuzzy climatic suitability model, Craig et al. (1999) obtained a malaria map for Africa based on biological relations of climate on parasite and vector development. Malaria vector niche models were applied to generate a high resolution map of malaria prevalence in northern Tanzania (Kulkarni et al. 2010) by estimating the interaction between mosquito niche space and climate. Further refinement of ecological approaches involved incorporating (historical) parasitological databases such as those compiled in the “Mapping Malaria Risk in Africa” (MARA) project (MARA/ARMA 1998) or Malaria Atlas Project (MAP) (Guerra et al. 2007) with climatic data (Omumbo et al. 1998; Snow et al. 1998a). Following advances in modeling techniques and use of RS data the Bayesian

geostatistical models were applied to produced smooth maps of malaria risk at different regions (Diggle et al. 2002; Gemperli et al. 2006a; Gosoniu et al. 2006) and the most recent global map of malaria endemicity (Hay et al. 2009). Other applications of mapping malaria using parasitological data, include Bayesian analysis of MIS databases adjusting for interventions, climatic and socio-economic factors which helped in measuring the malaria-related burden and produce smooth maps of malaria prevalence in African countries such as Zambia (Riedel et al. 2010); Angola (Gosoniu et al. 2010) and Liberia (Gosoniu and Vounatsou 2011). Model-based estimation is advantageous over GIS-based approaches as it can quantify the outcome-predictor relationship, but more importantly, it allows predictions in areas with no survey data with a measure of uncertainty (Banerjee et al. 2003; Patil et al. 2011).

Mapping malaria transmission using incidence data received less attention in Africa due to lack of reliable disease surveillance system in most of developing countries. Spatial generalized linear mixed models and model-based prediction were applied to produce smooth small-area maps of malaria incidence in South Africa (Kleinschmidt et al. 2001b, 2002). Employing malaria incidence data from several Southern African countries, maps showing spatial-temporal patterns and seasonality of malaria incidence in relation to climate were produced (Mabaso et al. 2006, 2007) and in another example, a Bayesian hierarchical model accounting for spatial variation was employed to link incidence of malaria and weather parameters in Mozambique but with no prediction (Zacarias and Andersson 2010, 2011).

EIR has been widely applied to map malaria transmission in Africa. The literature includes a work by Rogers et al. (2002) who produced a map of EIR in Africa after determine factors associated with the variation of malaria mosquito density and transmission. In The Gambia, entomological field data, satellite image analysis and GIS modeling techniques were combined to produce a smooth map of EIR (Bøgh et al. 2007). Employing the MARA database, Gosoniu et al. (2008) and Gemperli et al. (2006b) used transmission models to translate prevalence data to EIR parameters by fitting Bayesian geostatistical model adjusting for environmental and climatic factors and produce high resolution smooth map of EIR for Mali. However, estimating EIR from prevalence data might results into bias estimation of transmission due to small number of locations involved in collection of parasitological data.

1.7. Geostatistical modeling

Development and application of geostatistical models focused in estimation and mapping of malaria transmission exposures (i.e. EIR) within the study area at high spatial-temporal resolutions accounting for seasonal variations. The exposure surfaces are used to assess age-specific transmission-mortality relationships adjusting for control interventions while incorporating transmission uncertainty.

1.7.1. Bayesian modeling of large geostastical data

DSS collect data longitudinally (temporal) and over large number of fixed locations (geostatistical). Geographical proximity of locations, similarities in survey timing and common exposures introduce clustering, spatial and temporal correlation. Most of the classical statistical methods assume independence between observations. Ignoring space-time correlation might result in over/under-estimation of the significance of covariates (Ver Hoef et al. 2001; Cressie et al. 2009; Riedel et al. 2010). In modeling geostatistical data, random effect parameters are introduced at each location to account for a spatial correlation. The distribution of the random effects is assumed to be a zero-mean multivariate normal with a defined covariance matrix, referred to as Gaussian spatial process. The correlation between any pair of locations is considered to be a function of distance between them and modeled by the covariance matrix of the process (Gemperli et al. 2006a; Schur et al. 2011a). The number of these parameters increases with the number of survey locations.

Introduction of location-specific random effect parameters in geostatistical modeling results in highly parameterized models. It is difficult to apply maximum likelihood-based methods to estimate such a (very) large number of parameters. Bayesian approaches implemented via MCMC algorithms (Gelfand and Smith 1990) avoid the computational issues and can be used to estimate the outcome-predictor relation and spatial correlations (i.e. the random effects parameters) simultaneously (Gelfand and Smith 1990; Diggle et al. 1998). MCMC are iterative algorithms which estimate model parameters by drawing samples from their posterior distribution. The main challenge of MCMC estimation carried out in geostatistical modelling is that it requires repeated inversion of the covariance matrix of the spatial process. The inversion of this matrix is numerically not feasible when the number of locations (n) is very large as the computational expenses increase exponentially in terms of time and accuracy (Xia and Gelfand 2005).

1.7.2. Approximation of spatial processes using a subset of data

Several approaches have been suggested to overcome the computation problems involved in modeling large geostatistical data. These include the use Gaussian Markov fields (Rue and Tjelmeland 2002), sampling importance re-sampling methods (Smith and Gelfand 1992), parallelized computations (Yang et al. 2007), use of Kernel convolutions (Higdon 2001) and many more (see Chapter 3 for details). However, criticism related to adequacy (Jones and Zhang 1997; Stein et al. 2004), tuning (Gemperli and Vounatsou 2003) and performance (Gemperli and Vounatsou 2006) of some of the proposed algorithms have been reported. Recently, Banerjee et al. (2008) and Finley et al. (2009) proposed new strategies to fit large Gaussian geostatistical data by approximating the spatial process from a subset of locations, m ($m << n$) which reduces the dimension of the covariance matrix from $n*n$ to a smaller dimension $m*m$ which can be easily inverted. Direct application of the new proposed methods on field data is hindered by the underlying requirements including Gaussian outcomes and data observed on regular grids.

Selection of the subset for approximating the spatial process is an additional challenge. Random selection of the subset locations has been discussed (Banerjee et al. 2008) but does not perform well for clustered data. Finley et al. (2009) proposed selecting the subset of locations while minimizing the variance of the approximation, namely the spatially averaged predicted variance (SAPV). Compared with other sampling approaches including random sampling, balanced sampling (Deville and Tillé 2004) and minimax space filling design (Johnson et al. 1990), SAPV has been claimed to be the most efficient and precice method for sampling. However, this approach requires prior knowledge on spatial characteristics of the data and it is computationally expensive even for small subsets (Finley et al. 2009). In addition, further development of these approaches to non-Gaussian and irregularly distributed data, and exploration of less computation techniques to select subsets are essential. Moreover, evaluation of the performance of subsets to approximate the spatial process in real-life field data is also needed.

1.7.3. Modeling large sparse zero-inflated entomological data

Entomological data are usually sparse and with high variability, especially when collected over a large area and over a long period of time. The variability is caused by large number of zero outcomes, either in the SR or mosquito densities. This could be a result of environmental conditions for the mosquitoes but also related to logistic and infrastructure problems which make it infeasible to visit some areas. Excess zeros create over- or under-dispersion and zero-inflation which implies that the data contain too many zeros in relation to the underlying data distribution (e.g. Poisson for count data). In statistical analyses, EIR data or a logarithm logarithmic transformation of EIR is assumed to follow a Normal distribution and a linear regression model is applied. However, it is difficult to transform to normality sparse EIR data. In addition, EIR is the product of SR (binomial data) and density (count data), therefore a rigorous analysis should model SR and density data separately. Zero inflated analogues of the standard Poisson/negative binomial or binomial distributions are more appropriate to analyze sparse data (Greene 1994; Cheung 2002; Yau et al. 2003). A zero-inflated model is a two-component mixture model with a point mass at zero and a proper data distribution (Ridout et al. 2001; Yau et al. 2003). Zero inflated models have been applied in several studies however, applications in malaria transmission modeling are very limited (Barnes et al. 2005; Banerjee et al. 2008; Manh et al. 2011).

1.7.4. Modelling seasonality and temporal effect in malaria data

Malaria is an environmental disease with strong seasonal patterns. Accurate assessment of seasonality is important to understand the seasonal variability which is necessary for timely targeting of interventions and control programs. Assessing seasonal patterns in malaria has been often implemented by relating occurrence of events with climatic factors such as rain (Hay et al. 1998; Thomson et al. 1999, 2008; Mabaso 2007). This is done by comparing the outcome of interest among categories (e.g. wet and dry seasons) either through calculating seasonality indices (Mabaso et al. 2005; Rau 2006), summary statistics (such as frequency, rates, proportions), (Becker and Weng 1998; Pascual et al. 2008), or significance tests such as chi-square (Bailey et al. 1992; Hamad et al. 2002; Singh et al. 2007; Vitali et al. 2009).

However, seasonal data often are collected over a time series which introduces temporal correlation that is hardly captured using methods described earlier (Cameron and Trivedi 1990). Model-based approaches are recommended for accurate assessment of seasonal and temporal trends in the time series data (Zhang et al. 2007; Briët et al. 2008).

The use of harmonic functions (separately or within regression models) has been suggested to efficiently model seasonal trends (Stolwijk et al. 1999; Kynast-Wolf et al. 2006; Zeger et al. 2006; Ramroth et al. 2009; Griffin et al. 2010). Most of these formulations capture seasonal pattern but ignore temporal correlation. The correlation in time can be modeled by introducing random temporal parameters at each time point (such as week, month) which are considered to be either independent and follow a normal distribution with a common variance, or temporally correlated and modeled via an autoregressive process of a specific order. Selected trigonometric functions with extended terms accounting for temporal correlation can be employed to capture seasonal and temporal trends in malaria transmission and mortality, especially in the data coming from the DSS.

1.8. Aims and objectives

The aims of this thesis were i) to develop Bayesian geostatistical models to analyze large and sparse space-time count and binomial data by approximating the spatial process from a subset of survey locations, accounting for seasonality and temporal correlation ii) to validate and implement the developed methodology by analyzing the large point referenced entomological field data collected under the MTIMBA project, to estimate small scale heterogeneities and produce smooth maps of malaria transmission and lastly iii) to assess the malaria transmission attribution to mortality in children and adults accounting for the effect of interventions and demographic parameters.

1.8.1. Specific methodological objectives

- assess existing methods to capture seasonal pattern and temporal correlation in the data and propose a formulation to be incorporated in the spatio-temporal models (Chapter 2).
- develop geostatistical spatial-temporal model to analyze sparse non-Gaussian large data observed irregularly in space by extending existing methods developed for Gaussian outcomes and approximate the spatial process from a subset of locations (Chapter 3).
- propose a less computational strategy for selecting the subset of locations (knots) for approximating the spatial process preserving the variability of the outcome and the spatial process configuration (Chapter 3).
- formulate a rigorous approach for analyzing EIR data considering distributions generating EIR data, i.e. binomial (sporozoite rate) and negative binomial (density), (Chapter 4).

1.8.2. Epidemiological questions addressed by the developed methods

- estimate seasonal peaks of mortality at different stages of infancy in malaria endemic site.
- mapping seasonal and spatial variation of sporozoite rate of *An. gambiae* and *An. funestus* adjusting for environment and climate factors as covariates in Rufiji DSS, Tanzania.
- modeling spatial-temporal heterogeneity and mapping of malaria transmission while accounting for overdispersion and zero inflation and produce smooth seasonal and temporal maps of EIR for Rufiji DSS, Tanzania.
- assessing the relationship between malaria transmission and survival in children under fives and adults in the Rufiji DSS, Tanzania, correcting for demographic parameters and malaria control intervention (Chapter 5 and 6).

Chapter 2 Assessing seasonal variations and age patterns in mortality during the first year of life in Tanzania

Rumisha SF[1,2,4], Smith T[1,2], Abdulla S[3], Masanja H[3] and Vounatsou P[1,2‡]

[1]*Swiss Tropical and Public Health Institute, Basel, Switzerland,*

[2]*University of Basel, Basel, Switzerland,*

[3]*Ifakara Health Institute, Dar es Salaam, Tanzania,*

[4]*National Institute for Medical Research, Dar es Salaam, Tanzania,*

[‡]Corresponding author

This manuscript has been published in *Acta Tropica* **126** (2013) 28-36

Abstract

Lack of birth and death registries in most of developing countries, particularly those in sub-Saharan Africa led to the establishment of Demographic Surveillance Systems (DSS) sites which monitor large population cohorts within defined geographical areas. DSS collects longitudinal data on migration, births, deaths and their causes via verbal autopsies. DSS data provide an opportunity to monitor many health indicators including mortality trends. Mortality rates in Sub-Sahara Africa show seasonal patterns due to high infant and child malaria-related mortality which is influenced by seasonal features present in environmental and climatic factors. However, it is unclear whether seasonal patterns differ by age in the first few months of life. This study provides an overview of approaches to assess, capture and detect seasonality peaks and patterns in mortality using the infant mortality data from the Rufiji DSS, Tanzania. Seasonality was best captured using Bayesian negative binomial models with time and cycle dependent seasonal parameters and autoregressive temporal error terms. Seasonal patterns are similar among different age groups during infancy and timing of their mortality peaks do not differ. Seasonality in mortality rates with two peaks per year is pronounced which corresponds to rainy seasons. Understanding of these trends is important for public health preparedness.

Keywords: Seasonality modeling, harmonic models, mortality, Demographic Surveillance Systems, Bayesian inference, MCMC

2.1. Introduction

Demographic Surveillance Systems (DSS) sites established within the INDEPTH network in many Sub Saharan African countries continuously collect large amount of data on trends of disease morbidity, mortality, in- and out-migration (INDEPTH Network, 2002; Sankoh and Binka, 2005). Due to lack of efficient and reliable national vital registration systems to collect data on population and health in developing countries, these surveillance data are invaluable for evidence-based health planning and guide policy decisions (Snow et al., 1998; INDEPTH Network, 2002; Kynast-Wolf et al., 2002; Hammer et al., 2006; Lutambi et al., 2010). DSS data have been used to study number of health indicators such as mortality rates, produce life tables and estimating burden of diseases (Tollman and Zwi, 2000; Korenromp et al., 2004).

Most DSS sites are located in rural areas and in countries which are endemic for malaria (http://www.indepth-network.org) where environmental factors such as rain and temperature influence highly the patterns and seasonality of transmission (Snow et al., 1998). In these sites, just like in many parts of Sub-Saharan Africa, one-fifth of all deaths that occurred, including those of neonates, infants and children between 1-5 years is attributed to malaria alone (WHO, 2010). That is due to indirect effect of malaria including low birth weights (Steketee et al., 2001), high prevalence at early stage of life (Greenwood, 2006) and lack of proper determination of specific cause of death (Masanja et al., 2008; Ramroth et al., 2009; Adjuik et al., 2010; Shabani et al., 2010). Nevertheless, factors related to child immune development and passively transferred antibodies from mother to child, are believed to cause a relatively protection of neonates and infants under age of 3-6 months from severe consequences of malaria illness (Riley et al., 2000; 2001; Mutabingwa et al., 2005) hence malaria-related mortality would be expected to be low (Amaratunga et al., 2011; Kitua et al., 1996; Le Hesran et al., 2006; Snow et al, 1998). However, other differential factors such as age, quality of health services and genetics influences ultimately modify the mortality pattern, especially in young children (Riley et al., 2000; Poespoprodjo et al., 2010).

The complexity of the factors associated with mortality trends, especially in infants, pose a difficulty to predict the timing that mortality peaks and to assess whether these peaks are

age-dependent. Utilizing the richness of DSS databases, vigorous quantitative methodologies can be formulated to study and quantify association with risk factors and at the same time estimate seasonal peaks and temporal trends. Clarity in these variations is crucial to timely interventions (Lawn, Cousens, et al., 2005; Lawn, Shibuya, et al., 2005), prepare health system demand and for guiding proper allocation of resources (Fisman, 2007; Medina et al., 2007; Naumova et al., 2007; UNICEF, 2005).

Most epidemiological and longitudinal studies employ summary statistics, graphical presentation (Becker and Weng, 1998; Pascual et al., 2008) and statistical tests (Bailey et al., 1992; Yip and Yang, 2004; Singh et al., 2007; Vitali et al., 2009) while assessing seasonality. Selected applications extended the seasonality assessment with application of time series methodologies such as Seasonal Auto-Regression Integrated Moving Average (SARIMA) models (Tong et al., 2005; Hu et al., 2007; Zhang et al., 2007; Briët et al., 2008). However SARIMA models are mainly appropriate for Gaussian data (Zeger et al., 2006; Huang et al., 2011). Statistical techniques which incorporate harmonic functions with varying coefficients in traditional models have also been used to efficiently model seasonality, though claimed to introduce a large number of parameters and sometimes overfit the data (Stolwijk et al., 1999; Rau, 2006; Eilers et al., 2008). However, there are limited applications which involved assessing seasonality in tropical diseases or utilizing various DSS data specifically on mortality (Becher et al., 2008; Becker and Weng, 1998; Byass et al., 2002; Kynast-Wolf et al., 2006; Ramroth et al., 2009). The referred studies utilized Poisson regression models with trigonometric functions to capture seasonality proficiently (Kynast-Wolf et al., 2006; Becher et al., 2008). It is an observation that most of the modeling attempts ignore accounting for temporal correlation and overdispersion which are vital in longitudinal data analysis (Cameron and Trivedi, 1998).

This study aims to provide an overview of different approaches to assess seasonality in mortality data. Further extensions of existing measures are given to allow statistical inference and in contrary to previous application, negative binomial (NB) regression models with temporal random effects are used instead, to provide a rigorous but simplified approach for modeling seasonal patterns and detection of mortality peaks at different age groups in infancy.

Models are formulated in a Bayesian framework and accounted for excess zeros, temporal correlation and used various components to capture seasonal patterns. The methods are illustrated using the infant mortality data from the Rufiji DSS (RDSS) database and outputs are discussed to suggest best approaches that can be used to assess mortality peaks. The analysis was carried out for different age-groups during infancy and on pooled data (combining all age groups). This paper is organized as follows; Section 1 defines the data used, Section 2 describes methods considered in measuring seasonality with formulation of models. Results and discussion are presented in Sections 3 and 4 respectively.

2.2. Data

2.2.1 Infant Mortality

Mortality data were extracted from the Rufiji DSS database covering a period of October 2001 – September 2004. The RDSS, located in Rufiji District, Tanzania commenced in 1998. It extends from 7.470 to 8.030 south latitude and 38.620 to 39.170 east longitude. The DSS monitors 85,000 people, which is about 47% of the total population of the District (Source: INDEPTH Monogram). From the database we extracted dates of birth, entry and exit from a survey (given by day, month and year) and death status (Source: Rufiji DSS). The outcome of interest is a binary variable indicating a death status of an infant at exit of a specific calendar month during the study period. Infants were grouped in age intervals of thirty days (i.e. 0-30days, 31-60days..., 331-360days) and referred to age in month 0 to month 11. It is worth noting that, person-time (period) methods were used for data analysis, rather than cohort analysis based on number of live births, hence rates calculated on the basis of person-times are expected to be higher than those based on number of live-births. However, the approach facilitated analyses of multiple age groups.

Total death counts and *time at risk* were calculated by calendar month and by age group. An individual's "*time at risk*" is defined as number of days an infant was alive during a specific age group and/or a calendar month. Age-specific mortality rates (rate at i^{th} month of life, $i = 0,...,11$) were calculated by taking a ratio between total deaths counts and total *time at risk* (multiplied by 1000) and expressed as deaths per 1000 person-years. For this study, years are defined as Year 1 (October 2001-September 2002), Year 2 (October 2002-September 2003) and Year 3 (October 2003-September 2004).

2.2.2 Seasonality

Rufiji district is characterized by two main rainy seasons; short rains (October – December) and long/heavy rains (February – May) with the remaining months (January, June, July, August and September) remain relatively dry. In this study two seasons were considered, a dry season which comprised of the dry months and wet season which included both, the short and heavy rain months (season). This categorization was used in the calculation of the mortality indices and included in the regression models (described in the next section).

2.3. Methodology

In this section methods to assess seasonality patterns considered in this paper are described. These include mortality indices, statistical testing and modeling.

2.3.1 Seasonality Index

The ratio of wet season mortality rate (MR) to dry season MR is the most commonly index used to measure the strength of seasonality (Rau, 2007). Mathematically, a point estimate for this index (denoted as θ) is calculated as: $\theta = \sum_{i \in S_1} MR_i \Big/ \sum_{k \in S_2} MR_k$, where S_1 is a set of wet months and S_2 a set of dry months. The value $\theta = 1$ indicates no difference between the two seasons, $\theta > 1$ indicates higher mortality rate in wet season while $\theta < 1$ indicates higher rates during the dry season. However, a measure uncertainty for the parameter θ is not always considered hence limit making statistical inference for the index. To address this, a model based approach was used to estimate θ with its associated confidence intervals. Details of model formulation are described in the modeling section. In this study, point estimate values for θ denoted as θ_p were calculated for each age group on annual basis and for pooled data and while model based index denoted as θ_m were calculated only for pooled data. Results are presented and discussed.

2.3.2 Goodness-of-fit test

This is a form of a chi-square test which has been commonly applied to measure seasonality in time series data (Sogoba et al., 2007; Zhang et al., 2007; Mohorovic et al., 2010). Thinking of a time series as a process, the test determines whether the process systematically deviates from pre-defined expectation (Zeger et al., 2006). In this study the test was used to indicate whether it is reasonable to assume that mortality rates (of specific age group or pooled data) observed over a period of $t = 12$ months (a complete year) arised from a specific distribution, hypothesized as uniform distribution (Horn, 1977; Siegel and Castellan, 1988). With the null hypothesis being absence of seasonality, the test statistic, Z which has a chi-square distribution with $t-1$ (11) degrees of freedom, is calculated as $Z = \sum_{t=1}^{12} (Y_t - E_t)^2 / E_t$ where Y_t are the observed death rates at month t and E_t are the mean (average) of the rates over all 12 months. A small value (close to zero) of Z is interpreted as absence of seasonal pattern.

2.3.3 Modeling

This section explains formulation of NB regression models that were used to model monthly time series age-specific mortality rates to assess seasonality and estimating temporal patterns. Models were implemented in a Bayesian framework to allow flexible estimation of needed parameters (Diggle et al., 1998).

For each age group, let D_t and T_t represent the death counts and time at risk at a calendar month t respectively, and let the expected value of D_t denoted by $E(D_t) = \mu_t$. Then D_t is a random variable following a negative binomial distribution with parameters μ_t and r, i.e. $D_t \sim NB(\mu_t, r)$ where r is a dispersion parameter of the distribution. The general model with a log link is described as:

$$log(\mu_t) = log(T_t) + \mathbf{X}^T\boldsymbol{\beta} + f_T(t) + f_S(t)$$

$$log(D_t) = log(T_t) + \mathbf{X}^T\boldsymbol{\beta} + f_T(t) + f_S(t) + \varepsilon_t \qquad (1)$$

The $f_T(t)$ term models the time trend, $f_S(t)$ identifies the seasonal pattern, $\boldsymbol{\beta}$ is a vector of regression coefficients associated with the matrix of covariates $\mathbf{X}$, ε_t are time-specific error terms modeling temporal correlation and $log(T_t)$ is an offset term corresponds to the denominator of the mortality rate.

The time trend $f_T(t)$ considered annual and monthly time trends. Let $Y_j, j = 1,2,3$ and $M_t, t = 1,2,...12$ indicate time in years and months respectively. We then chose a discrete-time formulation for the annual effect, i.e. $f_T(t) = \sum_{k=1}^{2} \boldsymbol{\delta}_k I(Y_j = k+1)$ and a continuous trend for the monthly effect, i.e. $f_T(t) = \delta * t$ where δ and δ_k are estimated as normal regression coefficients. The seasonal pattern $f_S(t)$ was captured by (i) a binary variable defining wet and dry seasons, i.e. $f_S(t) = \psi I(S_t^{(W)} = 1)$ where $S_t^{(W)}$ is an indicator for wet season, (ii) a categorical variable defining calendar month, i.e., $f_S(t) = \sum_{k=1}^{11} \boldsymbol{\psi}_k I(M_t = k)$, where M_t indicates a calendar month k that an infant died and ψ, ψ_k are regression coefficients, and (iii) using harmonic cycles as described in the next section. The index θ (explained above) is equivalent to model (1) with $f_S(t)$ defined as a binary seasonal variable (wet and dry) and without the terms $\mathbf{X}^T\boldsymbol{\beta}$, $f_T(t)$, and ε_t in the model (1).

2.3.4 Description of harmonic cycle model

In this case the seasonal term $f_S(t)$ of the model is formulated via a cosine function expressed as:

$$f_S(t) = ACos\left(\frac{2\pi}{T}t - \varphi\right) = a * Cos\left(\frac{2\pi}{T}t\right) + b * Sin\left(\frac{2\pi}{T}t\right)$$

where t is the time (in months), T is the period of a cycle C, A defines the amplitude which estimates the peak mortality within the time period, expressed as $A = \sqrt{a^2 + b^2}$ and φ gives the phase, the point in time where the peak occurs (in radian scale), calculated as $\varphi = arctan(-a/b)$. Parameters a and b are estimated from a regression model.

To capture the rainy seasons in Rufiji and assess annual trend of the seasonal patterns, (see Figure 2.3), $f_S(t)$ was modeled considering two scenarios. The first case used a mixture of two cycles C, having periods of 6-months and 12-months and cycle-dependent parameters (i.e. a_c and $b_c, C = 1,2$). Two cycles are appropriate when two seasons are expected within a year. Hence, with this specification seasonal-specific amplitudes, A_c and phases, φ_c were estimated with an assumption that they do not vary from one year to another. In the second scenario, the model was similar to that expressed in the first case but with year-dependence, i.e. amplitudes and phases that vary annually, $f_S(t) = f_{Sj}(t)$. Parameters were therefore estimated for each cycle and year, i.e. a_{cj} and b_{cj} to estimate A_{Cj} and φ_{C_j} $j = 1,2,3, C = 1,2$.

The $f_{Sj}(t)$ and respective amplitudes and phases are expressed as follows,

$$f_{Sj}(t) = \sum_{c=1}^{C} \left\{ a_{Cj} * Cos\left(\frac{2\pi c}{T} t\right) + b_{Cj} * Sin\left(\frac{2\pi c}{T} t\right) \right\} \quad j = 1,2,3;\ C = 1,2$$

and, $$A_{Cj} = \sqrt{a_{Cj}^2 + b_{Cj}^2}\ ; \qquad \varphi_{C_j} = \arctan(-a_{Cj} / b_{Cj})$$

2.3.5 Bayesian model specification

Following Bayesian model formulation, prior distributions were specified for all model parameters. For the regression coefficients a non-informative normal prior distribution was assumed, a Gamma distribution with mean 1 and variance 100 was adopted for the parameter, r. Two specifications were adopted for the temporal error terms ε_t. In the first attempt, error terms $\varepsilon_t = \varepsilon_t^{(1)}$, $t = 1,...,12$ were considered to be independent variables arise from a normal distribution with mean 0 and a common variance σ_e^2. Alternatively, the error terms $\varepsilon_t = \varepsilon_t^{(2)}$, $t = 1,...,12$ were considered to be temporally correlated and modeled via an autoregressive process of first order i.e., $\varepsilon_t^{(2)} \sim AR(1)$, that is $\varepsilon_t^{(2)} = \gamma \varepsilon_{t-1}^{(2)} + \alpha_t$, $t \geq 2$, where α_t is a 'white noise' time series comprising of independent observations which are normally distributed with mean zero and variance $\sigma_w^2/(1-\gamma^2)$. The autocorrelation parameter, γ, which quantifies degree of dependence between time series points (Zeger et al., 2006) was assumed

to follow a Uniform distribution, $\gamma \sim U[-1,1]$. An inverse gamma prior distribution with mean 10 and variance 100 was used for the independent, σ_e^2, and autoregressive, σ_w^2 temporal variance parameters. Models were run in OpenBUGS and parameters were estimated using Markov Chain Monte Carlo (MCMC) simulations algorithm with 300,000 iterations and a burn-in period of the first 30,000 iterations with a thinning of 20 iterations.

2.3.6 Model assessment and selection

Convergence of the models was assessed by visual inspection of trace and density plots and the Gelman and Rubin convergence statistics (Gelman and Rubin, 1992). Model selection was based on the least square methods where error sum of squares (SS_E) were calculated and fitted values plotted against the original rates. The model giving the lowest SS_E was considered the best. The measure of variability for the NB distribution, r, was also assessed (Lloyd-Smith, 2007).

2.4. Results

2.4.1 Descriptive analysis

A total of 448 infants deaths were observed during the study period with an overall mortality rate of 49.4 deaths per 1000 person-years (p-y). The mortality rates were 58.3, 44.0 and 45.7 deaths per 1000 p-y for the 1st, 2nd and 3rd year of the study, respectively. Figure 2.1: Death counts and age-specific mortality rates (MR) of infants in Rufiji DSS, October 2001- September 2004, Figure 2.2 and Figure 2.3 present the pattern of the MR a) age (in months), b) by calendar months and c) over the entire study period (October 2001- September 2004). Infants of zero month of age presented the highest mortality rate of 185.4 deaths per 1000 p-y whilst the 10 months old had the lowest rate of 22.65 deaths per 1000 p-y. A total of 185 deaths (MR=49.0) were observed during the dry season while 263 deaths (MR=49.6) occurred during the wet season. About 32.6% of all deaths occurred in infants between 0-30 days of age. The remaining infants presented very similar rates with slight low rates observed after 6 months of age (Figure 2.1).

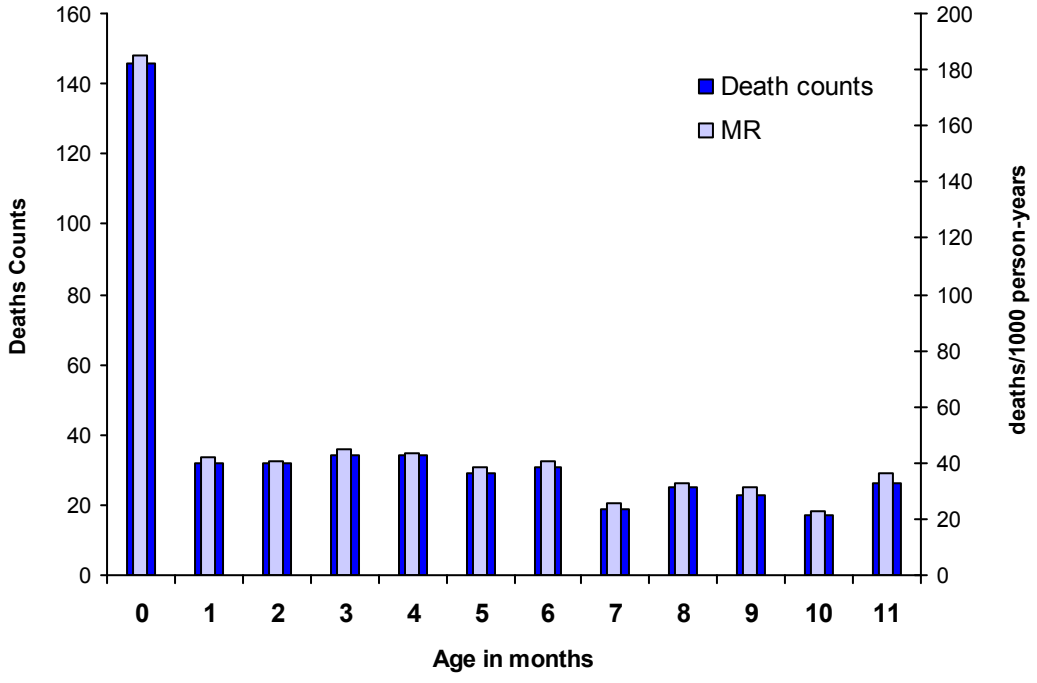


Figure 2.1: Death counts and age-specific mortality rates (MR) of infants in Rufiji DSS, October 2001- September 2004

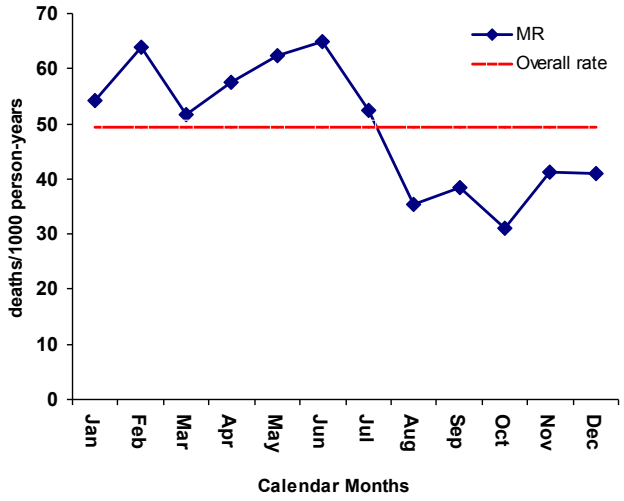


Figure 2.2: Age-specific mortality rates by calendar months

Comparing with the overall rate (red reference line), the mortality rates were higher during the first half of a calendar year showing a decreasing trend between months of July – December. Mortality peak on February and June (Figure 2.2). The mortality trend for the entire study period is shown in Figure 2.3.

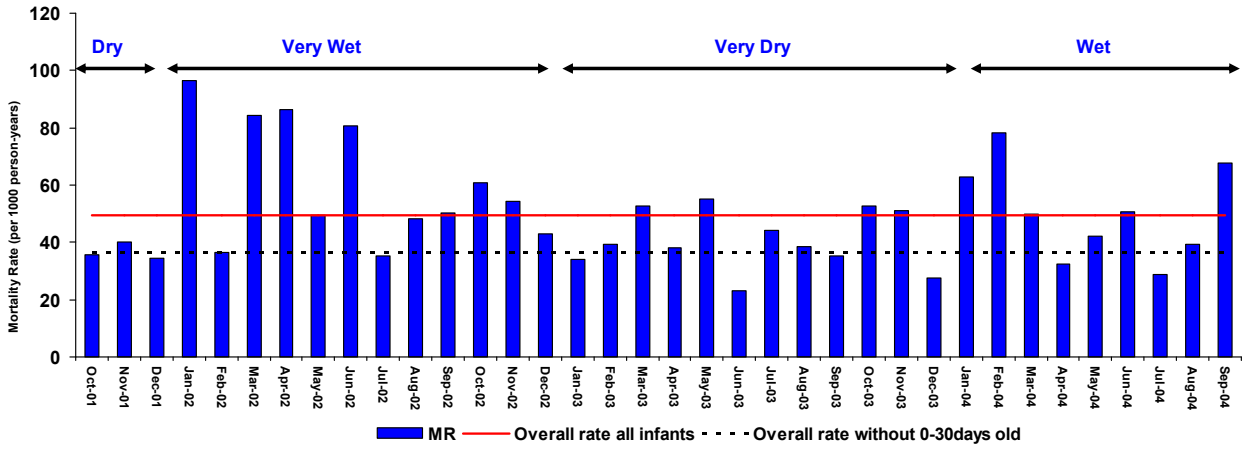


Figure 2.3: Mortality rates (MR) for the Infants during the period of October 01 – September 04, Rufiji DSS

The two reference lines indicated the overall MR when infants aged 0–30 days were included (solid) and excluded (dotted) from the data. The two rates differed by 12.8 deaths/1000p-y indicating high mortality in the neonates group. Figure 2.3 also elicit differences in dry and wet season mortality and between years.

2.4.2 Seasonality index

Point estimates, θ_p and model based θ_m seasonality indices estimated for each year and for the pooled are presented in Table 2.1. For the point estimates, most indices indicated higher mortality in the rainy season as compared to the dry season ($\theta_p > 1$). However, estimates were close to one indicating a small difference in mortality between the two seasons. Considerable variation was observed among the age groups and across years. During the study period, the highest value was estimated for the 4 months old infants during Year 2 (θ_p=2.857) indicating that, for this group, MR was 3 times higher in the wet season than in the dry season. The indices were consistently small in infants aged 11 months old suggesting this group to be affected more during dry season.

Table 2.1: Age-specific point estimate and model based (with 95% Bayesian Credible Interval (BCI)) seasonality indices by year and on pooled data

Age in days (Months)	Point Estimates, θ_p				†Model-based Estimates, θ_m (Pooled data)	
	Year 1	Year 2	Year 3	*All*	*Median*	95% BCI
0-30 (0)	0.907	0.913	1.015	*0.941*	*0.888*	(0.640, 1.238)
31-60 (1)	1.607	1.786	--- **	*2.143*	*2.113*	(0.985, 5.079)
61-90 (2)	2.619	1.429	0.714	*1.571*	*1.659*	(0.804, 3.681)
91-120 (3)	0.833	0.612	2.143	*0.905*	*0.931*	(0.471, 1.870)
121-150 (4)	1.286	2.857	0.714	*1.310*	*1.341*	(0.669, 2.814)
151-180 (5)	0.857	0.286	0.595	*0.580*	*0.568*	(0.267, 1.180)
181-210 (6)	0.833	0.476	0.714	*0.670*	*0.671*	(0.328, 1.367)
211-240 (7)	0.952	1.429	1.429	*1.224*	*1.215*	(0.483, 3.284)
241-270 (8)	1.000	0.179	--- **	*1.270*	*1.268*	(0.568,3.016)
271-300 (9)	0.952	2.143	1.190	*1.339*	*1.346*	(0.582, 3.381)
301-330 (10)	1.786	0.000	5.714	*2.321*	*2.398*	(0.837, 8.847)
331-360 (11)	0.238	0.286	0.595	*0.378*	*0.386*	(0.163, 0.853)

**** No deaths occurred during dry season †Model-based estimates correspond to posterior median**

Model based estimates were similar to the point estimates, however, all 95% BCIs (except for 11 months old) included one indicating non-significance difference of mortality rates between the two seasons. In addition, no significant difference of indices between age groups, years and their interaction could be found.

2.4.3 Goodness-of-fit test

The test statistic, Z, was calculated to test the null hypothesis which assumed presence of seasonality in the monthly mortality rates. The value of Z obtained was 254. Comparing $Z = 254$ with the critical value at 5% probability level which is $Z = 3.84$, the null hypothesis was rejected and it was concluded that there is a significant difference between the distribution of the observed and expected MR, which interpreted as existence of seasonality.

2.4.4 Modeling

Negative Binomial models were fitted for each age group separately. Following different formulations of seasonal term, the models (M) were grouped into 4 Sets. Set 1: models 1-4 (dry/wet), Set 2: models 5-8 (mixture of 2 cycles, year-independent), Set 3: models 9-12 (mixture of 2 cycles, year-dependent) and Set 4: models 13-16 (binary variable for months). Each set includes models that considered time trend as discrete or continuous and error terms independent or AR(1). For illustration, Table 2.2 shows results of performance of all models for the infants aged 10 months old.

Table 2.2: Model assessment: the error sum of squares, number of parameters, and variance and autocorrelation parameters from models fitted with data for infants of 10-months of age

	Model	¥Model description $log(\mu_t)$	SS_E	No. of parameters	r	γ
Set 1	M1	$\beta_0 + \psi\, I(S_t^{(W)} = 1) + \delta * t + \varepsilon_t^{(1)}$	25.953	42	8.23	----
	M2	$\beta_0 + \psi\, I(S_t^{(W)} = 1) + \sum_{k=1}^{2} \delta_k I(Y_t = k+1) + \varepsilon_t^{(1)}$	20.770	43	11.58	----
	M3	$\beta_0 + \psi\, I(S_t^{(W)} = 1) + \delta * t + \varepsilon_t^{(2)}$	27.744	44	6.48	0.23
	M4	$\beta_0 + \psi\, I(S_t^{(W)} = 1) + \sum_{k=1}^{2} \delta_k I(Y_t = k+1) + \varepsilon_t^{(2)}$	21.131	45	10.30	0.22
Set 2	M5	$\beta_0 + f_{S_C}(t) + \delta * t + \varepsilon_t^{(1)}$	26.595	47	5.58	----
	M6	$\beta_0 + f_{S_C}(t) + \sum_{k=1}^{2} \delta_k I(Y_t = k+1) + \varepsilon_t^{(1)}$	22.439	48	7.95	----
	M7	$\beta_0 + f_{S_C}(t) + \delta * t + \varepsilon_t^{(2)}$	27.793	49	5.09	0.45
	M8	$\beta_0 + f_{S_C}(t) + \sum_{k=1}^{2} \delta_k I(Y_t = k+1) + \varepsilon_t^{(2)}$	21.976	50	7.98	0.14
Set 3	**M9**	$\beta_0 + f_{Sj}(t) + \delta * t + \varepsilon_t^{(1)}$	**15.300**	**53**	**2.96**	**----**
	M10	$\beta_0 + f_{Sj}(t) + \sum_{k=1}^{2} \delta_k I(Y_t = k+1) + \varepsilon_t^{(1)}$	**13.414**	**54**	**3.72**	**----**
	M11	$\beta_0 + f_{Sj}(t) + \delta * t + \varepsilon_t^{(2)}$	**14.338**	**55**	**2.94**	**0.33**
	M12	$\beta_0 + f_{Sj}(t) + \sum_{k=1}^{2} \delta_k I(Y_t = k+1) + \varepsilon_t^{(2)}$	**13.007**	**56**	**3.59**	**0.41**
Set 4	M13	$\beta_0 + \sum_{k=1}^{11} \psi_k I(M_t = k) + \delta * t + \varepsilon_t^{(1)}$	16.450	52	3.50	----
	M14	$\beta_0 + \sum_{k=1}^{11} \psi_k I(M_t = k) + \sum_{k=1}^{2} \delta_k I(Y_t = k+1) + \varepsilon_t^{(1)}$	12.358	53	4.72	----
	M15	$\beta_0 + \sum_{k=1}^{11} \psi_k I(M_t = k) + \delta * t + \varepsilon_t^{(2)}$	24.773	54	2.60	0.044
	M 16	$\beta_0 + \sum_{k=1}^{11} \psi_k I(M_t = k) + \sum_{k=1}^{2} \delta_k I(Y_t = k+1) + \varepsilon_t^{(2)}$	18.506	55	3.93	0.065

¥ $t = 1,2,...12$; $j = 1,2,3$; **An offset** $log(T_t)$ **was included in all models.**

Generally, models with discrete time and autoregressive errors performed better than models which considered time as continuous or/ and independent errors terms. Models with harmonic cycles and year-dependent seasonal pattern (Set 3: M9-M12) presented lower SS_E as compared to other models followed by models in Set 4: M13-M16. Models in Set 1 and 2 had relatively high SSEs. Model sets 3 and 4 estimated comparative smaller value for the variance parameter r than Sets 1 and 2. Value for parameter r can be compared with the ratio of annual MR to its variance over a period of 12 months. For this age group this ratio was 22.65/35.0 = 1.54. Models in Set 4 estimated very low autocorrelation parameters, which could be explained as an effect of including each month as an independent variable. Due to that and a similar performance of Set 3 and Set 4, we excluded Set 4 for all future discussion. These

results were consistent in all age groups. Figure 2.4 depicts SSE for models from Set 1, 2 and 3 for all age groups with SSEs obtained from models in Set 3 circled.

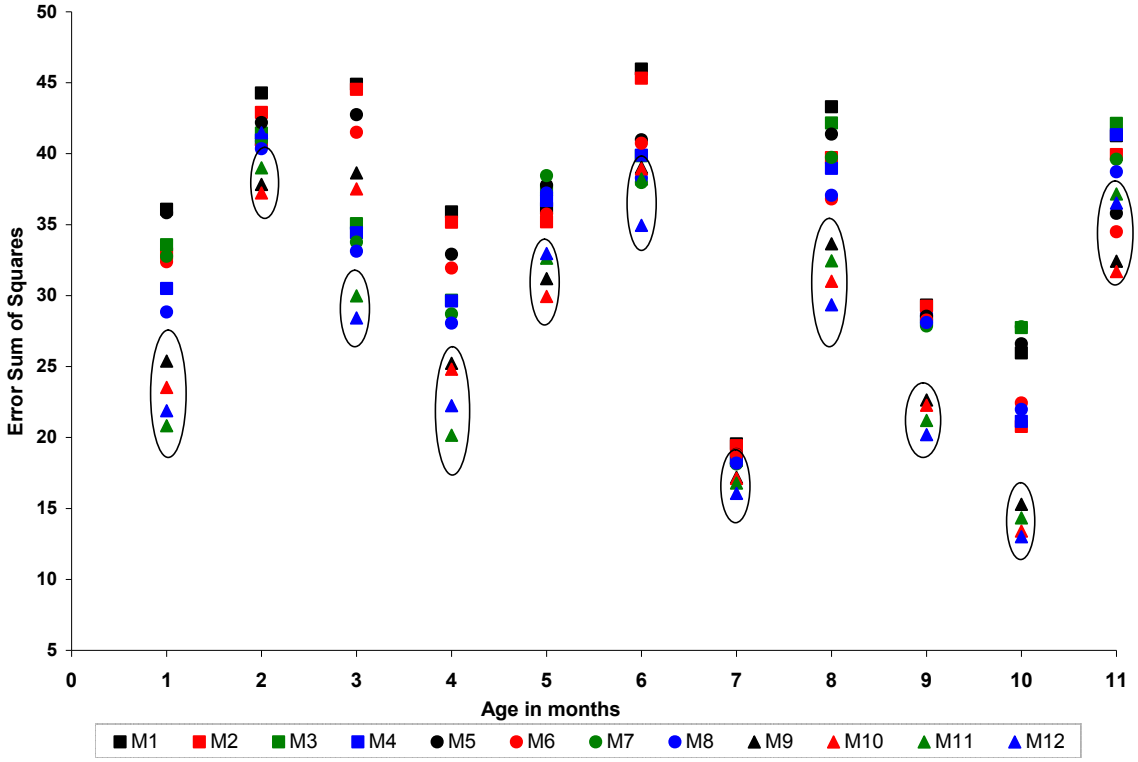


Figure 2.4: Error Sum of Squares for Models 1-12 for all age groups*

* Age group 1 excluded due to high deviation of estimates from other groups

M12 featured as the best model in 6 age groups (2,6,7,8,9,10) and second best in 2 age groups (1,4) hence selected as the best model. Parameters estimated from this model for selected age groups and pooled data are presented in Table 2.3. Effect of age was assessed when pooled data was used.

Table 2.3: Results of M12 for infants age 2, 5, 8 and 11 months and for pooled data

	Parameter	2 months Median (95% BCI)	5 months Median (95% BCI)	8 months Median (95% BCI)	11 months Median (95% BCI)	Pooled Data Median (95% BCI)
	β_1 (Age)	---	---	---	---	0.864 (0.836, 0.893)
				Seasonality terms		
Year 1	A_{11}	0.6 (0.03,2.099)	0.51 (0.025,1.73)	0.76 (0.038,2.825)	1.54 (0.123,3.778)	0.40 (0.03,0.95)
	A_{21}	0.39 (0.018,1.496)	0.40 (0.018,1.44)	0.69 (0.035,2.602)	0.78 (0.037,2.714)	0.13 (0.01,0.43)
	φ_{11}	8.40 (1.004,23.47)	7.97 (0.947,23.28)	6.41 (0.573,22.04)	7.93 (1.374,23.27)	5.56 (0.82,19.52)
	φ_{21}	7.0 (0.604,22.38)	7.29 (0.85,23.06)	6.20 (0.548,22.17)	7.46 (0.937,22.85)	6.35 (0.57,22.15)
Year 2	A_{12}	0.43 (0.02,1.628)	0.65 (0.028,2.61)	1.89 (0.118,5.383)	0.76 (0.035,2.504)	0.19 (0.01,0.7)
	A_{22}	0.43 (0.02,1.606)	1.02 (0.06,2.95)	0.86 (0.039,3.775)	0.69 (0.03,2.388)	0.14 (0.01,0.48)
	φ_{12}	7.29 (0.776,22.73)	5.93 (0.551,21.6)	8.27 (1.403,23.22)	8.02 (0.741,23.15)	7.29 (0.68,22.79)
	φ_{22}	7.09 (0.616,22.73)	5.38 (0.529,21.17)	7.39 (0.79,22.65)	7.43 (0.96,22.6)	6.75 (0.64,21.65)
Year 3	A_{13}	0.61 (0.028,2.298)	0.42 (0.019,1.56)	0.9 (0.041,3.397)	1.002 (0.064,2.672)	0.24 (0.01,0.77)
	A_{23}	0.64 (0.03,2.332)	0.80 (0.048,2.109)	0.78 (0.033,2.927)	0.43 (0.02,1.61)	0.16 (0.01,0.52)
	φ_{13}	7.29 (0.85,22.46)	7.15 (0.778,22.4)	8.21 (1.057,23.54)	6.26 (0.714,21.68)	5.66 (0.5,21.57)
	φ_{23}	7.11 (0.823,22.381)	8.37 (1.313,23.4)	6.23 (0.562,22.07)	7.47 (0.769,22.55)	7.99 (0.99,23.44)
				Annual time trend		
Year2	δ_1	0.85 (0.27,2.61)	0.48 (0.10, 1.65)	0.21 (0.017,1.36)	1.25 (0.24,8.33)	0.77 (0.58, 1.014)
Year3	δ_2	0.38 (0.093,1.35)	0.93 (0.29,2.80)	0.56 (0.10,2.66)	1.99 (0.42,12.06)	0.83 (0.63, 1.09)
				Other parameters		
	r	2.30 (0.46,18.57)	3.62 (0.67,22.60)	1.31 (0.21,14.95)	2.24 (0.40,18.69)	3.407 (2.106, 6.534)
	γ	0.19 (-0.80,0.94)	0.15 (-0.79,0.93)	0.28 (-0.72,0.97)	0.22 (-0.80,0.96)	0.18 (-0.994, 1.0)
	σ_w^2	0.49 (0.20,1.30)	0.41 (0.19,1.04)	0.79 (0.24,2.67)	0.44 (0.20,1.21)	0.042 (0.007, 0.253)

Among the groups no common pattern was observed for the temporal trend, however, the overall trend, estimated from pooled data, showed a reduction in mortality during the 2nd and 3rd year of the study as compared to year 1 (p-value >0.05). There were similarities in values of phases and amplitudes between groups, however, age-specific results showed wide CI for the variance and amplitude parameters. Mortality was seen to decrease as age increases (p-value<0.001).

Estimates from the pooled data fell within CIs of parameters estimated using age-specific data (some results not shown). Therefore, results from pooled data were used to calculate the amplitude (peak) and phase (point in time) for the MR of all infants. Average amplitudes for the

Year 1, 2 and 3 were 0.265, 0.165 and 0.20 respectively, indicating high mortality during the 1st year, a decrease in the 2nd year and an increase in the last year.

Using the formula suggested by Stolwijk et al. (2008), extreme time points (where minimum and maximum mortality occurred) were determined. The two extremes $\phi_{\min}$ and $\phi_{\max}$ were obtained by solving $\phi_{\min} = \varphi_{cj} * T / 2\pi$ (π=3.1415) and $\phi_{\max} = \phi_{\min} + T/2$ respectively. Based on the cycle of our data, which started on October, $\varphi_{\min} = 1.3$ and $\varphi_{\max} = 8.75$ are converted to a scale of month, thus values translate to a month of November and end of May. In Figure 2.5 monthly time series of observed mortality rates (fitted values and 95%BCI) are shown. The most visible peak is the one at the end of May.

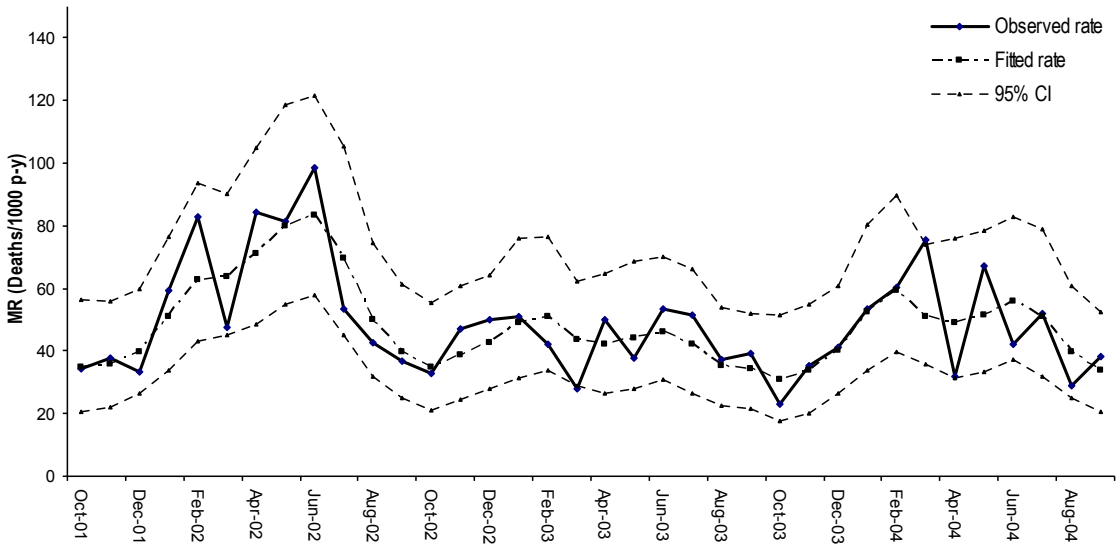


Figure 2.5: Monthly time series of mortality rates aligned with the fitted values, RDSS

2.5. Discussion

This study assessed different approaches for optimal capturing of seasonality and temporal trend in data obtained from longitudinal survey. Infant mortality data from Rufiji DSS were used as a case study to discuss performance of the methodologies. Extensions of existing methods were considered and different model formulations were presented. With the strength of the data obtained from the DSS in terms of spatial resolution and richness, we have developed and discussed easy-to-follow approaches to capture and quatify seasonality while considering pros and cons of each technique. These methodologies can be applied directly to data obtained from other longitudinal surveys focusing in epidemiology and public health.

Mortality indices and statistical tests are implemented easily and finely detect trends present in the data (Rau, 2006). However, such methods involved conservative prerequisite steps such as grouping and collapsing of data, which results into a significant loss of information, mask underlying features of the data and mostly rely on strong assumptions which could influence the results of the analysis (Rau, 2006; Eilers et al., 2008). In this paper, a reformulation of a seasonality index which was implemented in model framework and allows researchers to make statistical inferences on the estimated value was proposed. The advantage of model based approaches includes a possibility to incorporate other terms but also to estimate effect of predictors. Most rigorous techniques are hard to employ, but capture best information from the data enabling proper studying of trends and patterns. Our models took into consideration seasonality, temporal correlation using harmonic cycles and autoregressive temporal effects respectively and could be extended to account for confounders, effect modifiers and prediction.

We found an association between the peak of infant mortality and the climatic seasons. Higher mortality rate were observed in the wet seasons (months, year) as compared to dry ones. Changes in environment and climate influence seasonal trends of the disease and other health parameters (Lipp et al., 2002; Patz, 2002; Patz et al., 2008). However, to accurate determine presence of these trends, a longitudinal monitoring of events is required to assess repetitive behaviors. Interestingly, mortality was high in the first half of the year and dropped in the last

six months. Substantial behavioral, socio-economic status, food security or health system factors such as access to care could be associated with these patterns. Our model reported similar seasonal patterns and temporal trends of mortality among age groups during infancy. These results suggest that factors (e.g. environment, socio economic status) and causes (e.g. malaria, diarrhea) associated with mortality of infants can be hardly distinguished between such age categories in these data.

High proportion of deaths was observed in neonates, however, seasonal patterns could not be distinguished from one age group to another. Neonatal mortality accounts for 8.2% of the total mortality burden in RDSS (Source: District Health Profile Report, 2007); (Carneiro et al., 2010). Factors associated with neonate deaths such as place of birth, low birth weight, quality of obstetrical care and coverage of immunization program, are predominantly non-seasonal hence the overall seasonal trend could be hidden behind climatic and malaria-related factors (Rumisha et al., submitted for publication). A slight decrease in mortality observed after the age of six months could be explained by progressively immunity acquisition to specific diseases with age, especially malaria (Anderson and May, 1991).

By incorporating a harmonic function in a regression model we were able to determine the peaks of mortality which strongly mark a relation with environment seasons. The framework developed here can be extended to adjust for risk factors and accounting for other types of correlations e.g. spatial effects. This work is among few that present a step by step methodological framework on how best seasonality can be assessed (formulations are available upon request from the authors). Application of NB models with terms that capture seasonal patterns, temporal trend and correlation are rarely discussed in literature hence emphasis the novelty of this work. The implication of this work in malaria interventions and control is the timing and targeting. These are key elements for design of effective malaria control strategies as scarce resource and operational cost are of concern.

Funding: This work was supported by Swiss National Foundation (grant number Nr 325200-118379/1).
Acknowledgement: The authors thank the Ifakara Health Institute and the INDEPTH Network for facilitating the collection of the data and managing of the Rufiji DSS.
Conflicts of interest: None declared.

Chapter 3 Bayesian modeling of large geostatistical data to estimate seasonal and spatial variation of sporozoite rate

Rumisha SF[1,2,4], Kasasa S[1,2,5], Smith T[1,2], Abdulla S[3], Masanja H[3], and Vounatsou P[1,2‡]

[1]*Swiss Tropical and Public Health Institute, Basel, Switzerland,*

[2]*University of Basel, Basel, Switzerland,*

[3]*Ifakara Health Institute, Dar es Salaam, Tanzania,*

[4]*National Institute for Medical Research, Dar es Salaam, Tanzania,*

[5]*Makerere University School of Public Health/ Iganga-Mayuge DSS, Uganda*

‡Corresponding author:

This manuscript has been submitted to the *Journal of Spatial and Spatio-temporal Epidemiology*

Abstract

Understanding seasonal and spatial variations in malaria transmission is important for a targeted control strategy. The Malaria Transmission Intensity and Mortality Burdens across Africa project compiled the most comprehensive, large geo-referenced entomological database to study malaria seasonality. However, rigorous analyses have not been done due to limitations in statistical methodology. Recently, approaches have been developed for fitting large Gaussian geostatistical data observed on regular grids by approximating the spatial process from a subset of locations. Extending this methodology for non-Gaussian geostatistical longitudinal data, observed irregularly in space, we proposed ways of selecting the subset of locations for approximating the spatial process taking into account outcome variability, the underlying spatial process and location configuration. This methodology is implemented to study seasonality in the sporozoite rate as a proxy measure of malaria transmission. We fitted Bayesian geostatistical logistic models adjusting for environment and climate factors. Seasonal patterns were captured by a mixture of harmonic cosine functions with different cycles. Sporozoite rates were predicted using Bayesian kriging and monthly mosquito species-specific transmission maps were produced. Results indicated seasonal variation of the sporozoite rates that differ between species, location and time and suggest malaria control programs that target exact periods where transmission peaks.

Keywords: Africa, approximate spatial process, sporozoite rates, seasonality, malaria, Markov Chain Monte Carlo

3.1. Introduction

Malaria transmission is influenced by climatic and environmental conditions suitable for survival of the mosquito vector of the disease. Rainfall and temperature affect the breeding environment, vector species behavior and gonotrophic cycle (number of times mosquitoes lay eggs in a lifetime) hence the potential of malaria transmission. Seasonality in climatic factors introduces seasonal fluctuation in transmission. Climate also affects large areas leading to spatial variation of the disease burden. Malaria transmission has shown seasonal variations in many regions in Sub-Saharan Africa (Gemperli et al. 2006a; Romagosa et al. 2007; Abellana et al. 2008; Noor et al. 2008). Estimating seasonal and spatial variation helps in understanding disease dynamics, better designing and evaluating malaria interventions, and preparedness of the health system.

A number of entomological studies have looked at small area variations of the malaria transmission (Mbogo et al. 1995, 2003; Fontenille et al. 1997a; Robert et al. 2003; Shililu et al. 2003; Bøgh et al. 2007; Mboera et al. 2007, 2010; Sogoba et al. 2007; Atangana et al. 2009). However, very few were based on entomological data collected over a long period of time with high temporal resolution in order to capture seasonality taking into account year to year variation. Even more, to our knowledge there are no studies looking at spatial variation of seasonality using entomological data with high spatio-temporal scales. The Malaria Transmission Intensity and Mortality Burdens across Africa (MTIMBA) project has compiled the most comprehensive geo-referenced entomological database to study malaria seasonality. However rigorous analyses have not been carried out to date due to data characteristics and limitations in statistical methodology.

MTIMBA are large geostatistical data consisting of mosquito collections every two weeks during 2001-2004 over a number of household locations (up to 2500) within selected sites in Africa. Relevant analysis for such data are through geostatistical models which take into account spatial correlation via location-specific random effects which are viewed as latent data of an unobserved Gaussian spatial process (Diggle et al. 1998). Correlations between any pair of locations are considered to be a function of their distance modeled by the covariance

matrix of the process. Typical covariance functions are the exponential, matern, Gaussian, Cauchy, spherical and Bessel (Ecker and Gelfand 1999). Bayesian model formulations facilitate model fit via Markov chain Monte Carlo simulation. However, posterior calculations require repeated inversions of the covariance matrix of the process which for very large number of locations is not feasible.

A number of approaches have been suggested to overcome very large matrix computations involved in geostatistical model fit. Rue and Tjelmeland (2002) approximated a Gaussian spatial process by a Gaussian Markov field. Results showed that not all covariance models fitted well. Smith and Gelfand (1992) proposed replacing matrix inversion with simulation using sampling importance re-sampling as a first step in addressing the issue, however finding importance sampling densities is difficult and it requires tedious tuning (Gemperli and Vounatsou 2003). Stein et al. (2004) reduced the size of the covariance matrix by considering correlations only within the subset of "nearest" observations. The performance of this approximation depends on size of the subset and the choice of ordering the observations. Furrer et al. (2006) proposed covariance tapering to produce a sparse covariance matrix by forcing small covariances to be zero. However, the choice of the distance where correlation can be neglected is subjective and the misclassification is more severe when the study region is small as might lead to independency between observations. In geostatistical modeling, the decay parameter directly estimate this distance from the model. Gemperli and Vounatsou (2006) compared the performance of a number of numerical algorithms especially suited for sparse matrix inversion within the MCMC framework. Whiley and Wilson (2004) assessed various parallel MCMC algorithms for handling large geostatistical data. Similarly, Yang et al. (2007) parallelised computations related to large matrices involved in fitting spatial Gaussian models. Higdon et al. (2003)proposed a Metropolis coupled MCMC which runs multiple chains, a fine one estimating the spatial process at the whole set of locations and others which estimate the spatial process on a subset of locations chosen from coarse grids of various resolutions. The coarser faster running but less accurate chains “update” the slow fine chain is helpful for moderate rather than very large number of locations. Kernel convolutions, moving averages or basis functions (Wikle and Cressie 1999; Higdon 2001; Banerjee et al. 2003, 2008; Kammann and Wand 2003; Xia et al. 2006; Paciorek 2007; Cressie and Johannesson 2008) have all been applied to

approximate the spatial process. Approximations based on spectral domains received a lot of criticism and challenges due to inadequacy of the obtained approximation (Jones and Zhang 1997; Stein et al. 2004).

Recently, more approaches have been developed for fitting large Gaussian geostatistical data observed on regular grids by approximating the spatial process from a subset of locations (Xia and Gelfand 2005; Banerjee et al. 2008; Finley et al. 2009). However in many real time applications the outcomes of interest are neither Gaussian nor the locations of observed data are regularly distributed in space. Selection of the subset for approximating the spatial process is still a challenge. For Gaussian data, Finley et al. (2009) proposed selecting the subset of locations (knots) by minimizing the spatially averaged predicted variance. This variance measures how well the predictive process approximates the original process. However, the method requires large computation power even to select a small set of knots. For practical applications other methods like balance sampling can be employed. To date there are very few studies that have assessed the influence of using the subset of locations on estimation of spatial parameters in real life data.

This work extends the methodology of Banerjee et al. (2008) and proposes ways of selecting the subset of locations for approximating the spatial process preserving the variability of the outcome, the underlying spatial process and location configuration. It also assesses how well the knots represent the spatial characteristics of the observed data. This methodology is implemented to study malaria seasonality in the sporozoite rate data from the MTIMBA project. We fitted Bayesian geostatistical logistic models with environment and climate factors as covariates. Seasonal patterns were captured by a mixture of harmonic cosine functions with different cycles (Rumisha el al., 2013). The model assumes separable space-time variation with Gaussian spatial process and temporal random effects modeled with first order autoregressive correlation structure (Hay and Pettitt 2001). Sporozoite rates were predicted using Bayesian kriging and monthly transmission maps for two mosquito species were produced. The data used for analysis is presented in Section 2. The formulation of the Bayesian geo-statistical model is given in Section 3. The results of the application are in Section 4 and finally conclusion and discussion are presented in Section 5.

3.2. Data and Study site

The study was carried out in the Rufiji Demographic Surveillance Site (RDSS). The RDSS is located in the Rufiji District. The area extends from 7.470 to 8.030 south latitude and 38.620 to 39.170 east longitude and covers an area of 1813 square kms. The entomological data were collected as part of the MTIMBA project. Details of the protocol for mosquito collection are in the MTIMBA documentation (http://www.indepth-network.org). Shortly, anopheles mosquitoes caught in the light traps (or bed nets) over two consecutive nights every two weeks during the period of October 2001- September 2004 (http://www.indepth-network.org/dss_site_profiles/rufiji.pdf). Heads and thoraces were tested for *Plasmodium falciparum* circumsporozoite protein (CSP) by enzyme linked immuno-sorbent assay (ELISA). A total of 11,227 (482 positive) *An. gambiae* and 17,263 (405 positive) *An. funestus* mosquitoes were analyzed. Locations were geo-referenced with global positioning system (GPS) and date of collection was recorded. All repeated surveys of a specific location within the same month were collapsed to a single observation. A total of 670 data points were obtained for *An. gambiae* with 639 unique locations and 430 data points for *An. funestus* with 415 unique locations. RS climate and environment data were used as predictors. The data source, spatial and temporal resolutions are given in Table 3.1.

Table 3.1: Environment and Climate data

Variable	Spatial resolution	Temporal resolution	Source
Normalized Difference Vegetation Index (NDVI)	$250m^2$	16 days	MODIS
Temperature (Day and Night)	$1km^2$	8 days	MODIS
Rainfall	$8km^2$	10 days	ADDS
Water bodies	$1km^2$	---	Health Mapper

Lag analysis was used to link the RS to the mosquito data (Riedel et al. 2010). Lag refers to the time interval prior to data collection that was used to summarize the environmental/climatic covariates. The best lag was chosen based on the goodness of fit of the corresponding model. Lags considered include the current month (month of collection of mosquitoes), one month prior the collection, average of two previous months, average of

current month and one previous month, and the average of current and two previous months. The analysis took into account seasonality and distance from water bodies, and was carried out for each species separately.

3.3. Model formulation

Sporozoite rate is the proportion of positive mosquitoes to *P. falciparum* parasite calculated from entomological surveys. The data are typically binomial and modeled via the logistic regression. For a specific species, let N_{it} and Y_{it} be the number of mosquitoes tested and number infected, respectively, at location $s_i, i = 1,...,n,$ and calendar month $t = 1,...,36$ and $\boldsymbol{X}_{it} = (X_{it1}, X_{it2},..., X_{itp})^T$ be the vector of p associated predictors. We assume that Y_{it} is binomially distributed, $Y_{it} \sim Bin(N_{it}, \pi_{it})$ with parameter π_{it} measuring the sporozoite rate (SR) at location s_i and month t. A logistic regression model relates the sporozoite rate with the predictors via the equation $logit(\pi_{it}) = \mathrm{X}_{it}^{T}\boldsymbol{\beta} + f_S(t) + U_i + \varepsilon_t$ where $\boldsymbol{\beta} = \left(\beta_1, \beta_2,..., \beta_p\right)^T$ are the regression coefficients, $f_s(t)$ is a seasonality term, U_i, ε_t are parameters related with the spatial and temporal characteristics of the data, respectively. In our formulation, $f_s(t)$ captures seasonality through (i) a binary categorical variable (i.e. dry and wet) (ii) a cosine/sine function with one cycle C, every of 6 or 12 months, and (iii) a cosine/sine function with a mixture of two cycles corresponding to 6 and 12 months respectively (Stolwijk et al. 1999; Rau 2006), that is

$$f_S(t) = \sum_{c=1}^{C}\left\{\beta_{1c} * cos\left(\frac{2\pi}{T_c}t\right) + \beta_{2c} * sin\left(\frac{2\pi}{T_c}t\right)\right\}, C = 2; \quad t = 1,...,12/36$$

where T_c is the period or length of the season for cycle C (i.e. $T_1 = 12$ and $T_2 = 6$) and β_{1c} and β_{2c} describe the amplitude and phase within a period. Separate models were fitted assuming a constant seasonal pattern across the three years of the study taking $t = 1,...,12$. The ε_t model temporal correlation via an autoregressive AR(1) stationary process, i.e. $\varepsilon_1 \sim Normal(0, \sigma_1^2/(1-\gamma^2))$ and $\varepsilon_t \mid \varepsilon_{1,...t-1} \sim Normal(\gamma\varepsilon_{t-1}, \sigma_1^2), t \geq 2$. σ_1^2 is the temporal variance and γ is the autocorrelation parameter $|\gamma| < 1$ (Hay and Pettitt 2001).

3.3.1 Geostatistical modeling via Gaussian spatial processes

The U_i's are considered as observations from a latent isotropic Gaussian spatial process with covariance matrix Σ_{nxn} and elements defined by $\Sigma_{ij} = Cov(U_i, U_j)$. Under the assumption of stationarity, the spatial correlation is taken to be a function of distance between locations. The assumption is justified based a uniform movement of malaria vectors and human population in space and distribution of other transmission factors. We adopt an exponential correlation structure for the covariance matrix of the spatial process, that is $\Sigma_{ij} = \sigma^2 \exp(-d_{ij}\rho)$ where σ^2 is the spatial variance, d_{ij} is the distance between locations s_i and s_j. ρ measures the correlation decay and it estimates the effective range ($3/\rho$), the distance where the spatial correlation is less than 5%.

3.3.2 Geostatistical modeling via Gaussian spatial processes approximations

Computation of the Gaussian process requires the inversion of the covariance matrix, Σ, which for very large number of locations is not feasible. To facilitate model fit we approximate the spatial process by a subset of locations $\{s_i^*, i = 1,..,m\}$ ($m << n$) with latent observations $\boldsymbol{U}^* = \left(U(s_1^*),...U(s_m^*)\right)^T$. $\boldsymbol{U}^*$ is considered to arise from the same Gaussian process $\boldsymbol{U}$ and thus $\boldsymbol{U}^* \sim N(0, \Sigma^*)$ where Σ^* is the $\boldsymbol{m} \times \boldsymbol{m}$ covariance matrix of the sub-process. These latent observations $\boldsymbol{U}$ of the original process can be approximated by the mean of the Gaussian conditional distribution,

$$\boldsymbol{U}(\boldsymbol{s}) \mid \mathbf{U}^* \sim N(Q^T \Sigma^{*-1} \boldsymbol{U}^*, \sigma^2 - \boldsymbol{Q}^T \Sigma^{*-1} \boldsymbol{Q}) \qquad (1)$$

that is $\hat{\boldsymbol{U}} = \boldsymbol{Q}^T \Sigma^{*-1} \boldsymbol{U}^*$ where $\boldsymbol{Q} = Cov(\boldsymbol{U}^*, \boldsymbol{U})$ is an $\boldsymbol{m} \times \boldsymbol{n}$ matrix of the covariance functions between the full and the sub-process (Seeger 2003; Xia and Gelfand 2005; Banerjee et al., 2008).

3.3.3 Selection of knots

Banerjee et al (2008) proposed selecting the knots randomly from a regular grid, however this approach does not work well for clustered survey locations. Finley et al (2009) suggested selecting the subset of locations (knots) in such a way that the variance of the predictive process (1) described above becomes minimal. Although the later gives a good approximation of the spatial process, it is computationally very expensive and requires prior knowledge (a guess of the value) of the spatial variance of the full process, σ^2. Both methods of knot selections do not consider variation in the outcome. However, entomological data are sparse, clustered at household locations which are not evenly distributed within the area and have large variability. The balance sampling technique for selecting set of knots proposed by Deville and Tillé (2004) preserve the configuration of observed locations in space. The aforementioned strategy is extended here by additionally preserve the variability of the outcome and the location distance configuration. In particular, we overlay a grid over the study area and calculate the variability of the sporozoite rates within each tile. We then select the knots with inclusion probabilities equal to the proportion of tile-variability out of the total. The location coordinates (longitude and latitude in radian scale) were used as balancing variables. A sampling exercise was carried out to choose the sub-sample giving an empirical variogram as close as possible to the one obtained from the full set of observed locations. Several knot sizes have been selected equal to 50, 100, 150, 200, 250, 300 and 350. For a given knots size, 5000 samples were drawn using balance sampling and the respective empirical variogram $\xi_l(h), l = 1,...,5000$ was calculated. A distance measure D_l between the empirical variogram of the full data $\xi(h)$ and of the selected knots l was computed as,

$$D_l = \sum_{(i,j)\in N(h)} \left(\xi_l(h) - \xi(h)\right)^2$$

$\xi(h)$ is defined as $\xi(h) = \frac{1}{|N(h)|} \sum_{(i,j)\in N(h)} \left|Y_i - Y_j\right|^2$ where Y are residuals obtained after fitting a logistic regression model to the data. $N(h)$ denotes the set of pairs of observations i, j such that the distance between s_i and s_j is approximately equal to h and $|N(h)|$ is the number of pairs in the set (Cressie 1993). The subset with the minimum value of D_l across different knot sizes was considered to be the best for approximating the spatial process.

3.3.4 Bayesian kriging

Bayesian kriging was used to predict sporozoite rates at locations with no mosquito data. In particular at a set of k new un-sampled location $\mathbf{s}_0 = (s_{01}, s_{02}, ..., s_{0k})^T$ the number of infected mosquitoes $\mathbf{Y}_0 = (Y_{01}, Y_{02}, ..., Y_{0k})^T$ is estimated by the predictive distribution,

$$p(\mathbf{Y}_0 | \mathbf{Y}, N) = \int p(\mathbf{Y}_0 | \boldsymbol{\beta}, \hat{\mathbf{U}}_0) p(\hat{\mathbf{U}}_0 | \mathbf{U}^*, \sigma^2, \rho, \sigma_1^2, \gamma) \text{ x}$$
$$p(\boldsymbol{\beta}, \mathbf{U}^*, \sigma^2, \rho, \sigma_1^2, \gamma | \mathbf{Y}, N) d\boldsymbol{\beta} d\hat{\mathbf{U}}_0 d\mathbf{U}^* d\sigma^2 d\rho d\sigma_1^2 d\gamma$$

$p(\boldsymbol{\beta}, \mathbf{U}^*, \sigma^2, \rho, \sigma_1^2, \gamma | \mathbf{Y}, N)$ is the posterior distribution and $\hat{\mathbf{U}}_0$ are random effects at $\mathbf{s}_0$. $\hat{\mathbf{U}}_0 | \mathbf{U}^*$ are normally distributed, i.e. $\hat{\mathbf{U}}(\boldsymbol{s_0}) | \mathbf{U}^* \sim N(\boldsymbol{Q}_0^T \Sigma^{*-1} \boldsymbol{U}^*, \sigma^2 - \boldsymbol{Q}_0^T \Sigma^{*-1} \boldsymbol{Q_0})$ where $\boldsymbol{Q}_0 = Cov(s_0, s^*)$ is the covariance matrix between the new location and the knots.

$p(\mathbf{Y}_0 | \boldsymbol{\beta}, \hat{\mathbf{U}}_0) = \prod_{j=1}^{k} p(\mathbf{Y}_{0j} | \boldsymbol{\beta}, \hat{\mathbf{U}}_{0j})$ is binomially distributed with probability p_{0j} which estimates the sporozoite rate. At location j and time t $\text{logit}(p_{0jt}) = \boldsymbol{X}_{0jt}^T \boldsymbol{\beta} + f_s(t) + \hat{U}(s_{0j}) + \varepsilon_t$ where $\boldsymbol{X}_{0jt}$ represent the environmental covariates at the new location s_{0j} at time t. The above integral is estimated via simulation based inference (Gelfand and Smith 1990). The substantial advantage of this approach over classical methods of prediction is that it calculates the entire predictive distribution of the sporozoite rate at a given location hence allows estimation of the prediction error (Diggle et al. 1998).

3.3.5 Model validation

To assess the presence of temporal correlation in the data, spatial and spatial-temporal models were fitted. Models were applied on 90% of the data (training set) and their predictive abilities were assessed on the remaining 10% (test locations). Three approaches have been used for model validation i) the mean absolute error to measure the accuracy of the predictions by averaging the absolute difference between the observed sporozoite rates and the predicted values over the test locations, ii) Bayesian p-value calculated by the area of the predictive posterior distribution which is more extreme than the observed value, and iii) the proportion of test locations with SR included in the credible intervals (CI) of different probability coverage (ranging from 0 to 100% with interval of 5%) of the posterior predictive

distribution (Gosoniu et al. 2006). The model with the smallest mean absolute error, low Bayesian p-value and with higher proportion included in the low coverage credible intervals was considered the best.

3.3.6 Practical implementation

Bayesian model formulation requires specification of prior distributions for all parameters to be estimated. A non-informative uniform distribution was assumed for the regression coefficients β_p, i.e. β_p ~Unif(-1000,1000), inverse gamma for the variance parameters σ^2, σ_1^2 and gamma prior for the decay parameter ρ. The autocorrelation parameter was assumed to follow $\gamma \sim Unif[-1, 1]$. The model was fitted using the Gibbs sampling algorithm (Gelfand and Smith 1990). The spatial and temporal variance parameters were sampled directly from their inverse gamma full conditional distributions. The remaining parameters were simulated using Metropolis algorithm with a Normal proposal distribution. The mean of the proposal distribution was the parameter estimated from the previous Gibbs iteration with a fixed variance [Metropolis et al. 1953; Hastings 1970]. The variance was adjusted during the burn-in period to allow acceptance rates of parameters to be around 40%. Two separate chains were run in parallel with a total of 200,000 iterations each. A burn-in of 20,000 iterations was done and the last 1000 samples from each chain were used for posterior inference and prediction. The model was implemented in FORTRAN 90 using numerical libraries (The Numerical Algorithms Group Ltd). The packages *geoR* (Barry et al. 1997) and *Sampling* (Matei and Tillé 2005) in R (http://www.r-project.org) were used to calculate the empirical variogram and sampling of knots respectively.

3.4. Results

3.4.1. Descriptive analysis

A total of 11,227 *An. gambiae* mosquitoes (SR=4.29%) from 639 unique locations and 17,263 *An. funestus* mosquitoes (SR=2.35%) from 415 unique locations were analyzed. Figure 3.1 shows the crude SR for the two species with data aggregated by months and the summarized monthly rain and temperature data during the study period. Despite noise in the monthly observed SR, *An. gambiae* presents higher rates than *An. funestus*. Peaks for both species are observed in the months of May-June and Oct-Nov. It appears that the peaks of transmission are related with the trend of rain and temperature, as occurs mainly shortly after the rainy seasons. Few catches of *An. gambiae* in the months of May-July compared to that of *An. funestus* could explain the high fluctuation in the observed SR for this species.

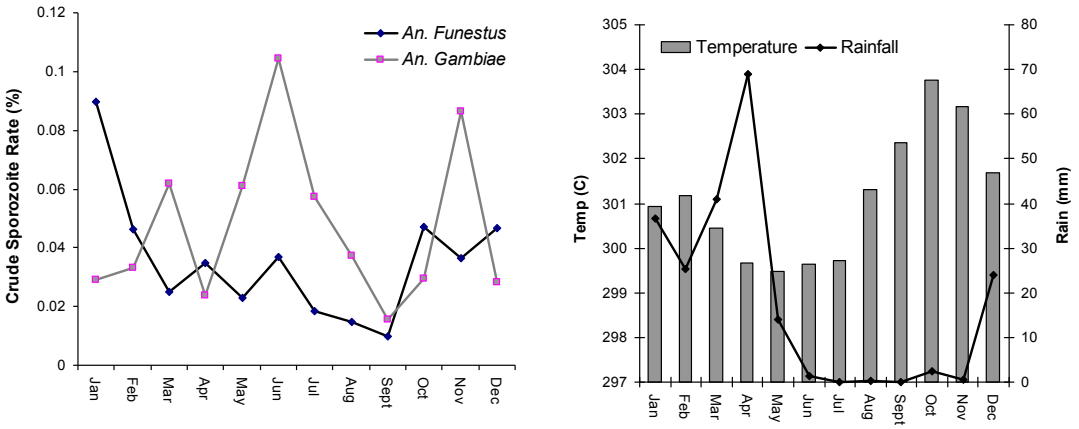


Figure 3.1: (a) Crude monthly sporozoite rate for *An. funestus* and *An. gambiae* (b) Monthly rain and temperature

Explorative analysis was carried out to select the best lag and best method to capture seasonal pattern using the Akaike Information Criteria (AIC) values. For both species, the best environmental lags were the average of 'current and one previous month' for day temperature and vegetation, and 'current month' for rainfall. Night temperature did not show any significant improvement in explaining the sporozoite rate hence was dropped. Seasonal pattern was best captured by the mixture of two cycles and period of 12 months and had a repeating pattern in

all years (p-value <0.001). Since data for 3 years (36 months) were available, all calendar months were considered the same regardless of the year. The seasonality terms included in our models were $cos(2\pi / T)$ and $sin(2\pi / T)$ for $T = 6$ and $T = 12$.

3.4.2. Selection of knots

The box plots in Figure 3.2 summarized the values of the distance measure D_l (in log scale) derived from the sampling exercise for each specific knot size for both species. The figures show the median, interquartile range, the 5th and 95th percentile intervals and extreme values beyond these intervals for 5000 estimates of D for a given knots size. Higher values of D_l indicate higher deviation of the spatial processes estimated by the knots and the original data hence a "bad" subset. For both species, *An. funestus* and *An. gambiae*, the proportion of what is referred as "bad" subset reduces significantly when the size of knots increases. The practical implication of this is that, as it is possible to obtain a good set regardless of the knots size, the chance of "landing" with a "bad" subset when small size of knots are used is higher given that the sampling is customarily done once. For model implementation, we select the subset with the minimum value of D from the knots size 200 for *An. funestus* (415 unique locations) and *An. gambiae* (639 unique).

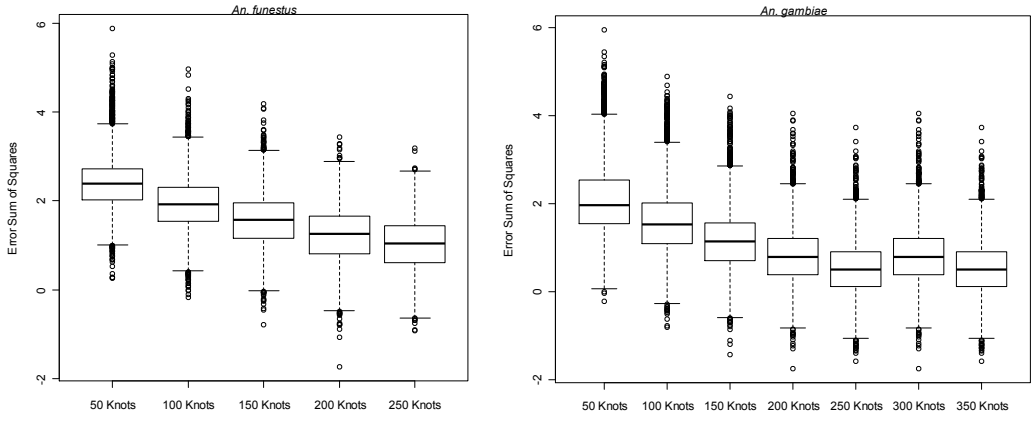


Figure 3.2: Box plots for the distance measure D for *An. funestus* (left) and *An. gambiae* (right) between the spatial parameters estimated by the empirical variables of the sub-sample

In Figure 3.3, the study area with surveyed locations for each species is plotted with red dots indicating the selected knots.

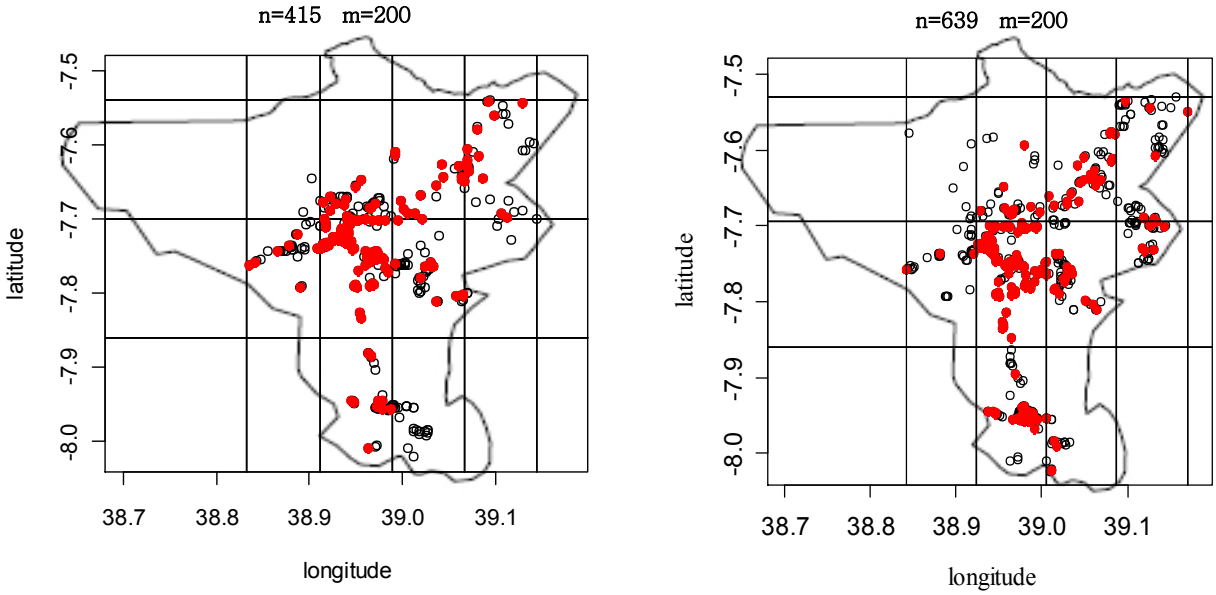


Figure 3.3: Grid showing original and sub-locations selected using a balance sampling for *An. funestus* (left) and *An. gambiae* (right)

3.4.3. Model validation

In Figure 3.4 the percentages of locations with SR included in each of the 20 credible intervals of the posterior predictive distribution, for both the space and spatial-temporal models are shown. Although the spatial model includes about 3.3% of the locations in the 5% CI, its performance was similar to that of spatial-temporal model up to 50% CI.

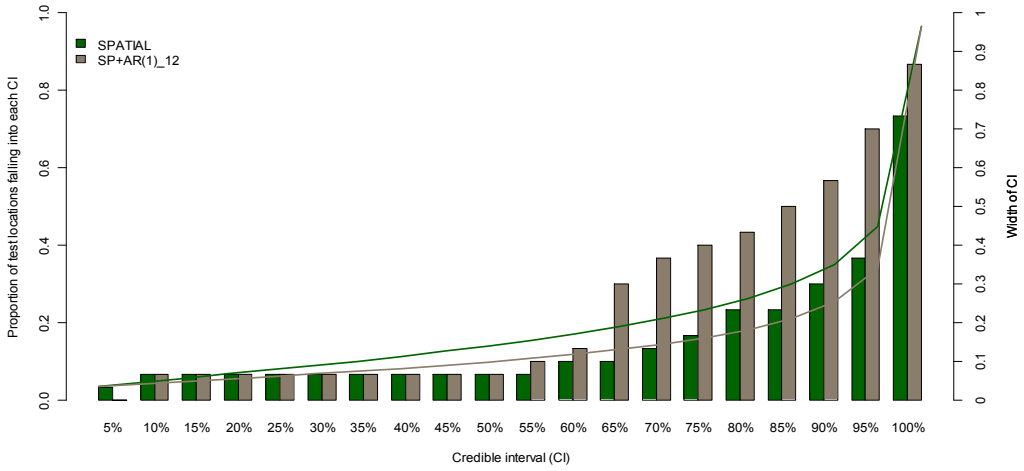


Figure 3.4: Proportion of points included in the CI of the posterior predictive distribution

At higher coverage CIs, the performance of spatial-temporal model was twice as good as the spatial model. At 95% CI, the spatial and spatial-temporal models were able to predict 36.7% and 70% of the locations respectively. The lines indicate the width of the CI. The spatial-temporal model also has the smallest CI width which indicates small uncertainty. The other two validation methods show similar results (not shown). These results suggest the spatial-temporal model to have high predictive power. Hence, in this paper, only results from this model are presented and discussed. In Table 3.2 estimates from this model are presented.

3.4.4. Model-based estimates

Table 3.2 depicts the model-based estimates for regression coefficients, spatial and temporal parameters for both species. Results obtained from a full model for *An. funestus* are presented to evaluate performance of the approximation.

Table 3.2: Space-time posterior estimates for predictors, space and time parameters

Variable	*An. funestus*				*An. gambiae*	
	Full		*Approximation*		*Approximation*	
	Median	**95% CI**	**Median**	**95% CI**	**Median**	**95% CI**
Intercept	-3.206	-4.940, -1.486	-3.738	-5.415, -2.766	-2.638	-4.031, 0.755
Environmental and Climatic Variables						
Day Temperature	-0.040	-0.211, 0.143	-0.081	-0.108, 0.078	-0.222	-0.466, -0.018
Rainfall	0.024	-0.20, 0.254	0.058	-0.069, 0.068	-0.32	-0.637, -0.029
NDVI	0.093	-0.130, 0.313	0.136	0.107, 0.151	-0.025	-0.282, 0.292
Distance to water bodies	0.210	0.051, 0.484	0.14	0.120, 0.253	0.009	-0.312, 0.362
Spatial parameters						
Spatial variance	1.051	0.490, 2.728	1.14	0.935, 1.162	0.946	0.274, 3.560
rho	35.13	13.05, 75.10	44.5	28.51, 62.73	21.61	5.78, 37.78
Range (in km)*	9.35	4.39, 25.30	7.44	5.33, 12.21	15.54	8.77, 57.72
Temporal parameters						
Temporal variance	0.303	0.133, 0.84	0.54	0.22,1.41	0.204	0.04, 1.061
Autocorrelation	0.371	-0.53, 0.87	0.63	0.42,0.72	0.356	-0.522, 0.786

***Conversion used radians to km: 1 = 110**

Effect of environment and climatic variables on the sporozoite rates differs between species. Temperature has a negative effect on the sporozoite rates of both species while vegetation and distance to the water bodies were only significant for *An. funestus*. Rain presents a negative

effect on the *An. gambiae* sporozoite rates. The effective range parameter indicates a moderate spatial correlation (less than 5%) of up to 7.4km and 15.5km for *An. funestus* and *An. gambiae* respectively. The temporal variance accounts for 32% of the total variance in the *An. funestus* and 18% for *An. gambiae* therefore most of the variability is due to spatial effect. Results of the approximation were within a good range when compared with results obtained from a model that utilized a full data, indicating good performance of the approximation.

Prediction of SR was carried out for 33,428 pixels over the Rufiji DSS area considering seasonality as well as environmental and climatic predictors. Smooth sporozoite rate maps for *An. funestus* and *An. gambiae* were produced for 12 calendar months to study seasonal differences (Figure 3.5 and Figure 3.6). Interestingly, the seasonal and spatial pattern of sporozoite rates differs between species. In particular, seasonal differences were observed in both species with the months of April and May showing the highest SR for *An. funestus* and March, September and October for *An. gambiae*.

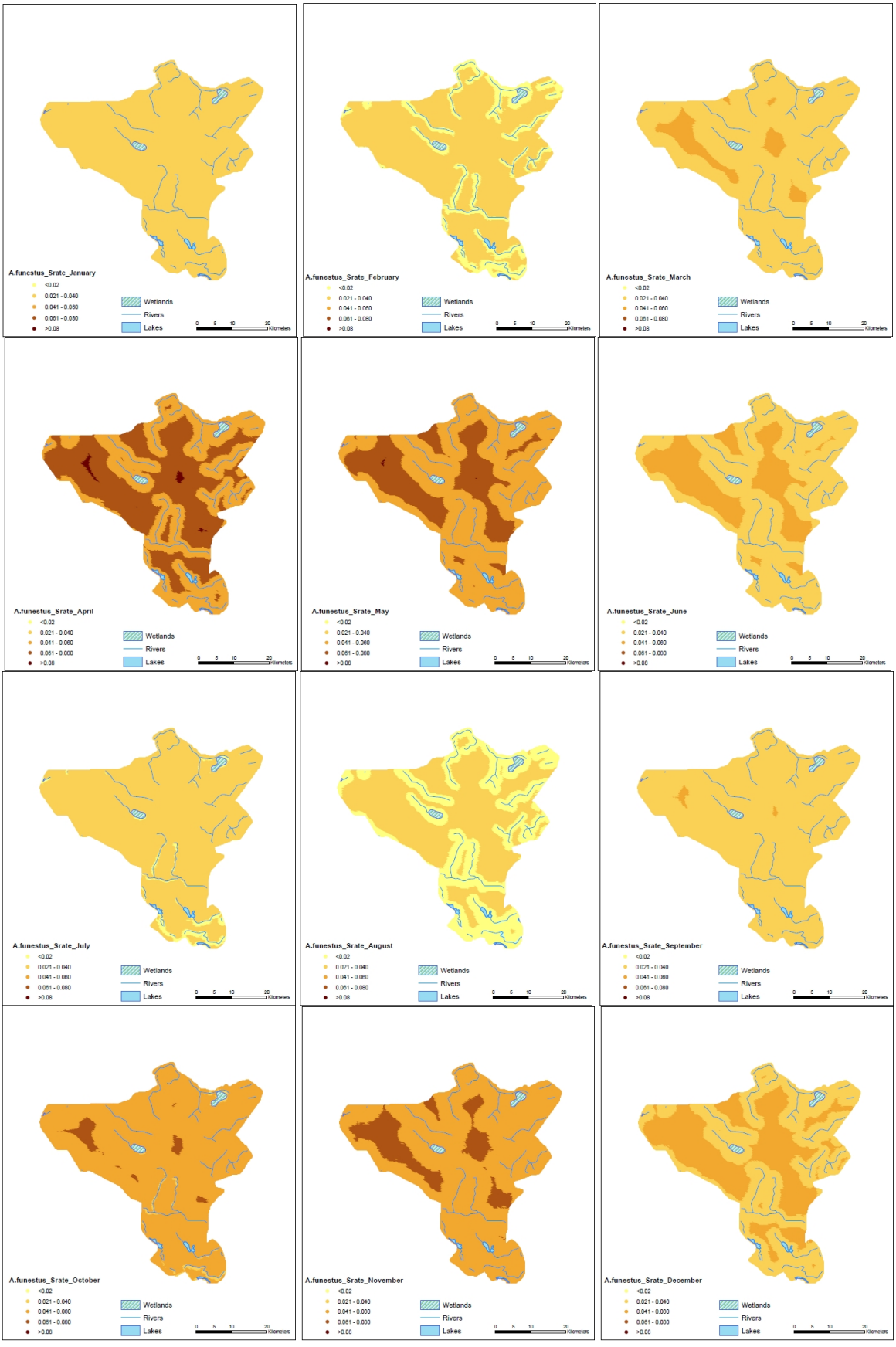


Figure 3.5: Predicted sporozoite rates *An. funestus* for January-December in Rufiji DSS, Tanzania

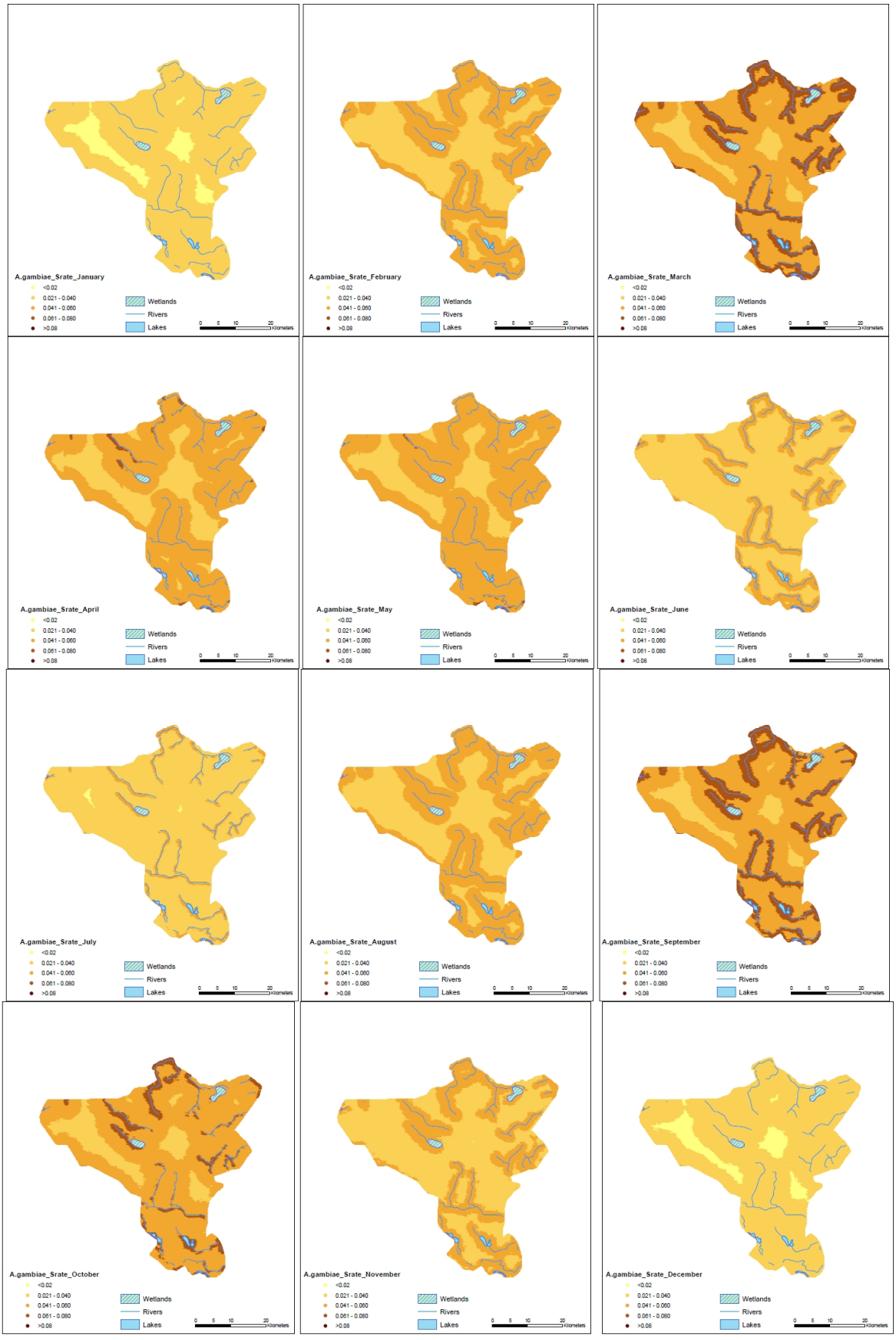


Figure 3.6: Predicted sporozoite rates *An. gambiae* for January-December in Rufiji DSS, Tanzania

Considering the heavy rains in Rufiji between March and May and the short rains during October-December, a clear relation between the rain seasons and transmission is profound in both species. The two species transmit in different regions of the study area and in different time periods which sustains the transmission over the area. *An. funestus* transmits more in the central and western parts while *An. gambiae* are in the north and south. Maps for the prediction error are found in Appendix.

The annual predicted SR was 3.7% and 4.1% for *An. funestus* and *An. gambiae* respectively. In Figure 3.7 the monthly predicted SR for *An. funestus* and *An. gambiae* averaged over the study area are plotted. The SR for *An. funestus* increased progressively starting March and peaks during April with a SR of about 6%, sharply decreasing in June and peaks again during the rainy season in October-November. The highest sporozoite rates for *An. gambiae* were observed in March and in the period of September-October. On average, the predicted sporozoite rates for *An. gambiae* were lower than that of *An. funestus*. Maximum SR predicted for *An. funestus* was about 9% while for *An. gambiae* was approximately 8%. Patterns for the observed (Figure 3.1) and predicted sporozoite rates are comparable which might suggest goodness fit of the model.

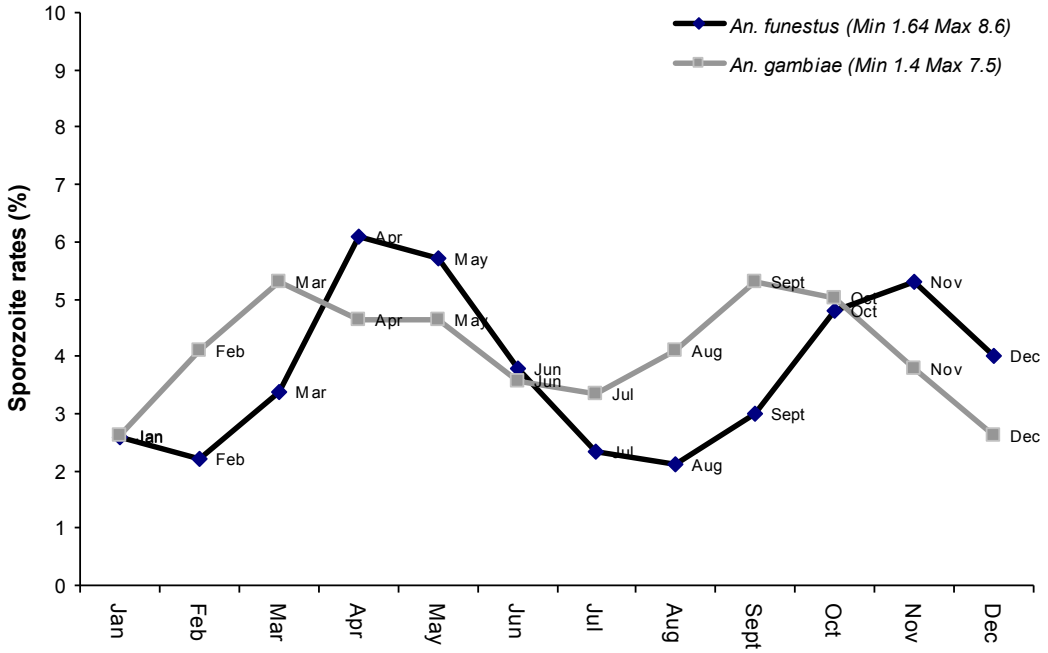


Figure 3.7: Mean predicted sporozoite rates for *An. funestus* and *An. gambiae*

3.5. Discussion

Recently, approaches have been developed for fitting large data, mainly Gaussian geostatistical data observed on regular grids by approximating the spatial process from a subset of locations. We extend this methodology for non-Gaussian geostatistical longitudinal data, observed irregularly in space and a proposed way of selecting the subset of locations for approximating the spatial process taking into account the variability of the outcome, the underlying spatial process and the distance configuration. This methodology is implemented to study seasonality and spatio-temporal patterns in malaria transmission extracting data from MTIMBA, the most comprehensive entomological database in Africa. The data have been available since 2004, however lack of appropriate methodologies delayed their analysis to date. They have been collected in 7 DSS sites with number of observed locations varying between 400 to 2800 locations.

Selection of knots for the approximation of the spatial process is still a challenging task (Finley et al. 2009). Previous suggested methodologies on selection of knots such as naive random selection would hardly provide optimal points to approximate the spatial process for this kind of data (Xia and Gelfand 2005; Xia et al. 2006; Banerjee et al. 2008). The spatially averaged predicted variance method proposed by Finley at al., (2009) can select the best possible sample regarding the spatial process but it is computationally expensive and ignores the variability of the outcome. Descriptive measures of spatial correlation, such as the variogram can provide guidance on the selection of the subset. Results of this work indicate that, it is not the size of the knots that determines the best sub-sample of location but rather its spatial representation. For the field data, with a lot of variability in the outcome and location configuration, the exercise becomes much more difficult. Balanced sampling approach with variance of the outcome as an inclusion probability and location coordinates as balancing variables was employed to select subset of location. The balance sampling algorithm was selected over other sampling techniques as it is a model-based approach which allows, unequal inclusion probability, adjusting over multiple covariates and it is easy to implement. The risk of obtaining a non representative subset when smaller knots sizes are used is quite high. To obtain a reliable configuration evaluation of different knots with different choices is required (Eidsvik et al. 2010). The computational cost gained by employing balance sampling in selecting multiple knots compared to other designs for knot selection such as SAPV (Finley et al., 2009) underscore the rationale of this choice. Selecting 200 knots out of 639 locations requires less that 5minutes in Intel® Core™ Duo Processor with 3GB of RAM.

Rufiji as most of the regions in the sub-Saharan Africa which are along the coast has stable perennial malaria transmission (Craig et al. 1999). The estimated seasonal pattern confirmed the actual situation and shows that the two species transmit malaria at different times in the year hence sustain the high transmission throughout the year. Seasonality was present in the sporozoite rates for both species and was driven mostly by environment and climate parameters. *An. gambiae* was observed to sustain transmission continuously for about three months contrary to *An. funestus* which peaks and drops quickly probably due to changes in climate. Pagès et al. (2008) reported the same situation in a two-year entomological study in Dakar concerning the sustainable transmission capacity of the *An. gambiae s.l.*. Effect of environment factors on sporozoite rates differ by species and this can be explained by the ecological behavior of these vectors for instance environment preferences on breeding areas (Wanji et al. 2003; Pagès et al. 2008; Kelly-Hope et al., 2009).

In areas that malaria is endemic, maps of the transmission pattern are useful tools to target control strategies and monitor success of interventions. Results of this study show spatial differences in malaria transmission within very small localities. The maps presented in this paper stimulate the discussion on why transmission of malaria can be sustained in a specific area despite several interventions and control available. In addition, understanding behavior of vector(s) responsible for transmission and how their response towards changes in different attributes is important. The two species *An. gambiae* and *An. funestus* indicated to be the main responsible vectors for the malaria transmission and play an important role in maintaining the transmission throughout. The species have been reported as principal malaria vectors in several parts in the sub-Saharan African (Wanji et al. 2003; Shililu et al., 1998; Lindblade et al., 1999; Atangana et al., 2009; Githeko et al., 2006). Following the modelling approach used, the model can be formulated jointly by species which will allow accounting for the correlation between species as the current model assumes independence between them.

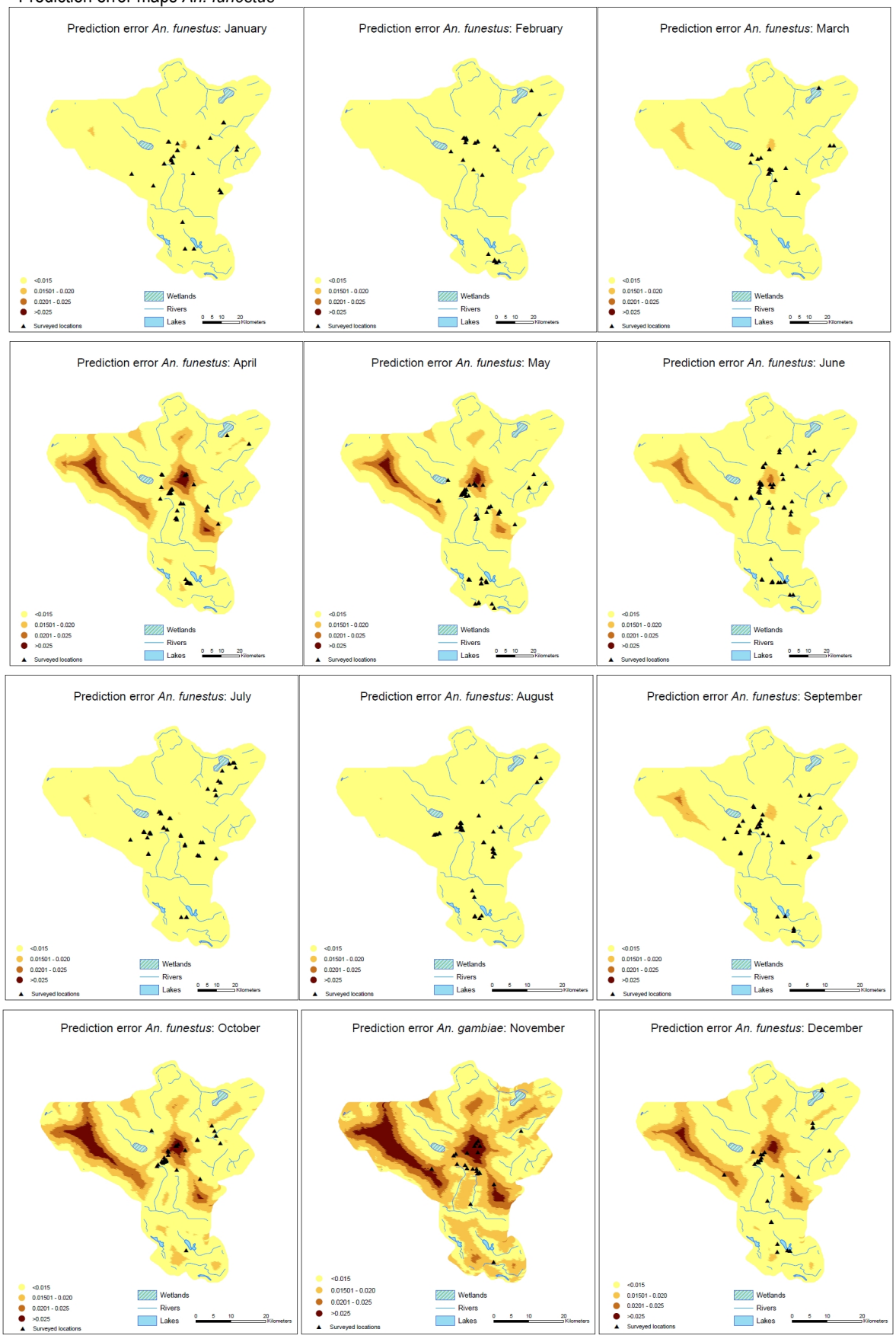
The results of this work can be extended to analyze important epidemiological questions like the relation between malaria transmission and mortality. In most of the DSS sites in sub-Saharan African countries and Asia routinely mortality data are collected. Model based prediction estimates of the malaria transmission can be obtained in location with mortality data. This study used sporozoite rates as a proxy measure for malaria transmission, however,

there are other measures of transmission intensity like Entomology Inoculation Rate which takes into account not only the infectivity of the mosquitoes but also the mosquito density. From the statistical point of view, the study confirmed that it is challenging to obtain a representative knots and the performance of the predictive process is sensitive to the choice made. Simulation based studies which leverage computational gains, assessing what important characteristics of the parent process are to be preserved and developing of a proper indicator to select knots with assured performance are among future work in this regard.

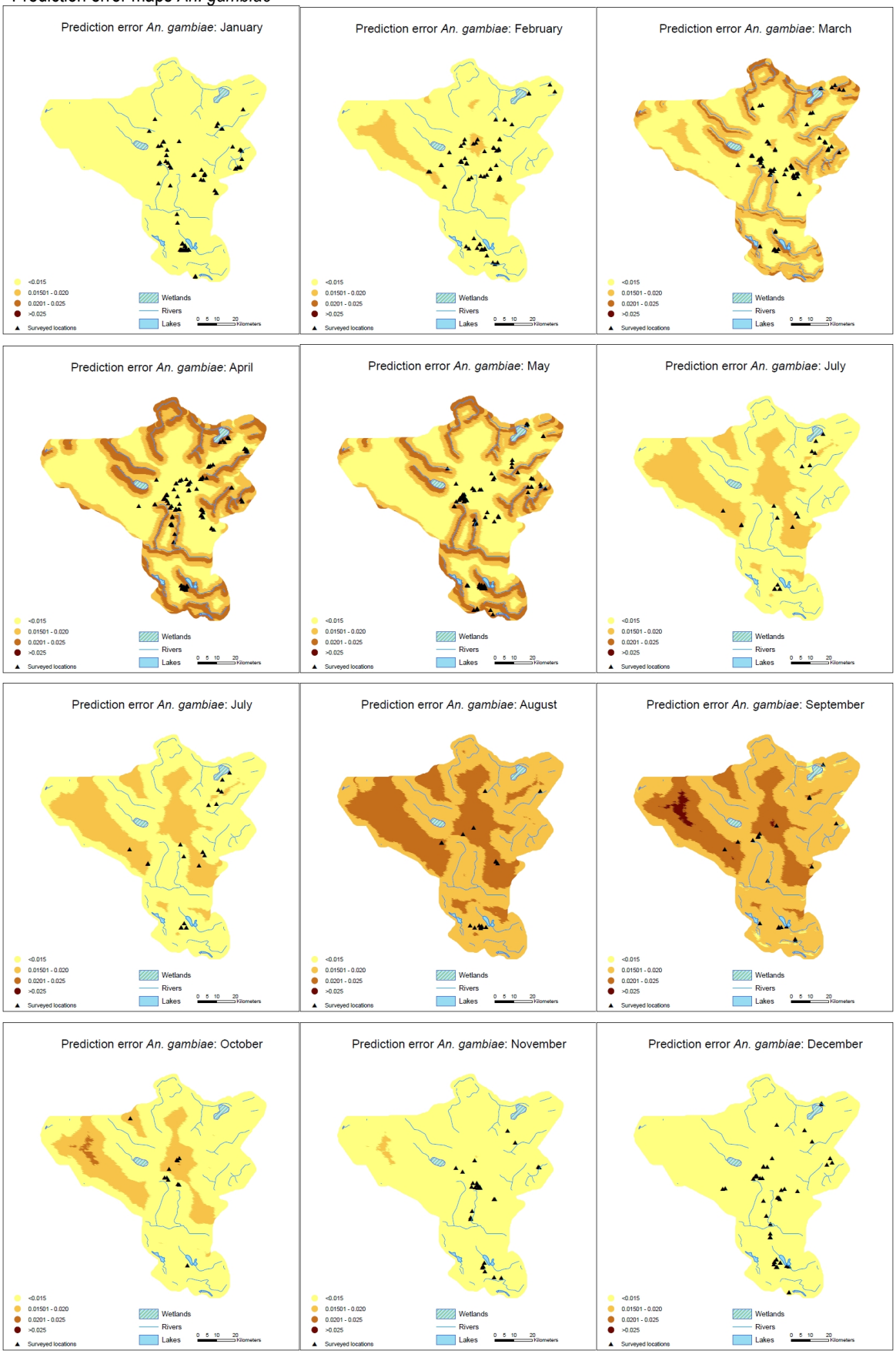
Acknowledgement: We would like to thank the whole INDEPTH Network team for facilitating the MTIMBA project and the Ifakara Health Institute for the management of the project and process of data collection and summarization. We specifically thank all field workers, data managers and supervisors in the Rufiji DSS. This study received financial support from the Swiss National Science Foundation NF. Project No. 325200-118379.

Appendix:

Prediction error maps *An. funestus*



Prediction error maps *An. gambiae*



Chapter 4 Modelling heterogeneity in malaria transmission using large sparse spatio-temporal entomological data

Rumisha SF[1,2,4], Smith T[1,2], Abdulla S[3], Masanja H[3], and Vounatsou P[1,2‡]

[1]*Swiss Tropical and Public Health Institute, Basel, Switzerland,*

[2]*University of Basel, Basel, Switzerland,*

[3]*Ifakara Health Institute, Dar es Salaam, Tanzania*

[4]*National Institute for Medical Research, Dar es Salaam, Tanzania*

‡Corresponding author

This manuscript has been submitted to the *Global Health Action*

Abstract

Malaria transmission is commonly measured using entomological inoculation rate (EIR), the number of infective mosquito bites per person per unit time. Understanding heterogeneity of transmission has been difficult due to lack of accurate data to conduct such analysis. The Malaria Transmission Intensity and Mortality Burden across Africa (MTIMBA) project has compiled a comprehensive entomological database at selected sites in Africa in a large number of household locations during 2001–2004 to study malaria transmission in relation to mortality. The data are sparse, large, with small-scale spatial-temporal variation and observed irregularly. We demonstrate the application of a more rigorous approach to analyze spatial heterogeneity of malaria transmission in Rufiji, Tanzania by employing Bayesian geostatistical models incorporating existing methods to approximate the spatial process from a subset of locations. The result of this work is essential in assessing the contribution of malaria transmission in relation to mortality and monitoring control and intervention strategies.

Keywords: approximate spatial process; malaria transmission; seasonality; MCMC; INDEPTH-MTIMBA

4.1. Introduction

Malaria remains among the serious tropical diseases with almost half of the world's population living in areas at risk of transmission. It is endemic in more than 100 countries worldwide, children and pregnant mothers being the most vulnerable groups for infections (WHO 2010b). In 2009, estimates of malaria cases were more than 200 million with about 800 thousands deaths most of these (~90%) occurring in Africa. The impact of the malaria burden on the achievement of Millennium Development Goals is enormous and its control is a potential contribution towards significant progress (WHO 2010b).

Malaria is transmitted by a female *Anopheles* mosquito. The transmission intensity is therefore highly sensitive to environmental variations that affect the densities of these vectors and their ability to transmit the infection (Hay and Snow 2006; Hay et al. 2009; Parham and Michael 2009; Gatton 2010). Up to 10-fold variations in transmission intensity have been observed within very small localities due to geographical, biological or socio-economic factors (Fontenille et al. 1997b; Beier et al. 1999; Hay et al. 2000b; Mboera et al. 2010). Understanding the heterogeneity in transmission and human exposure to malaria infection is critical for optimizing control programs and targeting interventions (Nedelman 1983; Alexander et al. 2000; Michael et al. 2001; Shaukat et al. 2010).

Malaria disease burden and transmission can be assessed using incidence or prevalence in human hosts. However, the Entomological Inoculation Rate (EIR) most directly quantifies the exposure of the human population to the infectious stages of the parasite (Smith et al. 1993; Killeen et al. 2000; Lee et al. 2001; Kelly-Hope and McKenzie 2009; Shaukat et al. 2010). EIR is the product of the human biting rate, e.g. mosquito bites/person/night (which can also be estimated using mosquito density) and the sporozoite rate (SR) which is the proportion of infective mosquitoes (Snow et al. 1999; Hay et al. 2000b). The measure expresses the average number of infective bites a person receives in a specified unit of time. It can be also used to predict other measures of transmission which are used to evaluate effectiveness of malaria control program (Woolhouse et al. 1997; Beier et al. 1999; Killeen et al. 2006). Uncertainty due to small sample, low values and variability in the SR and cost, complicate precise

estimation of EIR requiring standardized entomological surveys conducted over large areas (Fontenille et al. 1997b; Beier et al. 1999; Killeen et al. 2000; Kelly-Hope and McKenzie 2009). Accurate estimation of EIR requires longitudinal surveys within the study area to take into account spatio-temporal variations and seasonality trends. However, there is a paucity of this type of data due to cost and resources needed to collect them (Brogan and Zhao 1992; Smith et al. 2001; Thomson and Connor 2001).

The Malaria Transmission Intensity and Mortality Burden across Africa (MTIMBA) project was initiated by the INDEPTH Network (INDEPTH Network 2002) and conducted over a period of 2001 – 2004 in several countries in Africa including Tanzania, Kenya, Mozambique, Senegal, Ghana and Burkina Faso. The main objective of the initiative was to assess the relation between the intensity of malaria transmission and all-cause as well as malaria-specific mortality across Africa, taking into account the influence of malaria control activities. The MTIMBA entomological data have been collected fortnightly over large number of locations (households) and to date this is the only available entomological database appropriate to study space-time heterogeneity of malaria transmission in Africa. These data are sparse with seasonal variations and spatio-temporal correlations. Many of the survey locations had zero mosquitoes or proportion of infected ones. In standard modeling approaches, EIR is treated as a continuous outcome, logarithmically transformed to fulfill the assumption of normality (Gemperli et al. 2006a; Leisnham et al. 2007; Killeen et al. 2007b; Chase and Shulman 2009; Kweka et al. 2010). However, when EIR is estimated as a product of the SR and mosquito density which are generated from the binomial and a count distribution like Poisson or negative binomial, respectively, normality assumptions are void. In addition, due to the amount of zeros which is larger than what can be generated by the standard distributions, the data are over/under dispersed and zero inflated (Greene 1994; Cheung 2002; Yau et al. 2003; Ryan et al. 2004; Killeen et al. 2007b). Statistical analysis which accounts for these characteristic are essential to obtain unbiased estimates for the regression coefficients (Ridout et al. 2001; Agarwal et al. 2002; Warton 2005; Sogoba et al. 2007).

Moreover, the MTIMBA-EIR data have been collected at fixed locations and they are typically geostatistical data. Similar exposures of environmental and climatic conditions to locations

which are geographically close introduce spatial correlation between them. Geostatistical models take into account spatial correlation by introducing location-specific random effects as latent observations from a multivariate spatial Gaussian process (Cressie 1993). Spatial correlation between any pair of locations is considered often as a function of distance on the covariance matrix of the process. These models have large number of parameters. Bayesian formulations (Diggle et al. 1998) allow model fit via Markov Chain Monte Carlo (MCMC) simulation methods (Gelfand and Smith 1990). However, the estimation process involves covariance matrix computations which are infeasible when the number of locations is too large (Banerjee et al. 2008; Eidsvik et al. 2010). A computational flexible way to overcome this problem is the approximation of the spatial process from a subset of locations using properties of conditional multivariate Gaussian distribution of the process (Banerjee et al. 2008; Finley et al. 2009; Eidsvik et al. 2010). Most of these techniques have been applied in simulated data, observed in regular grid and mainly with Gaussian characteristics. In this study, selection of subset of locations is implemented using methods proposed in our previous work (Rumisha et al. submitted).

We now demonstrate a rigorous modeling way of analyzing large spatio-temporal EIR data and study the heterogeneity, space and temporal patterns of malaria transmission within one MTIMBA site, the Rufiji DSS area in Tanzania (Mwageni et al. 2002). The Gaussian process approximation proposed by Banerjee et al, (2008) is extended to binomial (sporozoite rates) and NB (density) data with zero inflation. The models are fitted using Bayesian MCMC simulation and assessed on the basis of their predictive ability. The EIR is computed as a product of model based predictions of SR and density. Model formulation details are given in the methodology section and selected results are presented afterwards. The discussion and conclusion of the findings consider the implications for timing and allocation of resources for malaria interventions.

4.2. Methodology

4.2.1. Study Site

The study utilized data collected from one of the MTIMBA sites in Tanzania, the Rufiji DSS (RDSS). The RDSS is located in Rufiji District, Coast Region, Tanzania about 178 kilometres south of Dar-es-Salaam. The RDSS area extends from 7.47^0 to 8.03^0 south latitude and 38.62^0 to 39.17^0 east longitude and operates in six contiguous wards and 31 villages. The surveillance area covers an area of 1,813 square kilometers and monitors 85,000 people, which is about 47% of the total population of the Rufiji District (INDEPTH Monogram). Rufiji District has an overall mean altitude of less than 500 metres. Its vegetation is mainly formed of tropical forests and grassland. The district has hot weather throughout the year and two rainy seasons; short rains (October to December) and long rains (February to May). The average annual precipitation in the district is between 800 to 1000 millimetres. A prominent feature in the District is the Rufiji River with its large flood plain and delta, the most extensive in the country (INDEPTH Monogram; Rufiji DSS Profile, 2000). The majority of the people in Rufiji District are subsistence farmers.

The main responsible malaria vectors in the area include *An. funestus,* and members of the *An. gambiae* complex, including *An. gambiae* (sensu stricto) and *An. arabiensis*. Mosquito populations usually peak during the rain seasons especially in areas where rice cultivation is taking place and during the dry months high population was usually observed in areas with permanent water bodies (INDEPTH Network 2002).

4.2.2. Mosquito Data

The entomological data were collected for the period of three years, October 2001-September 2004 (Source: http://www.indepth-network.org/dss_site_profiles/rufiji.pdf). The MTIMBA entomological protocol has been well defined in MTIMBA documentation (unpublished). In a snapshot, mosquitoes were captured at least twice every month using Centers for Disease Control (CDC) miniature light traps. The human population in the RDSS was classified into geographical clusters (100-1000 people) then for each round a simple random sampling (without replacement) was employed within clusters to select between 20-

100 "index" people (households) for set up of mosquito catches (traps). The traps were fitted indoors with incandescent bulbs and laid close to a human volunteer (randomly selected from members of the household) sleeping under an untreated bednet. Light traps operated from sundown to sunrise (i.e. 6pm–6am) for two consecutive nights in each household and bags were emptied every morning. In total mosquitoes were collected at 2479 unique locations (households).

All households where collection was done were geo-referenced. Collected mosquitoes were counted and sorted into vector species to allow separate assessment of transmission intensity. A total of 15983 *An. funestus* mosquitoes (obtained from 18% of the surveyed locations, n=447) and 17885 *An. gambiae* mosquitoes (obtained from 27.3% of the surveyed locations, n=678) were available for analysis of the sporozoite rate and mosquito density. This implies that most of the visited households, approximately 80% had zero mosquito collection. The crude sporozoite rates were 4.3% and 2.35% for *An. gambiae* (639 locations) and *An. funestus* (415 locations) respectively.

4.2.3. Environmental and Climatic Data

Remote sensing data were extracted from different sources with different spatial, Sp_R and temporal, T_R resolutions. These include normalized difference vegetation index (NDVI) [Sp_R:250m^2; T_R:16days; Source: MODIS], day and night temperature [Sp_R:1km^2; T_R:8days; Source: MODIS], rainfall [Sp_R:8km^2; T_R:10days; Source: ADDS] and distance to the nearest water bodies [Sp_R:1km^2; Source: Health Mapper].

4.2.4. Statistical analysis

Geostatistical zero inflated negative binomial and logistic regression models were fitted on the mosquito density and sporozoite rate data respectively. The models accounted for the effect of environmental and climatic predictors, seasonal patterns, spatial and temporal correlations. The predictive process was used to approximate the spatial process using a subset of locations. Model based prediction of SR and density were multiplied to obtain estimates of monthly and annual EIR. Details of the model formulation and its implementation are described in the subsections below.

i) Model formulation for density data

Let Y_{it} be the number of female mosquitoes and $\mathrm{X}_{it}^{(1)}$ be a vector of environmental predictors (extracted from satellite data) observed at location $\boldsymbol{s_i}, \boldsymbol{i} = 1,...,\boldsymbol{n},$ and calendar month $t = 1,...,36$ for a specific species. Y_{it} is assumed to follow a negative binomial distribution, $Y_{it} \sim NB(r, p_{it})$ where $p_{it} = r/(r+\mu_{it})$. $\boldsymbol{r}$ is an overdispersion parameter and μ_{it} is the mean mosquito density. Covariates $\mathrm{X}_{it}^{(1)}$, seasonal trends $f(t)^{(1)}$, spatial $U_i^{(1)}$, temporal $\varepsilon_t^{(1)} = (e_1, e_2, ..., e_t)$ and non-spatial $\phi_i^{(1)}$ random effects are introduced on the log scale of the mean count via the equation $log(\mu_{it}) = \mathbf{X}^{T^{(1)}}\boldsymbol{\beta}^{(1)} + f(t)^{(1)} + U_i^{(1)} + e_t^{(1)} + \phi_i^{(1)}$, where $\boldsymbol{\beta}^{(1)}$ is the vector of regression coefficients, $\phi_i^{(1)}$ is a residual error term capturing the remaining variability in the data. $f(t)^{(1)}$ is modeled via trigonometric function with a mixture of cycle, C

$$f(t) = \sum_{c=1}^{C}\left\{\delta_{1c}^{(1)} * cos\left(\frac{2\pi}{T_c}t\right) + \delta_{2c}^{(1)} * sin\left(\frac{2\pi}{T_c}t\right)\right\}, C = 2; \quad t = 1,...,12/36$$

T_c is the period of the season for cycle C (i.e. $T_1 = 12$ and $T_2 = 6$) and $\delta_{1c}^{(1)}$ and $\delta_{2c}^{(1)}$ are regression parameters used to describe the amplitude and phase within a period (Stolwijk et al. 1999; Rau 2006). Separate models were fitted assuming i) a constant seasonal pattern across the three years of the study by taking $t = 1,...,12$ or ii) a continuous time for the entire study period by taking $t = 1,...,36$. The seasonal pattern considering dry/wet categorization of the data was also assessed.

A zero inflated model formulation was adopted to take into account the excess zeros in the count data. The model is defined as a mixture of a degenerate distribution with mass at zero and a non-degenerate count distribution. The log-likelihood is therefore a sum of the log-likelihood for the non-zero and the zero counts. The distribution of the data is now defined as:

$$P(Y = 0 \mid p^*, \theta) = p^* + (1 - p^*)\pi(0 \mid \theta)$$
$$P(Y = y \mid p^*, \theta) = (1 - p^*)\pi(y \mid \theta), y > 0$$

where p^* is the probability for a count to arise from the zero mass and $1-p^*$ is the probability to observe a sample from a count distribution (i.e. $\pi(y|\theta) \equiv NB$ for our case, and θ is the vector of parameters associated with the distribution). This probability can be assigned a value between 0 and 1, usually approximates the proportion of zero counts in the sample or can be a function of covariates similar or different from those used in the full model (Lambert 1992; Agarwal et al. 2002; Killeen et al. 2007b; Lawal 2010). Involving possible sources of zero inflation (e.g. covariates) reduces bias in parameter estimation of p^* and other sources of uncertainty. In our case p^* is modeled with a logit link as a function of all climatic predictors X_i^* observed at location s_i, i.e. $logit(p_i^*) = X^{*T}\alpha$ where α is the corresponding vector of regression coefficients.

Bayesian model formulation requires the specification of prior distributions for all unknown parameters. For the regression coefficients, $\boldsymbol{\beta}^{(1)}$, $\boldsymbol{\delta}^{(1)}$ and α, a standard non-informative uniform prior is adopted, i.e. $\boldsymbol{\beta}^{(1)} \sim \text{Unif}(-\infty,\infty)$, $\delta^{(1)} \sim \text{Unif}(-\infty,\infty)$ and $\boldsymbol{\alpha} \sim \text{Unif}(-\infty,\infty)$ respectively. The latent observations $U_i^{(1)}$ introduced at each location s_i are assumed to derived from a multivariate normal distribution with a covariance matrix $\Sigma_{nxn}^{(1)}$, i.e. $\mathbf{U}^{(1)} \sim \text{MVN}(0, \Sigma_{nxn}^{(1)})$. The $\Sigma^{(1)}$ is a matrix with elements $\Sigma_{ij}^{(1)}$ and quantify the covariance $Cov(U_i, U_j)$ between pair of location s_i and s_j respectively. Its distribution defines the Gaussian spatial process. Under the assumption of stationarity, the spatial correlation is taken to be a function of distance between locations. An exponential correlation structure for the covariance matrix of the spatial process is adopted, that is $\Sigma_{ij}^{(1)} = \sigma_{sp}^{2^{(1)}} \exp(-d_{ij}\rho^{(1)})$ where $\sigma_{sp}^{2^{(1)}}$ is the spatial variance, d_{ij} is the distance between locations s_i and s_j and $\rho^{(1)}$ measuring the correlation decay and also known as the effective range ($3/\rho^{(1)}$) and estimates the distance where the spatial correlation is less than 5%. The decay parameter $\rho^{(1)}$ assumed to follows a gamma distribution.

Computation of the Gaussian process requires the inversion of the covariance matrix, $\Sigma^{(1)}$, which for very large number of locations is not feasible. To enable model fit we approximate

the spatial process by a subset of locations, knots, $\{s_i^*, i=1,..,m\}$ ($m<<n$) with latent observations $\boldsymbol{U}^{*(1)} = \left(U(s_1^*),...U(s_m^*)\right)^T$. $\boldsymbol{U}^{*(1)}$ is considered to arise from the same Gaussian process as $\boldsymbol{U}^{(1)}$ and thus $\boldsymbol{U}^{*(1)} \sim N(0,\Sigma^*)$ where Σ^* is the $\boldsymbol{m}\times\boldsymbol{m}$ covariance matrix of the sub-process. These latent observations $\boldsymbol{U}$ of the original process can be approximated by the "predictions" of the sub-process via the mean of Gaussian conditional distribution $\boldsymbol{U}^{(1)}(\boldsymbol{s})|\mathbf{U}^{*(1)} \sim N(Q^T\Sigma^{*-1}\boldsymbol{U}^{*(1)}, \sigma^2 - \boldsymbol{Q}^T\Sigma^{*-1}\boldsymbol{Q})$, that is $\hat{\boldsymbol{U}} = \boldsymbol{Q}^T\Sigma^{*-1}\boldsymbol{U}^{*(1)}$ where $\boldsymbol{Q} = Cov(\boldsymbol{U}^{*(1)}, \boldsymbol{U}^{(1)})$ is an $\boldsymbol{m}\times\boldsymbol{n}$ matrix of the covariance functions between the full and the sub-process (Seeger 2003; Xia and Gelfand 2005). Selection of subset of location was done using the minimax space filling design implemented in R software (Johnson et al. 1990). The approach optimizes the selection of the best subset by minimizing the maximum of the nearest-neighbor distance between the original survey and the subset locations.

The $\varepsilon_t^{(1)}$ model temporal correlation via a stationary autoregressive process of order one, i.e. $e_1 \sim Normal(0, \sigma_T^{2^{(1)}}/(1-\gamma^2))$ and $e_t | e_{1,\ldots t-1} \sim Normal(\gamma^{(1)}e_{t-1}, \sigma_T^{2^{(1)}}), t \geq 2$ where $\gamma^{(1)}$ is an autocorrelation parameter $|\gamma^{(1)}|<1$ which adopts a bounded uniform distribution, $\gamma^{(1)} \sim \text{Unif}[-1,1]$ and $\sigma_T^{2^{(1)}}$ is the temporal error (Hay and Pettitt 2001). The $\phi_i^{(1)}$ are assumed to follow a normal distribution with mean zero and a homoscedastic variance $\sigma_e^{2^{(1)}}$. Inverse gamma priors are adopted for the variance parameters $\sigma_{sp}^{2^{(1)}}$, $\sigma_T^{2^{(1)}}$ and $\sigma_e^{2^{(1)}}$.

ii) Model for SR

Let $\boldsymbol{N_{it}}$ and $\boldsymbol{Z_{it}}$ be the number of mosquitoes tested and number infected, respectively at location s_i and calendar month t. Z_{it} is assumed to arise from a Binomial distribution, $Z_{it} \sim Bin(N_{it}, \pi_{it})$ where $\boldsymbol{\pi_{it}}$ measure the SR at location s_i and time t. The regression function links the SR with other terms of the model (as shown for the density data) is given as $logit(\pi_{it}) = \mathrm{X}^{T(2)}\beta^{(2)} + f(t)^{(2)} + U_i^{(2)} + \varepsilon_t^{(2)} + \phi_i^{(2)}$. Similar specification described for the density model is followed in this model.

iii) Data management and environmental lags

To facilitate the assessment of the seasonal pattern, data was summarized by location and calendar month. That implies that all repeated surveys from a specific location within the same month were collapsed (sum of mosquito density/tested and positive) to a single observation.

To account for the environmental-lag-effect on mosquito density or sporozoite rate, non-spatial (negative) binomial models (with/without zero-inflation) were fitted and best lags were assessed. Lags refer to a climate/ environment value at different time intervals prior to the study date that might influence the amount of mosquitoes collected or the sporozoite rate. Lags considered include the current month (month of collection of mosquitoes); one/two/three month(s) prior to the collection; average of current and one previous month; average of one and two previous months; and lastly average of current, one and two previous months. The analysis took into account seasonality, distance from water bodies and time (annual effect) which was incorporated as a binary variable indicating the year of study. Analysis was conducted separately for each species. Fitted values from models with all possible combinations of the environmental lags were calculated and plotted against the observed values (mosquito counts or sporozoite rate). The combination which fits best the data was used for further analysis. This was implemented in STATA 10 (Stata Corps).

iv) Model validation and prediction

About 15% (test locations) randomly selected from the entire dataset was left out and models were fitted using the remaining 85% of the data. The predictive ability of the model was then assessed using the test locations. Specifically we calculated different credible intervals with different probability coverage of the posterior predictive distribution and compare the percentage of test locations which falls within these credible intervals (Gosoniu et al. 2006). The predicted power reported represents the percentage covered at 95% credible interval. Using the estimates obtained from the models, SR and mosquito density were predicted for the whole Rufiji site. The prediction was done at the 250m resolution.

v) Calculation of EIR

The EIR can be estimated as a product of the sporozoite rate and human-biting rate. Depending on the mosquitoes collection method used (human landing, light trap, etc), the human-biting rate can be correctly approximated either by the number of blood meals taken on humans/ mosquito/day or by the mosquito density. Established correlation between number of mosquitoes captured by light traps and human landing catches is usually used to adjust light trap collection to equivalence of biting catches and avoid collection bias (Lines et al., 1991). For this study EIR was calculated as a product of sporozoite rate (SR) and mosquito density and then adjusted using a correction factor of 1.605 to calibrate estimates obtained from light trap collection (Lines et al., 1991; Amek et al., 2012).

At a specific pixel j and month t the predicted values of SR, $\hat{\pi}_{jt}$ and mosquito density, $\hat{\mu}_{jt}$ were obtained for *An. funestus* and *An. gambiae* species. EIR estimates representing the infectious bite/person/day were calculated as:

$$\hat{EIR}_{jt} = 1.605 * \left((\hat{\pi}_{jt_{af}} * \hat{\mu}_{jt_{af}}) + (\hat{\pi}_{jt_{ag}} * \hat{\mu}_{jt_{ag}}) \right)$$

where 1.605 is the correction factor.

The $\hat{EIR}_{jt}$ were then multiplied by 30.5 and 365 to obtain monthly and annual estimates respectively. Monthly and annual maps were produced to show seasonal and temporal trend of the transmission.

vi) Geostatistical model implementation

The final model was implemented in OpenBUGS and parameters were estimated using the Gibbs sampler MCMC algorithm. The spatial variance parameter was sampled directly from its inverse gamma full conditional distributions using Gibbs sampling (Gelfand and Smith, 1990). The remaining parameters were simulated using Metropolis algorithm with a Normal proposal distribution. The mean of the proposal distribution was the parameter estimated from the previous iteration with a fixed variance (Hastings, 1970; Metropolis, 1987). Two separate chains were run in parallel with a total of 150,000 iterations each. A burn-in of 20,000 iterations was done and the last 5000 and 1000 samples were used for posterior inference and prediction respectively. The Gelman-Rubin model diagnostic tool (Gelman and Rubin, 1992)

was used to assess convergence of chains before summarizing the results. The package '*fields*' in R was used for selection of knots. For practical implementation of the geostatistical model 281 knots (2479 unique locations) were selected for the density data (both species), 177 (415 unique locations) for SR analysis of *An. funestus* and 219 (639 unique locations) for SR of *An. gambiae*. Predictions and calculation of EIR were done in Fortran 95 (Compaq Visual Fortran Professional 6.6.0).

4.3. Results

4.3.1. Data Description

In total of 2479 unique locations were visited for the collection of the mosquitoes. A total of 15983 and 17885 *An. funestus* and *An. gambiae* mosquitoes were captured respectively. About 83% and 74.3% of the visits for mosquito collection for *An. funestus* and *An. gambiae* received zero counts. The crude annual sporozoite rates, calculated as the percentage of infectious mosquitoes out of those tested, were 3.3%, 2.8% and 3.2% for Year 1 (October 01-September 02), Year 2 (October 02-September 03) and Year 3 (October 03-September 04) respectively. The observed EIR (i.e. observed SR multiplied by man-biting rate) were 507, 72.8 and 146 infectious bites/person/year for three years respectively. In Figure 4.1, the relation between rainfall, temperature and mosquito density is shown (data collapsed in a period of one calendar year).

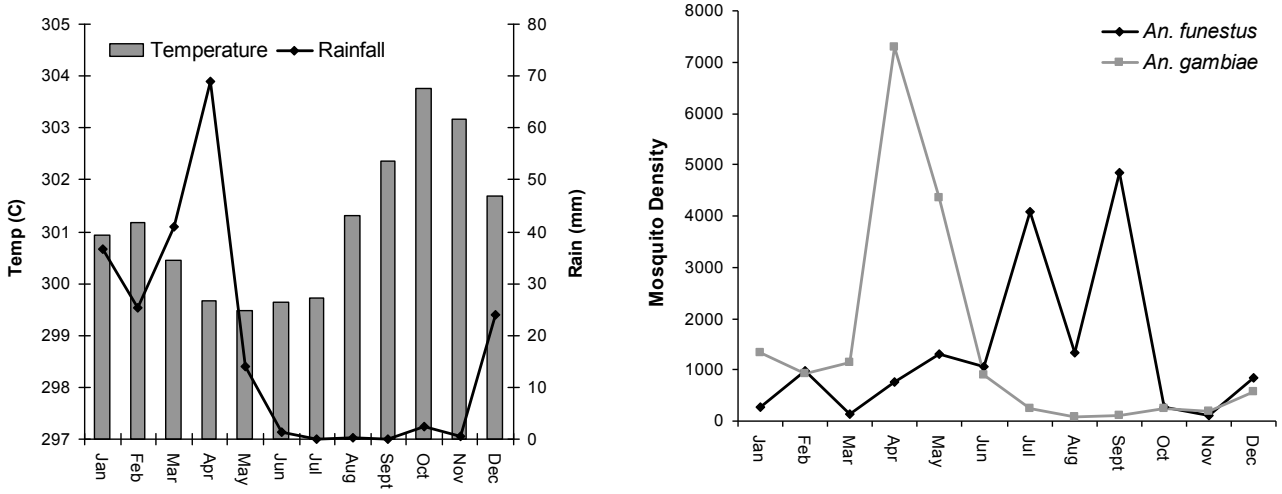


Figure 4.1: Seasonal variations of a) rainfall, temperature and b) mosquitoes densities of *An. gambiae* and *An. funestus* in the Rufiji DSS October 2001- September 2004

Most *An. gambiae* mosquitoes were captured during the months of April and May while most *An. funestus* were collected in the period of July – September. The number of *An. gambiae* collected was higher during the heavy rains while short rains with high temperature favor the population of *An. funestus* (Figure 4.1). Almost 74% of total *An. funestus* and 63% of *An. gambiae* mosquitoes caught for the entire study period were collected during the first Year. More than 50% of the all tested and positive mosquitoes were obtained in the 1st year of the study (data not shown).

4.3.2. Geostatistical model results

Table 4.1 summarizes the results of parameter estimations from a multivariate geostatistical models on sporozoite rates and mosquito density.

Table 4.1: Association of environment/climate variables on sporozoite rate and mosquito density and spatio-temporal parameters

	SR: Model: Binomial		**Density: Model: Zero Inflated NB**	
Parameter	**Median [95% CI[a]]**		**Median [95% CI[a]]**	
Seasonality	***An. funestus***	***An. gambiae***	***An. funestus***	***An. gambiae***
Cos 12	0.99 [0.41,2.41]	0.72 [0.29,1.66]	1.1 [0.54,2.3]	0.39 [0.2,0.86]
Sin 12	0.84 [0.31,2.53]	0.54 [0.19,1.32]	0.75 [0.4,1.55]	0.6 [0.32,0.96]
Cos6	1.27 [0.66,2.47]	0.81 [0.44,1.53]	0.75 [0.43,1.39]	0.76 [0.41,1.13]
Sin6	0.65 [0.34,1.25]	0.87 [0.45,1.68]	1.13 [0.58,2.08]	0.99 [0.53,2.43]
Environment and Climate				
NDVI	1.03 [0.85,1.25]	0.93 [0.79,1.1]	1.15 [0.87,1.6]	1.11 [0.92,1.35]
RAIN	0.96 [0.73,1.26]	**0.53 [0.36,0.79]**	**1.33 [1.06,1.68]**	1.26 [0.97,1.79]
Day temperature	**2.31 [1.06,6.97]**	0.92 [0.7,1.22]	1.23 [0.81,1.69]	**0.77 [0.64,0.89]**
Night temperature	1.04 [0.52,3.51]	0.96 [0.73,1.27]	**1.47 [1.02,2.02]**	0.84 [0.69,1.03]
Distance to the water bodies	0.93 [0.76,1.11]	0.97 [0.85,1.1]	0.96 [0.65,1.22]	0.94 [0.79,1.11]
Annual Trend				
Year 2	1.01 [0.61,1.67]	**0.48 [0.31,0.75]**	**0.13 [0.08,0.24]**	**0.17 [0.11,0.25]**
Year 3	**0.41 [0.2,0.79]**	**0.37 [0.24,0.57]**	**0.34 [0.2,0.61]**	**1.6 [1.04,2.53]**
Spatial Process				
Range (in km)[bc]	35.52 [11.1,78.81]	49.95 [15.54,81.03]	21.1 [12.2, 56.6]km	15.5 [8.9, 32.19]km
Variance σ^2_{sp}	0.9 [0.37,2.36]	0.45 [0.2,1.18]	11.35 [6.58,29.2]	5.04 [3.1,10.33]
Temporal Process				
Correlation γ	0.5 [-0.52,0.96]	0.5 [-0.51,0.96]	-0.15 [-0.79,0.67]	0.08 [-0.77,0.83]
Variance σ^2_T	0.34 [0.14,1.11]	0.33 [0.14,0.94]	0.61 [0.22,2.59]	0.51 [0.2,2.55]
Other parameters				
Non-spatial variance σ^2_e	0.31 [0.16,0.61]	0.34 [0.19,0.59]	2.88 [1.81,4.4]	2.59 [1.89,3.2]
Overdispersion r	---	---	2.64 [1.7,3.67]	1.16 [0.77,1.61]
Effect of covariates on the mixing probability				
NDVI	---	---	**0.3 [0.17,0.54]**	0.93 [0.7,1.29]
RAIN	---	---	1.3 [0.84,5.37]	0.65 [0.36,1.85]
Day temperature	---	---	**0.07 [0.01,0.64]**	**0.05 [0.02,0.18]**
Night temperature	---	---	0.53 [0.27,1.14]	0.71 [0.28,3.64]

[a]: Credible Intervals (or posterior intervals)

[b]: Based on spatial decay parameter, the Range is calculated as $3/\rho$ (x111km).

[c]: The spatial correlation is significant (>5%) within this distance

The effect of environmental variables differs significantly between species. Rain and temperature are the most influencing factors for density and sporozoite with higher effect on the *An. funestus* species. No significant effect of distance to the water bodies was obtained. Temporal effect (annual) is highly pronounced with a significant increase of mosquito population in the third year of the project as compared to first year. Spatial ranges are quite high especially for the sporozoite rates. The estimate of the over dispersion parameter of *An. funestus* is twice as large as that of *An. gambiae* which could be the influence of the amount of zeros in the data. However, the estimate of $\boldsymbol{r}$ is larger than 1 indicating that the data are not highly overdispersed. Day temperature significantly reduces the probability of observing zero mosquito counts. Spatial variability accounts more for the total variability in the data as compared to the non-spatial and temporal.

For a total of 63, 99, 368 and 368 test locations selected for validation of SR-*An. funestus*, SR-*An. gambiae*, Density-*An. funestus* and Density-*An. gambiae* models respectively, 68.3%, 63.6%, 84.1% and 89.9% of the locations were correctly predicted within 95% credible interval respectively.

4.3.3. Mapping EIR

Figure 4.2, presents selected EIR maps for the Rufiji DSS site for the *An. funestus* and *An. gambiae*.

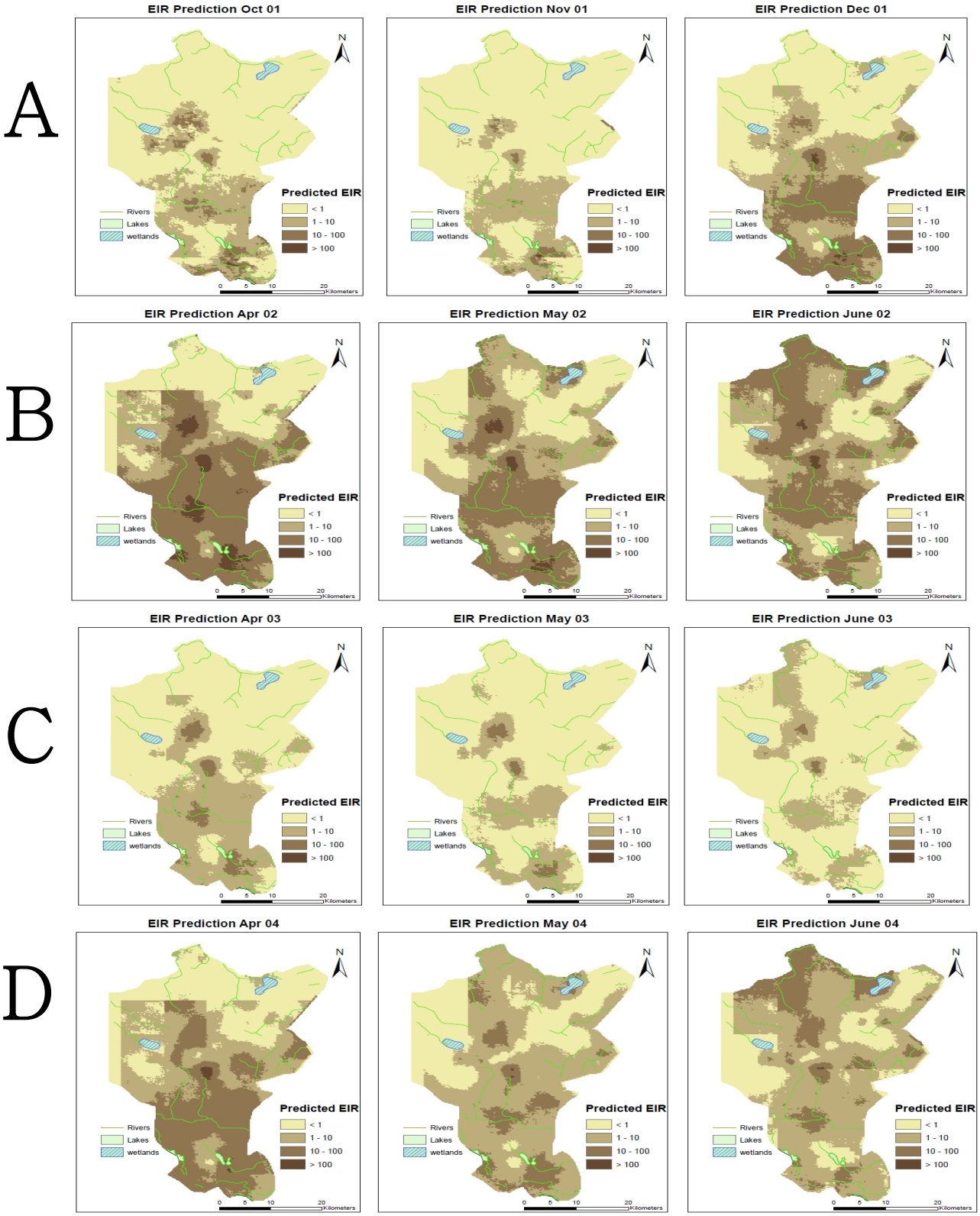


Figure 4.2: Selected EIR maps showing the spatial distribution and the seasonal pattern, for the period of Oct 2001- Sept 2004. A) Dry months followed by the period of short rains, B) Months immediately after the onset of heavy rains during the 1st year (very wet), C) Months immediately after the onset of heavy rains during the 2nd year (dry) and D) Months immediately after the onset of heavy rain season during the 3rd year (normal rains)

The southern part of the DSS showed high transmission throughout the years. High transmission was observed immediately at the onset of rains, especially during the heavy rain period. At the end of the rainy season (May-June) the transmission spread throughout the region.

In Figure 4.3 monthly time series (median) predicted EIR are plotted for the entire study period. Attribute of each species is also indicated.

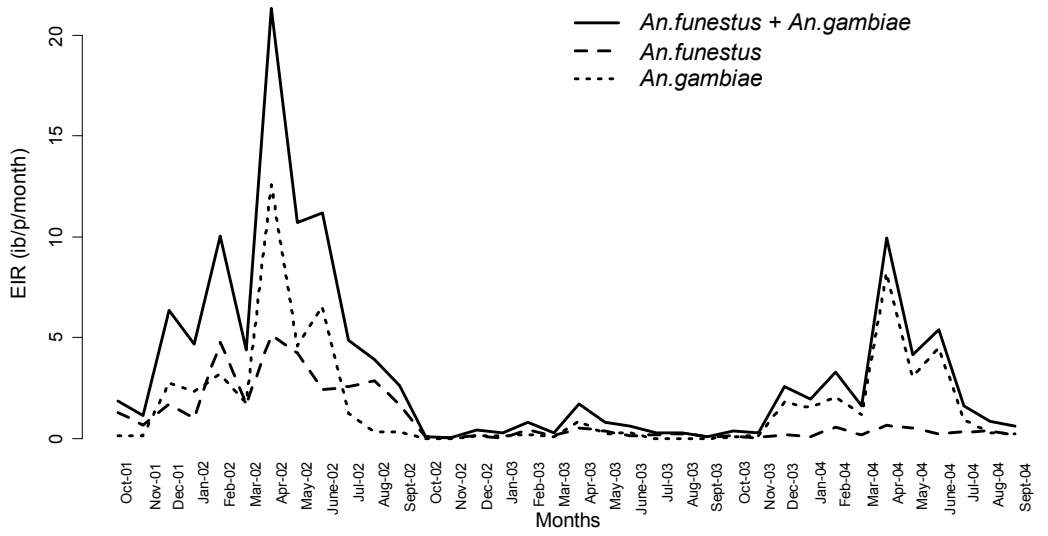


Figure 4.3: Predicted monthly EIR median and attribute of each species in Rufiji DSS

The transmission starts peaking at the month of April (just after rains) and gradually dropped in July (1st year of the study). There was a reduction in the second year of the study and EIR increased again during the last year. Similar monthly trend is observed across years which emphasize seasonality. *An. funestus* are more prominently during the dry months while *An. gambiae* responsible for the rainy periods. The spatial temporal distribution of year-by-year EIR is shown in Figure 4.4 with maps of prediction error. The prediction error for the EIR estimates was obtained my multiplying the prediction errors obtained from SR and Density models.

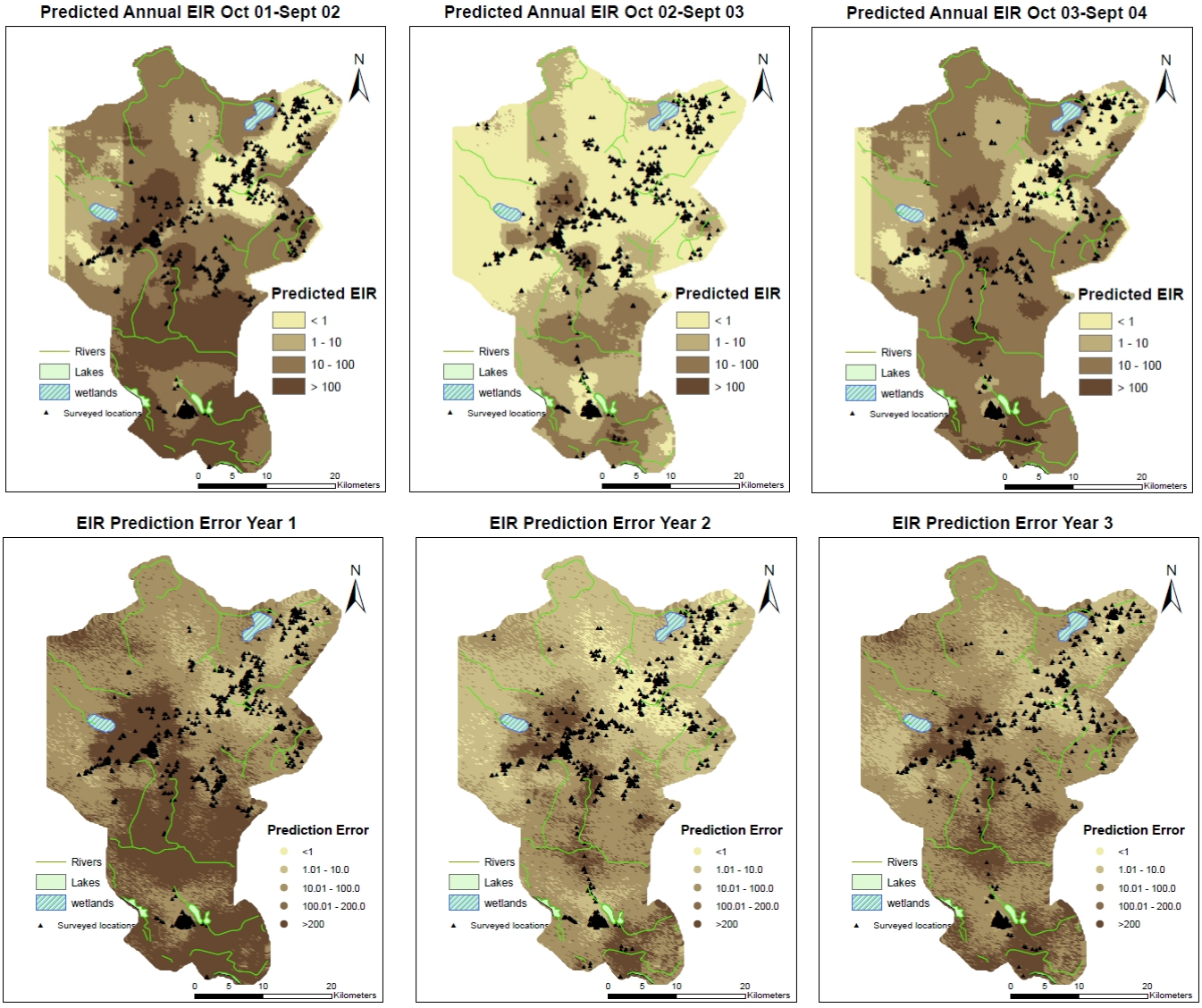


Figure 4.4: Spatial temporal distribution of annual EIR with prediction error maps.

Patterns in Figure 4.4 shows that few surveyed households are located in areas with EIR less than 1, however, a large proportion of household presented high transmission intensity. Higher prediction errors are seen in areas with no or few surveyed locations.

4.3.4. Population adjusted EIR

The annual and species-specific population-adjusted EIR were calculated by averaging predicted inoculation rates at all households within the RDSS (Figure 4.5) excluding all the other pixels. Results are presented in Table 4.2.

Table 4.2: Overall predicted EIR with the percent (%) attribute of each species

Period	*An. funestus* + *An. gambiae*	*An. funestus*	*An. gambiae*
October 2001 - September 2002	853.6	582.9 (68%)	270.7 (32%)
October 2002 - September 2003	113.7	88.8 (78%)	24.9 (22%)
October 2003 - September 2004	286.1	107.2 (37%)	178.9 (63%)

Overall transmission intensity reduced significantly during year 2 and 3 as compared to year 1 of the study. *An. funestus* was the main responsible vector for transmission in the first (68%) and second (78%) year while during the last year transmission was mainly driven by *An. gambiae* (63%).

In addition, we assessed the spatial shift (distribution) of transmission intensity over time, as illustrated in Table 4.3. EIR were categorized into five transmission intensities which were; no transmission (EIR=0), very low (EIR=>0.0-1), low (EIR=>1-10), average (EIR=>10-100), and high (EIR=>100). The change in the percentage of households exposed to a specific level of transmission was then studied.

Table 4.3: Distribution of predicted EIR over the RDSS area by Year, N*(%):

Category	EIR range	Year 1 N*(%)	Year 2	Year 3
No	**0**	4896 (27.5)	13124 (73.8)	4225 (23.8)
Very low	**>0.0-1**	704 (4.0)	1320 (7.4)	1286 (7.2)
Low	**>1-10**	4568 (25.7)	2081 (11.7)	6779 (38.1)
Average	**>10-100**	5377 (30.2)	1068 (6.0)	4781 (26.9)
High	**>100**	2238 (12.6)	190 (1.1)	712 (4.0)

*** The number of households within a specific transmission intensities category**

The proportion of households predicted with very low transmission intensity increased between the 1st year and the 3rd year of the study, from 4.0% to 7.2%. A significant reduction (over 68%) of locations with high transmission is seen during the last year of the study (i.e. 12.6% in the 1st year to 4% in the 3rd year).

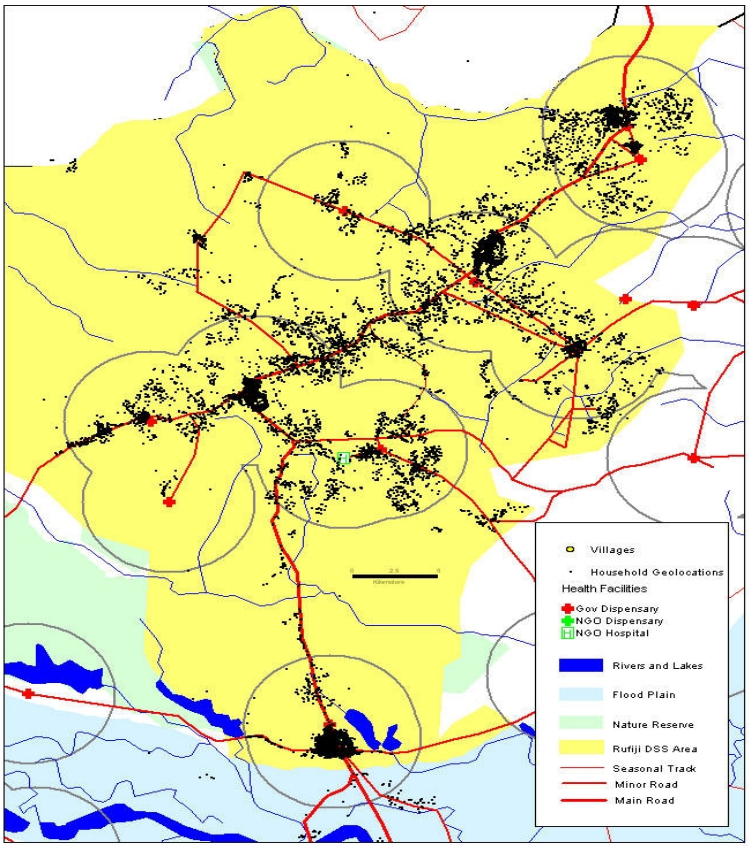


Figure 4.5: Distribution of households in the Rufiji DSS area (N=14,516), Source: TEHIP, 2002

4.4. Discussion

In this study we assessed spatial-temporal variation and heterogeneity of malaria transmission in the Rufiji DSS site using a large geo-referenced biweekly entomological dataset collected over three years, and rigorous Bayesian geostatistical models. Our work is amongst the few to address spatial modeling of EIR based on sparse data by applying current Bayesian methodologies approximating spatial processes for large data. The INDEPTH-MTIMBA data which was used in our application is the most comprehensive entomological database in Africa. Bayesian spatio-temporal binomial and zero inflated negative binomial regression models were developed to produce monthly maps of EIR taking into account the malaria-climate relation and seasonality in transmission (Thomas and Lindsay 2000; Sogoba et al. 2007; Gosoniu 2008; Reid et al. 2010; Stresman et al. 2010).

Geostatistical models have been widely used in malaria mapping in recent years ((Amek et al., 2012; Diggle et al. 1998; Gemperli et al. 2004; Gosoniu et al. 2006; Hay et al. 2009; Amek et al. 2011). Most of these analysis involved standard geostatistical models which are relevant for a moderate number of locations. Computation involved in these models is not feasible for data collected over large number of survey locations. In this study we used methods proposed by Barnejee et al. (2008) and Finley et al. (2009) to approximate the spatial process using a subset of survey locations selected via space filling design. Additive temporal correlations with autoregressive structure were incorporated in all models.

Changes in climate conditions, natural inhabitants and other human activities which depend on environment alter the intensity of malaria transmission (Snow et al. 1998b; Thomson and Connor 2001; Hay et al. 2002). Our results depict temporal and seasonal variation in EIR along the study period and study area. Transmission was higher during the rainy periods with high temperatures and very low during the dry season or year. Two species *An. funestus* and *An. gambiae* are mainly responsible for malaria transmission in this region. Differences on the effect of environmental factors on the mosquito abundance and sporozoite rates of the species were observed. The population of *An. gambiae* increases at the onset of heavy rains while that of *An. funestus* peaks during the short rains season. Similar results have been

reported in the Kilombero valley and other areas with similar climate in Africa and are associated with the preferential conditions of breeding sites of these species (Gillies and De Meillon 1968; Smith et al. 1993; Charlwood et al. 2000; Warrell and Gillies 2002; Guelbeogo et al. 2009; Adja et al. 2011; Mala et al. 2011). Highly significant effects of temperature on the SR and density of *An. funestus* was observed. Contrary to *An. gambiae* which has relatively exophilic behavior, this species is strictly endophilic, which could facilitate choice of conducive resting environment favoring the gonotrophic cycle resulting to higher survival hence longer infectivity (Charlwood et al. 2001; Kent et al. 2006; Atieli et al. 2009). Knowledge of these characteristics can be important for understanding disease dynamics and for efficient implementation of interventions (Beier et al. 1999; Thomson et al. 2005; Koudou et al. 2010; Adja et al. 2011).

There was considerable variation over short distances in intensity of transmission. Small scale variations in malaria transmission are commonly in sub Saharan Africa and create complexity in implementing strategies to combat malaria (Thomas and Lindsay 2000; Drakeley et al. 2005; Stewart et al. 2009; Bousema et al. 2010a; Mboera et al. 2010). The spatial correlation was still present over a substantial distance and the spatial variation comprised of about 90% of the total data variance. The spatial correlation arises partly due to spatial pattern in environmental drivers of transmission, partly due to effects of limited mosquito dispersion, and is also affected by human factors such as migration and human population densities (Finley et al. 2009; Eidsvik et al. 2010). We had an abundance of data on both mosquito and human populations, however, due to relative small DSS area, it is difficult to separate the contributions of these different factors to the spatial correlation which explains the higher spatial range.

The data available from Rufiji include comprehensive records of mortality in the human population at the time of the entomological surveillance. In our future work, the exposure surfaces estimated in this work will be used to assess the relation between malaria transmission and mortality. By allowing for both temporal variation and small area spatial variation in EIR, this analysis should provide much more accurate estimates of the benefits to

be gained by reducing malaria transmission than is possible from analyses that aggregate EIR over large areas and time periods.

Competing interests: The authors declare that they have no competing interests.

Authors' contributions: PV and TS conceived the idea and design the data analysis techniques, identified data sources, supervise the work and writing of the manuscript by providing intellectual contents. SFR carried out the analysis and interpret the results, draft the manuscript and coordinate the writing process; SA and HM participated in the design of the study, supervised data collection, its management and provided intellectual inputs for the manuscript. All authors read and approved the final manuscript.

Acknowledgements: We are grateful to the following individuals for their contributions to this paper; the community in the RDSS and staff from IHI who devoted their time to the acquisition of the data. The study received funding from the Swiss National Science Foundation (Nr 325200-118379/1).

Chapter 5 Assessing the relation between child survival and malaria transmission: an analysis of the MTIMBA data in Rufiji DSS, Tanzania

Rumisha SF[1,2,4], Smith T[1,2], Masanja H[3], Abdulla S[3], and Vounatsou P[1,2‡]

[1]*Swiss Tropical and Public Health Institute, Basel, Switzerland,*

[2]*University of Basel, Basel, Switzerland,*

[3]*Ifakara Health Institute, Dar es Salaam, Tanzania*

[4]*National Institute for Medical Research, Dar es Salaam, Tanzania*

‡Corresponding author

The manuscript has been submitted to *Malaria Journal*

Abstract

The precise nature of the relationship between malaria mortality and levels of transmission is unclear, since efforts to assess this have been inconclusive because of methodological limitations. The MTIMBA project was initiated by the INDEPTH Network to collect longitudinally entomological data within a number of DSS in sub-Saharan Africa to study this relationship. This work linked the MTIMBA entomology data, with the routinely collected vital events within the DSS to analyze the transmission-mortality relation in the Rufiji DSS. Bernoulli spatio-temporal regression models with village clustering, adjusted for age and ITNs, were fitted to assess the relation between mortality and malaria transmission measured by EIR. EIR was predicted at household locations using transmission models and it was incorporated in the model as a covariate with measurement error. Hazard ratios (HR) were estimated for predictors, spatial and temporal parameters. Separate analysis was carried out for neonates, infants and children 1-4 years of age. No significant relation between all-cause mortality and intensity of malaria transmission was indicated at any age in childhood. However, a strong age effect was shown. Comparing effects of ITN and EIR on mortality at different age categories, a decrease in protective efficacy of ITN was observed (i.e. neonates: HR=0.65; 95% BCI: 0.39-1.05; infants: HR=0.72; 95% BCI:0.48-1.07; children 1-4 years: HR=0.88; 95% BCI: 0.62-1.23) and reduction on the effect of malaria transmission exposure was detected (i.e. neonates: HR=1.15; 95% BCI:0.95-1.36; infants: HR=1.13; 95% BCI:0.98-1.25; children 1-4 years: HR=1.04; 95% BCI:0.89-1.18). A very strong spatial correlation was also observed. These results imply complexity in the malaria transmission-mortality relation. The clarity of this relation involved more than the knowledge on the performance of interventions and control measures from a single site. The relation depends on the levels of malaria endemicity which varies between sites, therefore, within and between sub-regions in SSA analyses should be conducted to assess reproducibility and validity of findings.

Key words: Child mortality, EIR, time varying covariates, DSS, malaria, MTIMBA

5.1. Introduction

In sub-Saharan African countries about 20 percent of all deaths occurring in under-fives are generally attributed to malaria (WHO, RBM 2002; Korenromp et al. 2003; Snow et al. 2005; Rowe et al. 2006). Malaria infection is a leading cause of deaths in children but also a main confounder of other causes such as low birth weights, malnutrition and anaemia (Pelletier et al. 1995; Snow et al. 1999). Recently, a decline in child mortality has been observed in most developing countries (Rajaratnam et al. 2010a; WHO 2010a; Lozano et al. 2011b). The drop in mortality is partly associated with success in interventions and control strategies targeting malaria transmission such as ITNs and efficacious antimalarial drugs (Schellenberg et al. 2001a; Lengeler 2004; Kleinschmidt et al. 2009a; WHO 2010b). In the Rufiji DSS, change of the first-line drug for the treatment of malaria from chloroquine to sulfadoxine pyremethamine and increasing coverage of ITNs are among contributing factors of the sharp decline in mortality and malaria transmission (Shabani et al. 2010; Kigadye et al. 2011). Thus, all-cause mortality in under-fives is an essential indicator of malaria control programs (RBM 2000; de Savigny and Binka 2004). Progress towards malaria eradication as the long term vision of RBM partnership, requires accurate knowledge of the transmission-mortality relation (Snow and Marsh 1995b; Trape and Rogier 1996; Smith et al. 2001; Ndugwa et al. 2008).

Lack of vital registration in developing countries, unreliable information on specific causes of deaths and problems related to malaria diagnosis, complicate the study of transmission-mortality relations (UNICEF/MICS 1995; Smith et al. 2001; Gemperli 2003; Gemperli et al. 2004; Hammer et al. 2006). Discrepant results which might be related to higher levels of indirect mortality attributable to malaria have been observed in attempts to relate transmission and mortality. In a review article, Smith et al. (2001) found an increase in infant mortality rate with increase in EIR in Africa. However, Gemperli (2003) linked the DHS and MARA databases to assess the effect of malaria transmission on mortality, and found no clear relationship. In a study conducted in Western Kenya, no difference in mortality rates could be observed between villages with and without ITNs intervention (Lindblade et al. 2004). These conclusions are based on reviews and/or analyses of aggregated data from studies conducted at different times (periods), regions, and designs, which might be not directly comparable. Many of these studies were not designed to assess the mortality attributed to malaria (Omumbo et al. 2004).

In malaria endemic areas, mortality is influenced not only by transmission but also factors related to poverty, control interventions and health systems performance (Snow et al. 2001; Wagstaff 2002; Wagstaff and Watanabe 2002; Müller et al. 2008; Ndugwa and Zulu 2008) and it is challenging to take these factors in to account.

The MTIMBA project initiated by the INDEPTH Network (INDEPTH Network 2002) was designed specifically to assess the malaria transmission-mortality relation. Integrated within the Demographic Surveillance Systems (DSS) which routinely monitor mortality, causes of death, and other demographic parameters, the MTIMBA project collected biweekly entomological data at a large number of geo-referenced household locations, using standardized methodology for a period of three years (Sankoh and Binka 2005a; Ramroth et al. 2009). The MTIMBA database has the epidemiological data required to study the above relation. However, data characteristics such as spatio-temporal correlations over large number of locations and lack of appropriate statistical methodologies delayed the data analysis to date. Banerjee et al. proposed modeling of large geostatistical data to approximate a spatial process from a subset of locations (Banerjee et al. 2008). These methods have been applied to model the MTIMBA entomological data in Rufiji (Rumisha et al. under review) and estimate monthly surfaces of the EIR malaria transmission measure during the three years of the project.

In this study, the Rufiji DSS-mortality databases are linked to EIR estimates to assess the relationship between malaria transmission and all cause mortality in children less than five years. The analysis is conducted using Bayesian geostatistical and temporal regression models applied on the mortality outcome, considering EIR as predictor and adjusting for malaria control interventions. The EIR is predicted from a spatio-temporal transmission model and the prediction uncertainty is incorporated during estimation of mortality risk (Rumisha et al. under review).

5.2. Methods

5.2.1. Study Area

Rufiji District is one of the six districts of the Coastal Region in the southeast part of Tanzania, with a population size of about 182,000 inhabitants. The Rufiji DSS is located in the Rufiji District (7.47° - 8.03° south latitude and 38.62° - 39.17° east longitude). The RDSS covers an area of 1,813 km^2 with 85,000 inhabitants under surveillance (Mwageni et al. 2002). The population density is 46 people per km^2 and the average household size is about five people (Bureau of Statistics, 1994). The major causes of mortality in the RDSS include acute respiratory infections, tuberculosis, acquired immunodeficiency syndrome (AIDS), perinatal causes, and malaria (Shabani et al. 2010). Malaria is endemic and seasonal throughout the region. Higher transmission occurs during and shortly after the rains (March – June). Prevalence of malaria was 28% in 2002 and 20% in 2004 which is approximately twenty-eight percent reduction in a period of two years (Source: INDEPTH Monograph, 2002).

5.2.2. Data

i) Mortality, demographic and malaria intervention data

Child all-cause mortality data were obtained from the RDSS database for the period of the MTIMBA project (i.e. October 2001- September 2004). We extracted individual-specific information comprising dates of birth, start and exit from the study, age, sex, and vital status (1 if death occurred during the study period and within the DSS and 0 otherwise). Other household-level information such as ITN possession, SES (INDEPTH Network 2005), travel time to health facilities and altitude were taken from other sources such as district health plans, and linked to the mortality database (Table 5.1). Time at risk (person-days) contributed by each child was calculated until exit, which could be migration (outside the study area), death or end of the study. In the case where a child migrated to a different location (within the study area), time at risk was computed separately for each location. The outcome of interest is the vital status and the mortality rates were expressed per 1000 person-years (py).

ii) Malaria transmission data

The entomological data from the MTIMBA project were analyzed using Bayesian geostatistical models to obtain EIR estimates at locations (households) and months where mortality data were available. In particular, separate geostatistical and temporal logistic regression and negative binomial models were fitted to sporozoite rate and mosquito density data, respectively. Using Bayesian prediction (kriging) and environmental factors as predictors, EIR was subsequently estimated by the product of the sporozoite rate and the man biting rate (MBR) predicted from the above models at the household locations. MBR was calculated from the mosquito density estimates (Lines et al. 1991a). Details of this work are available in (Rumisha et al. under review, under reviewb).

5.2.3. Linking mortality with other databases

The mortality database included information on 27049 children from 14847 locations. Linking with entomological, socio-economic and malaria interventions databases by household locations (latitude and longitude), the final dataset included 17717 children from 8144 locations. Although 45.1% of the locations with mortality data were lost after merging the databases, the proportion of deaths remained similar (Table 5.1).

Table 5.1: Number of individuals, deaths and locations after merging mortality database with entomological, socio-economic and malaria interventions databases

Merged database	Unique locations	Individuals	Deaths counts (%)
Mortality	14847	27049	831 (3.07%)
Entomological	11631	23905	768 (3.21%)
Socio-economic	9574	20341	651 (3.20%)
Malaria interventions	8144	17717	567 (3.20%)

5.2.4. Statistical analysis

Non-spatial survival models were fitted for different groups of child age (neonatals [0-28days]; postneonatal [29days to 11months]; infants [0days-11months]; children[1] [0 days to 59months]; children[2] [29days to 59 months] and, children[3] [12months-59months] to assess differences in mortality risks between the groups and decide whether separate analyses are required for each age subgroup. Mortality rates were calculated and compared between subgroups (see Table 5.2). For the selected groups we fitted bivariate models to assess potential non- linearity in the relation between EIR and mortality by considering the following transformations of EIR: i) categorical; ii) logarithmic and iii) fractional polynomials of different orders. The Akaike Information Criterion (AIC) was used to assess the model performance and select the best model including the one assuming linearity (Akaike 1973). The best model was the one with the smallest AIC value. These analyses were carried out in STATA v10 (Stata Corp., 2007).

In addition, bivariate and multivariate time dependent survival models with spatial and independent village level random effects were fitted for selected age groups. These models were approximated by a pooled logistic regression (D'Agostino et al. 1990; Singer and Willett 1993) and included monthly temporal random effects. The spatial random effects were considered to derive from a zero-mean multivariate normal distribution (Diggle et al. 1998) with covariance matrix assuming that spatial correlation decays exponentially with distance between villages. The temporal random effects were modeled by a first order autoregressive process. Following a Bayesian formulation, appropriate prior distributions for the parameters were adopted.

The data were disaggregated by months to incorporate a time varying trend of EIR. Bayesian prediction allowed estimation of the full posterior predictive distribution of EIR parameter at each location with a measure of uncertainty (Rumisha et al. under review; Diggle et al. 1998). EIR (logarithmic transformed) was then assumed to arise from a normal distribution with a mean and variance defined by median and standard deviation of the posterior predictive posterior, respectively. This allowed taking into account the measurement errors of predicted

EIR when estimating the regression parameters. The mortality events were related with a one-month lag EIR. Other predictors considered included age, sex, SES, ITN ownership, first-line malarial drug, travel time to the health facility and altitude. All geostatistical models were fitted in OpenBUGS (Spiegelhalter et al. 1999). Formulation of the geostatistical model is given in the Appendix.

5.3. Results

5.3.1. Mortality data

The complete mortality database included 27049 children from 32 villages which were followed up during the project period. The mean follow-up time was 1.6 years with a total time at risk of 44,286 py. Of these children, 31.5% (n= 8528) entered in the course of the study (via birth or in-migration). Among those available at the beginning of the study (n=15377), 2% (n=315) were neonatal, 21% (3207) postneonatal, and 77.1% (11855) were 1-4yr (59 months). At the end of the study, a total of 831 deaths were registered. The overall under-five mortality rate for the three year period was 18.7 per 1000py. For the year 1 (Oct 01- Sept 02), 2 (Oct 02- Sept 03) and 3 (Oct 03- Sept 04) of the study, the death counts (mortality rates per 1000py) were 321 (261.5), 237 (68.9) and 273 (7.1) respectively (bivariate analysis: p-value<0.001). The mean age at death was 2.36 years (±1.44). Almost a quarter (26%, n=213) of all deaths occurred in individuals before the age of one month. The numbers and proportion of post-neonatal deaths (n=304, 37%) and child deaths (12 months -59months) (n=314, 37%) were very similar. Figure 5.1 depicts the death rates by age and (calendar) month of death. The annual time series show higher mortality during the first half of the (calendar) year (corresponding to the rainy season) compared to the second half.

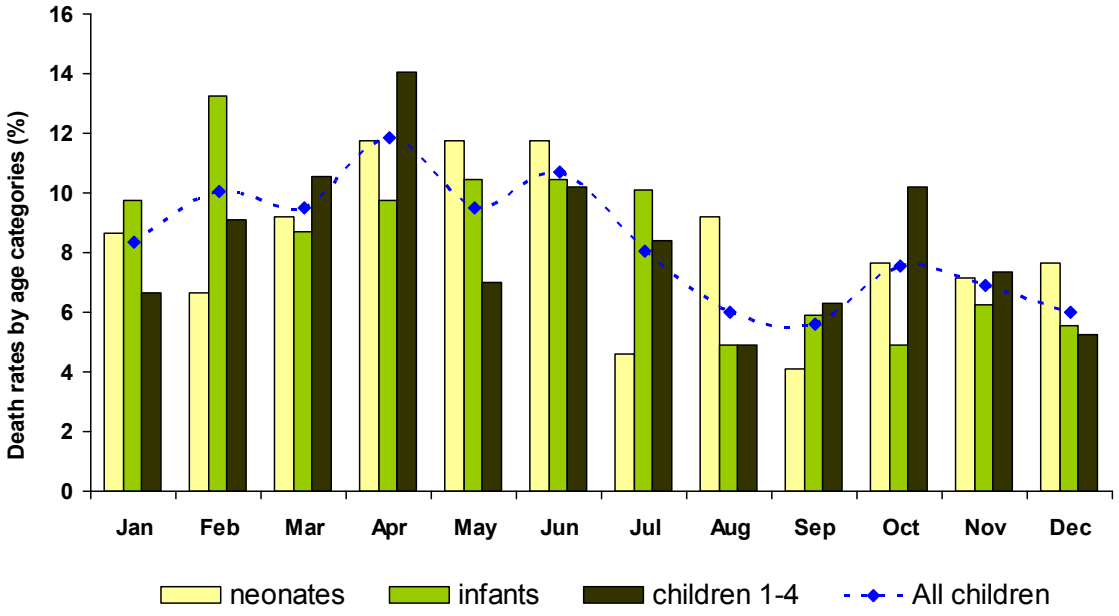


Figure 5.1: Child death rates by age categories and month of death, RDSS

In Table 5.2 the number of children, death counts and mortality rates at different age groups are shown.

Table 5.2: Descriptive statistics on mortality at different age groups of child health, Oct 2001- Sept 2004, Rufiji DSS

Group	Age range	n*	Death counts	Person years	Mortality rate (95% CI) (per 1000py)
1	Neonatal (0-28days)	9758	213	712	299.2 (265.71, 334.27)
2	Infants (0days-11months)	13228	517	9239	56.0 (51.35, 60.84)
3	Postneonatal (29days-11months)	13015	304	9236	32.9 (29.36, 36.75)
4	Children[1] (0days - 59months)	27049	831	44286	18.8 (17.52, 20.07)
5	Children[2] (29days-59months)	26836	618	44283	14.0 (12.88, 15.09)
6	Children[3] (12months-59months)	26539	314	44150	7.1 (6.35, 7.94)

***including migration within the study area**

The mortality rates declined with age. The highest rate was observed in neonates. Table 5.2 shows that mortality rates do differ significantly between the subsets of infants and children with and without neonates (groups 2 vs. 3 and 4 vs. 5) implying a need for a separate analyses for the neonates. In addition, mortality rates in children with and without postneonatal (groups 5 and 6) significantly differ. Taking all these descriptions into consideration, three separate analyses were conducted for the: i) neonates (0-28days), ii) postneonatal (29days-11months) and iii) children[3] (12months-59months).

5.3.2. Exploratory analysis

The results of the exploratory analysis carried out on the final datasets including all covariates (i.e. 17717 children and 567 deaths) are given in Table 5.3. EIR is categorized into five groups and mortality rates were estimated for each group. Bivariate analysis did not indicate strong associations of EIR with mortality, however, the mortality rates are considerably decreasing across age categories. No clear trend of mortality rates with increasing EIR intensity is observed, though this could be an effect of small death counts within categories. Low mortality is observed in individuals with ITNs compared to those without, though this was not statistically significant. The overall levels of ITNs ownership across SES quintiles (1st-5th), were 0.0%, 3.3%, 23.8%, 28.9% and 44.0% respectively. These proportions indicate a significant relationship between ITN possession and levels of income.

The overall poorest/least-poor mortality ratio was 1.49 and the ratios were 0.71, 1.33, 1.66 for neonate, infants and older children, respectively. This suggests that (except for the neonates) within this region, children living in poorest families have on average 50% higher risk of dying than those living in better-off families. Another observation (results not shown) was higher mortality rate marked in the households with many members (>10) than those with fewer individuals (≤5) (bivariate analysis: p-value <0.001). Differences in the mortality by family size could be highly influenced by the variation in the socioeconomic status of the families which might also alter the status of ITN possession. As the focus of this work was not to understand in detail the effect of these interactions, we opt to include only ITN in our models. Travel time to health facilities and altitude were not significant at the bivariate analysis hence not included in final models.

Table 5.3: Mortality rate according to ITN possession, SES and EIR levels (natural scale) in the Rufiji DSS

	Neonates			Infants			Children 1-4			All		
Variable	**n**	**deaths**	**MR (95%CI)**	**n**	**deaths**	**MR (95%CI)**	**n**	**deaths**	**MR (95%CI)**	**n**	**deaths**	**MR (95%CI)**
EIR												
0	4946	101	291.9 (244.5,342.9)	7852	132	27.1 (22.7,32)	14486	159	7.8 (6.6,9.1)	17455	392	15.1 (13.6,16.6)*
>0.0-1	809	18	367.3 (234.2,517.1)	2897	17	11.6 (6.8,18.5)*	7525	26	2.6 (1.7,3.8)	9225	61	4.6 (3.5,5.9)
>1-10	1323	20	238.1 (151.9,343.5)	3842	44	22.0 (16,29.4)	8883	27	2.4 (1.6,3.5)	10768	91	6.2 (5,7.6)
>10-100	442	8	285.7 (132.2,486.7)	1280	13	20.2 (10.8,34.3)	3243	14	4.1 (2.3,6.9)	4024	35	8.0 (5.6,11.1)
>100	34	2	1000 (158.1,1032.4)	118	2	35.7 (4.4,123.1)	327	0	0.0 (0,13.8)	417	4	11.5 (3.1,29.3)
ITN possesion												
No	5145	128	344.1 (295.9,394.8)	6801	177	40.1 (34.5,46.3)	12195	183	10.6 (9.1,12.2)	14704	488	22.1 (20.2,24.1)
Yes	1279	21	225.8 (145.5,324.2)	1731	31	27.9 (19.1,39.4)	3169	43	9.4 (6.8,12.7)	3775	95	16.5 (13.3,20.1)*
SES												
Poorest	1156	21	250.0 (161.9,356.4)	1560	44	44.0 (32.2,58.7)	2783	49	12.5 (9.3,16.5)	3345	114	22.8 (18.8,27.3)
Very poor	1351	37	377.6 (281.6,481.2)	1783	47	39.9 (29.5,52.7)	3237	57	12.2 (9.3,15.8)	3875	141	23.8 (20.1,28)
Poor	1514	42	381.8 (290.8,479.3)	1963	47	36.3 (26.8,48)	3496	45	9.0 (6.5,12.0)	4207	134	20.9 (17.5,24.7)
Less poor	1376	23	230.0 (151.7,324.9)	1841	41	34.9 (25.2,47)	3299	48	10.3 (7.6,13.6)	3971	112	18.9 (15.6,22.7)
Least poor	1027	26	351.4 (243.9,471.1)	1385	29	33.0 (22.2,47)	2549	27	7.5 (4.9,10.8)*	3081	82	17.9 (14.3,22.2)

*** significantly associated with mortality (bivariate analysis)**

5.3.3. Model-based results

The natural logarithmic transformation of EIR provided the lowest AIC value and was used for analysis (results not shown). Results of parameters estimated from multivariate spatial-temporal models for all groups, i.e. neonates, infants and children are described in Table 5.4. These include hazard ratio (HR) for predictors, spatial and temporal parameters.

Table 5.4: Parameter estimates obtained from Bayesian spatial-temporal models on neonates, infants and older children survival in the Rufiji DSS

	Neonates (0-28days)	Infants (1-11months)	Children (12-59 months)
Variable	**HR (95% BCI)**	**HR (95% BCI)**	**HR (95% BCI)**
Age	**0.79 (0.77,0.82)†**	**0.92 (0.88,0.96)‡**	**0.97 (0.96,0.98)‡**
ITN use	0.65 (0.39,1.05)	0.72 (0.48,1.07)	0.88 (0.62,1.23)
EIR (log scale)	1.15 (0.95,1.36)	1.13 (0.98,1.25)	1.04 (0.89,1.18)
Other parameters			
Spatial range (in km)	56.32 (16.12,82.15)	55.81 (17.2,82.08)	54.62 (15.68,82.06)
Spatial variance	0.28 (0.13,0.74)	0.29 (0.13,0.80)	0.30 (0.13,0.83)
Temporal variance	0.22 (0.11,0.57)	0.23 (0.11,0.52)	0.26 (0.12,0.70)
Non-spatial variance	0.22 (0.11,0.46)	0.21 (0.11,0.46)	0.20 (0.10,0.44)
Autocorrelation	0.32 (-0.67,0.98)	0.45 (-0.51,0.94)	0.99 (-0.09,1.00)

†Age in days; ‡ Age in months

In all categories, age is negatively related with the odds of dying and this is more prominent for the neonates (HR=0.79, 95% CI: 0.77, 0.82). No significant association was obtained between mortality and ITN possession or malaria transmission intensity. Nevertheless, comparing very young children and the older ones, a decrease in the protective effect of ITN (i.e. 35%, 28% and 12% for neonates, infants and older children, respectively) and, on the odds of mortality with levels of malaria transmission intensity were indicated (i.e. 15%, 13% and 4% for neonates, infants and older children, respectively) (Table 5.4). Similarly, pooled analysis (combining all age categories) did not indicate significant effect of neither EIR nor ITNs on mortality. The spatial range is similar in all groups and showed a strong correlation with a wide interval (almost the furthest distance between the two ends of the RDSS).

5.4. Discussion

Recent health statistics report reductions in mortality rates in some areas in the SSA region (You et al. 2010; WHO 2010b), and a drop in malaria infection is among the main factors that have been linked with this decline in the mortality. However, the relationship between child mortality and malaria transmission is not well understood (Snow and Marsh 1995b; Smith et al. 2001; Lengeler 2004). In this study we assessed the relation between malaria transmission and all-cause child mortality by linking the Rufiji DSS-mortality database and the malaria transmission database from the INDEPTH-MTIMBA project. This project is among the few initiatives aiming to understand the longitudinal effect of intervening malaria transmission on mortality in children and adults in different malaria endemic areas in SSA (INDEPTH Network 2002; de Savigny and Binka 2004). The intensity of malaria is measured by the EIR predicted at households which are routinely monitored for vital events hence create opportunity to precise estimation of the exposure and quantification of the relationship. Separate analyses were conducted for neonatal, infants and older children. The uncertainty of the EIR estimates is incorporated by including the measurement error of the predicted EIR. We adjusted for age, the effect malaria-related control strategies and took into account the spatial-temporal correlations.

Our study did not observed a significant relationship between malaria transmission and child mortality at any stage of childhood. Within the DSS areas, a number of malaria and non-malaria related interventions are routinely and effectively implemented. There could be therefore an optimal effect of the malaria interventions which reduces the malaria burden and its aftermath significantly (de Savigny and Binka 2004). Independently, the intensity of malaria transmission might play a role in the trend of mortality, however, the effect is minimized when other factors are put into consideration since its direct consequence which is the malaria infection is not significant. Therefore, it could be difficult to capture the actual transmission-mortality association within these settings (Greenwood et al. 1987; Ndugwa et al. 2008). However, there was a clear downwards trend of the effect of transmission with age which may be an effect of the cumulative malaria exposure (Doolan et al. 2009). It has been reported that high cumulative exposure reduces the risk of infection especially in older children (Henning et al. 2004; Maire et al. 2006).

A reduction on the protective efficacy of ITNs with age was indicated despite that the proportions of ITN possession (utilization) were very similar for the three subgroups. This observation might be associated with the acquisition on malaria immunity which is believed to increase with age or behavioral change of older children (Bejon et al. 2009). It is expected that ITNs reduce the exposure of an individual to mosquito hence reduce the chance for malaria transmission. Therefore, at early stages of life ITNs are beneficial as might lead to less maternal malaria and protect children with low (or no) immunity. With time the children build up the immunity and given that the malaria infection is significantly low, the effect of ITNs on their death risks becomes redundant which support the argument that other factors than malaria drives the mortality in these children (Alba 2010). However, a three fold reduction of households with high malaria transmission had been reported in this region in the area during that period (Rumisha et al. under review). Higher incidence of malaria is expected in older children in areas with intermediate transmission (Snow et al. 1997; Ross et al. 2006; Gardella et al. 2008). Monitoring of malaria incidence in young and older children over a period of time is required to evaluate if the reduction cause a substantial malaria risk to non-exposed individuals.

A significant drop in mortality was observed in the second and third year of the study period. Factors related to improvement in the health services, access to care and food security could explain the decline (Black et al. 2010; You et al. 2010). Our study also report less mortality for the children from the least poor families. Poverty leads to poor access to care and more exposure to diseases resulting into higher risk of death. Higher mortality risk presented for the households with more members could be an interaction between the family size and the income which regulate family expenditures (Lanjouw and Ravallion 1995; Mahfouz et al. 2009).

The spatial correlation was estimated at village level and the spatial range of more than 50kms was obtained. Factors related to spatial differences in child mortality include bio-demographic factors such as maternal age, place of birth and birth order (Kazembe et al. 2007; Sartorius et al. 2011). These risk factors are not expected to differ much within the DSS area which demonstrates wide dependency on occurrence of vital events. A significant amount of data could not be used for analysis due to missing coordinates which could cause bias in the output of our final analysis especially the linkage between malaria transmission and child mortality. However, bivariate models fitted at each stage assessing the transmission-mortality relationship reported similar results.

This study relates vital events to the closest measure of malaria exposure than previous approaches (Smith et al. 2001; Gemperli 2003; Lim et al. 2011), however few limitations accompanied the analysis. First, spatial effect was determined at the village level rather than location based. This may result to poor capturing of individuals' spatial variability and cause uncertainty in estimation of model parameters including significance level. Secondly, due to lack of cause-specific mortality data, we could not estimate the effect of transmission on direct malaria mortality. Verbal autopsy are used to ascertain causes of death in the DSS (Chandramohan et al. 2005). The method had been criticized as it tends to magnify the burden of malaria infection due to poor sensitivity and specificity in distinguishing fevers that caused by malaria and those which are not, especially in regions where transmission intensity has been reduced (Deressa et al. 2007; Dhingra et al. 2010; Mpimbaza et al. 2011). However, in endemic area malaria specific and all-cause mortality are highly related (Adjuik et al. 2006; Ndugwa et al. 2008) hence using all-cause mortality should be sufficient. However, this analysis will be performed in the next step of analysis of MTIMBA data.

Our analysis used the most comprehensive entomological database which has been linked with vital events to assess the site-specific relationship between malaria transmission and child mortality. The relation depends on the levels of endemicity which varies considerably from site to site. It is therefore difficult to generalize conclusion drawn from these results as they are valid for areas with comparable levels of transmission, coverage of intervention and control programs. However, the INDEPTH-MTIMBA project collected data in several DSS in SSA (Kasasa et al. in preparation). Our future works involve conducting multi-site comparison of the transmission-mortality relationship using mortality data from all sites and assessment of other measures of transmission than EIR. In the pooled analysis, child-specific cumulative exposure to malaria since birth which differentiate the degree of protection against malaria among children, (Baird et al. 1991; Snow et al. 1998b; Maire et al. 2006; Mayor et al. 2007; Ross et al. 2008; Doolan et al. 2009) will be calculated and evaluated on how it modifies the relationship. The meta analysis involving all MTIMBA sites should provide concrete evidence on the effect of interventions to the all- and cause-specific- mortality. Efforts to obtain information on cause-specific mortality are ongoing in most of the DSS sites.

Acknowledgement: This work was supported by the Swiss National Science Foundation (Nr 325200-118379/1). We thank the field and data staff of the Rufiji DSS for the continuing effort of collection of vital information and the support from the INDEPTH Network.

Competing interests: The authors declare that they have no competing interests.

Appendix: Geostatistical model specification

Let Y_{ijt} be the death status of a child i , $i = 1,2,3...n$ from a village j , $j = 1,2,3...v$ at a calendar month $t = 1,...,12$. Y_{ijt} is assumed to follow a Bernoulli distribution with a probability p_{ij} , i.e. $Y_{ijt} \sim Be(p_{ij})$ where p_{ij} is the probability of death. Let $\mathbf{X}_{ijt} = (X_{ijt}^{(1)}, X_{ijt}^{(2)}, X_{ijt}^{(3)})$ be a vector of covariates associated with Y_{ijt} which are age, ITN, and EIR, respectively. We modeled the relationship between p_{ij} and the covariates $\mathbf{X}_{ijt}$, the village-specific spatial U_j and non-spatial ϕ_j , and month-specific temporal $\varepsilon_t = (e_1, e_2, ..., e_t)$ random effects using a logit link via the equation $\text{logit}(p_{ij}) = \mathbf{X}^T\boldsymbol{\beta} + U_j + e_t + \phi_j$, where $\boldsymbol{\beta}$ is the vector of regression coefficients. Taking into account the measurement error of the EIR during estimation of the coefficient, the EIR (on logarithmic scale) was sampled from a normal distribution i.e. $X_{ijt}^{(3)} \sim \text{N}\left(\mu_x, \sigma_x^2\right)$ where μ_x and σ_x^2 are the mean and standard deviation of posterior predictive distribution of EIR at location i , $i = 1,2,3...n$ (a child was followed depending on the location), respectively.

The Bayesian model formulation requires specification of prior distributions for all unknown parameters. For the regression coefficients, $\boldsymbol{\beta}$ a non-informative normal prior is adopted, i.e. $\boldsymbol{\beta} \sim \text{N}\left(0, 10^3\right)$. The U_j's , i.e. $\boldsymbol{U} = \left(U(v_1), ...U(v_v)\right)^T$ are assumed to derive from a multivariate normal distribution with a covariance matrix Σ_{vxv} , i.e. $\mathbf{U} \sim \text{MVN}(0, \Sigma_{vxv})$ which defines the Gaussian spatial process. The Σ is a matrix with elements Σ_{ij} and quantify the covariance $Cov(U_i, U_j)$ between pair of villages v_i and v_j respectively. We assumed an isotropic spatial process where the spatial correlation is taken to be only a function of distance between the villages. An exponential correlation structure for the covariance matrix is adopted, that is $\Sigma_{ij} = \sigma_{sp}^2 \exp(-d_{ij}\rho)$ where σ_{sp}^2 is the spatial variance, d_{ij} is the Euclidean distance between villages v_i and v_j and ρ measuring the correlation decay and also known as the effective range $(3/\rho)$ and estimates the distance where the spatial correlation is less than 5%. The decay parameter ρ assumed to follows a gamma distribution. We modeled the ε_t via a stationary autoregressive process of order one, i.e. $e_1 \sim Normal(0, \sigma_T^2/(1-\gamma^2))$ and $e_t \mid e_{1,...t-1} \sim Normal(\gamma\, e_{t-1}, \sigma_T^2), t \geq 2$ where γ is an autocorrelation parameter $|\gamma| < 1$ which adopts a bounded uniform distribution, $\gamma \sim \text{Unif}[-1,1]$ and σ_T^2 is the temporal error (Hay and Pettitt 2001). The ϕ_j's are assumed to follow a normal distribution with mean zero and a homoscedastic variance σ_e^2. Inverse gamma priors are adopted for the variance parameters σ_{sp}^2, σ_T^2 and σ_e^2. The geostatistical models were implemented in OpenBUGS and parameters were estimated using the Gibbs sampler Markov Chain Monte Carlo (MCMC) algorithm (Gelfand and Smith 1990). Two parallel chains were run with a burn-in of 10000 initial samples, and the models were run till convergence before summarizing the results for statistical inferences.

Chapter 6 Malaria transmission intensity and mortality in older children and adults in Rufiji DSS, Tanzania

Rumisha SF[1,2,4], Smith T[1,2], Masanja H[3], Abdulla S[3], and Vounatsou P[1,2‡]

[1]*Swiss Tropical and Public Health Institute, Basel, Switzerland,*
[2]*University of Basel, Basel, Switzerland,*
[3]*Ifakara Health Institute, Dar es Salaam, Tanzania*
[4]*National Institute for Medical Research, Dar es Salaam, Tanzania*

‡Corresponding author

The manuscript has been submitted to the Journal of *Parasites and Vectors*

Abstract

The INDEPTH-MTIMBA project is among few initiatives aiming to study the association between malaria transmission intensity and mortality in SSA. The project includes the largest spatio-temporal, entomological database compiled across seven DSS sites for three years and linked to the routinely monitored vital events within the DSS. Information on how heterogeneity of plasmodium inoculation influences adult mortality is unclear. Potential interactions between human immunodeficiency virus (HIV) and malaria infections may increase the risk of malaria attributed mortality in adults. This work assesses the all-cause mortality-malaria transmission relationship in the population of the Rufiji DSS using data from the MTIMBA project. Mortality data were extracted for the period of project (October 2001-September 2004) and analyzed using Bayesian Bernoulli discrete time geostatistical models accounting for temporal, village-specific spatial and non-spatial correlations while adjusting for age and ITNs possession. The analyses were carried out separately for school-age children (5-14years), adults (15-60years) and older individuals (>60years). Malaria mortality was measured using EIR which was incorporated in models as a covariate with measurement error. Individual information were partitioned and spatially linked to monthly transmission intensities. Using model estimates, age-specific excess mortality was calculated and assessed how it varies at different levels of exposure. A total of 69,053 individuals (5-104years) were involved and 1,337 deaths were recorded during the study period. In overall, men showed higher mortality as compared to women individuals, with exception of those between 15-29years. Malaria transmission exposure increased the mortality risk for school-age children by 23% (95% BCI: 6%-36%), however, this was not a case in adults. Despite this, excess mortality for school-age children was not relatively higher than that observed in children 0-4yrs. Results observed in this study suggest that preventive interventions need to be promoted more in the school-age children, especially as malaria transmission has declined in formerly highly endemic areas. Poor association between adult mortality and malaria exposure, reinforce the conclusion from other studies that interaction between malaria and HIV does not contribute to an increase in mortality caused by other immunity suppressing diseases. A reduction in mortality in people over 60years with increasing malaria exposure may be explicable by selection effects. There is a need to examine data from other regions to better understand the reproducibility of these results.

Keywords: malaria transmission, mortality, school-age, adults, Rufiji DSS, Tanzania

6.1. Introduction

The MTIMBA project was initiated by the INDEPTH Network to evaluate the association between mortality and intensity of malaria transmission (INDEPTH Network 2002). The project includes the largest spatio-temporal and longitudinal entomological database compiled across seven DSS sites in Africa (Kasasa et al. in preparation). The MTIMBA database contains biweekly entomological data collected at the household level for a period of three years and linked to the routinely monitored vital events within the DSS. Most of the DSS are located in malaria endemic areas where a number of malaria transmission and control programs are successfully implemented (INDEPTH Network 2002; de Savigny and Binka 2004). In these settings a relation between malaria-specific deaths and all-cause death is therefore expected (Smith et al. 2004; Becher et al. 2008; Ndugwa et al. 2008). Weak health systems and poor diagnosis of malaria contribute to a large number of deaths attributed to malaria (de Savigny et al. 2004; WHO 2005; Deressa et al. 2007; Mboera et al. 2007). All-cause mortality has been used to evaluate malaria control interventions (Molineaux and Gramiccia 1980; McElroy et al. 2001; Lengeler 2004; Rowe and Steketee 2007; Eisele et al. 2010), however the relation between all-cause mortality and malaria transmission is not well understood (Snow and Marsh 1995b; Smith et al. 2001, 2004; Gemperli et al. 2006a; Lim et al. 2011). That knowledge is essential to facilitate monitoring targets set by the RBM initiative and MDGs aiming to reduce significantly the incidence and mortality of malaria by 2015 (RBM, WHO 1999; RBM:GMAP 2008).

The effect of malaria transmission intensity on morbidity and mortality has been well studied and documented in children under fives (Alonso et al. 1991; D'Alessandro et al. 1995; Snow et al. 1997; Ahmad et al. 2000; Abdulla et al. 2001; Aponte et al. 2009; Bejon et al. 2009). It is postulated that intervening levels of transmission alter malaria-immunity acquisition and could increase severe malaria infection and death especially at a later stage of a child life (Bradley 1991; Snow and Marsh 1995b, 2002; Beier et al. 1999; Bousema et al. 2010b). The cumulative exposure to malaria occurring during childhood might affect the malaria-related mortality risk during adulthood. HIV infection have been reported to increase frequency of clinical malaria,

severe malaria cases (Whitworth et al. 2000; Patnaik et al. 2005) and treatment failures (Shah et al. 2006; Martin-Blondel et al. 2007). However, some studies reported no evidence in such interactions (Kalyesubula et al. 1997; Quigley et al. 2005). Potential interaction between HIV/AIDS and malaria infections may increase the risk of mortality in adult population (Diallo et al. 2004; Cohen et al. 2005; Lopez et al. 2006; Becher et al. 2008; Saleri et al. 2009). Understanding the influence of plasmodium inoculation in adult mortality is important for long time evaluation of malaria control interventions (Yamano and Jayne 2004; Hill et al. 2007; Rajaratnam et al. 2010b). This study employed the MTIMBA-INDEPTH entomology and the mortality databases from the RDSS to assess the relation between all-cause mortality in adult (>5 years old) and malaria transmission measured by the EIR. The EIR parameters were predicted at monthly intervals from a Bayesian geostatistical transmission model accounting for seasonality, spatial and temporal correlations (Rumisha et al. under review, under reviewb). The mortality analysis is conducted using Bayesian Bernoulli discrete time models with temporal and spatial random effects, and adjusted for age and malaria control measures.

6.2. Methods

6.2.1. Study site and the data

The RDSS is located in Rufiji District, Tanzania about 178 kilometers south of Dar-es-Salaam (7.47° - 8.03° south latitude and 38.62° - 39.17° east longitude). Total population under surveillance is 85,000 individuals which is about 47% of the district total inhabitants (Mwageni et al. 2002). The main economic activity in the region is farming. Most of the agricultural areas are away from households, hence some families have temporary houses in farmland which are used for up to four months of the year (Mwageni et al. 2002).

Mortality data are routinely collected at household level within the DSS. The MTIMBA data collection activities were integrated within the DSS during Oct 2001- Sept 2004. Individual vital statistics were extracted for the period of the MTIMBA study. The initial follow up date for individuals already registered under the DSS is at the onset of the MTIMBA project, however, for those entered the DSS later (via birth or in-migration) their respective starting dates were recorded. Exit dates due to out-migration, death or the last date of the MTIMBA

study were also registered. All households were geo-referenced; therefore, a follow up time of an individual was computed based on the residing location. This implies that, the time at risk was re-initiated in case an individual migrated to a different location within the DSS. Time at risk for each individual was partitioned according to calendar months and spatially linked to the EIR parameter. Other information extracted including age, sex and ITNs possession at household level. Information on net ownership was collected as part of the RDSS general asset survey in October 2000 – January 2001 which is usually conducted in two years interval (INDEPTH Network 2005). The overall ITNs coverage was about 20%. The outcome of interest is the status of death (1 if death occurred during the study period and within the DSS and 0 otherwise). Using the total time at risk obtained from all individuals and the death counts at groups of interest, we computed mortality rates which were expressed per 1000 person years (py).

6.2.2. Statistical analysis

The analysis was carried out separately for three age categories, school-age (5-14years), adults (15-60) and older individuals (>60 years). We fitted Bayesian discrete time Bernoulli survival geostatistical models accounting for temporal (at month level) and spatial (at village level) correlations (Singer and Willett 1993; Diggle et al. 1998; Sartorius et al. 2011) (see Appendix). Non spatial analyses were also performed for comparative purposes. All models included exchangeable random effects at village level. The village-specific spatial random effects were assumed to follow a Gaussian distribution with zero mean and a covariance matrix with an isotropic exponential correlation function (Diggle et al. 1998) of distance between pairs of villages. The temporal correlation was captured using monthly dependent random effects modeled by an autoregressive process of first order. Following a Bayesian modeling formulation, a vague normal prior distribution was adopted for covariates coefficients and non-informative gamma distribution for the spatial parameters. The measurement error of predicted EIR was incorporated during parameter estimation and mortality events were linked to EIR with a one month-lag. Specifically, a natural logarithmic transformation of EIR estimates, assumed to follow a normal distribution with a mean and variance defined by the parameters from a posterior predictive distribution was employed (Rumisha et al. under review). The

transformation was selected after conducting a detailed analysis assessing best way to relate malaria transmission and mortality including non- linearity. All analyses were implemented in OpenBUGS (Spiegelhalter et al. 1999). Details of the model formulation are provided in Appendix.

6.3. Results

6.3.1. Descriptive statistics

A total of 69053 individuals (53.6% females) of age between 5 to 104 years were involved in the study. By the end of the study a total of 1337 deaths were recorded. Thirty-six percent among these were school-age children (5-14years). 53% were adults (15-60years) and the remaining 11% were older individuals above sixty years. The older adults (60+) contributed to almost 58% of all registered deaths. Other descriptions of the data are provided in Table 6.1.

Table 6.1: Number of individuals (n) and death counts (D) by age, gender and status of ITN possession

Age	Female		Males		ITN(No)		ITN(Yes)		ITN(Own)
(yrs)	N (%)	D (%)	N (%)	D (%)	N (%)	D (%)	N (%)	D (%)	%
5-14	12533 (33.8)	34 (4.8)	12655 (39.5)	50 (7.9)	19996 (36.6)	70 (6.2)	5192 (35.9)	14 (6.5)	0.21
15-29	11268 (30.4)	83 (11.8)	9381 (29.3)	58 (9.1)	15946 (29.2)	113 (10.1)	4703 (32.5)	28 (13.0)	0.23
30-44	5413 (14.6)	101 (14.4)	4211 (13.1)	93 (14.7)	7401 (13.6)	151 (13.5)	2223 (15.4)	43 (20.0)	0.23
45-60	3649 (9.9)	68 (9.7)	2724 (8.5)	73 (11.5)	5092 (9.3)	110 (9.8)	1281 (8.8)	31 (14.4)	0.20
60-70	1716 (4.6)	105 (14.9)	1144 (3.6)	84 (13.2)	2406 (4.4)	166 (14.8)	454 (3.1)	23 (10.7)	0.16
71-90	2232 (6)	247 (35.1)	1759 (5.5)	237 (37.4)	3425 (6.3)	427 (38.1)	566 (3.9)	57 (26.5)	0.14
90+	216 (0.6)	65 (9.2)	152 (0.5)	39 (6.2)	307 (0.6)	85 (7.6)	61 (0.4)	19 (8.8)	0.17
Total	**37027 (100)**	**703 (100)**	**32026 (100)**	**634 (100)**	**54573 (100)**	**1122 (100)**	**14480 (100)**	**215 (100)**	**0.21**

Overall, more deaths were observed in female than male individuals, exceptionally for school-age children where the number of deaths were higher in males. About 59% of deaths in the age of 15-29 years were attributed to females. Among all death recorded, only 16% were from population reporting ownership of ITNs. However, the bivariate analysis did not indicate any difference in the mortality rate among households claiming ITN possession and those who do not, in any of the subgroups. Figure 6.1 depicts the

age-specific rates mortality rates of females and males with error bars showing the 95% CI, for the three-year period of the INDEPTH-MTIMBA project.

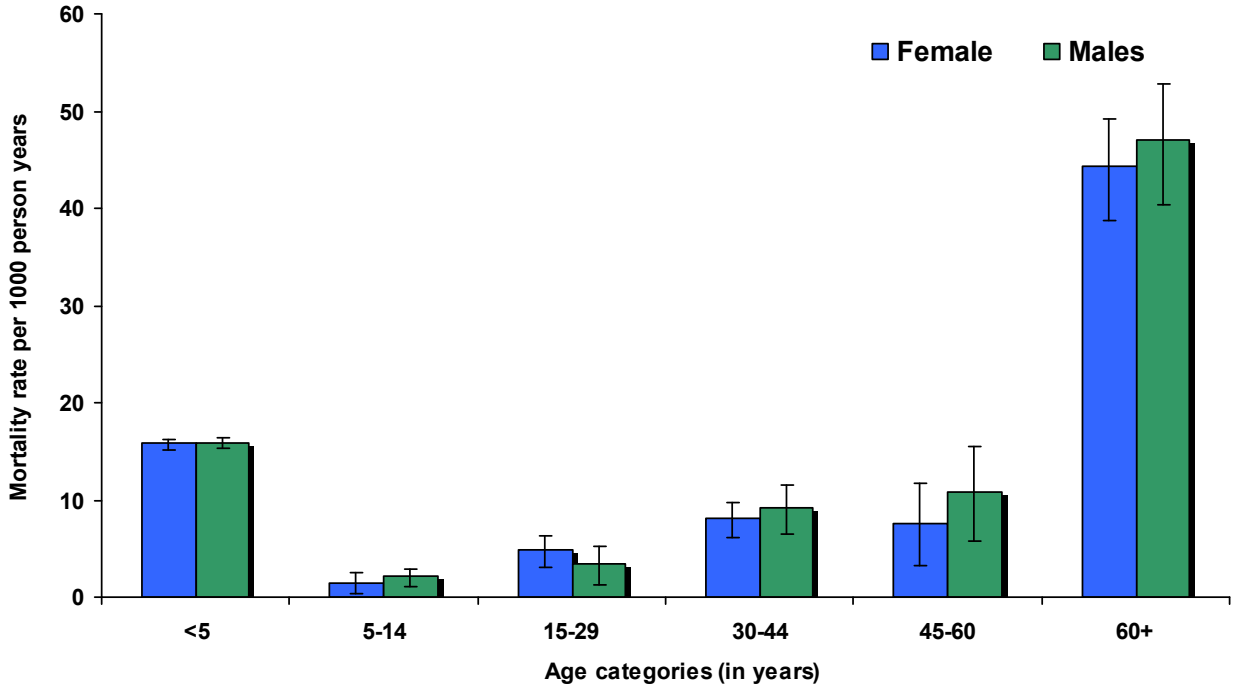


Figure 6.1: Gender and age specific estimates of mortality rates in children and adult population, RDSS, Oct 2001- Sept 2004

For both genders, the mortality rates increased consistently with age. The mortality rate lessened by approximately 90% from under fives to school age children. With an exception of those between 15-29years, the data in Figure 6.1 suggest that the mortality in men is higher compared to women (p-value<0.01, bivariate analysis). The higher mortality rates in male individuals (contrary to the death counts) could also be attributed to the imbalance population and total time of observation between male and female.

6.3.2. Model based results

Table 6.1 reports hazard ratio estimates of age, ITN and malaria transmission, and the spatial and temporal parameters obtained from the Bayesian geostatistical-temporal models fitted for the three age categories.

Table 6.1: Model estimates from Bayesian discrete time Bernoulli survival geostatistical models on school age children, adults and older individuals' survival in the Rufiji DSS, Oct 2001-Sept 2004

Variable	School-age (5-14years) HR (95% BCI)	Adults (15-60years) HR (95% BCI)	Older individuals (>60years) HR (95% BCI)
Age	0.97 (0.91,1.04)	**1.52 (1.39,1.65)**	**1.7 (1.6,1.81)**
ITN possession	0.85 (0.5,1.46)	0.93 (0.74,1.16)	0.83 (0.67,1.03)
Log EIR	**1.23 (1.06,1.36)**	0.97 (0.86,1.10)	0.95 (0.85,1.04)
Other parameters			
Spatial range (in km)	54.43 (14.29,81.94)	69.67 (24.81,82.87)	66.69 (24.3,82.59)
Spatial variance	0.33 (0.14,1.61)	0.33 (0.12,0.61)	0.23 (0.11,0.76)
Temporal variance	0.40 (0.16,1.34)	0.18 (0.09,0.41)	0.19 (0.09,0.45)
Non-spatial variance	0.26 (0.12,0.6)	0.15 (0.08,0.29)	0.13 (0.08,0.26)
Autocorrelation	0.97 (-0.74,0.99)	0.35 (-0.69,0.98)	0.87 (-0.61,0.99)

Age was positively associated with the mortality rate after the age of 15years (Adults: HR=1.52, 95% BCI: 1.39,1.65; Older individuals: HR=1.7 , 95% BCI: 1.6,1.81). Malaria transmission exposure increased the risk of dying for the school age children but had not substantial effect after an individual reached an age of 15years. No significant change in all-cause mortality rate as effect of ITNs was observed in any age category (p-value>0.05). The spatial range was highly significant suggesting that the spatial correlation existed up to a distance of 54kms. Spatial variability accounted for the 33%, 50.0% and 41.8% of the total data variability in the school age children, adults and older individuals, respectively. This was followed by temporal variability which was higher in school age data (40.4%) than in the adults (28.7%) or older individuals (34.5%).

6.3.3. Effect of EIR on mortality from birth to adulthood

In Figure 6.2 the HR of EIR on all-cause mortality estimated for different age categories are shown. Looking at the HR as the relative risk, decreasing pattern of risk was observed until the time a child reaches five years, and then a sharp peak was detected for school-age children.

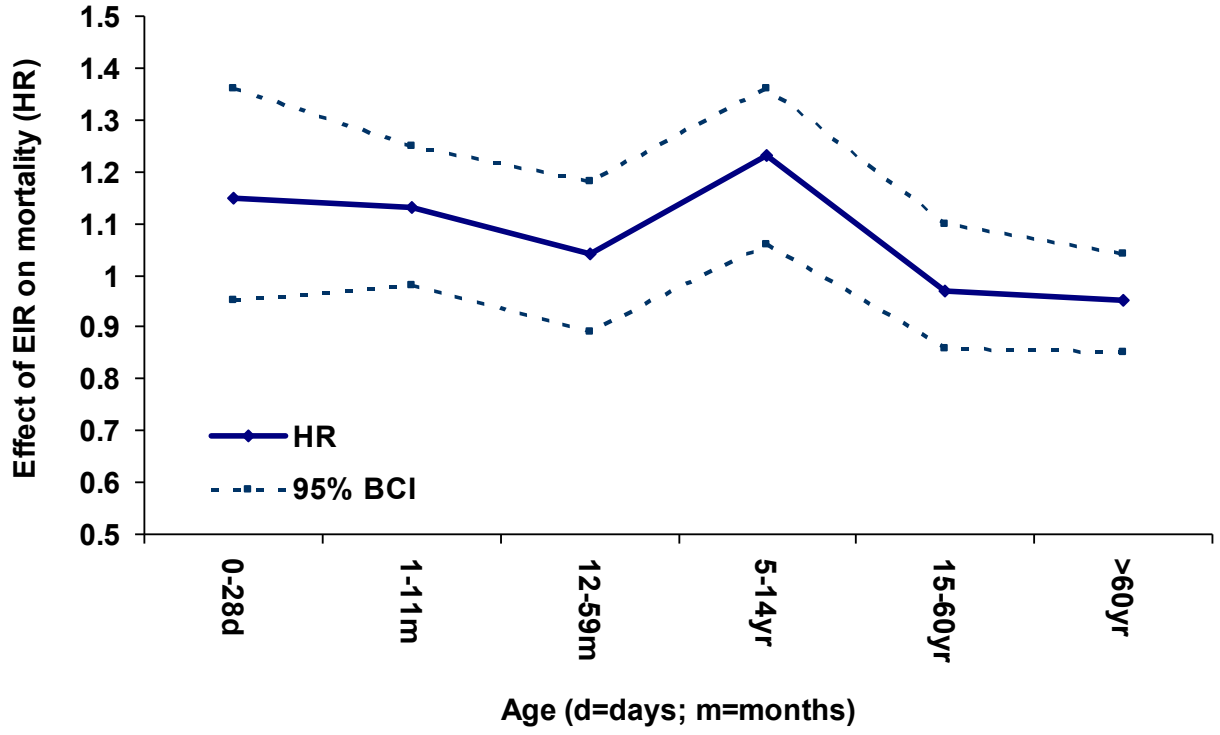


Figure 6.2: Estimates of HR of EIR on mortality (with 95% BCI) from a geostatistical model for different age categories. The under five years were added to allow clear presentation of the trend

6.3.4. Excess mortality attributed to malaria transmission

Using model coefficients, midpoint for each age category, ITN possession and EIR (infectious bites/month), probabilities of death, $prob$ were computed from the regression function over a range of EIR between 0.1 and 600 (with an interval of 0.1). The probabilities were converted to $rate$ using a function $rate = (-\ln(1 - prob))/t$ where $t = 1$ month. A Taylor series approximation was used to obtain probability at zero level of transmission. Excess mortality rate was then calculated as the difference between rate at any value of EIR and at zero, i.e. $rate_{ex} = rate(EIR > 0) - rate(EIR = 0)$ and expressed per 1000py. In Figure 6.3, age-specific excess mortality, $rate_{ex}$ are plotted against the EIR (note the difference in y-axis scale).

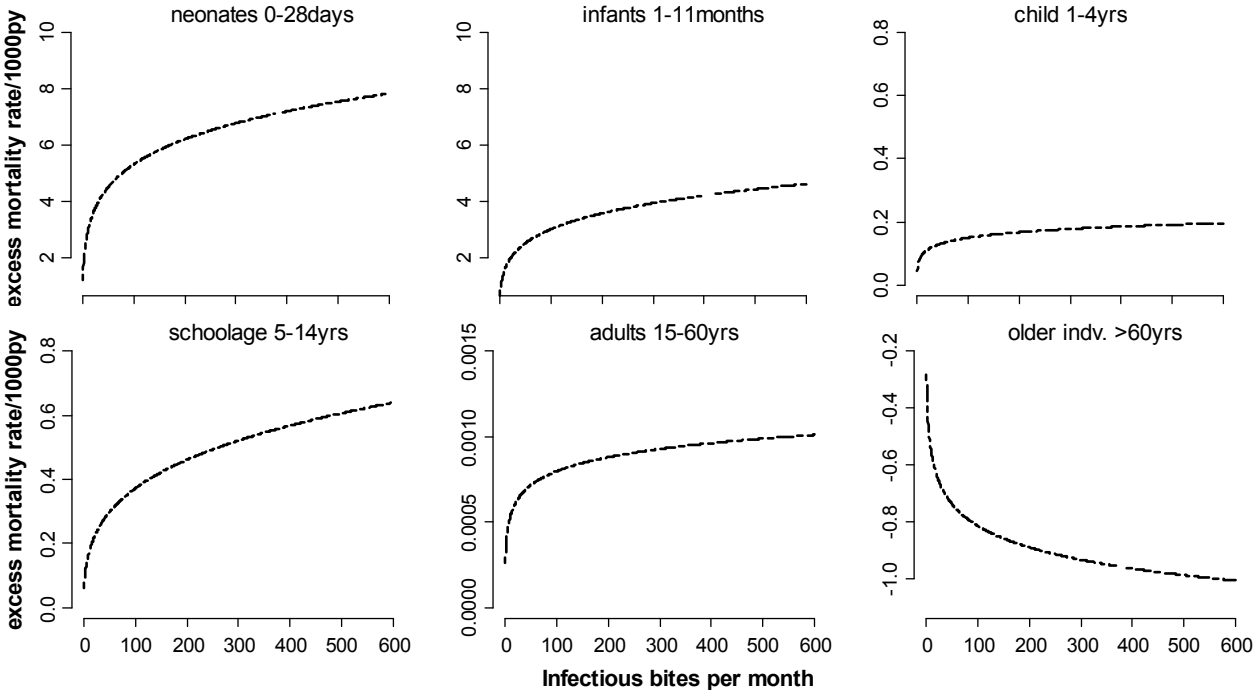


Figure 6.3: Age-specific patterns of excess mortality by transmission intensity

In children under five years of age, excess mortality shows a clear increase with malaria exposure. The highest rate is observed in neonates. School-age children demonstrate higher excess mortality than children 1-4 yrs. A protective effect is observed in older individuals and young adults showed a very low excess mortality as a consequence of transmission.

6.4. Discussion

This study linked the MTIMBA-INDEPTH and the mortality databases from Rufiji DSS to estimate the relationship between all cause mortality and malaria exposure. The analyses were implemented within the Bayesian framework, accounting for village clustering and variation between months, using as exposure measure, the EIR, predicted at the household level from a high resolution spatio-temporal model. Deaths and EIR were linked by month to allow for temporal variability in the malaria exposure during the whole study period. The prediction error of the EIR estimates was taken into account by modeling it as a covariate with measurement error. The analyses were conducted separately for school-age children, adults and older individuals. The overall age-trends in mortality were similar to those observed in other studies including Tanzania (Adjuik et al. 2006; Abdullah et al. 2007; TDHS 2010) and Zambia (Jamison et al. 2006). The mortality rate was higher in men than in women except in the 15-29 years age group.

During the study period the prevalence of malaria in the Rufiji DSS ranged between 20-35% and showed only small inter-annual fluctuations (Source: Rufiji Burden of Disease Profile, 1999; Tanzania Coastal District Health Profile, 2007); (Somi et al. 2008). In such settings, natural immunity to severe malaria is acquired early in life (Marsh and Snow 1999; Rogier 2000; Makani et al. 2003) in response to frequent exposure to the parasite (Jamison et al. 2006). In accordance with this, in younger children the mortality risk associated with a given EIR of inoculation of plasmodium exposure on child mortality reduced with age (i.e. unit increase in log EIR increased mortality risk by 15%, 13% and 4% for the neonates, infants and children between 1-4years, respectively) (Rumisha et al. in preparation). At high levels of malaria exposure there is effective clinical immunity in older children, with a consequent reduction in the effect of EIR on mortality (Abdullah et al. 2007), and field trials of ITNs have found protection in only the youngest children (Binka et al. 1998; Eisele et al. 2005). However, at the time of the study, Rufiji had much lower transmission than that experienced in either the Kisumu or Navrongo ITN trials, and so a different age-distribution as a consequence of lower exposure is plausible (Snow et al. 1997, 1998b).

In contrast to very small effect of EIR in children aged 1-4 years, we found associations between transmission exposure and mortality in children aged 5 to 14 years. In the school-age children a unit increase in log EIR was estimated to increase mortality risk by 23% (95% CI: 6%-36%)(Figure 6.2). This effect on mortality translates into a relatively small excess mortality rate (Figure 6.3), because the baseline mortality rate that is being multiplied is relatively low. However, unlike the ITN trials, the present study was not randomized, and it will be important both to see whether this result is reproduced in other sites, and to analyse malaria incidence in the whole age range from 1 month to 15 years over time to assess whether this shows patterns consistent with those in mortality (Mbogo et al. 1993; Jamison et al. 2006; Rowe and Steketee 2007; Carneiro et al. 2010).

Age dependence in human behaviour could explain why school age, rather than pre-school children seem to be more vulnerable to malaria exposure. Most malaria prevention campaigns (in particular, ITN promotion (Alaii et al. 2003)) have stressed the need to protect children under five and mothers. In children under 5, household ownership of ITNs appeared to be protective, but not in older ones, but ownership does not indicate the quality of the mosquito net and does not translate directly to consistent use (De La Cruz et al. 2006). Variations in net usage could easily bias the estimated EIR-mortality relationship, because it is likely that young children more often sleep under ITNs when mosquito densities are high. Protection measures are often relaxed as the child grows or when a new born arrives in the family (Alaii et al. 2003; Bejon et al. 2009; Eisele et al. 2010) and older children are more likely to stay out at late hours, increasing their exposure to infected mosquitoes and making net ownership irrelevant in this age group.

About twenty percent of the population possessed nets at the time of the project, but this has already increased. With massive distribution of ITNs and LLITNs ongoing in most SSA countries (WHO 2010b), higher levels of effect are expected in the future (Hawley et al. 2003; Killeen and Smith 2007), and this may accentuate any shift of mortality to school-age children, making it important both to stress the need for this age group to use nets, and to monitor changes in mortality patterns by age. Moreover, use of ITNs in this age group is in any case

important as they probably contribute a significant proportion of the parasite reservoir transmitting to mosquitoes (Killeen et al. 2007a; Atieli et al. 2011).

In contrast to the evidence we found for malaria attributable mortality in school-age children, we found no evidence of malaria mortality in adults, despite evidence that malaria and AIDS interact (Herrero et al. 2007), and the high prevalence of HIV, which alongside TB and maternal mortality remains a major cause of adult mortality in rural East African sites like Rufiji (Parise et al. 1998; Shulman 1999; Verhoeff et al. 1999; Francesconi et al. 2001). In the RDSS, HIV/AIDS prevalence ranges between 5-10% and contributed to about 17% of the total burden of the disease (Source: Tanzania Coastal District Health Profile, 2007; Global Fund Evaluation Report, 2009) increasing from 14% in 1999. Our results support other studies suggesting that the HIV-malaria interaction does not contribute substantially to increased mortality (Quigley et al. 2005; Van Geertruyden et al. 2006).

In the oldest age group, higher malaria exposure was linked to lower mortality. This result lends itself to interpretation as "selection effect", whereby individuals living at high exposure who are vulnerable to the disease because of genetic traits, die early, leaving only a selected group of resilient survivors in areas of high exposure (Jamison et al. 2006). This cannot be discounted as an explanation for age-dependence in malaria mortality since there is genetic variation (for instance in globin genes) with substantial effects on susceptibility to malaria mortality. However, it is not clear that this effect is reproducible, nor that individual-level exposures are consistent over long periods. Internal movement is common, especially of adults, who include people working in activities such as farming and fishing that can involve substantial mobility, and this needs to be considered in estimating long-term exposures. Nevertheless, selection effects on human mortality need carefully attention by assessing mortality across the whole age-range in other similar datasets.

In conclusion, this study found a strong association between malaria transmission and all-cause mortality in school-age children, suggesting that preventive interventions need to be promoted more strongly in this age group, especially as malaria transmission has declined in many formerly highly endemic areas. Adult mortality could not be attributed to malaria

exposure, reinforcing the conclusion from other studies that the interaction of malaria and HIV does not contribute to an increase in mortality caused by other immunity suppressing diseases. A reduction in mortality in people over the age of 60 with increasing malaria exposure may be explicable by selection effects. This supports the need to examine data from more sites to better understand the reproducibility of these results.

Acknowledgement: We thank field workers and data managers of RDSS for the continuing effort of collecting vital information. Support from the INDEPTH Network is highly appreciated. This work received a financial support from the Swiss National Science Foundation (Nr 325200-118379/1).

Competing interests: The authors declare that they have no competing interests.

Appendix: Geostatistical model specification

Let Y_{ijt} be the death status of an individual i , $i = 1,2,3...n$ residing at a location k , $k = 1,2,3...n$ in a village j , $j = 1,2,3...v$ at a calendar month $t = 1,...,12$ (i and k can be used interchangeably as individuals are monitored location based). Y_{ijt} follow a Bernoulli distribution with a probability p_{ij} , i.e. $Y_{ijt} \sim Be(p_{ij})$ where p_{ij} is the probability of death. Let $\mathbf{X}_{ijt} = (X_{ijt}^{(1)}, X_{ijt}^{(2)}, X_{ijt}^{(3)})$ denote a vector of covariates associated with Y_{ijt} which are age, ITN or EIR, respectively. Let U_j and $\boldsymbol{\phi}_j$ be village-specific spatial and non-spatial random effects, respectively, and $\varepsilon_t = (e_1, e_2, ..., e_t)$ introduce a month-specific temporal random effect. Using a binary logistic regression equation $\text{logit}(p_{ij}) = \mathbf{X}^T\boldsymbol{\beta} + U_j + e_t + \phi_j$ we can model the relationship between p_{ij} , the covariates and the remaining parameters, where $\boldsymbol{\beta}$ is the vector of regression coefficients. Taking into account the prediction error of the EIR estimates, EIR is modeled as a covariate with measurement error. Specifically, logarithmic transformed EIR_k value predicted at a location k are sampled from a normal distribution i.e. $X_{ijt}^{(3)} \sim \text{N}\left(\mu_x, \sigma_x^2\right)$ where μ_x and σ_x^2 are the mean and standard deviation of posterior predictive distribution of EIR at location i , $i = 1,2,3...n$, respectively.

Following a Bayesian model formulation (Diggle et al. 1998), vague normal prior distribution was adopted for the regression coefficients, $\boldsymbol{\beta}$, i.e. $\boldsymbol{\beta} \sim \text{N}\left(0, 10^3\right)$. The U_j's , i.e. $\boldsymbol{U} = \left(U(v_1), ...U(v_v)\right)^T$ are assumed to follow a zero-mean multivariate normal distribution with a covariance matrix Σ_{vxv} , i.e. $\mathbf{U} \sim \text{MVN}(0, \Sigma_{vxv})$ which defines the Gaussian spatial process. The Σ is a matrix with elements Σ_{ij} and quantify the covariance $Cov(U_i, U_j)$ between pair of villages v_i and v_j respectively. The spatial correlation is taken to be a function of distance and modeled with an exponential correlation structure for the covariance matrix, that is $\Sigma_{ij} = \sigma_{sp}^2 \exp(-d_{ij}\rho)$ where σ_{sp}^2 is the spatial variance and d_{ij} is the distance between villages v_i and v_j . ρ is a correlation decay parameter also known as the effective range ($3/\rho$) and used to estimate the distance where the spatial correlation is less than 5% and was assumed to follows a gamma distribution. The temporal effect, ε_t is modeled via a stationary autoregressive process of first order, i.e. $e_1 \sim Normal(0, \sigma_T^2/(1-\gamma^2))$ and $e_t \mid e_{1,...t-1} \sim Normal(\gamma e_{t-1}, \sigma_T^2), t \geq 2$ where γ is an autocorrelation parameter $|\gamma| < 1$ which adopts a bounded uniform distribution, $\boldsymbol{\gamma} \sim Unif[-1,1]$ and σ_T^2 is the temporal error (Hay and Pettitt 2001). The ϕ_j's are assumed to follow a normal distribution with mean zero and a constant variance σ_e^2. Inverse gamma priors are adopted for all variance parameters σ_{sp}^2, σ_T^2 and σ_e^2. Models were implemented in OpenBUGS and parameters were estimated using the Gibbs sampler Markov Chain Monte Carlo (MCMC) algorithm (Gelfand and Smith 1990). Two parallel chains were run with a burn-in of 10000 initial samples, and the models were run until convergence before results were summarized for inferences.

Chapter 7 General discussion and conclusion

7.1. A preamble

Wide spread implementation of malaria intervention and control strategies has reduced transmission intensities in much of SSA (Flaxman et al. 2010; WHO 2010b). More funding is being directed in development of new diagnostic tools, vaccines and preventive initiatives, which provide hope for achieving malaria elimination and eradication in the future (Source: www.theglobalfund.org: The Global Fund to Fight AIDS, Tuberculosis and Malaria), (WHO/UNICEF/PATH 2010; PMI 2011). However, global health statistics still show high burdens of malaria related morbidity and mortality in children and pregnant mothers (Hogan et al. 2010; WHO 2010a, 2010b). Various studies have collected data on malaria incidence, prevalence, transmission, interventions successes and trends of mortality, at national, regional or global scale aiming to track progress of targets set by RBM, WHO or MDG (Lengeler et al. 1998; Kobbe et al. 2007; Kleinschmidt et al. 2009b; Rajaratnam et al. 2010a). Specifically, efforts are done to understand the effect of control interventions in morbidity and mortality in children and adults (Lal et al. 2010; Atieli et al. 2011; Lim et al. 2011) and consequences of altering malaria transmission in child survival (D'Alessandro et al. 1995; Lengeler et al. 1995; Habluetzel et al. 1997; Smith et al. 2001; Schellenberg et al. 2001b; Fegan et al. 2007; Hamel et al. 2011). However, varying results and conclusion had been reported, resulting to lack of precise evidence on the nature of important public health parameters such as mortality-transmission relationship, correct burden of disease attributed to malaria, and significant risk factors. Aspects related to that include i) lack of adequate systems for registration of vital events, specifically statistics on causes of death, ii) weak and non-functional health systems in most of the developing countries and iii) inadequate statistical methodologies to accurate analyze available data and provide valid estimates (Sankoh and Binka 2005b). The poor understanding left open the key question of, "*how do we correctly assess if we are on track and sustain the progress that had been achieved if it is so difficult to collect and manage comprehensive data that can provide such evidence*?" This work discusses approaches to address specific statistical issues relevant to provide insights on a major public health concern on malaria transmission and its relation to mortality.

7.2. Justification for the research and goals

For the past twenty years, a number of developing countries established DSS to fill the information gaps, monitor population health and create a platform for generating fundamental information that would allow evidence-based decisions and effective policies in states with weak information systems (INDEPTH Network 2002, 2005; Sankoh and Binka 2005b). Within a DSS, rigorous monitoring of large population cohorts is done which gives opportunities to study a number of indicators and to understand small scale population dynamics (de Savigny and Binka 2004). Exploiting the existence of the DSS, MTIMBA initiative was established with the specific aim of compiling a standardized and comprehensive entomological database that will allow accurate and precise estimation of the malaria transmission-mortality relationship accounting for the effect of interventions and socio-demographic factors across multiple sites in Africa (INDEPTH Network 2002). The design, amount of data and regions involved made MTIMBA the richest entomological database ever compiled in the history of malaria research with high spatial-temporal resolutions and excellent basis for studying seasonal patterns of transmission in relation to mortality.

MTIMBA data are typical geostatistical data and require specific methodologies for accurate analysis. However, such methodologies are either not readily available or do not incorporate some data characteristics hindering direct application. A major statistical challenge in geostatistical modeling is estimation of the spatial processes when a number of surveyed locations is large. To overcome computation difficulties, recent developments focus on approximating the spatial process using a subset of survey locations (Banerjee et al. 2008; Finley et al. 2009). These strategies have been applied mainly for data with Gaussian characteristics observed over a grid, which is contrary to characteristic of many field data. There is a growing and continuing rate of acquisition of large and complex geostatistical data from DSS sites, regional and national representative surveys such as MIS, DHS (UNICEF/MICS 1995; RBM, WHO 2005) and other compiled databases such as MARA, MAP project (Guerra et al. 2007). Therefore new development, extension of existing strategies and formulation of rigorous statistical methods are required to allow wider application, to ensure efficient analysis and reliable inferences and precise prediction.

In this thesis a portion of the MTIMBA database extracted from the Rufiji DSS was employed to i) develop Bayesian geostatistical models to analyze very large and sparse geostatistical and temporal non-Gaussian data with seasonal patterns and (ii) apply the models to (a) estimate space-time heterogeneity in malaria transmission (b) assess mortality peaks between different stages of infancy age while adjusting for seasonality and (c) determine the relation between transmission intensity and risk of mortality in children and adult population after taking into account control interventions. The mortality events were not observed at similar locations as the mosquito data, hence, models developed were employed to predict malaria exposure (measured by EIR) at each household within the study area at monthly intervals at 250m by 250m spatial resolution. For the first time mortality events are linked with estimates of transmission intensity based on rigorous geostatistical analysis via models fitted on large spatial-temporal data. The uncertainty of the EIR predictions was incorporated while relating transmission with mortality.

The statistical methodologies and public health components discussed in this work provide opportunity for prompt and accurate analysis, disease mapping, identification and quantification of effect of potential risk factors, which are key elements for planning, implementation, and evaluation of control strategies but also for public health surveillance. Results of this work contributed to a better understanding on the consequences of altering malaria transmission in mortality of children and adults, and establish foundations for future research including estimation of burden of disease accounted for by the malaria transmission.

7.3. Structure

The work is split into five chapters. The first two focused on development of statistical model while addressing epidemiological questions (published in *Acta Tropica* and submitted to *Spatial and Spatio-temporal Epidemiology*, respectively); the third chapter applied the developed model to analyze and predict malaria transmission intensities over the study area and produce smooth seasonal maps of EIR (submitted to *Global Health Action*); the last two chapters utilized the predicted EIR to assess mortality-transmission relationship in children and adults (submitted to *Malaria Journal* and *Parasites and Vectors*, respectively). Detail discussions of the results from each chapter are provided within specific sections. This section provides a general discussion scrutinizing main findings, limitations, opportunities for future research and recommendations.

7.4. Statistical contribution

This thesis extended models for large geostatistical Gaussian data to large spatio-temporal non-Gaussian data while accounting for seasonality and zero-inflation (Chapters 2, 3 and 4). Strategies to analyze large geostatistical data, in particularly, approximation of the spatial processes had been discussed in detail for Gaussian outcomes (Banerjee et al. 2003; Finley et al. 2009). Zero-inflated (analogues of standard approaches) models are well developed and widely applied for count and binomial data (Lambert 1992; Welsh et al. 1996; Filipe et al. 2005). In addition, few applications of zero-inflated models formulated in Bayesian framework have been documented (Agarwal et al. 2002; Rathbun and Fei 2006; Ver Hoef and Jansen 2007; Vounatsou et al. 2009). However, none of the mentioned attempts discussed application in very large geostatistical data with seasonality and temporal effects. A significant proportion of field survey data are large, non-Gaussian, with complex characteristics, hence methods described in this thesis would be useful not only in entomological research but also in other fields generating data with similar characteristics. Introduction of techniques to approximate spatial processes for large data has created vast opportunities to handle several field data. Among available Bayesian algorithms which conduct Gaussian spatial process approximation and mapping include *BayesX* (Brezger et al. 2005) and *spBayes* (Finley et al. 2007) packages which are freely available and implemented in R software (www.r-project.org). However, these algorithms are limited to spatial analysis and do not allow accounting for temporal correlation and/or zero-inflation, which hindered direct application in the MTIMBA data. Specific modifications involving program writing are usually required to incorporate specific data characteristics to such statistical packages.

The spatial-temporal models with seasonality and temporal effects were applied to analyze MTIMBA entomological data and produce smooth temporal and seasonal maps of EIR in the Rufiji DSS (Chapters 3 and 4). The spatial process was approximated from a subset of location selected (in Chapter 3) using balance sampling algorithm, the cube method, (Deville and Tillé 2004; Chauvet and Tillé 2006). The algorithm was selected over other strategies as it allows stratification, unequal inclusion probability, balancing for more than one covariate and easy practical application. Stratification was introduced by splitting the region into a finite number

of tiles then the inclusion probabilities for location within tiles were computed while preserving the spatial configuration and outcome heterogeneity (see details in Chapter 3). Multiple subsets with different size were generated, and the criterion for best set was based on a distance measure comparing the variogram analysis of the full data and the samples (i.e. spatial parameters). The sampling exercise was performed separate for *An. funestus* and *An. gambiae* data and differences in efficiency of samples selected were observed. The number of zero observations differed between the two species with *An. funestus* having the highest proportion. In *An. gambiae* efficient samples was observed even at small subset size (include <25% of the full data) in contrary to *An. funestus* (Figure 3.3) suggesting influence of overdispersion in obtaining a "good" subset for approximation. Selection of balance sample from a dispersed data is challenging (Li et al. 2007) and deserves further research especially for geostatistic data. However, the procedure described in Chaper 3 can serve as a starting point. Other balance sampling algorithms include those described by Yates (1949) and Neyman (1934). However, most of these older methods are limited to equal probability sampling and allow only one balancing variable (Tillé 2006).

A rigorous approach of modeling EIR data is presented in Chapter 4. This is a substantial contribution, particularly in statistical analysis of entomological data. Classical approaches treat EIR data as a continuous outcome, performing logarithmic transformation to meet normality assumptions so linear regressions models can be applied (Shililu et al. 2003; Prakash et al. 2005; Githeko et al. 2006; Gemperli et al. 2006b; Kelly-Hope and McKenzie 2009). The main concern is the difficulty to obtain normality for sparse and zero-inflated data. Furthermore, EIR data is a product of sporozoite rate (binomial) and mosquito density (count) data of which each merit separate analysis. A step-by-step and easy-to-follow model description in Chapter 4 includes approximation of the spatial process, selection of covariates, model validation and calculation of EIR. Model-based estimates can significantly reduce the uncertainties of parameter estimation and improve model predictive ability.

7.5. Epidemiological contribution

Spatial mapping of disease risk is a useful tool for guiding control interventions, resource allocation and monitoring. Risk factors identification and 'hot spots' which requires more attention are additional gain from the application of mapping technologies. Using the transmission model and the environmental-transmission relation, high resolution maps (monthly and annual) for EIR were generated (Chapter 4) for the Rufiji DSS. The maps were plausible regarding the actual situation in the area which justified internal validity of the model estimates. A potential strength on these maps is that they account for small scale and species-specific spatial-temporal variation of malaria transmission within the RDSS area. Unique locations surveyed for this entomological work were 2479 which cover an area of 1813 km^2 and are widely distributed (Figure 4.4). Considering the size of the RDSS, these locations are relatively large than those used in national or regional malaria transmission mapping work such as in West and Central Africa (Gemperli et al. 2006a), Mali (Gemperli et al. 2006b), globally (Hay et al. 2009) and Kenya (Omumbo et al. 1998). Transmission maps based on climate suitability (Craig et al. 1999; Rogers et al. 2002; Tanser et al. 2003) and biological models (Martens et al. 1995) are generally good to establish link between climate and malaria transmission, however, temporal variations of key predictors (Gething et al. 2011) and practical validation are not always done.

Availability of remote sensing data at very high spatial and temporal resolution has resulted in development of accurate climate suitability models allowing efficient estimation of the outcome-exposure relationship, which is essential for prediction of health outcomes and forecasting of (malaria) epidemics, even in hard-to-reach areas. Efforts to ensure quality of public health research data, such as vital events within populations, should be complimented with routine monitoring of ecological risk factors. In this thesis environment data were obtained from MODIS and ADDS (now FEWS NET) databases. Other databases with refined and higher resolutions environment and climate data include the European Space Administration collected by Envisat satellite, National Aeronautics and Space Administration via the Aqua satellite, and Advanced Very High Resolution Radiometer. Utilization of most accurate outcome predictors can improve model efficiency significantly.

The relationship between malaria transmission and all-cause mortality in children and adults was assessed in this thesis (Chapters 5 and 6). This is not the first attempt to study this relation. Initial works include those by Smith et al (2001) and Snow and Marsh (2002) who conducted reviews to assess the transmission-mortality relation. Gemperli (2003) employed the MARA and DHS databases to study this link by fitting models adjusted for environmental and socio-demographic predictors. In specified approaches, ITNs randomized controlled trials had been designed to assess effect of net use on mortality rates in Ghana (Binka et al. 1998), Tanzania (Schellenberg et al. 2001a) and Kenya (Hawley et al. 2003; Fegan et al. 2007). Recently, Lim and colleagues conducted a multi-country analysis and observed a 23% reduction on child mortality attributed to ITN ownership (Lim et al. 2011). However, contradictory results have been reported (Lengeler 2004; Smith et al. 2004; Mathanga et al. 2005; Lim et al. 2011) making it difficult to draw a general conclusion concerning the pattern of the relationship. A major limitation in many of those efforts was inadequacy of the data, explicitly the small number of survey involved in some analyses and bias of study designs. Employing the MTIMBA database to determine the amount of mortality attributed to malaria transmission (adjusting for the uncertainty) might lead to robust and more informative estimates which allow generalization.

The age-specific trend of mortality-malaria transmission linkage obtained from this work is a potential step in clarifying and understanding of this relation. Similarity in the direction of our findings with those obtained in randomized trials, systematic analyses and quantitative models highlight areas which are well described and those that need further research to accurately establish the impact of malaria interventions (e.g. emphasize in school-age children). We accentuate the importance of evaluating how individual's history of exposure and malaria immunity modifies the impact of transmission and assessing if incidence-transmission relation agrees with that of mortality. However, this requires compilation of long term retrospective data from medical records. Strong association between malaria transmission and all-cause mortality in school-age children and decreasing risk in adults suggest consequences of malaria immunity. More research on duration of immune memory and timing of stabilization to the level that avert severe case is necessary (Filipe et al. 2007). However, examination of data from other sites is needed to evaluate reproducibility of these results.

A consistent non-significant protective effect of ITNs on mortality was detected in all ages with a slightly higher efficacy in very young children. Net ownership might not be the best proxy for net utilization, and the net-effect might be confounded by other factors such as socio-economic status. During the time of MTIMBA project, there was no program distributing nets free of charge in Rufiji district, hence net ownership was related to household's income and wealth (Mwageni et al. 2002). The poorest families therefore benefited less with net protection (poorest/least-poor overall mortality ratio = 1.49). (Table 7.1).

Table 7.1: ITN ownership (%) by age in the Rufiji DSS, Oct 2001- Sept 2004

	Age category (years)			
Socio Economic Status	**0-4**	**5-14**	**15-60**	**60+**
Poorest	0	0.4	0.3	0.2
Very poor	3.3	4.7	4.0	4.8
Poor	23.8	19.6	17.7	22.2
Less poor	28.9	29.6	28.2	35.9
Least poor	44.0	45.7	49.8	36.9

The interaction between SES and the ITNs possession justified the exclusion of SES in our mortality analysis. In doing so, power of the model was increased allowing optimal estimation of the effect size of the malaria transmission and reduces masking (Walker 1996). It remains to be established whether SES is on a causal pathway between transmission and mortality (Agha et al. 2007; Rowe et al. 2007; Bernard et al. 2009). We aim to address this issue in the second stage of analysis of MTIMBA data thorough SES quintile stratification analyses.

7.6. Limitations and challenges

While addressing MTIMBA objectives several challenges and bottlenecks which limit the performance of analysis were encountered. This section discussed some of these issues, underlined potential gaps and provides insights on possible considerations for future works.

7.6.1 Awareness and capacity in statistical data analysis

Timing between data acquisition, processing and feedback communication is an important aspects of research. Substantial intervals between these steps affect evidence-based decision, resource allocation and value of data. Availability of reliable data collection tools resulted acquiring and archiving of enormous surveys data in many (health) institutes. Despite the efforts invested to data collection, minimal priority is given to data analysis hence delayed optimal and quality gain of information which was the case for the MTIMBA database. The hindrances could be in the context of administrative and methodological constraints. However, most research work ignore involvement of statistical expertise hence the data are rarely gets into the hands of those who are capable to understand statistical issues involved to realize a need for detailed analysis. Without discounting substantial work done in previous analyses utilizing RDSS MTIMBA database (Abdullah et al. 2007; Kigadye et al. 2010, 2011), to date, no rigorous model-driven methods were used to address the key objective of the MTIMBA project. These underline a substantial shortage of human resources in biostatistics in many institutes in developing countries and emphasize a need for statistical capacity building through training and mentorship.

7.6.2 Epidemiological

i) Mortality trends: interventions, treatment efficacy or correct diagnosis

Marked reduction in malaria burden and mortality has been reported recently and is associated with widespread distribution of LLINs/ITNs and introduction of efficient anti-malarial drug such as artemisinin-based combination therapy (Flaxman et al. 2010). Accuracy in malaria diagnosis following use of rapid diagnostic tests (RDT) improved effective treatment of malaria patients in several regions in Africa and reduced unnecessary deaths attributed to false negatives (Hopkins et al. 2007; Msellem et al. 2009; WHO 2009; Lemma et al. 2010; D'Acremont et al. 2011). However, it is difficult to separate the effect of interventions, treatment and diagnosis on mortality trends. Clear understanding on the attributes of mortality will allow better resource investments and design effective interventions. The possibility that effects of control interventions have saturated in some regions should not be ignored and must be clarified to justify continuing implementation of specific interventions. Populations within DSSs are provided with optimal environment to access successfully a number of interventions implemented. In those settings it might be feasible to establish a distinction in mortality attributed to specific strategies. However, in a general population such assessments are far more difficult and not straightforward. Several countries in SSA including Tanzania are in the process of scaling up the use RDTs in peripheral health facilities (RBM 2009; Ishengoma et al. 2011) and wide distribution of LLINs is ongoing (WHO 2010b). These efforts should be accompanied by health system strengthening to allow thorough monitoring and evaluation of impact and consequences of interventions.

ii) Intervention and socio-demographic data in the MTIMBA database

The MTIMBA project was designed to collect comprehensive entomological data. However, the initiative overlooked routine acquisition of data on control interventions such as ITNs or IRS and information on SES at the time of the project. Integrating this information with scheduled entomological/vital registration collection rounds would allow precise estimation of their impact. During analysis of the MTIMBA-RDSS data, SES and ITN data were taken from household's asset survey which is conducted every two years (Mwageni et al. 2002). This source does not include all MTIMBA survey locations resulting to a considerable waste of data due to spatial misalignment. More critical, we assumed similar state of ITN possession and SES for the entire study period which may well not be the case and could affect model results.

iii) Malaria-specific mortality and transmission

A major concern in management of DSS mortality information is ascertaining causes of death. This is done using verbal autopsy which involves trained clinicians or computer-based algorithms (Setel et al. 2006; Lozano et al. 2011a; Riley 2011). Considering the amount of data generated in surveillance sites and time required to perform coding, this information is not always updated. Verbal autopsies are reported to be sensitive but unreliable in identifying malaria deaths (Snow et al. 1992; Dhingra et al. 2010). In malaria endemic regions, where accurate diagnosis is not practical, febrile illnesses are easily coded to malaria once no other etiology is confirmed (Abdullah et al. 2007). Additionally, malaria confounds other infections such as respiratory infections and deaths, specifically in children (Giglioli 1972; Molineaux 1985; Müller et al. 2003; Ramroth et al. 2009) contributing to over estimation of deaths attributed to malaria. Hospital-based death statistics could be used to relate malaria specific mortality and transmission intensity, however, they account for a small portion of the whole picture as many children deaths occurred outside health facilities (Rowe et al. 2007). The true incidence of malaria deaths lies between those reported in health facilities and community based deaths which in most developing countries are only available in DSS. Mechanisms to integrate such data sources are highly needed. In this thesis all-cause mortality were used to attribute the effect of malaria transmission on mortality in the population of Rufiji DSS. However, in endemic areas malaria-specific mortality rates are highly associated with all-cause mortality rate (Ndugwa et al. 2008) which justifies our approach.

7.6.3 Statistical

i) Lag analysis for environment and climate variables

The term “lag analysis” referred to a process of summarizing covariates to specific periods prior to survey dates, link the summarized data to the outcome of interest and assess which time interval explains the data better. In some instances the outcome variable has to be summarized to accord with the time series of the predictors. MTIMBA data were collected bi-weekly while the climate data were obtained in continuous scale at different temporal and spatial resolutions. A compromise aggregation time series was on monthly scale. Monthly data aggregation assumes similar temporal points between locations surveyed at the beginning of

the months and those visited towards the end. Collapsing of data might mask variation in the data lead to biased parameter estimates. Additionally, depending on the predictors-outcome relationship (e.g. temperature conditions and mosquito development), long temporal lags could result into a significant difference (Pampana 1969; Bradley 1987). Considering a wide application of remote sensing data, algorithms which calculate environmental/climatic time lags taking into account temporal and spatial variability should be considered.

Applications in this thesis observed variability in environmental lags () between species with long-term lags obtained in some cases (e.g. three months prior the collection). Apparent long-term effects of climate/environment on entomological parameters, might be influenced by similarity in climatic condition over time, however, might suggest existence of older mosquitoes and longevity. Analyses determine parity or age of mosquitoes would clarify this.

ii) Variable selection and modeling non-linearity

The choice of optimal and parsimonious models is a critical concern in statistical research and mainly involves selecting a set of covariates which explains the observed data best. For prediction models, choice of covariates is more sensitive to reduce prediction uncertainty and allow generalizability (Craig et al. 2007). In our analyses, selection of environmental and climatic variables was done by fitting non-spatial models using different combinations of variables (Gosoniu et al. 2006; Kristan et al. 2008; Riedel et al. 2010). The best set was discussed with local experts and supported by literature. Although these approaches are well justified, they do not optimize the choice of covariates as they ignore spatial and temporal dependence in the data. In a Bayesian framework, variable selection is done by obtaining the posterior probability of models derived from all possible combination of predictors. The approach gives an opportunity to assess all possible models and minimize arbitrary choice of covariates. Examples of MCMC based methods for variable selection include the stochastic search variable selection (George and McCulloch 1993) and Gibbs variable selection (Dellaportas et al. 2000, 2002). Most of these strategies can be formulated and implemented in BUGS software but are highly computationally intensive for large dataset (involving simultaneous estimation of spatial parameters) and when many (correlated) predictors are involved. The computations expense limits applications of Bayesian variable selection,

especially in disease mapping (Craig et al. 2007; Giardina et al. 2011; Gosoniu and Vounatsou 2011; Schur et al. 2011b). In practice, non-spatial selection can be used as an explorative step to identify candidate predictors for the Bayesian selection. Bootstrap simulation methods in combination with step-wise variable selection are also efficiently used to reduce the set of covariates (Harrell 2001; Austin and Tu 2004; Babyak 2004). Consideration of these less computational applications will improve model selection, convergence and computational efficiency (Nott and Leonte 2004; Eklund and Karlsson 2007; Li and Zhang 2010).

The variability in climatic factors and malaria transmission complicates the form of their relationship. Parametric models are not always able to capture complex relationships between variables, specifically non-linearity. To address this, logarithmic transformation, categorization of covariates and application of polynomial functions can be employed (Magalhães et al. 2011). Other alternative is use of semi- or nonparametric models which incorporate *spline* functions (Crainiceanu et al. 2005). Bayesian spline regression models were employed in mapping of malaria risk (Gosoniu et al. 2009) and schistosomiasis (Magalhães et al. 2011) in West Africa with good predictive performance. However, difficulties in interpreting coefficients estimated from spline regression models hinders their wide application (Marsh and Cormier 2002; Briand et al. 2004).

iii) Missing coordinates in geo-referenced data

Geostatistical modeling requires geo-reference information from all surveyed locations which is somewhat difficult to obtain in practice (Wieczorek et al. 2004). In some analyses, a vast amount of data was dropped due to lack of geographic coordinates. The geographical coordinates for each location were needed to align socio-demographic, intervention and malaria transmission databases. Coordinates obtained by averaging all locations observed from a higher level, such as village (representing cluster centroids) could be used instead to assign coordinates to locations belong to that cluster. Although that will save an enormous amount of data, has a consequence on the estimation of spatial parameters which could influence model prediction power. Sensitivity analysis assessing the magnitude of uncertainty would have been therefore necessary.

7.7. Future research and extension

This section highlights possible extensions and future analysis plans for this work.

7.7.1 Non-stationary spatial process and space-time interactions

Space and time correlations were modeled using separable models (i.e. additively effect of time and space), which assumes dependent processes. The model assumes a stationary spatial correlation over the whole time series and the temporal lags are generalized believed not to interact with space (Cressie 1993). Stationary spatial process implies that the correlation is only a function of distance between points. The assumption is critical for large region and might be feasible for small regions like DSS sites (Gosoniu et al. 2009; Vounatsou et al. 2009; Magalhães et al. 2011). However, seasonality in environmental factors might modify pattern of malaria transmission introducing effect of time and location characteristics on the spatial structure. For example, temporary water bodies (breeding sites) might appear and disappear on short time periods resulting to differences in mosquito production within close localities. Incorporation of variables which distinguish neighbouring locations such as soil types and agricultural systems might adequately account for non-stationary in small areas. However, non-stationary separable models do not necessarily capture space-time interactions. Correct formulation requires joint space-time process and introduce space-time random effects and interaction parameter (Gneiting 2002; Ma 2003; Stein et al. 2004; Stein 2005). The spatio-temporal models are highly computational for large data. The dimension of the covariance matrix increases with number of locations and time points and the inversion is more tedious (Genton 2007). Approximation of spatial process in space-time context has been described by Barnejee et al. (2008). The selection of subset shift from spatial only to space-time dimension resulting to signigicant computational cost. No practical applications utilizing these approaches have been reported.

7.7.2 Jointly modeling of malaria transmission and mortality accounting for spatial misalignment and species heterogeneity

The entomological and mortality data are collected at different locations within the DSS area. Mosquito species are also distributed differently over space. In statistic research this scenario is referred to as spatial misalignment (Banerjee and Gelfand 2002). Gemperli, (2003) formulated a Bayesian model to account misalignment in malaria risk and survival outcomes to produce a smooth mortality map for Mali. However, the spatial misalignment problem has not received enough attention in statistical research. In this thesis transmission models were built by fitting species-specific binomial (for SR) and negative binomial (for density) models and perform Bayesian kriging of each component separately before calculating EIR (Chapter 4). Computational inaccuracy occurred at each stage of the analysis could minimized by formulating a single model which approximate the spatial process, estimate covariates-outcome relation, calculating malaria transmission and predict at mortality locations. A major challenge is the computational load associated with it.

7.7.3 Improving EIR estimates and assessing other measures of transmission

EIR was used to measure the intensity of malaria transmission. This measure is widely applied and often used to assess effectiveness of control interventions such as ITN and IRS (Hii et al. 1993; Curtis et al. 1998; Maxwell et al. 2003; Lindblade et al. 2004). However, changes in environmental conditions and control intervention alter vector survival, making age and survival of the mosquito important parameters to consider when assessing variation in transmission and choosing a control strategy (Cook et al. 2008). Age of the mosquito is usually approximated using parity rate (Forattini et al. 1993; Smith and McKenzie 2004). Specifically, EIR can be estimated at different parity levels to determine any variation in transmission intensities between young and older mosquitoes. Following the change in malaria treatment policy in Tanzania, Huho et al (in preparation) observed that at the time of MTIMBA project, using the RDSS data, EIR levels and mosquito survival rate increased after introduction of combination therapy of Artesunate co-administered with sulfadoxine-pyrimethamine as compared to the period where Sulfadoxine-pyrimethamine (SP) were used. High vector survival could imply presence of adult mosquitoes in vector population.

7.7.4 Converting malaria transmission risk on mortality to disease burden

Combining within- and between-sites analyses of mortality – transmission relationship is the next step of this project. The pooled analysis will assess other malaria transmission measures than EIR and their relation with mortality in order to assess heterogeneity and similarities on the effect. The critical concern is the reproducibility and consistency of the impact of transmission on mortality between sites in SSA. Other components include determining the effect of malaria immunity and cumulative plasmodium exposure on the strength of infection and age-specific patterns of transmission on all-cause mortality as compared to that of malaria-specific mortality. Such understanding might allow translating the potential EIR risk to actual malaria burden at different age groups. Principally, quantification of deaths resulted from the *plasmodium* exposure. Harmonizing within- and between-site parameters a more refined model to relate different measures of transmission with mortality could be developed, following examples such as the Garki model (Dietz et al. 1974), Pull and Grab model (Pull and Grab 1974), the Lives Saved Tool (LiST) model (www.rbm.who.int) and a model proposed by Ross et al. (2006).

7.7.5 From DSS to vital registration system

Continuing establishment of new DSS sites in many regions in SSA and Asia has resulted to provision of key information to guide effective and evidence-based health policies decisions (Sankoh et al. 2003; Kamugisha et al. 2011). DSS are debated not to generate representative population parameters, however, DSS are sentinel vital registration system towards complete vital registration system (Figure 7.1) and should not be considered for national figures (Setel et al. 2005; WHO 2008b). Introduction of sample vital registration (SVR) and sample registration systems (SRS) in countries like India (Padmanabha 1982) and China (Lopez et al. 2006) have led to availability of national statistics on important public health indicators such as child and maternal mortality. Utilization of Sample Vital Statistics with Verbal Autopsy (SAVVY) methodologies in collection vital event has also shown substantial improvement (Mudenda et al. 2011) and should be a way forward for countries with poor systems. Following that, Tanzania is now in a process of establishing the Sentinel Panel District (SPD) which has a

complementary SAVVY with facility based information to generate national representative statistics (http://www.ihi.or.tz).

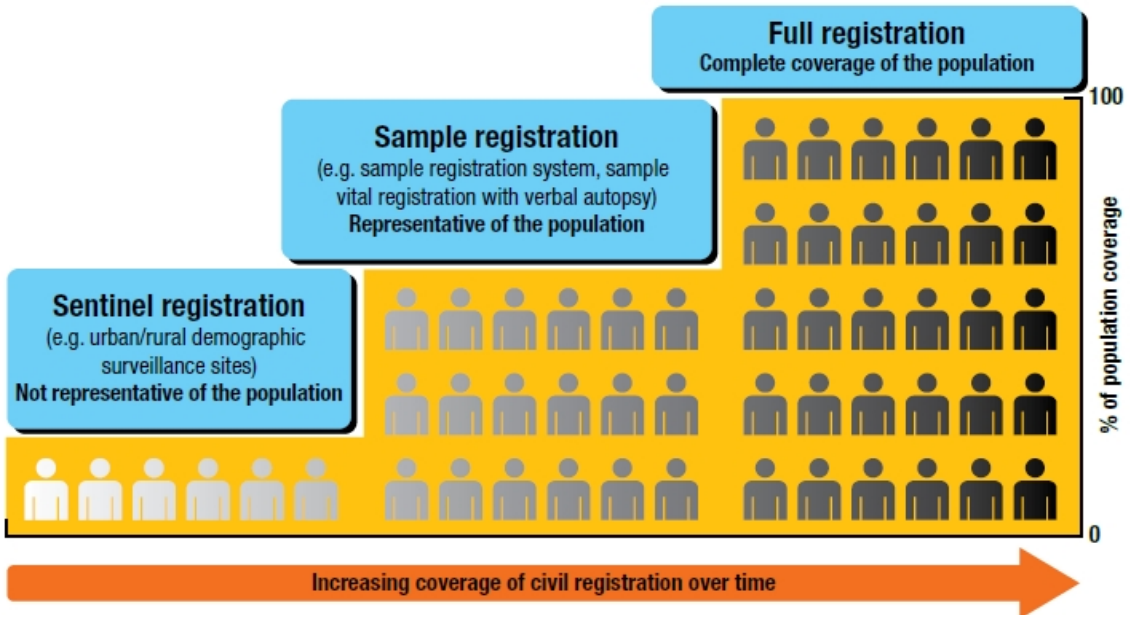


Figure 7.1: Age-specific patterns of excess mortality by transmission intensity (Source: WHO/Health Metrics Network)

However, implementation of such large survey requires integrated approaches and state commitment. Lessons learned in establishment and implementation of DSSs should be considered to ensure effective operation of SRS. Components such as identification of financial sources, evaluation, analysis plans, measurable indicators and potential stakeholders should not be ignoired. Possibilities to employ existing algorithms for selecting subsets of locations to select representative samples for these surveys should be explored.

7.8. Implication and concluding remark

The results of this thesis created a baseline for a better insight on the relation between malaria transmission and mortality which is essential for understanding the consequences of intervention strategies applied today. The clarity should steer targets set by RBM-WHO, Global funds and MDG and allow better evaluation of impact of interventions. Analyzing data from a single DSS site, a contextual modeling approaches to relate transmission and mortality in all age were established. The approaches can be employed to compare analyses within and between other sites in SSA to assess how consistent and reproducible the relation is.

References

Abdulla, S., Schellenberg, J. A., Nathan, R., Mukasa, O., Marchant, T., Smith, T., Tanner, M. and Lengeler, C. (2001). Impact on malaria morbidity of a programme supplying insecticide treated nets in children aged under 2 years in Tanzania: community cross sectional study. BMJ. 322(7281): 270–3.

Abdullah, S., Adazu, K., Masanja, H., Diallo, D., Hodgson, A., Ilboudo-Sanogo, E., Nhacolo, A., Owusu-Agyei, S., Thompson, R., Smith, T. and others (2007). Patterns of age-specific mortality in children in endemic areas of sub-Saharan Africa. The American journal of tropical medicine and hygiene. 77(6 Suppl): 99.

Abellana, R., Ascaso, C., Aponte, J., Saute, F., Nhalungo, D., Nhacolo, A. and Alonso, P. (2008). Spatio-seasonal modeling of the incidence rate of malaria in Mozambique. Malar. J. 7: 228.

Adja, A. M., N'goran, E. K., Koudou, B. G., Dia, I., Kengne, P., Fontenille, D. and Chandre, F. (2011). Contribution of Anopheles funestus, An. gambiae and An. nili (Diptera: Culicidae) to the perennial malaria transmission in the southern and western forest areas of Côte d'Ivoire. Ann Trop Med Parasitol. 105(1): 13–24.

Adjuik, M., Kanyomse, E., Kondayire, F., Wak, G. and Hodgson, A. (2010). Clustering of under-five mortality in the Navrongo HDSS in the Kassena-Nankana District of northern Ghana. Global Health Action. 3:

Adjuik, M., Smith, T., Clark, S., Todd, J., Garrib, A., Kinfu, Y., Kahn, K., Mola, M., Ashraf, A., Masanja, H., Adazu, K., Adazu, U., Sacarlal, J., Alam, N., Marra, A., Gbangou, A., Mwageni, E. and Binka, F. (2006). Cause-specific mortality rates in sub-Saharan Africa and Bangladesh. Bull. World Health Organ. 84(3): 181–8.

Adlaoui, E., Faraj, C., El Bouhmi, M., El Aboudi, A., Ouahabi, S., Tran, A., Fontenille, D. and El Aouad, R. (2011). Mapping Malaria Transmission Risk in Northern Morocco Using Entomological and Environmental Data. Malaria Research and Treatment. 2011: 1–9.

Afrane, Y. A., Lawson, B. W., Githeko, A. K. and Yan, G. (2005). Effects of microclimatic changes caused by land use and land cover on duration of gonotrophic cycles of Anopheles gambiae (Diptera: Culicidae) in western Kenya highlands. Journal of medical entomology. 42(6): 974–80.

Agarwal, D. K., Gelfand, A. E. and Citron-Pousty, S. (2002). Zero-inflated models with application to spatial count data. Environ Ecol Stat. 9(4): 341–55.

Agha, S., Van Rossem, R., Stallworthy, G. and Kusanthan, T. (2007). The impact of a hybrid social marketing intervention on inequities in access, ownership and use of insecticide-treated nets. Malar. J. 6: 13.

Ahmad, O. B., Lopez, A. D. and Inoue, M. (2000). The decline in child mortality: a reappraisal. Bulletin of the World Health Organization. 78: 1175–91.

Akaike (1973). Information theory and an extension of the maximum likelihood principle. 267–81.

Alaii, J. A., van den Borne, H. W., Kachur, S. P., Shelley, K., Mwenesi, H., Vulule, J. M., Hawley, W. A., Nahlen, B. L. and Phillips-Howard, P. A. (2003). Community reactions to the introduction of permethrin-treated bed nets for malaria control during a randomized controlled trial in western Kenya. Am. J. Trop. Med. Hyg. 68(4 Suppl): 128–36.

Alba, S. (2010). "An evaluation of integrated interventions to improve access to malaria treatment in Tanzania (access programme)." [Online]. Available: http://edoc.unibas.ch/1254/. (Accessed 4 Oct., 2011).

Alexander, N., Moyeed, R. and Stander, J. (2000). Spatial modelling of individual-level parasite counts using the negative binomial distribution. Biostat. 1(4): 453–63.

Alilio, M. S., Bygbjerg, I. C. and Breman, J. G. (2004). Are multilateral malaria research and control programs the most successful? Lessons from the past 100 years in Africa. Am. J. Trop. Med. Hyg. 71(2 Suppl): 268–78.

Alonso, P., Lindsay, S., Armstrong, J., de Francisco, A., Shenton, F., Greenwood, B., Conteh, M., Cham, K., Hill, A., David, P. and others (1991). The effect of insecticide-treated bed nets on mortality of Gambian children. The Lancet. 337(8756): 1499–502.

Amaratunga, C., Lopera-Mesa, T. M., Brittain, N. J., Cholera, R., Arie, T., Fujioka, H., Keefer, J. R. and Fairhurst, R. M. (2011). A Role for Fetal Hemoglobin and Maternal Immune IgG in Infant Resistance to Plasmodium falciparum Malaria. PLoS ONE. 6(4): e14798.

Amek, N., Bayoh, N., Hamel, M., Lindblade, K. A., Gimnig, J., Laserson, K. F., Slutsker, L., Smith, T. and Vounatsou, P. (2011). Spatio-temporal modeling of sparse geostatistical malaria sporozoite rate data using a zero inflated binomial model. Spatial and Spatio-temporal Epidemiology. 2(4): 283–90.

Anderson, R. M. and May, R. M. (1991). Infectious diseases of humans: dynamics and control. New York. 757.

Aponte, J. J., Schellenberg, D., Egan, A., Breckenridge, A., Carneiro, I., Critchley, J., Danquah, I., Dodoo, A., Kobbe, R. and Lell, B. (2009). Efficacy and safety of intermittent preventive treatment with sulfadoxine-pyrimethamine for malaria in African infants: a pooled analysis of six randomised, placebo-controlled trials. The Lancet. 374(9700): 1533–42.

Atangana, J., Fondjo, E., Fomena, A., Tamesse, J. L., Patchoké, S., Ndjemaï, H. N. M. and Ndong, P. A. B. (2009). Atangana et al. Seasonal variations of malaria transmission in Western Cameroon highlands: Entomological, parasitological and clinical investigations.J. Cell Anim. Biol. Journal of Cell and Animal Biology. 3(3): 033–8.

Atieli, H. E., Zhou, G., Afrane, Y., Lee, M.-C., Mwanzo, I., Githeko, A. K. and Yan, G. (2011). Insecticide-treated net (ITN) ownership, usage, and malaria transmission in the highlands of western Kenya. Parasit Vectors. 4: 113.

Atieli, H., Menya, D., Githeko, A. and Scott, T. (2009). House design modifications reduce indoor resting malaria vector densities in rice irrigation scheme area in western Kenya. Malar J. 8(1): 108.

Austin, P. C. and Tu, J. V. (2004). Bootstrap Methods for Developing Predictive Models. American Statistician. 58: 131–7.

Babyak, M. A. (2004). What you see may not be what you get: a brief, nontechnical introduction to overfitting in regression-type models. Psychosom Med. 66(3): 411–21.

Bailey, R. C., Jenike, M. R., Ellison, P. T., Bentley, G. R., Harrigan, A. M. and Peacock, N. R. (1992). The Ecology of Birth Seasonality Among Agriculturalists in Central Africa. Journal of Biosocial Science. 24(Special Issue 03): 393–412.

Baird, J. K., Jones, T. R., Danudirgo, E. W., Annis, B. A., Bangs, M. J., Basri, H., Purnomo and Masbar, S. (1991). Age-dependent acquired protection against Plasmodium falciparum in people having two years exposure to hyperendemic malaria. Am. J. Trop. Med. Hyg. 45(1): 65–76.

Banerjee, S. and Gelfand, A. (2002). Prediction, interpolation and regression for spatially misaligned data. Sankhyā: The Indian Journal of Statistics, Series A. 227–45.

Banerjee, S., Gelfand, A. E. and Carlin, B. P. (2003). Hierarchical Modeling and Analysis for Spatial Data. 1st ed. Chapman and Hall/CRC.

Banerjee, S., Gelfand, A. E., Finley, A. O. and Sang, H. (2008). Gaussian predictive process models for large spatial data sets. J R Stat Soc Series B Stat Methodol. 70(4): 825–48.

Barnes, K. I., Durrheim, D. N., Little, F., Jackson, A., Mehta, U., Allen, E., Dlamini, S. S., Tsoka, J., Bredenkamp, B., Mthembu, D. J., White, N. J. and Sharp, B. L. (2005). Effect of Artemether-Lumefantrine Policy and Improved Vector Control on Malaria Burden in KwaZulu–Natal, South Africa. PLoS Med. 2(11):

Barry, J., Crowder, M. and Diggle, P. (1997). Parametric estimation of the variogram. Department of Mathematics & Statistics, Lancaster University, UK.

Becher, H., Kynast-Wolf, G., Sié, A., Ndugwa, R., Ramroth, H., Kouyaté, B. and Müller, O. (2008). Patterns of malaria: cause-specific and all-cause mortality in a malaria-endemic area of west Africa. Am. J. Trop. Med. Hyg. 78(1): 106–13.

Becker, S. and Weng, S. (1998). Seasonal patterns of deaths in Matlab, Bangladesh. International Journal of Epidemiology. 27(5): 814–23.

Beier, J., Killeen, G. and Githure, J. (1999). Short report: entomologic inoculation rates and Plasmodium falciparum malaria prevalence in Africa. Am J Trop Med Hyg. 61(1): 109–13.

Bejon, P., Lusingu, J., Olotu, A., Leach, A., Lievens, M., Vekemans, J., Mshamu, S., Lang, T., Gould, J., Dubois, M. C. and others (2008). Efficacy of RTS, S/AS01E vaccine against malaria in children 5 to 17 months of age. New England Journal of Medicine. 359(24): 2521–32.

Bejon, P., Ogada, E., Peshu, N. and Marsh, K. (2009). Interactions between age and ITN use determine the risk of febrile malaria in children. PloS one. 4(12): e8321.

Bernard, J., Mtove, G., Mandike, R., Mtei, F., Maxwell, C. and Reyburn, H. (2009). Equity and coverage of insecticide-treated bed nets in an area of intense transmission of Plasmodium falciparum in Tanzania. Malar. J. 8: 65.

Billingsley, P. F., Charlwood, J. D. and Knols, B. G. . (2005). Rapid assessment of malaria risk using entomological techniques: taking an epidemiological snapshot. Frontis. 9(0): 51–68.

Binka, F. N., Indome, F. and Smith, T. (1998). Impact of spatial distribution of permethrin-impregnated bed nets on child mortality in rural northern Ghana. The American Journal of Tropical Medicine and Hygiene. 59(1): 80–5.

Birley, M. and Charlewood, J. (1987). Sporozoite rate and malaria prevalence. Parasitology Today. 3(8): 231–2.

Black, R. E., Cousens, S., Johnson, H. L., Lawn, J. E., Rudan, I., Bassani, D. G., Jha, P., Campbell, H., Walker, C. F., Cibulskis, R., Eisele, T., Liu, L. and Mathers, C. (2010). Global, regional, and national causes of child mortality in 2008: a systematic analysis. Lancet. 375(9730): 1969–87.

Bousema, T., Okell, L., Shekalaghe, S., Griffin, J. T., Omar, S., Sawa, P., Sutherland, C., Sauerwein, R., Ghani, A. C. and Drakeley, C. (2010a). Revisiting the circulation time of Plasmodium falciparum gametocytes: molecular detection methods to estimate the duration of gametocyte carriage and the effect of gametocytocidal drugs. Malar J. 9: 136.

Bousema, T., Roeffen, W., Meijerink, H., Mwerinde, H., Mwakalinga, S., van Gemert, G. J., van de Vegte-Bolmer, M., Mosha, F., Targett, G., Riley, E. M. and others (2010b). The Dynamics of Naturally Acquired Immune Responses to Plasmodium falciparum Sexual Stage Antigens Pfs230 & Pfs48/45 in a Low Endemic Area in Tanzania. PloS one. 5(11): e14114.

Box, G. E. P., Jenkins, G. M. and Reinsel, G. C. (2008). Time Series Analysis: Forecasting and Control. 4th ed. John Wiley & Sons, Hoboken, NJ. 784 pp.

Bradley, D. (1991). Morbidity and mortality at Pare-Taveta, Kenya and Tanzania, 1954-66: the effects of a period of malaria control. In: Feachem R, Jamison D, editors. Disease and Mortality in sub-Saharan Africa. New York: University Oxford Press. p. 248–63.

Bradley, T. (1987). Physiology of osmoregulation in mosquitoes. Annual review of entomology. 32(1): 439–62.

Brezger, A., Kneib, T. and Lang, S. (2005). BayesX: analyzing Bayesian structured additive regression models. J Stat Software. 14: 1–22.

Briand, L. C., Freimut, B. and Vollei, F. (2004). Using multiple adaptive regression splines to support decision making in code inspections. Journal of Systems and Software. 73(2): 205–17.

Briët, O. J. ., Vounatsou, P. and Amerasinghe, P. H. (2008). Malaria seasonality and rainfall seasonality in Sri Lanka are correlated in space. Geospatial Health. 2(2): 183–90.

Brogan, R. and Zhao, C. (1992). Designing a Longitudinal Study: Issues, Problems & Concerns. In: Annual Meeting of the Northeastern Educational Research Association. Annual Meeting of the Northeastern Educational Research Association.

Bryce, J., Boschi-Pinto, C., Shibuya, K. and Black, R. E. (2005). WHO estimates of the causes of death in children. Lancet. 365(9465): 1147–52.

Burkot, T. R. and Graves, P. M. (1995). The value of vector-based estimates of malaria transmission. Ann Trop Med Parasitol. 89(2): 125–34.

Byass, P., Berhane, Y., Emmelin, A., Kebede, D., Andersson, T., Högberg, U. and Wall, S. (2002). The role of demographic surveillance systems (DSS) in assessing the health of communities: An example from rural Ethiopia. Public Health. 116(3): 145–50.

Bøgh, C., Lindsay, S. W., Clarke, S. E., Dean, A., Jawara, M., Pinder, M. and Thomas, C. J. (2007). High spatial resolution mapping of malaria transmission risk in the Gambia, west Africa, using LANDSAT TM satellite imagery. Am. J. Trop. Med. Hyg. 76(5): 875–81.

CDC (2010). Malaria - Diagnosis & Treatment: Guidelines for clinicians for the treatment of malaria focusing on reporting cases and evaluation and diagnosis of cases.

Cameron, A. C. and Trivedi, P. K. (1990). Regression-based tests for overdispersion in the Poisson model* 1. Journal of Econometrics. 46(3): 347–64.

Cameron, A. C. and Trivedi, P. K. (1998). Regression analysis of count data. Cambridge University Press. 436 pp.

Carneiro, I., Roca-Feltrer, A., Griffin, J. T., Smith, L., Tanner, M., Schellenberg, J. A., Greenwood, B. and Schellenberg, D. (2010). Age-Patterns of Malaria Vary with Severity, Transmission Intensity and Seasonality in Sub-Saharan Africa: A Systematic Review and Pooled Analysis. PLoS One. 5(2):

Chandramohan, D., Soleman, N., Shibuya, K. and Porter, J. (2005). Editorial: Ethical issues in the application of verbal autopsies in mortality surveillance systems. Tropical Medicine & International Health. 10(11): 1087–9.

Charlewood, J. (1997). Vectorial capacity, species diversity and population cycles of anopheline mosquitoes (Diptera: Culicidae) from indoor light-trap collections in a house in southeastern Tanzania. African Entomology. 5(1): 93–101.

Charlwood, J. D. and Graves, P. M. (1987). The effect of permethrin-impregnated bednets on a population of Anopheles farauti in coastal Papua New Guinea. Med. Vet. Entomol. 1(3): 319–27.

Charlwood, J. D., Qassim, M., Elnsur, E. I., Donnelly, M., Petrarca, V., Billingsley, P. F., Pinto, J. and Smith, T. (2001). The impact of indoor residual spraying with malathion on malaria in refugee camps in eastern Sudan. Acta Trop. 80(1): 1–8.

Charlwood, J. D., Vij, R. and Billingsley, P. F. (2000). Dry season refugia of malaria-transmitting mosquitoes in a dry savannah zone of east Africa. Am. J. Trop. Med. Hyg. 62(6): 726–32.

Chase, J. M. and Shulman, R. S. (2009). Wetland isolation facilitates larval mosquito density through the reduction of predators. Ecol Entomol. 34(6): 741–7.

Chauvet, G. and Tillé, Y. (2006). A fast algorithm for balanced sampling. Computational Statistics. 21: 53–62.

Cheung, Y. B. (2002). Zero-inflated models for regression analysis of count data: a study of growth and development. Stat Med. 21(10): 1461–9.

Christopher, J. B., Le May, A., Lewin, S. and Ross, D. A. (2011). Thirty years after Alma-Ata: a systematic review of the impact of community health workers delivering curative interventions against malaria, pneumonia and diarrhoea on child mortality and morbidity in sub-Saharan Africa. Hum Resour Health. 9(1): 27.

Coetzee, M. and Fontenille, D. (2004). Advances in the study of Anopheles funestus, a major vector of malaria in Africa. Insect Biochemistry and Molecular Biology. 34(7): 599–605.

Cohen, C., Karstaedt, A., Frean, J., Thomas, J., Govender, N., Prentice, E., Dini, L., Galpin, J. and Crewe-Brown, H. (2005). Increased prevalence of severe malaria in HIV-infected adults in South Africa. Clin. Infect. Dis. 41(11): 1631–7.

Cook, P. E., McMeniman, C. J. and O'Neill, S. L. (2008). Modifying insect population age structure to control vector-borne disease. Transgenesis and the management of vector-borne disease. 126–40.

Craig, M. H., Sharp, B. L., Mabaso, M. L. H. and Kleinschmidt, I. (2007). Developing a spatial-statistical model and map of historical malaria prevalence in Botswana using a staged variable selection procedure. Int J Health Geogr. 6: 44.

Craig, M. H., Snow, R. W. and le Sueur, D. (1999). A climate-based distribution model of malaria transmission in sub-Saharan Africa. Parasitol. Today (Regul. Ed.). 15(3): 105–11.

Crainiceanu, C. M., Ruppert, R. and Wand, M. P. (2005). Bayesian analysis for penalised spline regression using WinBUGS. J. Stat. Softw. 14: 1–24.

Cressie, N. (1993). Statistics for Spatial Data. Revised Edition. Wiley-Interscience.

Cressie, N. and Johannesson, G. (2008). Fixed rank kriging for very large spatial data sets. Journal of the Royal Statistical Society: Series B (Statistical Methodology). 70(1): 209–26.

Cressie, N., Calder, C. A., Clark, J. S., Hoef, J. M. V. and Wikle, C. K. (2009). Accounting for uncertainty in ecological analysis: the strengths and limitations of hierarchical statistical modeling. Ecological Applications. 19(3): 553–70.

Curtis, C. F., Maxwell, C. A., Finch, R. J. and Njunwa, K. J. (1998). A comparison of use of a pyrethroid either for house spraying or for bednet treatment against malaria vectors. Trop. Med. Int. Health. 3(8): 619–31.

Davies, C. (1999). The Multilateral Initiative on Malaria: co-ordination and co-operation in international malaria research. Parassitologia. 41(1-3): 497.

De La Cruz, N., Crookston, B., Dearden, K., Gray, B., Ivins, N., Alder, S. and Davis, R. (2006). Who sleeps under bednets in Ghana? A doer/non-doer analysis of malaria prevention behaviours. Malar. J. 5: 61.

Dellaportas, P., Forster, J. J. and Ntzoufras, I. (2000). Bayesian variable selection using the Gibbs sampler. Biostatistics-Basel-. 5: 273–86.

Dellaportas, P., Forster, J. J. and Ntzoufras, I. (2002). On Bayesian model and variable selection using MCMC. Statistics and Computing. 12(1): 27–36.

Deressa, W., Fantahun, M. and Ali, A. (2007). Malaria-related mortality based on verbal autopsy in an area of low endemicity in a predominantly rural population in Ethiopia. Malaria Journal. 6: 128.

Detinova, T. S. (1962). Age-grouping Methods in Diptera of Medical Importance, with Special Reference to some Vectors of Malaria. World Health Organization, Geneva.

Deville, J.-C. and Tillé, Y. (2004). Efficient balanced sampling: The cube method. Biometrika. 91(4): 893–912.

Dhingra, N., Jha, P., Sharma, V. P., Cohen, A. A., Jotkar, R. M., Rodriguez, P. S., Bassani, D. G., Suraweera, W., Laxminarayan, R. and Peto, R. (2010). Adult and child malaria mortality in India: a nationally representative mortality survey. Lancet. 376(9754): 1768–74.

Diallo, A. H., Ki-Zerbo, G., Sawadogo, A. B. and Guiguemde, T. R. (2004). Severe malaria and HIV in adult patients in Bobo-Dioulasso, Burkina Faso. Med Trop (Mars). 64(4): 345–50.

Dietz, K., Molineaux, L. and Thomas, A. (1974). A malaria model tested in the African savannah. Bull. World Health Organ. 50(3-4): 347–57.

Diggle, P. J., Tawn, J. A. and Moyeed, R. A. (1998). Model-based geostatistics. J R Stat Soc Ser C Appl Stat. 47(3): 299–350.

Diggle, P., Moyeed, R., Rowlingson, B. and Thomson, M. (2002). Childhood Malaria in the Gambia: A Case-Study in Model-Based Geostatistics. Journal of the Royal Statistical Society. Series C (Applied Statistics). 51(4): 493–506.

Dominici, F., McDermott, A., Zeger, S. L. and Samet, J. M. (2002). On the Use of Generalized Additive Models in Time-Series Studies of Air Pollution and Health. American Journal of Epidemiology. 156(3): 193–203.

Doolan, D. L., Dobano, C. and Baird, J. K. (2009). Acquired Immunity to Malaria. Clin. Microbiol. Rev. 22(1): 13–36.

Drakeley, C. J., Carneiro, I., Reyburn, H., Malima, R., Lusingu, J. P. A., Cox, J., Theander, T. G., Nkya, W. M. M. M., Lemnge, M. M. and Riley, E. M. (2005). Altitude-dependent and -independent variations in Plasmodium falciparum prevalence in northeastern Tanzania. J. Infect. Dis. 191(10): 1589–98.

Drakeley, C., Akim, N., Sauerwein, R., Greenwood, B. and Targett, G. (2000). Estimates of the infectious reservoir of Plasmodium falciparum malaria in The Gambia and in Tanzania. Transactions of the Royal Society of tropical Medicine and Hygiene. 94(5): 472–6.

Drakeley, C., Schellenberg, D., Kihonda, J., Sousa, C. A., Arez, A. P., Lopes, D., Lines, J., Mshinda, H., Lengeler, C., Schellenberg, J. A., Tanner, M. and Alonso, P. (2003). An estimation of the entomological inoculation rate for Ifakara: a semi-urban area in a region of intense malaria transmission in Tanzania. Tropical Medicine & International Health. 8(9): 767–74.

Dutta, H. M. and Dutt, A. K. (1978). Malarial ecology: a global perspective. Social Science & Medicine. Part D: Medical Geography. 12(2): 69–84.

Dye, C. (1986). Vectorial capacity: must we measure all its components? Parasitol. Today (Regul. Ed.). 2(8): 203–9.

D'Acremont, V., Kahama-Maro, J., Swai, N., Mtasiwa, D., Genton, B. and Lengeler, C. (2011). Reduction of anti-malarial consumption after rapid diagnostic tests implementation in Dar es Salaam: a before-after and cluster randomized controlled study. Malar J. 10: 107–107.

D'Acremont, V., Lengeler, C., Mshinda, H., Mtasiwa, D., Tanner, M. and Genton, B. (2009). Time To Move from Presumptive Malaria Treatment to Laboratory-Confirmed Diagnosis and Treatment in African Children with Fever. PLoS Med. 6(1): e252.

D'Agostino, R. B., Lee, M., Belanger, A. J., Cupples, L. A., Anderson, K. and Kannel, W. B. (1990). Relation of pooled logistic regression to time dependent cox regression analysis: The framingham heart study. Statistics in Medicine. 9(12): 1501–15.

D'Alessandro, U., Olaleye, B., Langerock, P., Aikins, M., Thomson, M., Cham, M., Greenwood, B., McGuire, W., Bennett, S. and Cham, B. (1995). Mortality and morbidity from malaria in Gambian children after introduction of an impregnated bednet programme. The lancet. 345(8948): 479–83.

Ecker, M. D. and Gelfand, A. E. (1999). Bayesian Modeling and Inference for Geometrically Anisotropic Spatial Data. MATHEMATICAL GEOLOGY. 31: 67--83.

Eidsvik, J., Finley, A. O., Banerjee, S. and Rue, H. (2010). Approximate Bayesian inference for large spatial datasets using predictive process models. Technical Report.

Eijk, A. M., Ayisi, J. G., Kuile, F. O., Slutsker, L., Otieno, J. A., Misore, A. O., Odondi, J., Rosen, D. H., Kager, P. A., Steketee, R. W. and others (2004). Implementation of intermittent preventive treatment with sulphadoxine–pyrimethamine for control of malaria in pregnancy in Kisumu, western Kenya. Tropical Medicine & International Health. 9(5): 630–7.

Eijk, A., Blokland, I., Slutsker, L., Odhiambo, F., Ayisi, J., Bles, H., Rosen, D., Adazu, K. and Lindblade, K. (2005). Use of intermittent preventive treatment for malaria in pregnancy in a rural area of western Kenya with high coverage of insecticide-treated bed nets. Tropical Medicine & International Health. 10(11): 1134–40.

Eilers, P. H. C., Gampe, J., Marx, B. D. and Rau, R. (2008). Modulation models for seasonal time series and incidence tables. Statistics in Medicine. 27(17): 3430–41.

Eisele, T. P., Larsen, D. and Steketee, R. W. (2010). Protective efficacy of interventions for preventing malaria mortality in children in Plasmodium falciparum endemic areas. International Journal of Epidemiology. 39: i88–101.

Eisele, T. P., Lindblade, K. A., Wannemuehler, K. A., Gimnig, J. E., Odhiambo, F., Hawley, W. A., Ter Kuile, F. O., Phillips-Howard, P., Rosen, D. H., Nahlen, B. L., Vulule, J. M. and Slutsker, L. (2005). Effect of sustained insecticide-treated bed net use on all-cause child mortality in an area of intense perennial malaria transmission in western Kenya. Am. J. Trop. Med. Hyg. 73(1): 149–56.

Eklund, J. and Karlsson, S. (2007). Computational Efficiency in Bayesian Model and Variable Selection. Central Bank of Iceland, Economics Department.

Fegan, G. W., Noor, A. M., Akhwale, W. S., Cousens, S. and Snow, R. W. (2007). Effect of expanded insecticide-treated bednet coverage on child survival in rural Kenya: a longitudinal study. Lancet. 370(9592): 1035–9.

Fernandez-Salas, I., Rodriguez, M. H. and Roberts, D. R. (1994). Gonotrophic cycle and survivorship of Anopheles pseudopunctipennis (Diptera: Culicidae) in the Tapachula foothills of southern Mexico. Journal of medical entomology. 31(3): 340–7.

Filipe, J. A. N., Boussinesq, M., Renz, A., Collins, R. C., Vivas-Martinez, S., Grillet, M.-E., Little, M. P. and Basáñez, M.-G. (2005). Human infection patterns and heterogeneous exposure in river blindness. Proc. Natl. Acad. Sci. U.S.A. 102(42): 15265–70.

Filipe, J. A. N., Riley, E. M., Drakeley, C. J., Sutherland, C. J. and Ghani, A. C. (2007). Determination of the Processes Driving the Acquisition of Immunity to Malaria Using a Mathematical Transmission Model. PLoS Comput Biol. 3(12): e255.

Finley, A. O., Banerjee, S. and Carlin, B. P. (2007). spBayes: An R Package for Univariate and Multivariate Hierarchical Point-referenced Spatial Models. Journal of Statistical Software. 19(4): 1–24.

Finley, A. O., Sang, H., Banerjee, S. and Gelfand, A. E. (2009). Improving the performance of predictive process modeling for large datasets. Comput Stat Data Anal. 53(8): 2873–84.

Fisman, D. N. (2007). Seasonality of Infectious Diseases. Annu. Rev. Public. Health. 28(1): 127–43.

Flaxman, A. D., Fullman, N., Otten, M. W., Menon, M., Cibulskis, R. E., Ng, M., Murray, C. J. L. and Lim, S. S. (2010). Rapid Scaling Up of Insecticide-Treated Bed Net Coverage in Africa and Its Relationship with Development Assistance for Health: A Systematic Synthesis of Supply, Distribution, and Household Survey Data. PLoS Med. 7(8): e1000328.

Fontenille, D., Lochouarn, L., Diagne, N., Sokhna, C., Lemasson, J. J., Diatta, M., Konate, L., Faye, F., Rogier, C. and Trape, J. F. (1997a). High annual and seasonal variations in malaria transmission by anophelines and vector species composition in Dielmo, a holoendemic area in Senegal. Am. J. Trop. Med. Hyg. 56(3): 247–53.

Fontenille, D., Lochouarn, L., Diatta, M., Sokhna, C., Dia, I., Diagne, N., Lemasson, J.-J., Ba, K., Tall, A., Rogier, C. and Trape, J.-F. (1997b). Four years' entomological study of the transmission of seasonal malaria in Senegal and the bionomics of Anopheles gambiae and A. arabiensis. Trans R Soc Trop Med Hyg. 91(6): 647–52.

Forattini, O. P., Kakitani, I., Massad, E. and Gomes, A. C. (1993). Studies on mosquitoes (Diptera: Culicidae) and anthropic environment: 1-Parity of blood seeking Anopheles (Kerteszia) in South-Eastern Brazil. Revista de Saúde Pública. 27(1): 1–8.

Francesconi, P., Fabiani, M., Dente, M. G., Lukwiya, M., Okwey, R., Ouma, J., Ochakachon, R., Cian, F. and Declich, S. (2001). HIV, malaria parasites, and acute febrile episodes in Ugandan adults: a case-control study. AIDS. 15(18): 2445–50.

Furrer, R., Genton, M. G. and Nychka, D. (2006). Covariance tapering for interpolation of large spatial datasets. JOURNAL OF COMPUTATIONAL AND GRAPHICAL STATISTICS. 15: 502--523.

Gamble, C., Ekwaru, P. J., Garner, P. and Ter Kuile, F. O. (2007). Insecticide-treated nets for the prevention of malaria in pregnancy: a systematic review of randomised controlled trials. PLoS medicine. 4(3): e107.

Garcia, L. S. (2010). Malaria. Clin. Lab. Med. 30(1): 93–129.

Gardella, F., Assi, S., Simon, F., Bogreau, H., Eggelte, T., Ba, F., Foumane, V., Henry, M.-C., Kientega, P. T., Basco, L., Trape, J.-F., Lalou, R., Martelloni, M., Desbordes, M., Baragatti, M., Briolant, S., Almeras, L., Pradines, B., Fusai, T. and Rogier, C. (2008). Antimalarial drug use in general populations of tropical Africa. Malar. J. 7: 124.

Garrett-Jones, C. (1964). The human blood index of malaria vectors in relation to epidemiological assessment. Bull World Health Organ. 30(2): 241–61.

Gatton, M. (2010). Environmental factors and malaria transmission risk. Lancet Infect Dis. 10(1): 15–6.

Gelfand, A. E. and Smith, A. F. M. (1990). Sampling Based Approaches to Calculating Marginal Densities. Journal Amer. Stat. Assoc. 85: 398–409.

Gelman, A. and Rubin, D. B. (1992). Inference from Iterative Simulation Using Multiple Sequences. Statist. Sci. 7(4): 457–72.

Gemperli, A. (2003). Development of spatial statistical methods for modelling point-referenced spatial data in malaria epidemiology. Thesis.

Gemperli, A. and Vounatsou, P. (2003). Fitting generalized linear mixed models for pointreferenced spatial data. J Modern Appl Stat Meth. 2: 497–511.

Gemperli, A. and Vounatsou, P. (2006). Strategies for Fitting Large, Geostatistical Data in MCMC Simulation. Communications in Statistics - Simulation and Computation. 35(2): 331–45.

Gemperli, A., Sogoba, N., Fondjo, E., Mabaso, M., Bagayoko, M., Briët, O. J. T., Anderegg, D., Liebe, J., Smith, T. and Vounatsou, P. (2006a). Mapping malaria transmission in West and Central Africa. Tropical Medicine & International Health. 11(7): 1032–46.

Gemperli, A., Vounatsou, P., Kleinschmidt, I., Bagayoko, M., Lengeler, C. and Smith, T. (2004). Spatial patterns of infant mortality in Mali: the effect of malaria endemicity. Am. J. Epidemiol. 159(1): 64–72.

Gemperli, A., Vounatsou, P., Sogoba, N. and Smith, T. (2006b). Malaria mapping using transmission models: application to survey data from Mali. Am. J. Epidemiol. 163(3): 289–97.

Genton, M. G. (2007). Separable approximations of space-time covariance matrices. Environmetrics. 18(7): 681–95.

George, E. I. and McCulloch, R. E. (1993). Variable selection via Gibbs sampling. Journal of the American Statistical Association. 881–9.

Gething, P. W., Van Boeckel, T. P., Smith, D. L., Guerra, C. A., Patil, A. P., Snow, R. W. and Hay, S. I. (2011). Modelling the global constraints of temperature on transmission of Plasmodium falciparum and P. vivax. Parasit Vectors. 4: 92.

Giardina, F., Gosoniu, L., Konate, L., Diouf, M. B., Perry, R., Gaye, O., Faye, O. and Vounatsou, P. (2011). Estimating the burden of malaria in Senegal: Bayesian Zero-Inflated Binomial geostatistical modeling of the MIS 2008 data. Plos One. Submitted.

Giglioli, G. (1972). Changes in the pattern of mortality following the eradication of hyperendemic malaria from a highly susceptible community. Bull. World Health Organ. 46(2): 181–202.

Gillies, M. T. and De Meillon, B. (1968). The Anophelinae of Africa south of the Sahara (Ethiopian zoogeographical region), 2nd edition. South African Institute for Medical Research, Johannesburg. 54: 1–343.

Githeko, A. K., Ayisi, J. M., Odada, P. K., Atieli, F. K., Ndenga, B. A., Githure, J. I. and Yan, G. (2006). Topography and malaria transmission heterogeneity in western Kenya highlands: prospects for focal vector control. Malar J. 5: 107–107.

Githeko, A., Brandling-Bennett, A., Beier, M., Atieli, F., Owaga, M. and Collins, F. (1992). The reservoir of Plasmodium falciparum malaria in a holoendemic area of western Kenya* 1. Transactions of the Royal Society of Tropical Medicine and Hygiene. 86(4): 355–8.

Gneiting, T. (2002). Nonseparable, stationary covariance functions for space-time data. Journal of the American Statistical Association. 97(458): 590–600.

Gosoniu, L. (2008). Development of Bayesian geostatistical models with applications in malaria epidemiology. Thesis.

Gosoniu, L. and Vounatsou, P. (2011). Geostatistical variable selection in malaria risk mapping: an application to the Liberia Malaria Indicator Survey data. Parasites and Vectors. sumbitted.

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

Gosoniu, L., Vounatsou, P., Sogoba, N. and Smith, T. (2006). Bayesian modelling of geostatistical malaria risk data. Geospat Health. 1: 127--139.

Gosoniu, L., Vounatsou, P., Sogoba, N., Maire, N. and Smith, T. (2009). Mapping malaria risk in West Africa using a Bayesian nonparametric non-stationary model. Computational Statistics & Data Analysis. 53(9): 3358–71.

Greene, W. H. (1994). Accounting for Excess Zeros and Sample Selection in Poisson and Negative Binomial Regression Models. New York University, Leonard N. Stern School of Business, Department of Economics.

Greenwood, B. (2006). Review: Intermittent preventive treatment – a new approach to the prevention of malaria in children in areas with seasonal malaria transmission. Trop Med Int Health. 11(7): 983–91.

Greenwood, B. M., Bojang, K., Whitty, C. J. M. and Targett, G. A. T. (2005). Malaria. Lancet. 365(9469): 1487–98.

Greenwood, B. M., Fidock, D. A., Kyle, D. E., Kappe, S. H. ., Alonso, P. L., Collins, F. H. and Duffy, P. E. (2008). Malaria: progress, perils, and prospects for eradication. The Journal of clinical investigation. 118(4): 1266.

Greenwood, B. M., Greenwood, A. M., Bradley, A. K., Tulloch, S., Hayes, R. and Oldfield, F. S. (1987). Deaths in infancy and early childhood in a well-vaccinated, rural, West African population. Ann Trop Paediatr. 7(2): 91–9.

Greenwood, B. and Mutabingwa, T. (2002). Malaria in 2002. Nature. 415(6872): 670–2.

Griffin, J. T., Hollingsworth, T. D., Okell, L. C., Churcher, T. S., White, M., Hinsley, W., Bousema, T., Drakeley, C. J., Ferguson, N. M., Basáñez, M.-G. and Ghani, A. C. (2010). Reducing Plasmodium falciparum Malaria Transmission in Africa: A Model-Based Evaluation of Intervention Strategies. PLoS Med. 7(8): e1000324.

Gu, W., Regens, J. L., Beier, J. C. and Novak, R. J. (2006). Source reduction of mosquito larval habitats has unexpected consequences on malaria transmission. Proceedings of the National Academy of Sciences. 103(46): 17560–3.

Guelbeogo, W., Sagnon, N., Grushko, O., Yameogo, M., Boccolini, D., Besansky, N. and Costantini, C. (2009). Seasonal distribution of Anopheles funestus chromosomal forms from Burkina Faso. Malar J. 8(1): 239.

Guerra, C. A., Hay, S. I., Lucioparedes, L. S., Gikandi, P. W., Tatem, A. J., Noor, A. M. and Snow, R. W. (2007). Assembling a global database of malaria parasite prevalence for the Malaria Atlas Project. Malar. J. 6: 17.

Habluetzel, A., Diallo, D., Esposito, F., Lamizana, L., Pagnonl, F., Lengeler, C., Traore, C. and Cousens, S. (1997). Do insecticide-treated curtains reduce all-cause child mortality in Burkina Faso? Tropical Medicine & International Health. 2(9): 855–62.

Haddow, A. (1943). Measurements of temperature and light in artificial pools with reference to the larval habitat of Anopheles (Myzomyia) gambiae, Giles, and A. (M.) funestus, Giles. Bull Entomol Res. 34: 89–93.

Hamad, A. A., Nugud, A. E. H. D., Arnot, D. E., Giha, H. A., Abdel-Muhsin, A. M. A., Satti, G. M. H., Theander, T. G., Creasey, A. M., Babiker, H. A. and Elnaiem, D. E. A. (2002). A marked seasonality of malaria transmsission in two rural sites in eastern Sudan. Acta tropica. 83(1): 71–82.

Hamel, M. J., Adazu, K., Obor, D., Sewe, M., Vulule, J., Williamson, J. M., Slutsker, L., Feikin, D. R. and Laserson, K. F. (2011). A reversal in reductions of child mortality in western Kenya, 2003-2009. Am. J. Trop. Med. Hyg. 85(4): 597–605.

Hammer, G. P., Kouyaté, B., Ramroth, H. and Becher, H. (2006). Risk factors for childhood mortality in sub-Saharan Africa: A comparison of data from a Demographic and Health Survey and from a Demographic Surveillance System. Acta Tropica. 98(3): 212–8.

Harrell, F. E. (2001). Regression modeling strategies: with applications to linear models, logistic regression, and survival analysis. Springer Verlag.

Harrison, G. (1978). Mosquitoes, Malaria and Man: A History of the Hostilities Since 1880. John Murray, London.

Hastings, W. K. (1970). Monte Carlo Sampling Methods Using Markov Chains and Their Application. Biometrika. 57: 97–109.

Hawley, W. A., Phillips-Howard, P. A., ter Kuile, F. O., Terlouw, D. J., Vulule, J. M., Ombok, M., Nahlen, B. L., Gimnig, J. E., Kariuki, S. K., Kolczak, M. S. and Hightower, A. W. (2003). Community-wide effects of permethrin-treated bed nets on child mortality and malaria morbidity in western Kenya. Am. J. Trop. Med. Hyg. 68(4 Suppl): 121–7.

Hay, J. L. and Pettitt, A. N. (2001). Bayesian analysis of a time series of counts with covariates: an application to the control of an infectious disease. Biostatistics. 2(4): 433–44.

Hay, S. I. and Snow, R. W. (2006). The Malaria Atlas Project: Developing Global Maps of Malaria Risk. PLoS Med. 3(12): e473.

Hay, S. I., Cox, J., Rogers, D. J., Randolph, S. E., Stern, D. I., Shanks, D. G. D., Myers, M. F. and Snow, R. W. (2002). East African Highland Malaria Resurgence Independent of Climate Change. Directions in Science. 1: 82–5.

Hay, S. I., Guerra, C. A., Gething, P. W., Patil, A. P., Tatem, A. J., Noor, A. M., Kabaria, C. W., Manh, B. H., Elyazar, I. R. F., Brooker, S., Smith, D. L., Moyeed, R. A. and Snow, R. W. (2009). A World Malaria Map: Plasmodium falciparum Endemicity in 2007. Mueller, I. (ed). Plos Med. 6(3): e48.

Hay, S. I., Omumbo, J. A., Craig, M. H. and Snow, R. W. (2000a). Earth observation, geographic information systems and Plasmodium falciparum malaria in sub-Saharan Africa. In: Remote Sensing and Geographical Information Systems in Epidemiology. Vol. Volume 47. Academic Press. 173–215.

Hay, S. I., Rogers, D. J., Toomer, J. F. and Snow, R. W. (2000b). Annual Plasmodium falciparum entomological inoculation rates (EIR) across Africa: literature survey, internet access and review. Trans R Soc Trop Med Hyg. 94(2): 113–27.

Hay, S. I., Shanks, G. D., Stern, D. I., Snow, R. W., Randolph, S. E. and Rogers, D. J. (2005). Climate variability and malaria epidemics in the highlands of East Africa. Trends in parasitology. 21(2): 52–3.

Hay, S. I., Snow, R. W. and Rogers, D. J. (1998). Predicting malaria seasons in Kenya using multitemporal meteorological satellite sensor data. Transactions of the Royal Society of Tropical Medicine and Hygiene. 92(1): 12–20.

Henning, L., Schellenberg, D., Smith, T., Henning, D., Alonso, P., Tanner, M., Mshinda, H., Beck, H.-P. and Felger, I. (2004). A prospective study of Plasmodium falciparum multiplicity of infection and morbidity in Tanzanian children. Trans. R. Soc. Trop. Med. Hyg. 98(12): 687–94.

Herrero, M. D., Rivas, P., Rallón, N. I., Ramírez-Olivencia, G. and Puente, S. (2007). HIV and malaria. AIDS Rev. 9(2): 88–98.

Higdon, D. (2001). Space and space time modeling using process convolutions. Institute of Statistics and Decision Sciences, Duke University, Durham.

Higdon, D., Lee, H. and Holloman, C. (2003). Markov chain Monte Carlo-based approaches for inference in computationally intensive inverse problems. In: Bernardo JM, Bayarri, MJ, Berger, JO, Dawid, AP, Heckerman, D, Smith, AFM, West, M (Eds.), Bayesian Statistics 7. Oxford University Press. 181–97 pp.

Hii, J. L., Kanai, L., Foligela, A., Kan, S. K., Burkot, T. R. and Wirtz, R. A. (1993). Impact of permethrin-impregnated mosquito nets compared with DDT house-spraying against malaria transmission by Anopheles farauti and An.punctulatus in the Solomon Islands. Med. Vet. Entomol. 7(4): 333–8.

Hii, J., Alexander, N., Chuan, C. K., Rahman, H. A., Safri, A. and Chan, M. (1995). Lambdacyhalothrin impregnated bednets control malaria in Sabah, Malaysia. Southeast Asian J. Trop. Med. Public Health. 26(2): 371–4.

Hill, K., Thomas, K., AbouZahr, C., Walker, N., Say, L., Inoue, M. and Suzuki, E. (2007). Estimates of maternal mortality worldwide between 1990 and 2005: an assessment of available data. The Lancet. 370: 1311–9.

Hogan, M. C., Foreman, K. J., Naghavi, M., Ahn, S. Y., Wang, M., Makela, S. M., Lopez, A. D., Lozano, R. and Murray, C. J. L. (2010). Maternal mortality for 181 countries, 1980-2008: a systematic analysis of progress towards Millennium Development Goal 5. Lancet. 375(9726): 1609–23.

Hommerich, L., Von Oertzen, C., Bedu-Addo, G., Holmberg, V., Acquah, P., Eggelte, T., Bienzle, U. and Mockenhaupt, F. (2007). Decline of placental malaria in southern Ghana after the implementation of intermittent preventive treatment in pregnancy. Malaria journal. 6(1): 144.

Hopkins, H., Kambale, W., Kamya, M. R., Staedke, S. G., Dorsey, G. and Rosenthal, P. J. (2007). Comparison of HRP2- and pLDH-based rapid diagnostic tests for malaria with longitudinal follow-up in Kampala, Uganda. Am. J. Trop. Med. Hyg. 76(6): 1092–7.

Horn, S. D. (1977). Goodness-of-fit tests for discrete data: a review and an application to a health impairment scale. Biometrics. 237–47.

Hu, W., Tong, S., Mengersen, K. and Connell, D. (2007). Weather Variability and the Incidence of Cryptosporidiosis: Comparison of Time Series Poisson Regression and SARIMA Models. Annals of Epidemiology. 17(9): 679–88.

Hu, W., Tong, S., Mengersen, K. and Oldenburg, B. (2006). Rainfall, mosquito density and the transmission of Ross River virus: A time-series forecasting model. Ecological Modelling. 196(3-4): 505–14.

Huang, F., Zhou, S., Zhang, S., Wang, H. and Tang, L. (2011). Temporal correlation analysis between malaria and meteorological factors in Motuo County, Tibet. Malaria Journal. 10(1): 54.

Huho, B. J., Killeen, G. F., Ferguson, H. M., Tami, A., Lengeler, C., Charlewood, J., Kihonda, A., Kihonda, J., Kachur, S. P., Smith, T. A. and Abdulla, S. M. K. (in preparation). The introduction of artemisinin-based combination therapy did not lead to measurable reductions in human infectiousness to vectors in a setting of intense malaria transmission.

INDEPTH Network (2002). Population and health in developing countries—population, health and survival at INDEPTH sites. Ottawa: The International Development Research Centre.

INDEPTH Network (2005). Measuring Health Equity In Small Areas: Findings From Demographic Surveillance Systems. Ashgate Pub Co.

INDEPTH Network- Homepage (2011).[Online]. Available: http://www.indepth-network.org. (Accessed 14 Sep., 2011).

Idro, R., Aloyo, J., Mayende, L., Bitarakwate, E., John, C. and Kivumbi, G. (2006). Severe malaria in children in areas with low, moderate and high transmission intensity in Uganda. Tropical Medicine & International Health. 11(1): 115–24.

Ishengoma, D. S., Francis, F., Mmbando, B. P., Lusingu, J. P. A., Magistrado, P., Alifrangis, M., Theander, T. G., Bygbjerg, I. C. and Lemnge, M. M. (2011). Accuracy of malaria rapid diagnostic tests in community studies and their impact on treatment of malaria in an area with declining malaria burden in north-eastern Tanzania. Malar. J. 10: 176.

Jamison, D. T. and Bank, W. (2006). Disease control priorities in developing countries.

Jamison, D. T., Feachem, R. G., Makgoba, M. W., Bos, E. R., Baingana, F. K., Hofman, K. J. and Rogo, K. O. (2006). Disease And Mortality in Sub-saharan Africa. 2nd ed. World Bank Publications. 416 pp.

Jeffery, G. M. and Eyles, D. E. (1955). Infectivity to mosquitoes of Plasmodium falciparum as related to gametocyte density and duration of infection. The American Journal of Tropical Medicine and Hygiene. 4(5): 781.

Johnson, M., Moore, L. M. and Ylvisaker, D. (1990). Minimax and maximin distance designs. J Stat Plan Inference. 26(2): 131–48.

Jones, R. H. and Zhang, Y. (1997). Models for continuous stationary space-time processes. In: Modelling Longitudinal and Spatially Correlated Data: Methods, Applications and Future Directions, P. J. Diggle, W.G.Warren, and R. D. Wolfinger , eds. New York: Springer.

Jongwutiwes, S., Putaporntip, C., Iwasaki, T., Sata, T. and Kanbara, H. (2004). Naturally acquired Plasmodium knowlesi malaria in human, Thailand. Emerging infectious diseases. 10(12): 2211.

Kalyesubula, I., Musoke-Mudido, P., Marum, L., Bagenda, D., Aceng, E., Ndugwa, C. and Olness, K. (1997). Effects of malaria infection in human immunodeficiency virus type 1-infected Ugandan children. Pediatr. Infect. Dis. J. 16(9): 876–81.

Kammann, E. E. and Wand, M. P. (2003). Geoadditive models. Journal of the Royal Statistical Society: Series C (Applied Statistics). 52(1): 1–18.

Kamugisha, M. L., Mmbando, B. P., Francis, F., Ishengoma, D. R. S., Challe, D. P. and Lemnge, M. M. (2011). Establishing and implementing Demographic Surveillance System as a tool for monitoring health interventions in Korogwe District, north-eastern Tanzania. Tanzania Journal of Health Research. 13(4):

Kasasa, S., Rumisha, S. F., Nyaguara, O. A., Diboulo, E., Huho, B. J., Smith, T. and Vounatsou, P. (in preparation). Malaria transmission intensity and mortality burden across Africa project; statistical issues and approaches to data analysis.

Kazembe, L. N., Appleton, C. C. and Kleinschmidt, I. (2007). Spatial analysis of the relationship between early childhood mortality and malaria endemicity in Malawi. Geospatial Health. 2(1): 41–50.

Kelly-Hope, L. and McKenzie, F. E. (2009). The multiplicity of malaria transmission: a review of entomological inoculation rate measurements and methods across sub-Saharan Africa. Malar J. 8(1): 19.

Kent, R. J., Coetzee, M., Mharakurwa, S. and Norris, D. E. (2006). Feeding and indoor resting behaviour of the mosquito Anopheles longipalpis in an area of hyperendemic malaria transmission in southern Zambia. Med Vet Entomol. 20(4): 459–63.

Kigadye, E. S. P., Nkwengulila, G., Magesa, S. M. and Abdulla, S. (2010). Diversity, spatial and temporal abundance of Anopheles gambiae complex in the Rufiji River basin, south-eastern Tanzania. Tanzan J Health Res. 12(1): 68–72.

Kigadye, E. S. P., Nkwengulila, G., Magesa, S. M. and Abdulla, S. (2011). Spatial variability in the density, distribution and vectorial capacity of anopheline species in Rufiji district, south-eastern Tanzania. Tanzan J Health Res. 13(4):

Killeen, G. F. and Smith, T. A. (2007). Exploring the contributions of bed nets, cattle, insecticides and excitorepellency to malaria control: a deterministic model of mosquito host-seeking behaviour and mortality. Trans. R. Soc. Trop. Med. Hyg. 101(9): 867–80.

Killeen, G. F., McKenzie, F. E., Foy, B. D., Schieffelin, C., Billingsley, P. F. and Beier, J. C. (2000). A simplified model for predicting malaria entomologic inoculation rates based on entomologic and parasitologic parameters relevant to control. Am. J. Trop. Med. Hyg. 62(5): 535–44.

Killeen, G. F., Ross, A. and Smith, T. (2006). Infectiousness of malaria-endemic human populations to vectors. Am. J. Trop. Med. Hyg. 75(2 Suppl): 38–45.

Killeen, G. F., Smith, T. A., Ferguson, H. M., Mshinda, H., Abdulla, S., Lengeler, C. and Kachur, S. P. (2007a). Preventing childhood malaria in Africa by protecting adults from mosquitoes with insecticide-treated nets. PLoS Medicine. 4(7): e229.

Killeen, G. F., Tami, A., Kihonda, J., Okumu, F. O., Kotas, M. E., Grundmann, H., Kasigudi, N., Ngonyani, H., Mayagaya, V., Nathan, R., Abdulla, S., Charlwood, J. D., Smith, T. A. and Lengeler, C. (2007b). Cost-sharing strategies combining targeted public subsidies with private-sector delivery achieve high bednet coverage and reduced malaria transmission in Kilombero Valley, southern Tanzania. BMC Infect. Dis. 7: 121.

Kitron, U. and Spielman, A. (1989). Suppression of transmission of malaria through source reduction: antianopheline measures applied in Israel, the United States, and Italy. Rev. Infect. Dis. 11(3): 391–406.

Kitua, A. Y., Smith, T., Alonso, P. L., Masanja, H., Urassa, H., Menendez, C., Kimario, J. and Tanner, M. (1996). Plasmodium falciparum malaria in the first year of life in an area of intense and perennial transmission. Trop. Med. Int. Health. 1(4): 475–84.

Kleinschmidt, I., Bagayoko, M., Clarke, G., Craig, M. and Le Sueur, D. (2000). A spatial statistical approach to malaria mapping. International Journal of Epidemiology. 29(2): 355.

Kleinschmidt, I., Omumbo, J., Briet, O., Van De Giesen, N., Sogoba, N., Mensah, N. K., Windmeijer, P., Moussa, M. and Teuscher, T. (2001a). An empirical malaria distribution map for West Africa. Tropical Medicine & International Health. 6(10): 779–86.

Kleinschmidt, I., Schwabe, C., Benavente, L., Torrez, M., Ridl, F. C., Segura, J. L., Ehmer, P. and Nchama, G. N. (2009a). Marked Increase in Child Survival after Four Years of Intensive Malaria Control. The American Journal of Tropical Medicine and Hygiene. 80(6): 882–8.

Kleinschmidt, I., Schwabe, C., Shiva, M., Segura, J. L., Sima, V., Mabunda, S. J. A. and Coleman, M. (2009b). Combining Indoor Residual Spraying and Insecticide-Treated Net Interventions. The American Journal of Tropical Medicine and Hygiene. 81(3): 519–24.

Kleinschmidt, I., Sharp, B. L., Clarke, G. P., Curtis, B. and Fraser, C. (2001b). Use of generalized linear mixed models in the spatial analysis of small-area malaria incidence rates in Kwazulu Natal, South Africa. Am. J. Epidemiol. 153(12): 1213–21.

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. Am. J. Epidemiol. 155(3): 257–64.

Kobbe, R., Kreuzberg, C., Adjei, S., Thompson, B., Langefeld, I., Thompson, P. A., Abruquah, H. H., Kreuels, B., Ayim, M., Busch, W., Marks, F., Amoah, K., Opoku, E., Meyer, C. G., Adjei, O. and May, J. (2007). A Randomized Controlled Trial of Extended Intermittent Preventive Antimalarial Treatment in Infants. CLIN INFECT DIS. 45(1): 16–25.

Korenromp, E. L., Arnold, F., Williams, B. G., Nahlen, B. L. and Snow, R. W. (2004). Monitoring trends in under-5 mortality rates through national birth history surveys. Int J Epidemiol. 33(6): 1293–301.

Korenromp, E. L., Miller, J., Cibulskis, R. E., Kabir Cham, M., Alnwick, D. and Dye, C. (2003). Monitoring mosquito net coverage for malaria control in Africa: possession vs. use by children under 5 years. Tropical Medicine & International Health. 8(8): 693–703.

Koudou, B. G., Doumbia, M., Janmohamed, N., Tschannen, A. B., Tanner, M., Hemingway, J. and Utzinger, J. (2010). Effects of seasonality and irrigation on malaria transmission in two villages in Côte d'Ivoire. Ann Trop Med Parasitol. 104(2): 109–21.

Kristan, M., Abeku, T. A., Beard, J., Okia, M., Rapuoda, B., Sang, J. and Cox, J. (2008). Variations in entomological indices in relation to weather patterns and malaria incidence in East African highlands: implications for epidemic prevention and control. Malar. J. 7: 231.

Kulkarni, M. A., Desrochers, R. E. and Kerr, J. T. (2010). High Resolution Niche Models of Malaria Vectors in Northern Tanzania: A New Capacity to Predict Malaria Risk? PLoS ONE. 5(2): e9396.

Kweka, E. J., Mwang'onde, Beda J and Mahande, A. M. (2010). Optimization of odour-baited resting boxes for sampling malaria vector, Anopheles arabiensis Patton, in arid and highland areas of Africa. Parasit Vectors. 3: 75–75.

Kynast-Wolf, G., Hammer, G. P., Müller, O., Kouyaté, B. and Becher, H. (2006). Season of death and birth predict patterns of mortality in Burkina Faso. International Journal of Epidemiology. 35(2): 427–35.

Kynast-Wolf, G., Sankoh, O. A., Gbangou, A., Kouyaté, B. and Becher, H. (2002). Mortality patterns, 1993–98, in a rural area of Burkina Faso, West Africa, based on the Nouna demographic surveillance system. Tropical Medicine & International Health. 7(4): 349–56.

Lal, S., Lahariya, C. and Saxena, V. K. (2010). Insecticide treated nets, antimalarials and child survival in India. Indian J Pediatr. 77(4): 425–30.

Lambert, D. (1992). Zero-inflated Poisson regression, with an application to defects in manufacturing. Technometrics. 34(1): 1–14.

Lanjouw, P. and Ravallion, M. (1995). Poverty and household size. The Economic Journal. 1415–34.

Lardeux, F. J., Tejerina, R. H., Quispe, V. and Chavez, T. K. (2008). A physiological time analysis of the duration of the gonotrophic cycle of Anopheles pseudopunctipennis and its implications for malaria transmission in Bolivia. Malar. J. 7: 141.

Lawal, B. H. (2010). Zero-inflated count regression models with applications to some examples. Qual Quant.

Lawn, J. E., Cousens, S., Zupan, J. and others (2005a). 4 million neonatal deaths: When? Where? Why? The Lancet. 365(9462): 891–900.

Lawn, J., Shibuya, K. and Stein, C. (2005b). No cry at birth: global estimates of intrapartum stillbirths and intrapartum-related neonatal deaths. Bulletin of the World Health Organization. 83(6): 409–17.

Le Hesran, J.-Y., Fiévet, N., Thioulouse, J., Personne, P., Maubert, B., M'bidias, S., Etye'ale, D., Cot, M. and Deloron, P. (2006). Development of cellular immune responses to Plasmodium falciparum blood stage antigens from birth to 36 months of age in Cameroon. Acta Trop. 98(3): 261–9.

Lee, H.-I., Lee, J.-S., Shin, E.-H., Lee, W.-J., Kim, Y.-Y. and Lee, K.-R. (2001). Malaria transmission potential by Anopheles sinensis in the Republic of Korea. Korean J Parasitol. 39(2): 185–92.

Leisnham, P. T., Slaney, D. P., Lester, P. J., Weinstein, P. and Heath, A. C. G. (2007). Mosquito density, macroinvertebrate diversity, and water chemistry in water-filled containers: Relationships to land use. New Zealand Journal of Zoology. 34(3): 203.

Lemma, H., Byass, P., Desta, A., Bosman, A., Costanzo, G., Toma, L., Fottrell, E., Marrast, A.-C., Ambachew, Y., Getachew, A., Mulure, N., Morrone, A., Bianchi, A. and Barnabas, G. A. (2010). Deploying artemether-lumefantrine with rapid testing in Ethiopian communities: impact on malaria morbidity, mortality and healthcare resources. Trop. Med. Int. Health. 15(2): 241–50.

Lengeler, C. (2004). Insecticide-treated bed nets and curtains for preventing malaria. status and date: Edited (no change to conclusions), published in.

Lengeler, C., Armstrong-Schellenberg, J., D'Alessandro, U., Binka, F. and Cattani, J. (1998). Relative versus absolute risk of dying reduction after using insecticide-treated nets for malaria control in Africa. Trop. Med. Int. Health. 3(4): 286–90.

Lengeler, C., Schellenberg, J. A. and D'Alessandro, U. (1995). Will reducing Plasmodium falciparum malaria transmission alter malaria mortality among African children? Parasitol. Today (Regul. Ed.). 11(11): 425; author reply 426.

Li, F. and Zhang, N. R. (2010). Bayesian variable selection in structured high-dimensional covariate spaces with applications in genomics. Journal of the American Statistical Association. 105(491): 1202–14.

Li, P., Church, K. W. and Hastie, T. J. (2007). Conditional random sampling: A sketch-based sampling technique for sparse data. Advances in neural information processing systems. 19: 873.

Lim, S. S., Fullman, N., Stokes, A., Ravishankar, N., Masiye, F., Murray, C. J. L. and Gakidou, E. (2011). Net Benefits: A Multicountry Analysis of Observational Data Examining Associations between Insecticide-Treated Mosquito Nets and Health Outcomes. PLoS Med. 8(9): e1001091.

Lindblade, K. A., Eisele, T. P., Gimnig, J. E., Alaii, J. A., Odhiambo, F., ter Kuile, F. O., Hawley, W. A., Wannemuehler, K. A., Phillips-Howard, P. A., Rosen, D. H., Nahlen, B. L., Terlouw, D. J., Adazu, K., Vulule, J. M. and Slutsker, L. (2004). Sustainability of Reductions in Malaria Transmission and Infant Mortality in Western Kenya With Use of Insecticide-Treated Bednets. JAMA: The Journal of the American Medical Association. 291(21): 2571–80.

Lindblade, K. A., Walker, E. D., Onapa, A. W., Katungu, J. and Wilson, M. L. (1999). Highland malaria in Uganda: prospective analysis of an epidemic associated with El Niño. Trans. R. Soc. Trop. Med. Hyg. 93(5): 480–7.

Lines, J. D., Curtis, C. f., Wilkes, T. j. and Njunwa, K. j. (1991a). Monitoring Human-Biting Mosquitoes (Diptera: Culicidae) in Tanzania with Light-Traps Hung Beside Mosquito Nets. Bulletin of Entomological Research. 81(01): 77–84.

Lines, J. D., Wilkes, T. J. and Lyimo, E. O. (1991b). Human Malaria Infectiousness Measured by Age-Specific Sporozoite Rates in Anopheles Gambiae in Tanzania. Parasitology. 102(02): 167–77.

Lipp, E. K., Huq, A. and Colwell, R. R. (2002). Effects of global climate on infectious disease: the cholera model. Clinical microbiology reviews. 15(4): 757.

Lloyd-Smith, J. O. (2007). Maximum likelihood estimation of the negative binomial dispersion parameter for highly overdispersed data, with applications to infectious diseases. PLoS One. 2(2): e180.

Lopez, A. D., Mathers, C. D., Ezzati, M., Jamison, D. T. and Murray, C. J. L. (2006). Global and regional burden of disease and risk factors, 2001: systematic analysis of population health data. The Lancet. 367(9524): 1747–57.

Lozano, R., Lopez, A. D., Atkinson, C., Naghavi, M., Flaxman, A. D. and Murray, C. J. (2011a). Performance of physician-certified verbal autopsies: multisite validation study using clinical diagnostic gold standards. Popul Health Metr. 9: 32.

Lozano, R., Wang, H., Foreman, K. J., Rajaratnam, J. K., Naghavi, M., Marcus, J. R., Dwyer-Lindgren, L., Lofgren, K. T., Phillips, D., Atkinson, C., Lopez, A. D. and Murray, C. J. L. (2011b). Progress towards Millennium Development Goals 4 and 5 on maternal and child mortality: an updated systematic analysis. Lancet. 378(9797): 1139–65.

Lutambi, A. M., Alexander, M., Charles, J., Mahutanga, C. and Nathan, R. (2010). Under-five mortality: spatial—temporal clusters in Ifakara HDSS in South-eastern Tanzania. Glob Health Action. 3:

Lysenko, A. J. and Semashko, I. N. (1968). Geography of malaria. A medico-geographic profile of an ancient disease. In: Medicinskaja Geografija. Lebedew AW ed. Moscow, USSR: Academy of Sciences. 25–146.

MARA/ARMA (1998). Towards an Atlas of Malaria Risk in Africa. First technicalreport of the MARA/ARMA collaboration (www.mara.org.za) South Africa.

MMV (2009). Medicines for Malaria Venture: Organizational Report.

Ma, C. (2003). Spatio-temporal stationary covariance models. Journal of multivariate analysis. 86(1): 97–107.

Mabaso, M. L. H. (2007). "Temporal variations in malaria risk in Africa." [Online]. Available: http://edoc.unibas.ch/650/. (Accessed 5 Sep., 2011).

Mabaso, M. L. H., Kleinschmidt, I., Sharp, B. and Smith, T. (2007). El Niño Southern Oscillation (ENSO) and annual malaria incidence in Southern Africa. Trans. R. Soc. Trop. Med. Hyg. 101(4): 326–30.

Mabaso, M. L. H., 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. Int J Health Geogr. 5: 20.

Mabaso, M., Craig, M., Vounatsou, P. and Smith, T. (2005). Towards empirical description of malaria seasonality in southern Africa: the example of Zimbabwe. Tropical Medicine & International Health. 10(9): 909–18.

MacDonald, G. (1957). The Epidemiology and Control of Malaria. Oxford University Press, Oxford.

Magalhães, R. J. S., Clements, A. C. A., Patil, A. P., Gething, P. W. and Brooker, S. (2011). The Applications of Model-Based Geostatistics in Helminth Epidemiology and Control. Adv Parasitol. 74: 267–96.

Mahande, A., Mwang'onde, B., Msangi, S., Kimaro, E., Mnyone, L., Mazigo, H., Mahande, M. and Kweka, E. (2010). Is aging raw cattle urine efficient for sampling Anopheles arabiensis Patton? BMC Infect Dis. 10(1): 172.

Mahfouz, M. S., Surur, A. A., Ajak, D. A. A. and Abdelrahim, E. (2009). Level and Determinants of Infant and Child Mortality in Malakal Town–Southern Sudan. Sudanese Journal of Public Health. 4(2): 250–5.

Maire, N., Smith, T., Ross, A., Owusu-Agyei, S., Dietz, K. and Molineaux, L. (2006). A model for natural immunity to asexual blood stages of Plasmodium falciparum malaria in endemic areas. Am. J. Trop. Med. Hyg. 75(2 Suppl): 19–31.

Makani, J., Matuja, W., Liyombo, E., Snow, R. W., Marsh, K. and Warrell, D. A. (2003). Admission diagnosis of cerebral malaria in adults in an endemic area of Tanzania: implications and clinical description. QJM. 96(5): 355–62.

Mala, A., Irungu, L., Shililu, J., Muturi, E., Mbogo, C., Njagi, J. and Githure, J. (2011). Dry season ecology of Anopheles gambiae complex mosquitoes at larval habitats in two traditionally semi-arid villages in Baringo, Kenya. Parasit Vectors. 4(1): 25.

Manh, B. H., Clements, A. C. A., Thieu, N. Q., Hung, N. M., Hung, L. X., Hay, S. I., Hien, T. T., Wertheim, H. F. L., Snow, R. W. and Horby, P. (2011). Social and environmental determinants of malaria in space and time in Viet Nam. International Journal for Parasitology. 41(1): 109–16.

Manzi, F., Schellenberg, J., Hamis, Y., Mushi, A. K., Shirima, K., Mwita, A., Simba, A., Rusibamayila, N., Kitambi, M., Tanner, M. and others (2009). Intermittent preventive treatment for malaria and anaemia control in Tanzanian infants; the development and implementation of a public health strategy. Transactions of the Royal Society of Tropical Medicine and Hygiene. 103(1): 79–86.

Marsh, K. and Snow, R. (1999). Malaria transmission and morbidity. Parassitologia. 41(1-3): 241.

Marsh, L. and Cormier, D. R. (2002). Spline regression models. Vol. 137. Sage Publications.

Martens, W. J., Niessen, L. W., Rotmans, J., Jetten, T. H. and McMichael, A. J. (1995). Potential impact of global climate change on malaria risk. Environ Health Perspect. 103(5): 458–64.

Martin-Blondel, G., Barry, M., Porte, L., Busato, F., Massip, P., Benoit-Vical, F., Berry, A. and Marchou, B. (2007). Impact of HIV infection on malaria in adults. Med Mal Infect. 37(10): 629–36.

Masanja, H., de Savigny, D., Smithson, P., Schellenberg, J., John, T., Mbuya, C., Upunda, G., Boerma, T., Victora, C., Smith, T. and Mshinda, H. (2008). Child survival gains in Tanzania: analysis of data from demographic and health surveys. The Lancet. 371(9620): 1276–83.

Matei, A. and Tillé, Y. (2005). Evaluation of variance approximations and estimators in maximum entropy sampling with unequal probability and fixed sample size. Journal of Official Statistics. 21(4): 543–70.

Mathanga, D. P., Campbell, C. H., Taylor, T. E., Barlow, R. and Wilson, M. L. (2005). Reduction of childhood malaria by social marketing of insecticide-treated nets: a case-control study of effectiveness in Malawi. Am. J. Trop. Med. Hyg. 73(3): 622–5.

Maxwell, C., Chambo, W., Mwaimu, M., Magogo, F., Carneiro, I. and Curtis, C. (2003). Variation of malaria transmission and morbidity with altitude in Tanzania and with introduction of alphacypermethrin treated nets. Malaria journal. 2(1): 28.

Maxwell, C., Msuya, E., Sudi, M., Njunwa, K., Carneiro, I. and Curtis, C. (2002). Effect of community-wide use of insecticide-treated nets for 3–4 years on malarial morbidity in Tanzania. Tropical Medicine & International Health. 7(12): 1003–8.

Mayor, A., Aponte, J. J., Fogg, C., Saúte, F., Greenwood, B., Dgedge, M., Menendez, C. and Alonso, P. L. (2007). The epidemiology of malaria in adults in a rural area of southern Mozambique. Malar. J. 6: 3.

Mboera, L. (2006). Sampling techniques for adult Afrotropical malaria vectors and their reliability in the estimation of entomological inoculation rate. Tanzania Journal of Health Research. 7(3): 117–24.

Mboera, L. E. G., Makundi, E. A. and Kitua, A. Y. (2007). Uncertainty in malaria control in Tanzania: crossroads and challenges for future interventions. The American journal of tropical medicine and hygiene. 77(6 Suppl): 112.

Mboera, L. E. G., Senkoro, K. P., Mayala, B. K., Rumisha, S. F., Rwegoshora, R. T., Mlozi, M. R. S. and Shayo, E. H. (2010). Spatio-temporal variation in malaria transmission intensity in five agro-ecosystems in Mvomero district, Tanzania. Geospat Health. 4(2): 167–78.

Mbogo, C. M., Mwangangi, J. M., Nzovu, J., Gu, W., Yan, G., Gunter, J. T., Swalm, C., Keating, J., Regens, J. L., Shililu, J. I., Githure, J. I. and Beier, J. C. (2003). Spatial and temporal heterogeneity of Anopheles mosquitoes and Plasmodium falciparum transmission along the Kenyan coast. Am. J. Trop. Med. Hyg. 68(6): 734–42.

Mbogo, C. N. M., Snow, R. W., Khamala, C. P. M., Kabiru, E. W., Ouma, J. H., Githure, J. I., Marsh, K. and Beier, J. C. (1995). Relationships between Plasmodium Falciparum Transmission by Vector Populations and the Incidence of Severe Disease at Nine Sites on the Kenyan Coast. The American Journal of Tropical Medicine and Hygiene. 52(3): 201–6.

Mbogo, C. N., Snow, R. W., Kabiru, E. W., Ouma, J. H., Githure, J. I., Marsh, K. and Beier, J. C. (1993). Low-level Plasmodium falciparum transmission and the incidence of severe malaria infections on the Kenyan coast. Am. J. Trop. Med. Hyg. 49(2): 245–53.

McElroy, P. D., Ter Kuile, F. O., Hightower, A. W., Hawley, W. A., Phillips-Howard, P. A., Oloo, A. J., Lal, A. A. and Nahlen, B. L. (2001). All-cause mortality among young children in western Kenya. VI: the Asembo Bay Cohort Project. The American journal of tropical medicine and hygiene. 64(1 suppl): 18.

Medina, D. C., Findley, S. E., Guindo, B. and Doumbia, S. (2007). Forecasting Non-Stationary Diarrhea, Acute Respiratory Infection, and Malaria Time-Series in Niono, Mali. PLoS ONE. 2(11): e1181.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H. and Teller, E. (1953). Equations of State Calculations by Fast Computing Machines. Journal of Chemical Physics. 21: 1087–92.

Michael, E., Ramaiah, K. D., Hoti, S. L., Barker, G., Paul, M. R., Yuvaraj, J., Das, P. K., Grenfell, B. T. and Bundy, D. A. (2001). Quantifying mosquito biting patterns on humans by DNA fingerprinting of bloodmeals. Am. J. Trop. Med. Hyg. 65(6): 722–8.

Mockenhaupt, F. P., Reither, K., Zanger, P., Roepcke, F., Danquah, I., Saad, E., Ziniel, P., Dzisi, S. Y., Frempong, M., Agana-Nsiire, P. and others (2007). Intermittent preventive treatment in infants as a means of malaria control: a randomized, double-blind, placebo-controlled trial in northern Ghana. Antimicrobial agents and chemotherapy. 51(9): 3273.

Mohorovic, L., Petrovic, O., Haller, H. and Micovic, V. (2010). Pregnancy Loss and Maternal Methemoglobin Levels: An Indirect Explanation of the Association of Environmental Toxics and Their Adverse Effects on the Mother and the Fetus. International Journal of Environmental Research and Public Health. 7(12): 4203–12.

Molineaux, L. (1985). The pros and cons of modelling malaria transmission. Trans. R. Soc. Trop. Med. Hyg. 79(6): 743–7.

Molineaux, L. (1988). The epidemiology of human malaria as an explanation of its distribution, including some implications for its control. In: Malaria: Principles and and Practice of Malariology. (eds. Wernsdorfer WH & McGregor I). Churchill Livingstone, Edinburgh. 913–98.

Molineaux, L. and Gramiccia, G. (1980). The Garki Project: Research on the Epidemiology and Control of Malaria in the Sudan Savanna of West Africa. Geneva: WHO.

Mpimbaza, A., Filler, S., Katureebe, A., Kinara, S. O., Nzabandora, E., Quick, L., Ratcliffe, A., Wabwire-Mangen, F., Chandramohan, D. and Staedke, S. G. (2011). Validity of verbal autopsy procedures for determining malaria deaths in different epidemiological settings in Uganda. PLoS ONE. 6(10): e26892.

Msellem, M. I., Mårtensson, A., Rotllant, G., Bhattarai, A., Strömberg, J., Kahigwa, E., Garcia, M., Petzold, M., Olumese, P., Ali, A. and Björkman, A. (2009). Influence of Rapid Malaria Diagnostic Tests on Treatment and Health Outcome in Fever Patients, Zanzibar—A Crossover Validation Study. PLoS Med. 6(4): e1000070.

Mudenda, S. S., Kamocha, S., Mswia, R., Conkling, M., Sikanyiti, P., Potter, D., Mayaka, W. C. and Marx, M. A. (2011). Feasibility of using a World Health Organization-standard methodology for Sample Vital Registration with Verbal Autopsy (SAVVY) to report leading causes of death in Zambia: results of a pilot in four provinces, 2010. communities. 1(8): 19.

Muturi, E. J., Burgess, P. and Novak, R. J. (2008). Malaria Vector Management: Where Have We Come From and Where Are We Headed? The American Journal of Tropical Medicine and Hygiene. 78(4): 536–7.

Mwageni, E., Momburi, D., Juma, Z., Irema, M., Masanja, H., TEHIP and AMMP teams (2002). Rufiji DSS, Tanzania. INDEPTH Monograph: Volume 1.

Müller, O., De Allegri, M., Becher, H., Tiendrebogo, J., Beiersmann, C., Ye, M., Kouyate, B., Sie, A. and Jahn, A. (2008). Distribution Systems of Insecticide-Treated Bed Nets for Malaria Control in Rural Burkina Faso: Cluster-Randomized Controlled Trial. PLoS ONE. 3(9):

Müller, O., Traoré, C., Becher, H. and Kouyaté, B. (2003). Malaria morbidity, treatment-seeking behaviour, and mortality in a cohort of young children in rural Burkina Faso. Tropical Medicine & International Health. 8(4): 290–6.

Naumova, E. N., Jagai, J. S., Matyas, B., De Maria, A., MacNeill, I. B. and Griffiths, J. K. (2007). Seasonality in six enterically transmitted diseases and ambient temperature. Epidemiol Infect. 135(2): 281–92.

Ndugwa, R. P. and Zulu, E. M. (2008). Child morbidity and care-seeking in Nairobi slum settlements: the role of environmental and socio-economic factors. Journal of Child Health Care. 12(4): 314.

Ndugwa, R., Ramroth, H., M\üller, O., Jasseh, M., Sié, A., Kouyaté, B., Greenwood, B. and Becher, H. (2008). Comparison of all-cause and malaria-specific mortality from two West African countries with different malaria transmission patterns. Malaria journal. 7(1): 15.

Nedelman, J. (1983). A negative binomial model for sampling mosquitoes in a malaria survey. Biometrics. 39(4): 1009–20.

Neyman, J. (1934). On the two different aspects of the representative method: the method of stratified sampling and the method of purposive selection. Journal of the Royal Statistical Society. 97: 558–606.

Noor, A. M., Clements, A. C., Gething, P. W., Moloney, G., Borle, M., Shewchuk, T., Hay, S. I. and Snow, R. W. (2008). Spatial prediction of Plasmodium falciparum prevalence in Somalia. Malar J. 7: 159–159.

Noor, A. M., Mutheu, J. J., Tatem, A. J., Hay, S. I. and Snow, R. W. (2009). Insecticide-treated net coverage in Africa: mapping progress in 2000–07. Lancet. 373(9657): 58–67.

Nott, D. J. and Leonte, D. (2004). Sampling schemes for Bayesian variable selection in generalized linear models. Journal of Computational and Graphical Statistics. 13(2): 362–82.

Nájera, J. A. (2001). Malaria control: achievements, problems and strategies. Parassitologia. 43(1-2): 1–89.

Okiro, E., Al-Taiar, A., Reyburn, H., Idro, R., Berkley, J. and Snow, R. (2009). Age patterns of severe paediatric malaria and their relationship to Plasmodium falciparum transmission intensity. Malaria journal. 8(1): 4.

Omeara, W., Bejon, P., Mwangi, T., Okiro, E., Peshu, N., Snow, R., Newton, C. and Marsh, K. (2008). Effect of a fall in malaria transmission on morbidity and mortality in Kilifi, Kenya. The Lancet. 372(9649): 1555–62.

Omumbo, J. A., Hay, S. I., Guerra, C. A. and Snow, R. W. (2004). The relationship between the Plasmodium falciparum parasite ratio in childhood and climate estimates of malaria transmission in Kenya. Malar J. 3: 17–17.

Omumbo, J., Ouma, J., Rapuoda, B., Craig, M. H., le Sueur, D. and Snow, R. W. (1998). Mapping malaria transmission intensity using geographical information systems (GIS): an example from Kenya. Ann Trop Med Parasitol. 92(1): 7–21.

PATH, Program for Appropriate Technology in Health (2007). Accelerating Progress Toward Malaria Vaccines. Bethesda, MD: PATH.

PMI (2011). The President's Malaria Initiative, AnnualReport.

Paciorek, C. J. (2007). Computational techniques for spatial logistic regression with large data sets. Computational Statistics & Data Analysis. 51(8): 3631–53.

Padmanabha, P. (1982). Sample Registration System in India. Sample Regist Bull. 16(2): 45–50.

Pagès, F., Texier, G., Pradines, B., Gadiaga, L., Machault, V., Jarjaval, F., Penhoat, K., Berger, F., Trape, J.-F., Rogier, C. and Sokhna, C. (2008). Malaria transmission in Dakar: a two-year survey. Malar. J. 7: 178.

Pampana, E. (1969). A textbook of malaria eradication. Oxford University Press, London, New York, Toronto.

Pankratz, A. (1983). Forecasting with univariate Box-Jenkins models:concepts and cases. John Wiley & Sons, Hoboken, NJ. 576 pp.

Parham, P. E. and Michael, E. (2009). Modeling the Effects of Weather and Climate Change on Malaria Transmission. Environ Health Perspect. 118(5): 620–6.

Parise, M. E., Ayisi, J. G., Nahlen, B. L., Schultz, L. J., Roberts, J. M., Misore, A., Muga, R., Oloo, A. J. and Steketee, R. W. (1998). Efficacy of sulfadoxine-pyrimethamine for prevention of placental malaria in an area of Kenya with a high prevalence of malaria and human immunodeficiency virus infection. Am. J. Trop. Med. Hyg. 59(5): 813–22.

Pascual, M., Cazelles, B., Bouma, M. ., Chaves, L. . and Koelle, K. (2008). Shifting patterns: malaria dynamics and rainfall variability in an African highland. Proceedings of the Royal Society B: Biological Sciences. 275(1631): 123–32.

Patil, A. P., Gething, P. W., Piel, F. B. and Hay, S. I. (2011). Bayesian geostatistics in health cartography: the perspective of malaria. Trends Parasitol. 27(6): 246–53.

Patnaik, P., Jere, C. S., Miller, W. C., Hoffman, I. F., Wirima, J., Pendame, R., Meshnick, S. R., Taylor, T. E., Molyneux, M. E. and Kublin, J. G. (2005). Effects of HIV-1 serostatus, HIV-1 RNA concentration, and CD4 cell count on the incidence of malaria infection in a cohort of adults in rural Malawi. J. Infect. Dis. 192(6): 984–91.

Patz, J. A. (2002). A human disease indicator for the effects of recent global climate change. Proceedings of the National Academy of Sciences. 99(20): 12506.

Patz, J. A., Vavrus, S. J., Uejio, C. K. and McLellan, S. L. (2008). Climate change and waterborne disease risk in the Great Lakes region of the US. American journal of preventive medicine. 35(5): 451–8.

Pelletier, D. L., Frongillo Jr, E. A., Schroeder, D. G. and Habicht, J. P. (1995). The effects of malnutrition on child mortality in developing countries. Bulletin of the World Health Organization. 73(4): 443.

Pethleart, A., Prajakwong, S., Suwonkerd, W., Corthong, B., Webber, R. and Curtis, C. (2004). Infectious reservoir of Plasmodium infection in Mae Hong Son province, north-west Thailand. Malar J. 3(1): 34.

Phillips-Howard, P. A., Nahlen, B. L., Alaii, J. A., ter Kuile, F. O., Gimnig, J. E., Terlouw, D. J., Kachur, S. P., Hightower, A. W., Lal, A. A., Schoute, E., Oloo, A. J. and Hawley, W. A. (2003). The efficacy of permethrin-treated bed nets on child mortality and morbidity in western Kenya I. Development of infrastructure and description of study site. Am. J. Trop. Med. Hyg. 68(4 Suppl): 3–9.

Poespoprodjo, J. R., Hasanuddin, A., Fobia, W., Sugiarto, P., Kenangalem, E., Lampah, D. A., Tjitra, E., Price, R. N. and Anstey, N. M. (2010). Severe Congenital Malaria Acquired in utero. The American Journal of Tropical Medicine and Hygiene. 82(4): 563–5.

Prakash, A., Bhattacharyya, D., Mohapatra, P. and Mahanta, J. (2005). Malaria transmission risk by the mosquito Anopheles baimaii (formerly known as An. dirus species D) at different hours of the night in North-east India. Medical and veterinary entomology. 19(4): 423–7.

Pull, J. H. and Grab, B. (1974). A simple epidemiological model for evaluating the malaria inoculation rate and the risk of infection in infants. Bull. World Health Organ. 51(5): 507–16.

Quigley, M. A., Hewitt, K., Mayanja, B., Morgan, D., Eotu, H., Ojwiya, A. and Whitworth, J. A. G. (2005). The effect of malaria on mortality in a cohort of HIV-infected Ugandan adults. Trop. Med. Int. Health. 10(9): 894–900.

Quiñones, M. L., Lines, J. D., Thomson, M. C., Jawara, M., Morris, J. and Greenwood, B. M. (1997). Anopheles gambiae gonotrophic cycle duration, biting and exiting behaviour unaffected by permethrin-impregnated bednets in The Gambia. Medical and Veterinary Entomology. 11(1): 71–8.

RBM (2000). Roll Back Malaria, The Abuja Declaration.

RBM (2009). Tracking Progress in Scaling-Up Diagnosis and Treatment for Malaria: A Compilation of Data on African Malaria Endemic Countries' Estimates of their Commodity Needs and Funding Available.

RBM, WHO (1999). Malaria control achievements, problems and strategies. World Health Organization, Geneva.

RBM, WHO (2005). Malaria Indicator Survey: Basic documentation for survey designand implementation.

RBM:GMAP (2008). The Global Malaria Action Plan (GMAP).

Rabinovich, N. R. (2002). Beyond Our Borders: Are we there yet? The road to a malaria vaccine. Western journal of medicine. 176(2): 82.

Rajaratnam, J. K., Marcus, J. R., Flaxman, A. D., Wang, H., Levin-Rector, A., Dwyer, L., Costa, M., Lopez, A. D. and Murray, C. J. L. (2010a). Neonatal, postneonatal, childhood, and under-5 mortality for 187 countries, 1970-2010: a systematic analysis of progress towards Millennium Development Goal 4. The Lancet. 375(9730): 1988–2008.

Rajaratnam, J. K., Marcus, J. R., Levin-Rector, A., Chalupka, A. N., Wang, H., Dwyer, L., Costa, M., Lopez, A. D. and Murray, C. J. (2010b). Worldwide mortality in men and women aged 15-59 years from 1970 to 2010: a systematic analysis. The Lancet. 375(9727): 1704–20.

Ramroth, H., Ndugwa, R. P., Müller, O., Yé, Y., Sié, A., Kouyaté, B. and Becher, H. (2009). Decreasing childhood mortality and increasing proportion of malaria deaths in rural Burkina Faso. Glob Health Action. 2:

Rathbun, S. L. and Fei, S. (2006). A spatial zero-inflated Poisson regression model for oak regeneration. Environmental and Ecological Statistics. 13(4): 409–26.

Rau, R. (2006). Seasonality in Human Mortality: A Demographic Approach. 1st ed. Springer.

Reid, H., Vallely, A., Taleo, G., Tatem, A., Kelly, G., Riley, I., Harris, I., Henri, I., Iamaher, S. and Clements, A. (2010). Baseline spatial distribution of malaria prior to an elimination programme in Vanuatu. Malar J. 9(1): 150.

Ridout, M., Hinde, J. and Demétrio, C. G. B. (2001). A Score Test for Testing a Zero-Inflated Poisson Regression Model against Zero-Inflated Negative Binomial Alternatives. Biometrics. 57(1): 219–23.

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

Riley, E. M., Wagner, G. E., Ofori, M. F., Wheeler, J. G., Akanmori, B. D., Tetteh, K., McGuinness, D., Bennett, S., Nkrumah, F. K., Anders, R. F. and Koram, K. A. (2000). Lack of Association between Maternal Antibody and Protection of African Infants from Malaria Infection. Infect. Immun. 68(10): 5856–63.

Riley, I. (2011). Computer-based analysis of verbal autopsies: revolution or evolution? Popul Health Metr. 9: 26.

Robert, V., Macintyre, K., Keating, J., Trape, J.-F., Duchemin, J.-B., Warren, M. and Beier, J. C. (2003). Malaria transmission in urban sub-Saharan Africa. Am. J. Trop. Med. Hyg. 68(2): 169–76.

Roberts, L. and Enserink, M. (2007). Did They Really Say ... Eradication? Science. 318(5856): 1544–5.

Rodriguez, M. H., BOWN, D. N., ARREDONDO-JIMENEZ, J. I., VILLARREAL, C., LOYOLA, E. G. and FREDERICKSON, C. E. (1992). Gonotrophic cycle and survivorship of Anopheles albimanus (Diptera: Culicidae) in southern Mexico. Journal of medical entomology. 29(3): 395–9.

Rogers, D. J., Randolph, S. E., Snow, R. W. and Hay, S. I. (2002). Satellite imagery in the study and forecast of malaria. Nature. 415(6872): 710–5.

Rogier, C. (2000). Natural history of Plasmodium falciparum malaria and determining factors of the acquisition of antimalaria immunity in two endemic areas, Dielmo and Ndiop (Senegal). Bull. Mem. Acad. R. Med. Belg. 155(5-6): 218–26.

Romagosa, C., Ordi, J., Saute, F., Quintó, L., Machungo, F., Ismail, M. R., Carrilho, C., Osman, N., Alonso, P. L. and Menendez, C. (2007). Seasonal variations in maternal mortality in Maputo, Mozambique: the role of malaria. Trop. Med. Int. Health. 12(1): 62–7.

Ross, A., Maire, N., Molineaux, L. and Smith, T. (2006). An epidemiologic model of severe morbidity and mortality caused by Plasmodium falciparum. Am. J. Trop. Med. Hyg. 75(2 Suppl): 63–73.

Ross, A., Penny, M., Maire, N., Studer, A., Carneiro, I., Schellenberg, D., Greenwood, B., Tanner, M. and Smith, T. (2008). Modelling the Epidemiological Impact of Intermittent Preventive Treatment against Malaria in Infants. Hay, S.I. (ed). PLoS ONE. 3(7): e2661.

Rowe, A. K. and Steketee, R. W. (2007). Predictions of the impact of malaria control efforts on all-cause child mortality in sub-Saharan Africa. Am. J. Trop. Med. Hyg. 77(6 Suppl): 48–55.

Rowe, A. K., Rowe, S. Y., Snow, R. W., Korenromp, E. L., Schellenberg, J. R. ., Stein, C., Nahlen, B. L., Bryce, J., Black, R. E. and Steketee, R. W. (2006). The burden of malaria mortality among African children in the year 2000. International Journal of Epidemiology. 35(3): 691.

Rowe, A. K., Steketee, R. W., Arnold, F., Wardlaw, T., Basu, S., Bakyaita, N., Lama, M., Winston, C. A., Lynch, M., Cibulskis, R. E., Shibuya, K., Ratcliffe, A. A. and Nahlen, B. L. (2007). Viewpoint: evaluating the impact of malaria control efforts on mortality in sub-Saharan Africa. Trop. Med. Int. Health. 12(12): 1524–39.

Rue, H. and Tjelmeland, H. (2002). Fitting Gaussian Markov random fields to Gaussian fields. Scandinavian Journal of Statistics. 29(1): 31–49.

Rumisha, S. F., Kasasa, S., Smith, T., Abdulla, S., Masanja, H. and Vounatsou, P. (submitted). Bayesian Modeling of Large Geostatistical Data to Estimate Seasonal and Spatial Variation of Sporozoite Rate. Spatial and Spatio-temporal Epidemiology.

Rumisha, S. F., Smith, T. A., Masanja, H., Abdulla, S. and Vounatsou, P. (submitted). Assessing the relation between child survival and malaria transmission: an analysis of the MTIMBA data in Rufiji DSS, Tanzania, Malaria Journal.

Rumisha, S.F., Smith, T., Abdulla, S., Masanja, H. & Vounatsou, P. (2013) Assessing seasonal variations and age patterns in mortality during the first year of life in Tanzania. Acta Tropica, 126:28– 36.

Rumisha, S. F., Smith, T., Abdulla, S., Masanja, H. and Vounatsou, P. (submitted). Modeling heterogeneity in malaria transmission using large sparse spatio-temporal entomological data. Global Health Action.

Ryan, P. A., Lyons, S. A., Alsemgeest, D., Thomas, P. and Kay, B. H. (2004). Spatial statistical analysis of adult mosquito (Diptera: Culicidae) counts: an example using light trap data, in Redland Shire, Southeastern Queensland, Australia. J. Med. Entomol. 41(6): 1143–56.

Sachs, J. and Malaney, P. (2002). The economic and social burden of malaria. Nature. 415(6872): 680–5.

Sachs, J. and McArthur, J. (2005). The Millennium Project: a plan for meeting the Millennium Development Goals. The Lancet. 365: 347–53.

Saleri, N., Capone, S., Pietra, V., De Iaco, G., Del Punta, V., Rizzi, M., Levi, M., Rouamba, S., Somé, E., Simporé, J., Patroni, A., Carosi, G. and Castelli, F. (2009). Outcome and predictive factors of mortality in hospitalized HIV-patients in Burkina Faso. Infection. 37(2): 142–7.

Sankoh, O. A. and Binka, F. (2005a). INDEPTH Network: a viable platform for the assessment of malaria risk in developing countries. In: Takken, Martens and Bogers (eds), Environmental Change and Malaria Risk-Global and Local Implications. Berlin, Heidelberg: Springer Verlag.

Sankoh, O. A. and Binka, F. N. (2005b). INDEPTH network: generating empirical population and health data in resource-constrained countries in the developing world. In: H. Becher and B. Kouyaté, Editors, Health Research in Developing Countries. A Collaboration Between Burkina Faso and Germany. Springer-Verlag, Heidelberg. 21–32.

Sankoh, O. A., Kynast-Wolf, G., Kouyaté, B. and Becher, H. (2003). Patterns of adult and old-age mortality in rural Burkina Faso. Journal of Public Health. 25(4): 372.

Sartorius, B., Kahn, K., Collinson, M. A., Vounatsou, P. and Tollman, S. M. (2011). Survived infancy but still vulnerable: spatial-temporal trends and risk factors for child mortality in the Agincourt rural sub-district, South Africa, 1992-2007. Geospatial Health. 5(2): 285–95.

Schellenberg, D., Menendez, C., Kahigwa, E., Aponte, J., Vidal, J., Tanner, M., Mshinda, H. and Alonso, P. (2001a). Intermittent treatment for malaria and anaemia control at time of routine vaccinations in Tanzanian infants: a randomised, placebo-controlled trial. The Lancet. 357(9267): 1471–7.

Schellenberg, J. R., Abdulla, S., Nathan, R., Mukasa, O., Marchant, T. J., Kikumbih, N., Mushi, A. K., Mponda, H., Minja, H., Mshinda, H., Tanner, M. and Lengeler, C. (2001b). Effect of large-scale social marketing of insecticide-treated nets on child survival in rural Tanzania. Lancet. 357(9264): 1241–7.

Schur, N., Hürlimann, E., Garba, A., Traoré, M. S., Ndir, O., Ratard, R. C., Tchuenté, L. A. T., Kristensen, T. K., Utzinger, J. and Vounatsou, P. (2011a). Geostatistical Model-Based Estimates of Schistosomiasis Prevalence among Individuals Aged≤ 20 Years in West Africa. PLoS Neglected Tropical Diseases. 5(6): e1194.

Schur, N., Hürlimann, E., Stensgaard, A.-S., Chimfwembe, K., Mushinge, G., Simoonga, C., Kabatereine, N. B., Kristensen, T. K., Utzinger, J. and Vounatsou, P. (2011b). Spatially explicit Schistosoma infection risk in eastern Africa using Bayesian geostatistical modelling. Acta Tropica.

Seeger, M. (2003). "Bayesian Gaussian Process Models: PAC-Bayesian Generalisation Error Bounds and Sparse Approximations." [Online]. Available: http://www.era.lib.ed.ac.uk/handle/1842/321. (Accessed 14 Aug., 2011).

Service, M. W. (1977). A critical review of proceduresfor sampling populations of adult mosquitoes. Bulletin of Entomological Research. 67: 343–82.

Setel, P. W., Sankoh, O., Rao, C., Velkoff, V. A., Mathers, C., Gonghuan, Y., Hemed, Y., Jha, P. and Lopez, A. D. (2005). Sample registration of vital events with verbal autopsy: a renewed commitment to measuring and monitoring vital statistics. Bull. World Health Organ. 83(8): 611–7.

Setel, P. W., Whiting, D. R., Hemed, Y., Chandramohan, D., Wolfson, L. J., Alberti, K. G. M. M. and Lopez, A. D. (2006). Validity of verbal autopsy procedures for determining cause of death in Tanzania. Trop. Med. Int. Health. 11(5): 681–96.

Shabani, J., Lutambi, A. M., Mwakalinga, V. and Masanja, H. (2010). Clustering of under-five mortality in Rufiji Health and Demographic Surveillance System in rural Tanzania. Glob Health Action. 3:

Shah, S. N., Smith, E. E., Obonyo, C. O., Kain, K. C., Bloland, P. B., Slutsker, L. and Hamel, M. J. (2006). HIV immunosuppression and antimalarial efficacy: sulfadoxine-pyrimethamine for the treatment of uncomplicated malaria in HIV-infected adults in Siaya, Kenya. J. Infect. Dis. 194(11): 1519–28.

Shaukat, A., Breman, J. and McKenzie, F. E. (2010). Using the entomological inoculation rate to assess the impact of vector control on malaria parasite transmission and elimination. Malar J. 9(1): 122.

Shililu, J. I., Maier, W. A., Seitz, H. M. and Orago, A. S. (1998). Seasonal density, sporozoite rates and entomological inoculation rates of Anopheles gambiae and Anopheles funestus in a high-altitude sugarcane growing zone in Western Kenya. Trop. Med. Int. Health. 3(9): 706–10.

Shililu, J., Ghebremeskel, T., Mengistu, S., Fekadu, H., Zerom, M., Mbogo, C., Githure, J., Novak, R., Brantly, E. and Beier, J. C. (2003). High seasonal variation in entomologic inoculation rates in Eritrea, a semi-arid region of unstable malaria in Africa. Am. J. Trop. Med. Hyg. 69(6): 607–13.

Shulman, C. E. (1999). Malaria in pregnancy: its relevance to safe-motherhood programmes. Ann Trop Med Parasitol. 93 Suppl 1: S59–66.

Siegel, S. and Castellan, N. J. (1988). Nonparametric statistics for thebehavioral science. (2nd ed.). New York: McGraw-Hill.

Singer, J. D. and Willett, J. B. (1993). It's about Time: Using Discrete-Time Survival Analysis to Study Duration and the Timing of Events. Journal of Educational Statistics. 18(2): 155–95.

Singh, B., Sung, L. K., Matusop, A., Radhakrishnan, A., Shamsul, S. S. G., Cox-Singh, J., Thomas, A. and Conway, D. J. (2004). A large focus of naturally acquired Plasmodium knowlesi infections in human beings. The Lancet. 363(9414): 1017–24.

Singh, G. P., Chavan, B. S., Arun, P. and Sidana, A. (2007). Seasonal pattern of psychiatry service utilization in a tertiary care hospital. Indian J Psychiatry. 49(2): 91–5.

Smith, A. F. M. and Gelfand, A. E. (1992). Bayesian statistics without tears: A sampling-resampling perspective. American Statistician. 84–8.

Smith, D. L. and McKenzie, F. E. (2004). Statics and dynamics of malaria infection in Anopheles mosquitoes. Malar. J. 3: 13.

Smith, D., Dushoff, J., Snow, R. and Hay, S. (2005). The entomological inoculation rate and Plasmodium falciparum infection in African children. Nature. 438(7067): 492.

Smith, T. A., Charlwood, J. D., Kihonda, J., Mwankusye, S., Billingsley, P., Meuwissen, J., Lyimo, E., Takken, W., Teuscher, T. and Tanner, M. (1993). Absence of seasonal variation in malaria parasitaemia in an area of intense seasonal transmission. Acta Trop. 54(1): 55–72.

Smith, T. A., Leuenberger, R. and Lengeler, C. (2001). Child mortality and malaria transmission intensity in Africa. Trends in Parasitology. 17(3): 145–9.

Smith, T., Charlwood, J. D., Takken, W., Tanner, M. and Spiegelhalter, D. J. (1995). Mapping the densities of malaria vectors within a single village. Acta Trop. 59(1): 1–18.

Smith, T., Killeen, G., Lengeler, C. and Tanner, M. (2004). Relationships between the outcome of Plasmodium falciparum infection and the intensity of transmission in Africa. The American journal of tropical medicine and hygiene. 71(2 suppl): 80.

Snow, R. W. and Marsh, K. (1995a). Will reducing Plasmodium falciparum transmission alter malaria mortality among African children? Parasitology Today. 11(5): 188–90.

Snow, R. W. and Marsh, K. (1995b). Will reducing Plasmodium falciparum transmission alter malaria mortality among African children? Parasitology Today. 11(5): 188–90.

Snow, R. W. and Marsh, K. (2002). The consequences of reducing transmission of Plasmodium falciparum in Africa. Adv. Parasitol. 52: 235–64.

Snow, R. W. and Marsh, K. (2010). Malaria in Africa: progress and prospects in the decade since the Abuja Declaration. Lancet. 376(9735): 137–9.

Snow, R. W., Armstrong, J. R., Forster, D., Winstanley, M. T., Marsh, V. M., Newton, C. R., Waruiru, C., Mwangi, I., Winstanley, P. A. and Marsh, K. (1992). Childhood deaths in Africa: uses and limitations of verbal autopsies. Lancet. 340(8815): 351–5.

Snow, R. W., Craig, M. H., Deichmann, U. and le Sueur, D. (1999). A preliminary continental risk map for malaria mortality among African children. Parasitol. Today (Regul. Ed.). 15(3): 99–104.

Snow, R. W., Gouws, E., Omumbo, J., Rapuoda, B., Craig, M. H., Tanser, F. C., le Sueur, D. and Ouma, J. (1998a). Models to predict the intensity of Plasmodium falciparum transmission: applications to the burden of disease in Kenya. Trans. R. Soc. Trop. Med. Hyg. 92(6): 601–6.

Snow, R. W., Guerra, C. A., Noor, A. M., Myint, H. Y. and Hay, S. I. (2005). The global distribution of clinical episodes of Plasmodium falciparum malaria. Nature. 434(7030): 214.

Snow, R. W., Nahlen, B., Palmer, A., Donnelly, C. A., Gupta, S. and Marsh, K. (1998b). Risk of Severe Malaria among African Infants: Direct Evidence of Clinical Protection during Early Infancy. Journal of Infectious Diseases. 177(3): 819–22.

Snow, R. W., Omumbo, J. A., Lowe, B., Molyneux, C. S., Obiero, J.-O., Palmer, A., Weber, M. W., Pinder, M., Nahlen, B., Obonyo, C., Newbold, C., Gupta, S. and Marsh, K. (1997). Relation between severe malaria morbidity in children and level of Plasmodium falciparum transmission in Africa. The Lancet. 349(9066): 1650–4.

Snow, R. W., Trape, J. F. and Marsh, K. (2001). The past, present and future of childhood malaria mortality in Africa. TRENDS in Parasitology. 17(12): 593–7.

Sogoba, N., Vounatsou, P., Doumbia, S., Bagayoko, M., Touré, M. B., Sissoko, I. M., Traore, S. F., Touré, Y. T. and Smith, T. (2007). Spatial analysis of malaria transmission parameters in the rice cultivation area of Office du Niger, Mali. Am. J. Trop. Med. Hyg. 76(6): 1009–15.

Somi, M. F., Butler, J. R., Vahid, F., Njau, J. D., Kachur, S. P. and Abdulla, S. (2008). Use of proxy measures in estimating socioeconomic inequalities in malaria prevalence. Tropical Medicine & International Health. 13(3): 354–64.

Spiegelhalter, D. J., Thomas, A. and Best, N. G. (1999). Winbugs version 1.2 user manual. MRC Biostatistics Unit.

Stein, M. L. (2005). Space-time covariance functions. Journal of the American Statistical Association. 100(469): 310–21.

Stein, M. L., Chi, Z. and Welty, L. J. (2004). Approximating likelihoods for large spatial data sets. J Royal Statistical Soc B. 66(2): 275–96.

Stewart, L., Gosling, R., Griffin, J., Gesase, S., Campo, J., Hashim, R., Masika, P., Mosha, J., Bousema, T., Shekalaghe, S., Cook, J., Corran, P., Ghani, A., Riley, E. M. and Drakeley, C. (2009). Rapid Assessment of Malaria Transmission Using Age-Specific Sero-Conversion Rates. PLoS ONE. 4(6): e6083.

Stolwijk, A. M., Straatman, H. and Zielhuis, G. A. (1999). Studying seasonality by using sine and cosine functions in regression analysis. J Epidemiol Community Health. 53(4): 235–8.

Stresman, G. H., Kamanga, A., Moono, P., Hamapumbu, H., Mharakurwa, S., Kobayashi, T., Moss, W. J. and Shiff, C. (2010). A method of active case detection to target reservoirs of asymptomatic malaria and gametocyte carriers in a rural area in Southern Province, Zambia. Malar J. 9: 265–265.

Suhartono (2011). Time Series Forecasting by using Seasonal Autoregressive Integrated Moving Average: Subset, Multiplicative or Additive Model. Journal of Mathematics and Statistics. 7(1): 20–7.

Tanzania Demographic and Health Survey (2010). National Bureau of Statistics.

Tanner, M. and de Savigny, D. (2008). Malaria eradication back on the table. Bull. World Health Organ. 86(2): 82.

Tanser, F. C., Sharp, B. and le Sueur, D. (2003). Potential effect of climate change on malaria transmission in Africa. The Lancet. 362(9398): 1792–8.

The malERA Consultative Group on Vector Control (2011). A Research Agenda for Malaria Eradication: Vector Control. PLoS Medicine. 8: e1000401.

Thomas, C. J. and Lindsay, S. W. (2000). Local-scale variation in malaria infection amongst rural Gambian children estimated by satellite remote sensing. Trans. R. Soc. Trop. Med. Hyg. 94(2): 159–63.

Thomson, M. C. and Connor, S. J. (2000). Environmental information systems for the control of arthropod vectors of disease. Medical and Veterinary Entomology. 14(3): 227–44.

Thomson, M. C. and Connor, S. J. (2001). The development of malaria early warning systems for Africa. Trends Parasitol. 17(9): 438–45.

Thomson, M. C., Connor, S. J., D'Alessandro, U., Rowlingson, B., Diggle, P., Cresswell, M. and Greenwood, B. (1999). Predicting malaria infection in Gambian children from satellite data and bed net use surveys: the importance of spatial correlation in the interpretation of results. Am. J. Trop. Med. Hyg. 61(1): 2–8.

Thomson, M. C., Herrera, R. G. and Beniston, M. (2008). Seasonal forecasts, climatic change and human health: health and climate. Springer. 232 pp.

Thomson, M. C., Mason, S. J., Phindela, T. and Connor, S. J. (2005). Use of rainfall and sea surface temperature monitoring for malaria early warning in Botswana. Am. J. Trop. Med. Hyg. 73(1): 214.

Thomson, M., Connor, S., Milligan, P. and Flasse, S. (1997). Mapping malaria risk in Africa: What can satellite data contribute? Parasitology Today. 13(8): 313–8.

Tillé, Y. (2006). Sampling algorithms. Springer. 222 pp.

Tollman, S. M. and Zwi, A. B. (2000). Health system reform and the role of field sites based upon demographic and health surveillance. Bull World Health Organ. 78(1): 125–34.

Tong, S., Hu, W., Nicholls, N., Dale, P., MacKenzie, J. S., Patz, J. and McMichael, A. J. (2005). Climatic, high tide and vector variables and the transmission of Ross River virus. Internal Medicine Journal. 35(11): 677–80.

Trape, J. F. and Rogier, C. (1996). Combating malaria morbidity and mortality by reducing transmission. Parasitology today. 12(6): 236–40.

Tseng, F. M., Yu, H. C. and Tzeng, G. H. (2002). Combining neural network model with seasonal time series ARIMA model. Technological Forecasting and Social Change. 69(1): 71–87.

UNICEF (2005). Model IMCI Handbook: Integrated management of childhood illness, WHO, Geneva, 2005.

UNICEF/MICS (1995). Monitoring progress toward the goals of the World Summit for Children. A practical handbook for multiple-indicator surveys (mimeo). New York: Programme Division, UNICEF.

Usher, P. K. (2010). Modelling Malaria Transmission Potential for Climate Scenarios in West Africa and Europe. Earth & E-nvironment. 5: 40–65.

Van Geertruyden, J.-P., Mulenga, M., Mwananyanda, L., Chalwe, V., Moerman, F., Chilengi, R., Kasongo, W., Van Overmeir, C., Dujardin, J.-C., Colebunders, R., Kestens, L. and D'Alessandro, U. (2006). HIV-1 immune suppression and antimalarial treatment outcome in Zambian adults with uncomplicated malaria. J. Infect. Dis. 194(7): 917–25.

Ver Hoef, J. M. and Jansen, J. K. (2007). Space—time zero-inflated count models of Harbor seals. Environmetrics. 18(7): 697–712.

Ver Hoef, J. M., Cressie, N., Fisher, R. N. and Case, T. J. (2001). Uncertainty and spatial linear models for ecological data. Spatial uncertainty for ecology: implications for remote sensing and GIS applications. 214–37.

Verhoeff, F. H., Brabin, B. J., Hart, C. A., Chimsuku, L., Kazembe, P. and Broadhead, R. L. (1999). Increased prevalence of malaria in HIV-infected pregnant women and its implications for malaria control. Trop. Med. Int. Health. 4(1): 5–12.

Vitali, A., Segnalini, M., Bertocchi, L., Bernabucci, U., Nardone, A. and Lacetera, N. (2009). Seasonal pattern of mortality and relationships between mortality and temperature-humidity index in dairy cows. Journal of Dairy Science. 92(8): 3781–90.

Vounatsou, P., Raso, G., Tanner, M., N'goran, E. K. and Utzinger, J. (2009). Bayesian geostatistical modelling for mapping schistosomiasis transmission. Parasitology. 136(13): 1695–705.

WHO (1957). Expert Committee on Malaria. Sixth Report. Geneva: World Health Organization.

WHO (1984). Malaria control as part of primary health care: Report of a WHO Study Group.

WHO (2000). The world health report - Health systems: improving performance.

WHO (2005). World Malaria Report. WHO, Geneva.

WHO (2006). The Africa Malaria Report.

WHO (2007). The world health report - A safer future: global public health security in the 21st century.

WHO (2008a). World Malaria Report.

WHO (2009). Malaria rapid diagnostic test performance results.

WHO (2010a). World Health Statistics.

WHO (2010b). World Malaria Report.

WHO, RBM (2002). What is Malaria.

WHO (2008b). Framework and Standards for Country Health Information Systems. Vol. 2nd Editio. World Health Organization.

WHO/UNICEF/PATH (2010). Malaria Funding and Resource Utilization: The First Decade of Roll Back Malaria.

Wagstaff, A. (2002). Inequality aversion, health inequalities and health achievement. Journal of health economics. 21(4): 627–41.

Wagstaff, A. and Watanabe, N. (2002). What difference does the choice of SES make in health inequality measurement. Technical note, World Bank, Washington, DC.

Walker, A. M. (1996). Confounding by indication. Epidemiology. 7(4): 335–6.

Wanji, S., Tanke, T., Atanga, S. N., Ajonina, C., Nicholas, T. and Fontenille, D. (2003). Anopheles species of the mount Cameroon region: biting habits, feeding behaviour and entomological inoculation rates. Tropical Medicine & International Health. 8(7): 643–9.

Warrell, D. A. and Gillies, H. M. (2002). Essential Malariology. fourth ed. Arnold, Oxford.

Warton, D. I. (2005). Many zeros does not mean zero inflation: comparing the goodness-of-fit of parametric models to multivariate abundance data. Environmetrics. 16(3): 275–89.

Welsh, A. H., Cunningham, R. B., Donnelly, C. F. and Lindenmayer, D. B. (1996). Modelling the abundance of rare species: statistical models for counts with extra zeros. Ecological Modelling. 88(1-3): 297–308.

Wernsdorfer, W. H. and McGregor, I. (1988). The epidemiology of human malaria as an explanation of its distributions including some implications for its control. In: Malaria: Principles and Practice of Malariology. Churchill Livingstone, London. 950–1.

Wheeler, C. and Berkley, S. (2001). Initial lessons from public-private partnerships in drug and vaccine development. BULLETIN-WORLD HEALTH ORGANIZATION. 79(8): 728–34.

Whiley, M. and Wilson, S. P. (2004). Parallel algorithms for Markov chain Monte Carlo methods in latent spatial Gaussian models. Statistics and Computing. 14(3): 171–9.

Whitworth, J., Morgan, D., Quigley, M., Smith, A., Mayanja, B., Eotu, H., Omoding, N., Okongo, M., Malamba, S. and Ojwiya, A. (2000). Effect of HIV-1 and increasing immunosuppression on malaria parasitaemia and clinical episodes in adults in rural Uganda: a cohort study. Lancet. 356(9235): 1051–6.

Wieczorek, J., Guo, Q. and Hijmans, R. (2004). The point-radius method for georeferencing locality descriptions and calculating associated uncertainty. International journal of geographical information science. 18(8): 745–67.

Wikle, C. and Cressie, N. (1999). A dimension-reduced approach to space-time Kalman filtering. Biometrika. 86(4): 815–29.

Woolhouse, M. E. J., Dye, C., Etard, J.-F., Smith, T., Charlwood, J. D., Garnett, G. P., Hagan, P., Hii, J. L. K., Ndhlovu, P. D., Quinnell, R. J., Watts, C. H., Chandiwana, S. K. and Anderson, R. M. (1997). Heterogeneities in the transmission of infectious agents: Implications for the design of control programs. Proc Natl Acad Sci U S A. 94(1): 338–42.

Xia, G. and Gelfand, A. E. (2005). Stationary Process Approximation for the Analysis of Large Spatial Datasets. ISDS, Duke University.

Xia, G., Miranda, M. L. and Gelfand, A. E. (2006). Approximately optimal spatial design approaches for environmental health data. Environmetrics. 17(4): 363–85.

Yamano, T. and Jayne, T. S. (2004). Measuring the Impacts of Working-Age Adult Mortality on Small-Scale Farm Households in Kenya. World Development. 32(1): 91–119.

Yamey, G. (2000). African heads of state promise action against malaria. BMJ. 320(7244): 1228.

Yang, M. T., Sniadecki, N. J. and Chen, C. S. (2007). Geometric Considerations of Micro- to Nanoscale Elastomeric Post Arrays to Study Cellular Traction Forces. Advanced Materials. 19(20): 3119–23.

Yates, F. (1949). Sampling Methods for Censuses and Surveys. London: Griffin.

Yau, K. K. W., Wang, K. and Lee, A. H. (2003). Zero-Inflated Negative Binomial Mixed Regression Modeling of Over-Dispersed Count Data with Extra Zeros. Biometrical Journal. 45(4): 437–52.

Yip, P. S. F. and Yang, K. C. T. (2004). A comparison of seasonal variation between suicide deaths and attempts in Hong Kong SAR. J Affect Disord. 81(3): 251–7.

You, D., Jones, G., Hill, K., Wardlaw, T. and Chopra, M. (2010). Levels and trends in child mortality, 1990-2009. Lancet. 376(9745): 931–3.

Zacarias, O. P. and Andersson, M. (2010). Mapping malaria incidence distribution that accounts for environmental factors in Maputo Province--Mozambique. Malar. J. 9: 79.

Zacarias, O. P. and Andersson, M. (2011). Spatial and temporal patterns of malaria incidence in Mozambique. Malar. J. 10: 189.

Zeger, S. L., Irizarry, R. and Peng, R. D. (2006). On Time Series Analysis of Public Health and Biomedical Data. Annu. Rev. Public. Health. 27(1): 57–79.

Zhang, Y., Bi, P. and Hiller, J. (2007). Climate variations and salmonellosis transmission in Adelaide, South Australia: a comparison between regression models. Int J Biometeorol. 52(3): 179–87.

de Savigny, D. and Binka, F. (2004). Monitoring future impact on malaria burden in sub-Saharan Africa. The American journal of tropical medicine and hygiene. 71(2 suppl): 224.

de Savigny, D., Mayombana, C., Mwageni, E., Masanja, H., Minhaj, A., Mkilindi, Y., Mbuya, C., Kasale, H. and Reid, G. (2004). Care-seeking patterns for fatal malaria in Tanzania. Malaria Journal. 3(1): 27.

Curriculum Vitae

1. Personal Details

Name	**Rumisha**, Susan Fred
Date/ Place of birth	30TH January 1977/ Kahama, Tanzania
Nationality	Tanzanian
Permanent Home Address (Tanzania)	P.O. Box 681 Tanga, Tanzania Telephone: +255 27 2646032 Mobile: +255 68 395 59 02; +255 71 895 59 01
Official Address (Tanzania)	National Institute for Medical Research (NIMR) 2448, Barack Obama/Ocean Road, P.O.BOX 9653 Dar es Salaam, Tanzania Telephone: +255 22 2121400/ Fax: +255 22 2121360 Email: srumisha@nimr.or.tz; siaeli@gmail.com
Official Address (Switzerland)	Swiss Tropical and Public Health Institute Socinstrasse 57, 4051 Basel, Switzerland Telephone:Office: +41 61 284 86 00 / Mobile: +41 76 258 28 95 E-mail: susan.rumisha@unibas.ch; siaeli@gmail.com
Profession	Biostatistics/ Health Information
Broad interests	Health Information Systems, Disease Surveillance and Mapping, Monitoring and Evaluation

2. Education Background:

	College/University	Diploma/ Degree	Year Attended
i.	UNIVERSITÄT BASEL, SWITZERLAND (Bayesian Statistics for large spatio-temporal data)	PHD, EPIDEMIOLOGY	2007 – 2011
ii.	UNIVERSITEIT HASSELT, BELGIUM (MSc. Biostatistics)	MASTERS OF SCIENCE	2006 – 2007
iii.	UNIVERSITEIT HASSELT, BELGIUM (MSc. Applied Statistics)	MASTERS OF SCIENCE	2005 – 2006
iv.	UNIVERSITY OF DAR ES SALAAM (BSc. Mathematics and Statistics)	BACHELOR	1997 – 2000

3. **Language skill:** Kiswahili and English
4. **Computing skills and Software experience:** STATA, EpiInfo/Data, SAS, R, OpenBUGS, Fortran

5. Professional/Work Experience

1. Statistician/Health Information, 2001 –2002: East African Integrated Disease Surveillance Network Project, NIMR, Tanzania
2. Research Scientist (Medical Statistics), 2002 – 2004: Integrated Disease Surveillance and Response Programme, NIMR, Tanzania
3. Research Scientist/Biostatistics, July 2004 – To date: NIMR, Dar es Salaam, Tanzania

6. Publications

6.1. Dissertation / Theses

1. Modelling the seasonal and spatial variation of malaria transmission in relation to mortality in Africa, (2011) PhD Thesis, Swiss Tropical and Public Health Institute, an associated institute of the University of Basel, Basel, Switzerland, 195pp
2. Dealing with Missing Data in Cross Sectional Data on Transport. (2007). Thesis for Master of Science in Biostatistics, Universiteit Hasselt, Belgium, 87pp
3. Disease Mapping: A Spatial-Temporal Analysis of Coronary Artery Bypass Graft at the District level in Belgium. (2006). Thesis for Master of Science in Applied Statistics, Universiteit Hasselt, Belgium, 41pp
4. Identification and Prioritization of Reasons for Low Use of Operations Research, (2000) BSc. Dissertation, University of Dar es Salaam, Tanzania, 38 pp.

6.2. Journal Publications

1. **Rumisha,** S.F., Kasasa, S., Smith, T., Abdulla, S., Masanja, H. & Vounatsou, P. Bayesian Modeling of Large Geostatistical Data to Estimate Seasonal and Spatial Variation of Sporozoite Rate. Spatial and Spatio-temporal Epidemiology, (submitted).
2. Mwanga J.R., Lwambo N.J.S., **Rumisha, S.F.**, Vounatsou P. & Utzinger J. (2013) Dynamics of people's socio-economic status in the face of schistosomiasis control interventions in Ukerewe district, Tanzania. Acta Tropica, (DOI: http://dx.doi.org/10.1016/j.actatropica.2013.01.004)
3. **Rumisha, S.F**., Smith, T., Abdulla, S., Masanja, H. & Vounatsou, P. (2013) Assessing seasonal variations and age patterns in mortality during the first year of life in Tanzania. Acta Tropica; 126:28– 36.
4. **Rumisha**, S.F., Smith, T., Abdulla, S., Masanja, H. & Vounatsou, P. Modeling heterogeneity in malaria transmission using large sparse spatio-temporal entomological data. Global Health Action, (submitted)
5. **Rumisha**, S.F., Smith, T., Masanja, H. , Abdulla, S. & Vounatsou, P. Assessing the relation between child survival and malaria transmission: an analysis of the MTIMBA data in Rufiji DSS, Tanzania. Malaria Journal, (submitted).
6. **Rumisha**, S.F., Smith, T., Masanja, H. , Abdulla, S. & Vounatsou, P. Malaria transmission intensity and mortality after the age of five years, Rufiji DSS, Tanzania, Parasites and Vectors, (submitted).
7. Mboera, L.E., Senkoro, K.P., **Rumisha, S.F.,** Mayala, B.K., Shayo, E.H. & Mlozi, M.R. (2011). Plasmodium falciparum and helminth coinfections among schoolchildren in relation to agro-ecosystems in Mvomero District, Tanzania. Acta Tropica; 120:95-102.
8. Minja H., Nsanzabana C., Maure C., Hoffmann A., **Rumisha** S., Ogundahunsi O., Zicker F., Tanner M. & Launois P. (2011) Impact of health research capacity strengthening in low- and middle-income countries: the case of WHO/TDR programmes, PLoS Negl Trop Dis.;5(10):e1351
9. Mboera, L.E., Shayo, E.H., Senkoro, K.P., **Rumisha, S.F.,** Mlozi, M.R. & Mayala, B.K. (2010) Knowledge, perceiptions and practices of farming communities on linkage of malaria and agriculture in Mvomero district, Tanzania. Acta Tropica.;113(2):139–144.
10. Mboera, L.E., Senkoro, K.P., Mayala, B.K**., Rumisha, S.F**., Rwegoshora, R.T., Mlozi, M.R. & Shayo, E.H. (2010) Spatio-temporal variation in malaria transmission intensity in five agro-ecosystems in Mvomero district, Tanzania. Geospat Health.; 4(2):167-78.

11. Mwingira, V.S., Mayala, B.K, Senkoro, K.P., **Rumisha, S.F**., Shayo, E.H, Mlozi, M.R.S & Mboera, L.E.G. (2009) Mosquito larval productivity in rice-fields infested with Azolla in Mvomero District, Tanzania. Tanzania Journal of Health Research; 11(1):17-22.
12. Mboera, L.E.G., Kamugisha, M.L., **Rumisha, S.F**., Kisinza, W.N., Senkoro, K.P. & Kitua, A.Y. (2008) Malaria and mosquito net utilisation among schoolchildren in villages with or without healthcare facilities at different altitudes in Iringa District, Tanzania. African Health Science 8, 114-119.
13. Jones C.P., Truman BI, Elam-Evans L.D., Jones C.A., Jones C.Y., Jiles R., **Rumisha, S.F**. & Perry GS. (2008) Using "socially assigned race" to probe white advantages in health status. Ethn Dis.; 18(4):496-504.
14. Mboera, L.E.G., **Rumisha, S.F**., Senkoro, K.P., Mayala, B.K., Shayo, E.H. & Kisinza, W.N. (2007) Knowledge and health information communication in Tanzania, East Afr J Public Health.;4 (1):33-9
15. **Rumisha, S.F**., Mboera, L.E.G., Senkoro, K.P., Gueye, D. & Mmbuji, P.K. (2007). Monitoring and evaluation of Integrated Disease Surveillance and Response in selected Districts in Tanzania. Tanzania Health Research Bulletin, 9, 1-11.
16. Mboera, L.E.G., Kamugisha, M.L., **Rumisha, S.F**., Msangeni, H.A., Barongo, V., Molteni, F. & Kitua, A.Y. (2006) The relationship between malaria parasitaemia and availability of healthcare facility in Mpwapwa District, central Tanzania. Tanzania Health Research Bulletin 8, 22-27.
17. **Rumisha, S.F**., Senkoro, K.P. Ngadaya, E., Shayo, E.H., Mayala, B.K., Mtandu, R. & Mboera, L.E.G. (2006) Community knowledge and information communication gaps on HIV/AIDS in Iringa Municipality, Tanzania. Tanzania Health Research Bulletin, 8, 101-108.
18. Mboera, L.E.G., **Rumisha, S.F**., Mwanemile, E.J., Mziwanda, E. & Mmbuji, P. (2005) Enhancing disease surveillance reporting using public transport in Dodoma District, central Tanzania. Tanzania Health Research Bulletin 7, 201-205.
19. Mghamba, J.M., Mboera, L.E.G., Krekamoo, W., Senkoro, K.P., **Rumisha, S.F**., Shayo, E.H. & Mmbuji, P. (2004) Challenges of implementing Integrated Disease Surveillance and Response strategy using the current Health Management Information System in Tanzania. Tanzania Health Research Bulletin 6, 57-63.
20. Mboera, L.E.G. & **Rumisha, S.F**. (2004) The role of mass media in disease outbreak reporting in the United Republic of Tanzania. Tanzania Health Research Bulletin 6, 73-77.
21. **Rumisha, S.F**. (2004). Child violence and sexual abuse in Tanzania. Tanzania Health Research Bulletin.;6(1) 30 – 35.
22. Miller, L.F., Fields, R., Mmbuji, P.K.L., Posner, S., Mboera, L.E.G., Jimmerson, A., Senkoro, K.P., **Rumisha, S.F**., Shayo, E.H. & Mwami, J.A. (2003) Situation Analysis of Infectious Disease Surveillance in two Districts in Tanzania, 2002. Bethesda, MD: The Partners for Health Reformsplus Project, Abt Associates Inc.
23. Shayo, E., Mboera, L.E.G., Mmbuji, P., **Rumisha, S.F**., Senkoro, K.P. & Mwami, A.J. (2003) The role of community and traditional healers in communicable disease surveillance and management in Babati and Dodoma Districts, Tanzania. Tanzania Health Research Bulletin.; 5(2) 48-55.
24. **Rumisha, S.F**., Mboera, L.E.G., Kisinza, W.N., Chuwa, G.J., Kitua, A.Y. & Molteni, F. (2003) Community perceptions of malaria and its management in Iringa District, south-western Tanzania. Tanzania Health Research Bulletin.; 5(2) 41-47.
25. **Rumisha, S.F**., Mboera, L.E.G. Kamugisha, M.L., Kalumuna, A., Amri, M., Asila, E. & Kitua, A.Y. (2003) Pattern and distribution of communicable diseases in border districts of Bukoba and Tanga, Tanzania. Tanzania Health Research Bulletin.; 5, 19-23
26. Mboera, L.E.G., **Rumisha, S.F**., Magesa, S.M. & Kitua, A.Y. (2001) Utilisation of Health Management Information System in disease surveillance in Tanzania. Tanzania Health Research Bulletin.; 3(2), 15-18.
27. Mboera, L.E.G., **Rumisha, S.F**. & Kitua, A.Y. (2001) Strategic approach for strengthening national and regional disease surveillance system: The East African example. Tanzania Health Research Bulletin.; 3(2), 6-9.