CLIMATIC, ENVIRONMENTAL AND SOCIO-ECONOMIC FACTORS FOR MALARIA TRANSMISSION MODELLING IN KWAZULU-NATAL, SOUTH AFRICA

by

OSADOLOR OBIAHON EBHUOMA

Submitted in fulfilment of the academic requirements for the degree of

Doctor of Philosophy

in Environmental Sciences

School of Agricultural, Earth and Environmental Sciences.

College of Agriculture, Engineering and Science

University of KwaZulu-Natal

Westville

South Africa

March 2018

PREFACE

The research contained in this study was completed by the candidate while based in the Discipline of Environmental Sciences, School of Agricultural, Earth and Environmental Sciences of the College of Agriculture, Engineering and Science, University of KwaZulu-Natal, Westville Campus, South Africa. The research was financially supported by College of Agriculture, Engineering and Science, University of KwaZulu-Natal.

The contents of this work have not been submitted in any form to another university and, except where the work of others is acknowledged in the text, the results reported are due to investigations by the candidate.

Signed: Prof. Michael T. Gebreslasie

Date: 12/03/2018

DECLARATION 1: PLAGIARISM

I, Osadolor Obiahon Ebhuoma, declare that:

(i) the research reported in this dissertation, except where otherwise indicated or acknowledged, is my original work;

(ii) this dissertation has not been submitted in full or in part for any degree or examination to any other university;

(iii) this dissertation does not contain other persons' data, pictures, graphs or other information, unless specifically acknowledged as being sourced from other persons;

(iv) this dissertation does not contain other persons' writing, unless specifically acknowledged as being sourced from other researchers. Where other written sources have been quoted, then:

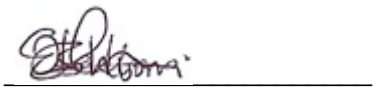
a) their words have been re-written but the general information attributed to them has been referenced;

b) where their exact words have been used, their writing has been placed inside quotation marks, and referenced;

(v) where I have used material for which publications followed, I have indicated in detail my role in the work;

(vi) this dissertation is primarily a collection of material, published as journal articles or presented as a poster and oral presentations at conferences. In some cases, additional material has been included;

(vii) this dissertation does not contain text, graphics or tables copied and pasted from the Internet, unless specifically acknowledged, and the source being detailed in the dissertation and in the References sections.



Signed: Osadolor O. Ebhuoma

Date: 12/03/2018

DECLARATION 2: PUBLICATIONS

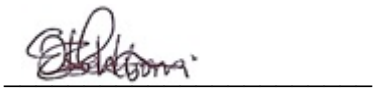
1. **Ebhuoma O, Gebreslasie M** (2016). Remote sensing-driven climatic/environmental variables for modelling malaria transmission in sub-Saharan Africa: a review. *International Journal of Environmental Research and Public Health 13(6): 584.*

2. **Ebhuoma O, Gebreslasie M, Magubane L** (2017). Modelling malaria control intervention effect in KwaZulu-Natal, South Africa using intervention time series analysis. *Journal of Infection and Public Health 10(3): 334-338.*

3. **Ebhuoma O, Gebreslasie M, Magubane L** (In press). A seasonal autoregressive integrated moving average (SARIMA) forecasting model to predict monthly malaria cases in KwaZulu-Natal, South Africa. *South African Medical Journal*

4. **Ebhuoma O, Gebreslasie M, Ropo O** (Submitted). Socio-economic determinants of malaria transmission risk in KwaZulu-Natal, South Africa: a Bayesian inference approach.

5. **Ebhuoma O, Gebreslasie M, Arab A** (Ready for submission). Climatic and environmental determinants for modelling malaria disease risk in a province with low malaria transmission using Bayesian zero-inflated models in INLA.



Signed: Osadolor O. Ebhuoma

Date: 12/03/2018

DECLARATION 3: CONFERENCE PARTICIPATIONS

1. **Ebhuoma O, Gebreslasie M, Manda S** (2015). Remote sensing technology and malaria epidemiological studies in sub-Saharan Africa. Oral presentation at the college of Agricultural, Engineering and Science Research and Innovation Day, 22 September 2015, Pietermaritzburg Campus, UKZN, Pietermaritzburg, South Africa.

2. **Ebhuoma O, Gebreslasie M, Magubane L** (2016). Modelling the efficacy of malaria control strategy in KwaZulu-Natal, South Africa using intervention time series analysis. Oral presentation at the college of Agricultural, Engineering and Science Research and Innovation Day, 29 November 2016, Howard College Campus, UKZN, Durban, South Africa.

3. **Ebhuoma O, Gebreslasie M, Ropo O** (2016). Socio-economic determinants of malaria transmission in a province with low malaria incidence employing a Bayesian modelling approach. Oral presentation at the 3rd National Conference on Global Change, 05 – 08 December 2016, Durban, South Africa.

4. **Ebhuoma O, Gebreslasie M, Ropo O** (2017). Socio-economic determinants of malaria transmission risk in KwaZulu-Natal, South Africa: A Bayesian modelling approach. Oral presentation at the 3rd Annual Congress on Infectious Diseases, 21 – 23 August 2017, San Francisco, USA.

Signed: Osadolor O. Ebhuoma

Date: 12/03/2018

ABSTRACT

Sub-Saharan Africa (SSA) largely bears the burden of the global malaria disease, with the transmission and intensity influenced by the interaction of a variety of climatic, environmental, socio-economic, and human factors. Other factors include parasitic and vectoral factors. In South Africa (SA) in general and KwaZulu-Natal (KZN) in particular, the change of the malaria control intervention policy in 2000, may be responsible for the significant progress over the past two decades in reducing malaria case report to near zero. Currently, malaria incidence in KZN is less than 1 case per 1000 persons at risk placing the province in the malaria elimination stage. To meeting the elimination target, it is necessary to study the dynamics of malaria transmission in KZN employing various analytical/statistical models. Thus, the aim of this study was to explore the factors that influence malaria transmission by employing different analytical models and approaches in a setting with low malaria endemicity and transmission. This involves a sound appraisal of the existing literature on the contribution of remote sensing technology in understanding malaria transmission, evaluation of existing malaria control intervention; delineation of empirical map of malaria risk; provide information on the climatic, environmental and socio-economic factors that influences malaria risk and transmission; and formulation of a relevant malaria forecast and surveillance models. The investigator started with a systemic review of studies in chapter two. The studies were aimed at identifying significant remotely-sensed climatic and environmental determinants of malaria transmission for modelling malaria transmission and risk in SSA via a variety of statistical approaches. Normalised difference vegetation index (NDVI) was identified as the most significant remotely-sensed climatic/environmental determinants of malaria transmission in SSA. Majority of the studies employed the generalised linear modelling approach compared to the Bayesian modelling approach. In the third chapter, malaria cases from the endemic areas of KZN with remotely-sensed climatic and environmental data were used to model the climatic and environmental determinants of malaria transmission and develop a malaria risk map in KZN. The spatiotemporal zero inflated Poisson model formulated indicates that at 95% Bayesian credible interval (BCI) NDVI (0.91; 95% BCI = 0.71, -1.12), precipitation (0.11; 95% BCI = 0.08, 0.14), elevation (0.05; 95% BCI = 0.032, 0.07) and night temperature (0.04; 95% BCI = 0.03, 0.04) are significantly related to malaria transmission in KZN, SA. The area with the highest risk of malaria morbidity in KZN was identified as the north-eastern part of the province. The fourth chapter was to establish the socio-economic status (SES) that influence

malaria transmission in the endemic areas of KZN, by employing a Bayesian inference approach. The obtained posterior samples revealed that, significant association existed between malaria disease and low SES such as illiteracy, unemployment, no toilet facilities and no electricity at 95% BCI Lack of toilet facilities (odds ration (OR) =12.54; 95% BCI = 0.63, 24.38) exhibited the strongest association with malaria and highest risk of malaria disease. This was followed by no education (OR =11.83; 95% BCI = 0.54, 24.27) and lack of electricity supply (OR =10.56; 95% BCI = 0.43, 23.92) respectively. In the fifth chapter, the seasonal autoregressive integrated moving average (SARIMA) intervention time series analysis (ITSA) was employed to model the effect of the malaria control intervention, dichlorodiphenyltrichloroethane (DDT) on confirmed monthly malaria cases. The result is an abrupt and permanent decline of monthly malaria cases ($w0 = -1174.781$, p-value = 0.003) following the implementation of the intervention policy. Finally, the sixth chapter employed a SARIMA modelling approach to predict malaria cases in the endemic areas of KZN. Three plausible models were identified, and based on the goodness of fit statistics and parameter estimation, the SARIMA $(0,1,1)(0,1,1)_{12}$ model was identified as the best fit model. The SARIMA $(0,1,1)(0,1,1)_{12}$ model was used to forecast malaria cases during 2014, and it was observed to fit closely with the reported malaria cases during January to December 2014. The models generated in this study demonstrated the need for the KZN malaria program, relevant policy makers and stakeholders to further strengthen the KZN malaria elimination efforts. The required malaria elimination fortification are not limited to the implementation of additional sustainable developmental approach that combines both improved malaria intervention resources and socio-economic conditions, strengthening of existing community health workers, and strengthening of the already existing cross-border collaborations. However, more studies in the area of statistical modelling as well as practical applications of the generated models are encouraged. These can be accomplished by exploring new avenues via cross-sectional survey to understand the impact of community and social related structures in malaria burden; strengthening of existing community health workers; knowledge, attitude and practices in malaria control and intervention; and the likely effects of temporal/seasonal and spatial variations of malaria incidence in neighbouring endemic countries should be explored.

ACKNOWLEDGMENTS

My unspeakable gratitude goes to The Almighty God, The author and finisher of my faith, my director, my provider, my shield and my buckler for granting me the grace, courage and strength to complete this study. I am nothing without you my Lord.

To my amicable supervisor, Prof Michael Gebreslasie, your efforts, intellectual and academic professionalism were pivotal in making the study a huge success. This was attained through your motivation, tolerance, patience and constructive criticism. Thank you for the knowledge and skills acquired under your guidance.

I appreciate the School of Geography and Environmental Science, College of Agriculture, Engineering and Science, University of KwaZulu-Natal and her staff for providing funding and a conducive environment to conduct my research.

I also appreciate the KwaZulu-Natal malaria program, Jozini and the Statistic South Africa, Durban, for providing the malaria case data and socio-economic data used in this research.

I wish to acknowledge Prof Ali Arab, Lethumusa Magubane and Ropo Ogunsakin for their patience and guidance in the statistical modelling employed in my research.

To my dearest parents, Mr and Mrs G. E. Ebhuoma and siblings, Eromose, Ese, Nosa and Osalumense, words are not enough to express my heart-felt gratitude for the unequivocal and unconditional love, words of encouragements, spiritual, emotional and financial support you lavished on me in the course of the program. Thank you being my pillar and strength.

I am also grateful to my friends and research colleagues, Dr Eze Ideozu, Mr Ighodalo Ebhuoma, Mr Dandyxon Imogirie, Meryl Jagarnath, Lindani Mavimbela, Naeem Agjee, Sifiso Xulu, Nasiphi Ntshanga, Ashlyn Padayachee, Mr. Henry Nwaokoro, Solikume Joel, Mr. Ibijola Stephen Olukayode, Dr. Ekemena Oseghe, Dr. Samson Akpotu, Dr Lawal Isiaka, Dr Ekemini Akpan, Dr Tawanda Manyangadze, and Mr Ochuko Erukainure. Thanks a bunch for having my back at all times and providing a shoulder to lean on in trying moments. Your supports and encouragements were instrumental in the course of this journey.

Finally, my unreserved gratitude in no particular order goes to the families of Bar and Mrs G.O.K. Ebowe, Mr and Mrs Francis Ebhuoma, Mr and Mrs Siemode, Mr Efosa Ebowe and family, Ms Justina Ebowe, Mr Victor Nnadozie and family, The Thompson-Edosah's, Ms Justina Ebhuoma, Mr Patrick Ebhuoma and family, Mrs Imaria Ebhuoma and family, Mr and Mrs Ogedegbe, Mr and Mrs Akpesiri Emorhokpor, Mr and Mrs Jeffery Umukoro for their prayers, emotional and financial support.

TABLE OF CONTENT

Page

LIST OF TABLES

Table **Page**

LIST OF FIGURES

Figure **Page**

LIST OF ABBREVIATIONS

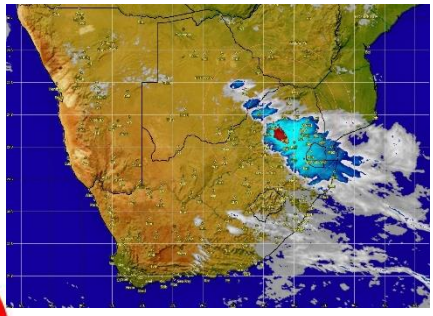
Abbreviations used in this dissertation:

ACF	Autocorrelation functions
ACT	Artemisinin-Based Combination Therapy
ADDS	Africa Data Dissemination Service
AIC	Akaike Information Criterion
AR	Autoregressive
ARIMA	Autoregressive integrated moving average
AVHRR	Advanced Very High Resolution Radiometer
BCI	Bayesian credible interval
BIC	Bayesian information criterion
CAR	Conditional autoregressive
CCD	Cold cloud duration
CL	Confidence level
CMAP	Climate Prediction Centre Merged Analysis of Precipitation
Coef.	Coefficient
DDT	Dichlorodiphenyltrichloroethane
DEM	Digital Elevation Model
DIC	Deviance information criteria
DMI	Dipole model index
DWP	Nocturnal dew point
E8	Elimination 8
EIR	Entomological inoculation rate
ENSO	El Niño-Southern oscillation
ETa	Actual evapotranspiration
ETa	Actual evapotranspiration
ETM+	Enhanced Thematic Mapper plus
EVI	Enhanced vegetation index
GLM	Generalised linear model
GLMM	Generalised linear mixed model
GLOBE	Global Land One-km Base Elevation

GMRF	Gaussian Markov Random Field
GPM	Global Precipitation Measurement mission
INLA	Integrated nested Laplace approximation
ISI	Institute for Scientific Information
ITSA	Intervention time series analysis
JAXA	Japan aerospace exploration agency
KZN	KwaZulu-Natal
LL	Log likelihood
LP DAAC	Land Processes Distributed Active Archive Centre
LSDI	Lubombo Spatial Development Initiative
LST	Land surface temperature
MA	Moving average
MARA/ARMA	Mapping Malaria Risk in Africa
MBG	Model-based geostatistical
MC	Monte Carlo
MCMC	Markov Chain Monte Carlo
MIR	Mid-infrared
MODIS	Moderate-resolution Imaging Spectrometer
MOSASWA	Mozambique, South Africa and Swaziland
MOZIZA	Mozambique, Zimbabwe and South Africa
NCDC	National Climate Data Centre
NCEP	National Centres for Environmental Prediction
NDOH	National Department of Health
NDVI	Normalized difference vegetation index
NDWI	Normalized difference water index
NINO3	Niño 3 region
NOAA	National Oceanic and Atmospheric Administration
OR	Odds ratio
PACF	Partial autocorrelation functions
RDT	Rapid diagnostic test
RFE	Rainfall estimates
RS	Remote sensing
S.E.	Standard error

SA	South Africa
SAR	Seasonal autoregressive
SARIMA	Seasonal autoregressive integrated moving average
SD	Standard deviation
SES	Socio-economic status
SMA	Seasonal moving average
SMAP	Soil Moisture Active/Passive mission
SPDE	Stochastic partial differential equations
SRTM	Shuttle Radar Topography Mission
SSA	Sub-Saharan Africa
SST	Sea surface temperature
SWS	Soil water storage index
TAMSAT	Tropical Application of Meteorology using satellite data and ground-based observations
TFA	Temporal Fourier analysis
TRMM	Tropical rainfall measuring mission
TSI	Temperature suitability index
TWI	Topographic wetness index
USGS	United States Geological Service
WHO	World Health Organisation
ZIB	Zero inflated binomial
ZINB	Zero-inflated negative binomial
ZIP	Zero-inflated Poisson

CHAPTER 1: GENERAL INTRODUCTION



1.1 Background

The mortality and morbidity attributed to malaria disease across sub-Saharan Africa (SSA) region is alarming. Malaria is responsible for majority of the infectious disease related deaths in the region. The latest World Health Organisation (WHO) reports suggest that SSA accounted for 90% of the 212 million malaria cases, and 92% of the estimated 429,000 malaria deaths globally. With respect to the five medically important malaria parasites of the *Plasmodium* species, the infectious stages are transmitted to humans by an infected female *Anopheles* mosquito [1]. However, other factors have been suggested to contribute directly and indirectly to the spatial and temporal heterogeneity of malaria risk and transmission. Some of these factors are ecological, climatic, environmental and socio-economical [2-9]. Considering the spatial and temporal heterogeneity of malaria risk and transmission in endemic areas, it is crucial to investigate and have an in-depth understanding of malaria [10-14]. This in turn can provide insight into how relevant malaria prevention and control parastatals can develop and implement tailor-made interventions.

In South Africa (SA), malaria endemicity is limited to areas in the east bordering Mozambique and the north bordering Zimbabwe. Malaria is endemic in regions lying at altitudes lower than 1 000m above sea level of KwaZulu-Natal (KZN), Limpopo and Mpumalanga provinces [15, 16]. This pre-disposes the approximately 5.4 million people living in these regions to malaria due to little or no immunity they possess [15]. Infants and vulnerable people such as children below 5 years old, the elderly above 65 years old, pregnant and post-partum women, splenectomised people, immunocompromised people including people living with HIV/ AIDS, non-immune travellers and migrants are mainly at high risk of contracting malaria disease [15].

Malaria transmission in KZN is confined to uMkhanyakude, uThungulu and Zululand district municipalities. These areas are bordered by Swaziland and Mozambique to the north and the Indian Ocean stretching from the east down to the southeast [15-17]. The parasite responsible for over 99% of malaria cases in the province is *Plasmodium falciparum*, while *Anopheles arabiensis* and *Anopheles funestus* of the *Anopheles gambiae* complex are the major malaria vector species [16, 17]. The malaria incidence in KZN is low with cases from 0.01 to 0.10 per 1000 persons [18, 19], positioning the province within the malaria elimination trigger in the WHO malaria elimination continuum.

The observed low malaria transmission and burden caused by *Plasmodium falciparum* in KZN can be attributed to the change of SA's malaria vector control policy which included the re-introduction of dichlorodiphenyltrichloroethane (DDT) for malaria vector control purposes after it was discontinued in 1995. DDT is a chemical with insecticide properties but was earlier believed to be unsafe to humans and hence the discontinuation in most countries. In SA, after DDT got reintroduced, it was up-scaled and sustained as an investment in malaria control and intervention in the peak of the 1999/2000 malaria epidemic [15, 20-22]. As an implication, Indoor residual spraying (IRS) with DDT, became the mainstay of KZN's malaria vector control i.e., intervention. Other malaria interventions responsible for the low malaria transmission and burden in the province alongside the use of DDT are active and passive malaria case surveillance using rapid diagnostic test (RDT) or microscopy and treatment with Artemisinin-Based Combination Therapy (ACT)) [16, 17]. In addition, cross-border malaria initiatives and collaborations like the Lubombo Spatial Development Initiative (LSDI) and MOZIZA (Mozambique, Zimbabwe and South Africa) cross-Border malaria initiative were also instrumental in accelerating and sustaining the low local malaria transmission recorded in KZN and across the endemic regions in SA [16, 18]. Malaria cross-border initiative like the MOSASWA (Mozambique, SA and Swaziland) malaria cross-border initiative [18, 23] and Elimination 8 (E8) malaria initiative are regional collaborative ventures currently aiding the sustained low malaria in KZN and SA by addressing imported cases of malaria in addition to already existing malaria intervention resources [19, 23].

Irrespective of the limited malaria transmission in KZN and across SA, the South Africa National Department of Health (SA NDOH) considers malaria as one of the priority disease, because of its propensity to result in an epidemic with a related high illness and death [15, 16]. Accordingly, the KZN malaria programme aims to eliminate malaria in 2020, i.e., no cases of public health importance, and prevent the resurgence of malaria transmission in subsequent years [16, 19]. Pivotal to meeting the elimination target, are sound assessments and understanding of the multiple factors that are influencing and sustaining the low malaria transmission in KZN, and delineate high risk areas of malaria transmission. In addition, formulating surveillance models and systems that provides information on the spatial and temporal distribution of malaria cases are important.

One group of factors that have been suggested by the WHO and SA NDOH that needs to be assessed and monitored in relation to malaria transmission and elimination are climatic and environmental factors [1, 16, 24]. This cannot be overemphasised because previous malaria outbreaks in KZN have been attributed to climatic variations and climatic suitability for malaria vector proliferation [20]. Substantial studies across SSA and other malaria endemic regions of the world have shown variations in the types of climatic and environmental factors that influence malaria transmission [25-27]. With the advent of remote sensing technology, a wide range of climatic and environmental proxy data such as precipitation, rainfall estimate (RFE), day land surface temperature (LST), night LST, vegetation indices such as normalized difference vegetation index (NDVI) and enhanced vegetation index (EVI), actual evapotranspiration (ETa), elevation, landuse and landcover can be obtained with ease for defined locations instead of interpolating and extrapolating data obtained from weather stations [28-33]. In turn this can lead to improved understanding of the dynamics of malaria transmission by collecting climatic and environmental data specific to each location [4, 34-36].

Another group of factors that may have an influential role in the malaria transmission in KZN are socio-economic status (SES) and factors, and yet less attention has been given to understand their influence on malaria transmission in the province [16]. SES and factors suggested to influence malaria transmission across different settings are not limited to age, gender, literacy, knowledge and awareness of vector control measures, sanitary facilities, electricity, employment type and house type [2, 7, 8, 37, 38]. Thus, it is important to evaluate the implications of SES on malaria transmission in KZN. It can equip the relevant authorities and policy makers with the necessary information to improve socio-economic conditions on one hand [2, 8], on the other hand, adopt appropriate malaria intervention strategies [2, 38] in addition to the already existing ones to fortify KZN malaria elimination efforts.

An aspect that also requires close attention is the evaluation of the long-term use of DDT on malaria transmission in KZN [21]. In the face of seemingly DDT ineffectiveness over time due to the significant reduction of malaria cases to consistently low but not zero, it is necessary to determine whether DDT use is still viable while other factors may be needing to be adjusted to ultimately lead to zero malaria transmission in the province. The outcome of this evalution will serve as a validation of the substantive significance of DDT on malaria in KZN. This is vital to the province's malaria control and elimination efforts, because it will bring to light the necessity

of identifying other practical ways that may be required to upscale the existing malaria vector control strategies to achieve zero malaria transmission.

Once the factors that influence malaria transmission in KZN and the areas at highest risk of malaria transmission have been identified, it is vital to continuously monitor the progress of the malaria elimination efforts and develop robust and reliable predictive models [16]. The development of predictive models will strengthen the public health service in decision making for effective targeted malaria transmission combating and elimination strategies. In addition, predictive models constitue a vital tool for malaria surveillance essential to policy makers and public health workers to project the future occurrence of the disease and act proactively [39, 40].

1.2 Aims

As KZN seeks to eliminate malaria in 2020, it is important to explore various mechanisms related to malaria transmission, and evaluate their implications towards continued transmission and risk. This will serve as a road map for the relevant stakeholders (malaria control program, the community health workers, SA NDOH and SA department of social development) to channel intervention programmes and resources sustainably and efficiently in space and time. It will also aid policy makers in refining already laid down policies or generate new policies that will fortify the already existing policies towards actualising KZN's malaria elimination goals. This is in line with the SA malaria elimination strategy drafted by SA NDOH [16]. Thus, the aim of this study was to explore the factors that influence malaria transmission by employing different analytical models and approaches in a setting with low malaria endemicity and transmission.

1.3 Objectives

The specific objectives were:

- To systematically appraise the existing body of literature on RS-derived climatic and environmental determinants of malaria transmission in SSA by identifying determinants peculiar to regions, appraise modelling approaches, and current research shortfalls.
- To identify the significant climatic and environmental determinants of malaria transmission, and delineate the malarious areas in KZN, SA.

- To determine the socio-economic factors that influence malaria transmission at local municipality level in the malarious regions in KZN, SA.
- To evaluate the significance of the malaria control intervention (the use of DDT) on malaria transmission in the malarious regions in KZN, SA.
- To develop a forecasting model to predict malaria in the malarious regions in KZN, SA.

1.4 Outline of study

This study comprises of seven chapters adopting the manuscript format. It has a general introduction to the study (Chapter 1), literature review (Chapter 2), four research chapters (Chapters 3 – 6) and general conclusions and recommendations (Chapter 7). Five of the seven chapters were prepared in a peer-reviewed publication format. Chapters 2 and 5 are published in peer-reviewed journals, Chapter 6 has been accepted for publication in a peer-reviewed journal, Chapter 5 is under review in a peer-reviewed journal and Chapter 3 is ready for submission to a relevant peer-reviewed journal.

Chapter 1, 'General Introduction' provides an introduction of the malaria situation and burden in KZN, SA, and the justification for the study.

Chapter 2, 'Literature review' explores the vast array of studies conducted in SSA with a view to appraise the utilisation and applications of RS technology in enhancing the understanding of malaria transmission dynamics in the region with a focus on RS-driven climatic and environmental variables. Detailed assessment, evaluation and understanding of this technology in relation to malaria were discussed in detail to harness its potential, which in turn, would enhance spatio-temporal risk modelling and identification of reliable malaria transmission predictor variables.

Chapter 3, 'Climatic and environmental determinants for modelling malaria disease risk in a province with low malaria transmission using Bayesian zero-inflated models in INLA' describes and evaluates the application of remote sensing derived variables such as, precipitation, day and night LST, EVI, NDVI, and elevation in identifying relevant climatic and environmental determinants of malaria transmission and develop a malaria risk map in the malarious areas of KZN using a Bayesian spatiotemporal zero inflated model.

Chapter 4, 'Socio-economic determinants of malaria transmission risk in KwaZulu-Natal, SA: a Bayesian approach' describes the influence of socio-economic risk factors such as for malaria incidence at local municipality level in KZN, SA. The demographic and socio-economic variables explored were: gender, children (less than 5 years old), elderly (above 65 years old), no education, no electricity, no toilet facilities, unemployment. In addition, the possible ways the relevant agencies and policy makers can improve socio-economic conditions as a means of malaria control intervention alongside adopting appropriate malaria intervention strategies in addition to the already existing ones were assessed.

Chapter 5, 'Modelling Malaria Control Intervention effect in KwaZulu-Natal, SA using intervention time series analysis' evaluates if the long-term use of DDT significantly lead to low and sustained malaria transmission in KZN. In addition, the possible ways the province can strengthen her already existing malaria control and elimination efforts, to achieve zero malaria transmission were assessed.

Chapter 6, 'A Seasonal Autoregressive Integrated Moving Average (SARIMA) forecasting model to predict monthly malaria cases in KwaZulu-Natal, South Africa' develops a SARIMA temporal model using long-term historical malaria case data and predict malaria monthly cases.

Chapter 7, 'General conclusions and recommendations' provides a brief analysis of the implications of the study outcomes and suggestions for future studies.

1.5 References

1. World Health Organisation. *World malaria report 2016*. 2016. Accessed 10 January 2017; Available from: http://www.who.int/malaria/publications/world-malaria-report-2016/en/.

2. Tusting, L.S., B. Willey, H. Lucas, J. Thompson, H.T. Kafy, R. Smith, and S.W. Lindsay, *Socioeconomic development as an intervention against malaria: a systematic review and meta-analysis.* The Lancet, 2013. **382**(9896): p. 963-972.

3. Tatem, A.J., P.W. Gething, D.L. Smith, and S.I. Hay, *Urbanisation and the global malaria recession.* Malaria Journal, 2013. **12**(1): p. 133.

4. Ebhuoma, O. and M. Gebreslasie, *Remote Sensing-Driven Climatic/Environmental Variables for Modelling Malaria Transmission in Sub-Saharan Africa.* International Journal of Environmental Research and Public Health, 2016. **13**(6): p. 584.

5. Matthys, B., E.K. N'Goran, M. Kone, B.G. Koudou, P. Vounatsou, G. Cisse, A.B. Tschannen, M. Tanner, and J. Utzinger, *Urban agricultural land use and characterisation of mosquito larval habitats in a medium-sized town of Cote d'Ivoire.* Journal of Vector Ecology, 2006. **31**(2): p. 319-33.

6. Snow, R., M. Craig, C. Newton, and R. Steketee, *The Public Health Burden of Plasmodium Falciparum Malaria in Africa: Deriving the Numbers (The Disease Control Priorities Project (DCPP) Working Paper Number 11, Washington DC., 2003)*. 2008, Working Paper.

7. Ayele, D.G., T.T. Zewotir, and H.G. Mwambi, *Prevalence and risk factors of malaria in Ethiopia.* Malaria Journal, 2012. **11**.

8. Monteiro, T.H.A., T.d.S.S. Chaves, H.J.d. Matos, N.F.d.L. Sofffiatti, R.J.d.P.S.e. Guimarães, L.H.R. Guimarães, A.M.R. Ventura, and R.L.D. Machado, *Basic sanitation, socioeconomic conditions, and degree of risk for the presence and maintenance of malaria in a low-transmission area in the Brazilian Amazon.* Revista da Sociedade Brasileira de Medicina Tropical, 2015. **48**: p. 573-579.

9. Tatem, A.J., C.A. Guerra, C.W. Kabaria, A.M. Noor, and S.I. Hay, *Human population, urban settlement patterns and their impact on Plasmodium falciparum malaria endemicity.* Malaria Journal, 2008. **7**.

10. Walker, M., P. Winskill, M.-G. Basanez, J.M. Mwangangi, C. Mbogo, J.C. Beier, and J.T. Midega, *Temporal and micro-spatial heterogeneity in the distribution of Anopheles vectors of malaria along the Kenyan coast.* Parasites & Vectors, 2013. **6**.

11. Alegana, V.A., P.M. Atkinson, J.A. Wright, R. Kamwi, P. Uusiku, S. Katokele, R.W. Snow, and A.M. Noor, *Estimation of malaria incidence in northern Namibia in 2009 using Bayesian conditional-autoregressive spatial-temporal models.* Spatial and Spatio-temporal Epidemiology, 2013. **7**: p. 25-36.

12. Shimaponda-Mataa, N.M., E. Tembo-Mwase, M. Gebreslasie, T.N.O. Achia, and S. Mukaratirwa, *Modelling the influence of temperature and rainfall on malaria incidence in four endemic provinces of Zambia using semiparametric Poisson regression.* Acta Tropica, 2017. **166**(Supplement C): p. 81-91.

13. Kasasa, S., V. Asoala, L. Gosoniu, F. Anto, M. Adjuik, C. Tindana, T. Smith, S. Owusu-Agyei, and P. Vounatsou, *Spatio-temporal malaria transmission patterns in Navrongo demographic surveillance site, northern Ghana.* Malaria Journal, 2013. **12**: p. 63.

14. Moiroux, N., A. Djenontin, A.S. Bio-Bangana, F. Chandre, V. Corbel, and H. Guis, *Spatio-temporal analysis of abundances of three malaria vector species in southern Benin using zero-truncated models.* Parasites & Vectors, 2014. **7**: p. 103.

15. South Africa National Department of Health, *Guidelines For The Treatment Of Malaria In South Africa-2016*. 2016: Pretoria, South Africa.

16. South Africa National Department of Health, *Republic Of South Africa Malaria Elimination Strategy 2011–2018*. 2012: Pretoria, South Africa

17. World Health Organisation. *Malaria country profile 2016*. 2016. Accessed 11/09/2017; Available from: http://www.who.int/malaria/publications/country-profiles/en/.

18. Raman, J., N. Morris, J. Frean, B. Brooke, L. Blumberg, P. Kruger, A. Mabusa, E. Raswiswi, B. Shandukani, E. Misani, et al., *Reviewing South Africa's malaria elimination strategy (2012–2018): progress, challenges and priorities.* Malaria Journal, 2016. **15**(1): p. 438.

19. Elimination 8. *Annual report 2016*. 2016. Accessed 11/9/2017; Available from: https://malariaelimination8.org/wp-content/uploads/2017/05/e8-annual-report-2016.pdf.

20. Coetzee, M., P. Kruger, R. Hunt, D. Durrheim, J. Urbach, and C. Hansford, *Malaria in South Africa: 110 years of learning to control the disease.* South African Medical Journal, 2013. **103**(10): p. 770-778.

21. Maharaj, R., D. Mthembu, and B. Sharp, *Impact of DDT re-introduction on malaria transmission in KwaZulu-Natal.* South African Medical Journal, 2005. **95**(11): p. 871.

22. Ebhuoma, O., M. Gebreslasie, and L. Magubane, *Modeling malaria control intervention effect in KwaZulu-Natal, South Africa using intervention time series analysis.* Journal of Infection and Public Health, 2017. **10**(3): p. 334-338.

23. Moonasar, D., R. Maharaj, S. Kunene, B. Candrinho, F. Saute, N. Ntshalintshali, and N. Morris, *Towards malaria elimination in the MOSASWA (Mozambique, South Africa and Swaziland) region.* Malaria Journal, 2016. **15**(1): p. 419.

24. World Health Organisation. *Holistic research reducing vector-borne diseases like malaria, dengue and Zika*. 2016 Accessed 10 January 2017; Available from: http://www.who.int/tdr/about/holistic-res-reduc-vb-diseases/en/.

25. Rumisha, S.F., T. Smith, S. Abdulla, H. Masanja, and P. Vounatsou, *Modelling heterogeneity in malaria transmission using large sparse spatio-temporal entomological data.* Global Health Action, 2014. **7**: p. 22682.

26. Xia, J., S. Cai, H. Zhang, W. Lin, Y. Fan, J. Qiu, L. Sun, B. Chang, Z. Zhang, and S. Nie, *Spatial, temporal, and spatiotemporal analysis of malaria in Hubei Province, China from 2004–2011.* Malaria Journal, 2015. **14**(1): p. 145.

27. Zinszer, K., R. Kigozi, K. Charland, G. Dorsey, T.F. Brewer, J.S. Brownstein, M.R. Kamya, and D.L. Buckeridge, *Forecasting malaria in a highly endemic country using environmental and clinical predictors.* Malaria Journal, 2015. **14**(1): p. 1.

28. Gaudart, J., O. Toure, N. Dessay, A.L. Dicko, S. Ranque, L. Forest, J. Demongeot, and O.K. Doumbo, *Modelling malaria incidence with environmental dependency in a locality of Sudanese savannah area, Mali.* Malaria Journal, 2009. **8**.

29. Hay, S., M. Renshaw, S.A. Ochola, A.M. Noor, and R.W. Snow, *Performance of forecasting, warning and detection of malaria epidemics in the highlands of western Kenya.* Trends in Parasitology, 2003. **19**(9): p. 394-9.

30. Graves, P.M., D.E. Osgood, M.C. Thomson, K. Sereke, A. Araia, M. Zerom, P. Ceccato, M. Bell, J. Del Corral, S. Ghebreselassie, et al., *Effectiveness of malaria control during changing climate conditions in Eritrea, 1998-2003.* Tropical Medicine & International Health, 2008. **13**(2): p. 218-28.

31. Nygren, D., C. Stoyanov, C. Lewold, F. Mansson, J. Miller, A. Kamanga, and C.J. Shiff, *Remotely-sensed, nocturnal, dew point correlates with malaria transmission in Southern Province, Zambia: a time-series study.* Malaria Journal, 2014. **13**: p. 231.

32. Gosoniu, L., P. Vounatsou, N. Sogoba, and T. Smith, *Bayesian modelling of geostatistical malaria risk data.* Geospatial Health, 2006. **1**(1): p. 127-39.

33. Gosoniu, L., A.M. Veta, and P. Vounatsou, *Bayesian geostatistical modeling of Malaria Indicator Survey data in Angola.* PLoS One, 2010. **5**(3): p. e9322.

34. Palaniyandi, M., *The role of Remote Sensing and GIS for spatial prediction of vector-borne diseases transmission: A systematic review.* Journal of Vector Borne Diseases, 2012. **49**(4): p. 197.

35. Palaniyandi, M., P. Anand, and R. Maniyosai, *Spatial cognition: a geospatial analysis of vector borne disease transmission and the environment, using remote sensing and GIS.* International Journal of Mosquito Research, 2014. **1**(3): p. 39-54.

36. Onyiri, N., *Estimating malaria burden in Nigeria: a geostatistical modelling approach.* Geospatial Health, 2015. **10**(2): p. 163-170.

37. Amoran, O.E., *Impact of health education intervention on malaria prevention practices among nursing mothers in rural communities in Nigeria.* Nigerian Medical Journal, 2013. **54**(2): p. 115-122.

38. Obaldia, N., *Determinants of low socio-economic status and risk of Plasmodium vivax malaria infection in Panama (2009–2012): a case–control study.* Malaria Journal, 2015. **14**(1): p. 14.

39. Cunha, G.B.d., J.F. Luitgards-Moura, E.L.M. Naves, A.O. Andrade, A.A. Pereira, and S.T. Milagre, *Use of an artificial neural network to predict the incidence of malaria in the city of Canta, state of Roraima.* Revista da Sociedade Brasileira de Medicina Tropical, 2010. **43**(5): p. 567-570.

40. Zinszer, K., A.D. Verma, K. Charland, T.F. Brewer, J.S. Brownstein, Z. Sun, and D.L. Buckeridge, *A scoping review of malaria forecasting: past work and future directions.* BMJ Open, 2012. **2**(6).

CHAPTER 2: REMOTE SENSING-DRIVEN CLIMATIC AND ENVIRONMENTAL VARIABLES FOR MODELLING MALARIA TRANSMISSION IN SUB-SAHARAN AFRICA: A REVIEW

This chapter is based on:

Ebhuoma O, Gebreslasie M (2016). Remote sensing-driven climatic/environmental variables for modelling malaria transmission in sub-Saharan Africa: a review. International Journal of Environmental Research and Public Health 13(6): 584.

2.1 Abstract

Malaria is a serious public health threat in Sub-Saharan Africa (SSA), and its transmission risk varies in space and time. Modelling its geographic characteristics is essential for identifying the spatial and temporal risk of malaria transmission. Remote sensing (RS) has been serving as an important tool in providing and assessing a variety of potential climatic and environmental malaria transmission variables in diverse areas. This review focuses on the utilisation of RS-driven climatic and environmental variables in determining malaria transmission in SSA. A systematic search on Google Scholar and the Institute for Scientific Information (ISI) Web of Knowledge^{SM} databases namely PubMed, Web of Science and Science Direct was carried out. The investigator identified thirty-five peer-reviewed articles that studied the relationship between remotely-sensed climatic variable(s) and malaria case data in the SSA sub-regions. The relationship between malaria case data and different climatic and environmental proxies was examined using different statistical methods. Across the SSA sub-region, the normalised difference vegetation index (NDVI) derived from either the National Oceanic and Atmospheric Administration (NOAA) Advanced Very High Resolution Radiometer (AVHRR) or Moderate-resolution Imaging Spectrometer (MODIS) satellite sensors was most frequently returned as a statistically-significant variable to model both spatial and temporal malaria transmission. Furthermore, generalised linear models such as linear regression, logistic regression and Poisson regression were the most frequently-employed methods of statistical analysis in determining malaria transmission predictor variables in East, Southern and West Africa. By contrast, multivariate analysis was used in Central Africa. The investigator stress that the utilisation of RS in determining reliable malaria transmission predictor variables, and climatic and environmental monitoring variables would require a tailored approach that will have cognisance of the geographical and climatic setting, the stage of malaria elimination continuum, the characteristics of the RS variables and the analytical approach, which in turn, would support the channeling of intervention resources sustainably.

Keywords: *remote sensing, climatic variables, environmental variables, epidemiology, Sub-saharan Africa*

2.2 Introduction

Malaria, of all infectious diseases, remains the number one killer in sub Saharan Africa (SSA) [1]. In 2013, an estimated 198 million malaria cases and 584,000 malaria deaths were recorded. About 90% of the malaria deaths recorded were from the SSA region [2]. Out of the five medically important malaria parasite species identified so far, viz. *Plasmodium falciparum*, *P. vivax*, *P. ovale*, *P. malariae* and *P. knowlesi* [2], *P. falciparum* is the most prevalent of the human malaria parasites in SSA, while *P. vivax* is more common across the Horn of Africa [3]. The spatial and temporal variation of malaria disease is known to be influenced by socio-economic and human, ecological or environmental and climatic factors [4, 5]. The climatic variables suggested to possess a direct and indirect influence on malaria transmission are rainfall, temperature, altitude and humidity [6-12]. Rainfall, in conducive amounts, expands mosquito breeding habitats, which in turn increases mosquito population densities and risk of malaria transmission [13]. While rainfall can increase vector densities, excessive rainfall is also capable of flushing the breeding sites [14]. Accordingly, temperature needs to be conducive, ranging between 15°C and 40°C for the completion of the malaria life cycle, and between 16°C and 33°C for the development and survival of mosquitoes [13, 15]. Regarding bite rates and feeding habits, at 17°C, mosquitoes take a human blood meal every four days, while at 25°C, they feed on humans every two days [13]. Altitude has an indirect relationship with temperature, and as such, areas above 1500 m in Africa have little or no risk of malaria transmission [16]. Relative humidity above 60% does not substantially affect the longevity of mosquitoes, but relative humidity lower than 10% results in death within hours [17], while malaria parasites develop between 55% and 80% humidity [18]. The aforementioned climatic variables have been shown to be important malaria transmission indicators that can be used to determine and predict the spatial and temporal distribution of the disease. Consequently, this can guide malaria control managers in decision and policy making in distributing cost-effective intervention resources in time and space [19].

In line with modelling and prediction practicalities, the availability of socio-economic or human and ecological, environmental and climate data has aroused a wide interest in the development of reliable malaria risk maps, forecast models or integrated malaria early warning systems. These data combine with historical malaria case data [6, 8-12, 20] to guide proper channelling of intervention resources before an epidemic occurs [19]. This further indicates the importance of acquiring and applicability of utilising historical climatic variables to adequately study and

understand the role they play in the temporal and spatial heterogeneity of malaria. Meteorological stations have been a long standing source of historical climatic data useful for identifying and modelling malaria transmission [21]. However, they are not only sparsely locate but also malfunctioning and limited numbers of meteorological stations make it challenging to obtain historical and spatially-continuous observations of climatic and environmental variables on a wider geographical scale in SSA [22]. Therefore, there is a need to search for and acquire alternative, indirect or proxy data as remote sensing (RS) becomes essential [20, 22].

With the emergence of RS satellites, a wide array of environmental variables at different spatial and temporal scales [23-28], are now easily accessible facilitating the analytical processes for the association of these factors and malaria [4, 29, 30]. Detailed assessment, evaluation and understanding of this technology in relation to malaria is needed to adequately harness its potential, which in turn, would enhance spatial risk modelling and identification of reliable malaria transmission predictor variables. Therefore, the aim of this review is to appraise the utilisation and applications of RS technology and to discuss its contribution in enhancing the understanding of malaria transmission dynamics in SSA with a focus on RS-driven climatic and environmental variables. This paper will serve as a framework for health practitioners and researchers aiming to identify relevant climatic and environmental variables that are highly related to malaria in particular localities and regions in SSA.

2.3 Methodology

2.3.1 Search strategy

A systematic search to retrieve relevant literature and referenced articles began in September 2014, and the final search was conducted in March 2015. The search was aimed at identifying epidemiological studies in SSA that utilised RS-derived climatic and environmental variables in mapping, modelling or forecasting malaria by carrying out a search on Google Scholar and the ISI Web of Knowledge^{SM} databases: PubMed, Web of Science and ScienceDirect. The database queries were formulated using Boolean operators to combine two or more keywords. The keywords were identified and selected from public and environmental health studies, epidemiological studies and subject headings. The keywords were “remote sensing”, “geographical information system”, “Earth observation”, “spatial techniques”, “geo-spatial analysis”, “geo-spatial techniques”, “malaria”, “forecasting”, “modelling”, “mapping”,

"prediction", "epidemic", "climate change", "climatic factors", "climatic variables", "environmental proxies", "temperature", "rainfall", "normalized difference vegetation index (NDVI)", "humidity", "EL Nino Southern Oscillation", "West Africa", "Central Africa", "East Africa", "Southern Africa" and "Sub-Saharan Africa". Titles and abstracts were initially examined to determine their relevance. Thereafter, the full texts were downloaded to ascertain if they met the selection criteria listed below. Finally, the reference list of each relevant article was assessed to identify other relevant article(s). The search strategy, screening and selection processes are illustrated in Figure 2.1.

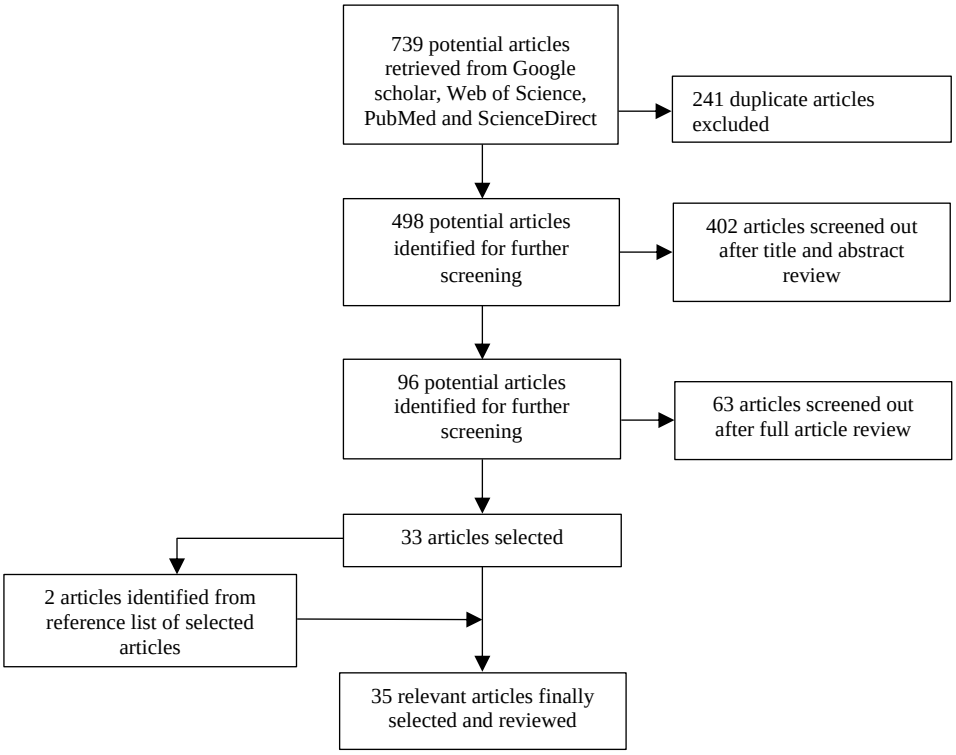


Figure 2.1 Flow chart of publication screening and selection processes for the use of remotely sensed variables for malaria modelling and mapping in sub-Saharan Africa

2.3.2 Selection criteria

The selection criteria involved *post hoc* inclusion and exclusion criteria suggested by Arksey *et al.* [31] and Levac *et al.* [32]. They were developed based on familiarity with the subject matter through reading articles and reviews around malaria epidemiology. The authors discussed and agreed on the study inclusion and exclusion criteria at the beginning of the selection process, and various stages of the conceptual review stages and the selection criteria were refined until

the final selection criteria were accepted. This enabled us to eliminate studies that were outside the scope of our study aim and ensured consistency. The articles finally selected were:

1. Original peer-reviewed articles published in English between 1 January 2000 and 31 December 2014. The search period was selected because since 2000, robust appreciation and application of RS in malaria studies occurred, which can be attributed to the easy access of RS data and the emergence of improved remote sensing sensors. Furthermore, this period coincides with the availability of moderate resolution imaging spectrometer (MODIS) data [33].

2. Articles that applied RS-derived climatic and environmental variables and/or climatic proxy indicators in evaluating malaria risk, distribution, transmission and mapping.

3. Studies that assessed the impact of inter-annual climate variability on malaria transmission. Studies in which climate change projections were used to estimate future malaria distribution were excluded.

4. Publications that used malaria incidence and/or prevalence data in their epidemiological study design (descriptive/explorative, spatial and/or temporal analysis and time series analysis). Studies that used only entomological data were excluded.

5. Studies conducted in Sub-Saharan Africa. Continental-wide studies were excluded, because many African countries have made significant progress in fighting malaria, and malaria is clustered in small areas.

2.3.3 Description of the study region

SSA can be sub-divided into four regions (East, West, Central and Southern Africa) as shown in Figure 2.2. Malaria is endemic in a substantial part of SSA where the climate supports 20%–100% suitability (Figure 2.3) [34]. At the fringes of this region, there are areas where malaria rarely occurs, because the climate is not always suitable. Nevertheless, variation in weather or climatic conditions could instigate an epidemic. The changes in climatic conditions are normally due to higher than normal rainfall and temperature in desert and highland fringes, respectively [34, 35].

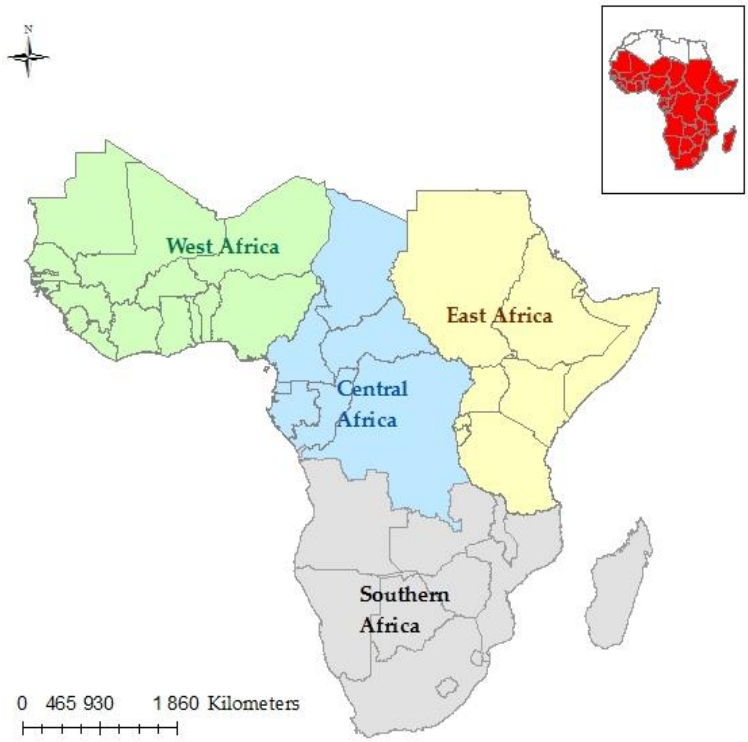


Figure 2.2 Map of SSA showing study region

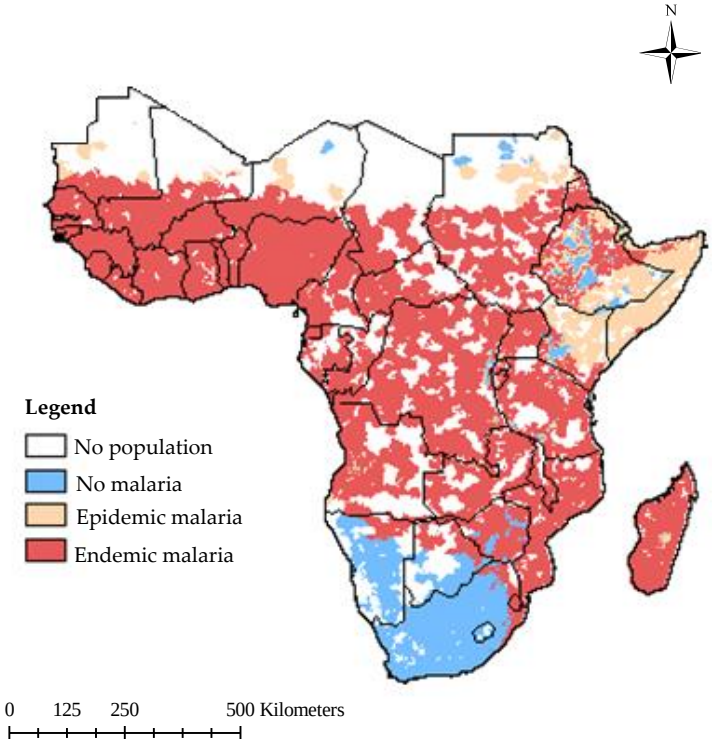


Figure 2.3 Malaria risk stratification of SSA (MARA/ARMA [34])

The climate of the endemic parts of East Africa favours a seasonal and perennial transmission of malaria, while in substantial parts of Kenya and parts of the Horn of Africa, malaria transmission is strongly seasonally prone to epidemics, and the duration of malaria transmission is between 1–3 months owing to low rainfall and inter-annual variability [34]. The White Nile River, Blue Nile River, Lake Victoria, Lake Albert, Lake Tanganyika and Lake Malawi, coupled with the varying climate, are significant risk factors for endemic and epidemic malaria in the region.

Monthly temperature variations, which peak in the rainy season [36], coupled with rainfall that increases towards the Equator, may be responsible for the highly seasonal and varied malaria suitability in Southern Africa [37]. The climate in the malaria-free within Southern Africa, does not totally support malaria endemicity. However, environmental factors, such as the Orange River, which runs through Lesotho, Namibia and South Africa, and the Zambezi River in Botswana, can potentially support malaria epidemicity, while the parts of Southern Africa endemic to malaria can be supported by suitable climate and water bodies (Limpopo and Zambezi Rivers) that favour seasonal malaria transmission [34].

With regards to West Africa, the endemicity of malaria spans across the whole region substantially, excluding only the desert and semi-desert areas. The region is characterised by the Sahelian, Sudanian, tropical humid and equatorial climates. In this region, temperature increases northwards while rainfall increases southwards [37]. The region supports seasonal (4–6 months) and perennial (7–12 months) malaria transmission [34]. In addition, major water bodies like rivers (Benue, Niger, Volta and Senegal) and lakes (Volta and Chad) can sustain malaria transmission in the region.

In Central Africa is significantly endemic to malaria. Suitable climatic conditions (relatively high and reliable rainfall over the coastal and central parts and a temperature range between 19°C and 28°C) [37] coupled with the Congo River, Lake Tanganyika and Lake Albert, contributes to the perennial transmission of malaria experienced in the region [34].

2.4 Results

Initially, 739 related publications were identified. After thoroughly assessing them according to the search strategy and selection criteria, 35 articles were finally selected. From the selected

articles, 14 out of the 35 study sites were located in East Africa (Table 2.1), followed by Southern Africa with 11 studies (Table 2.2) and then West Africa with nine studies (Table 2.3). The only study that covered the Central Africa region utilised datasets covering both Central and West Africa (Table 2.4). The study area(s), malaria case data, climatic variables and their sources, the statistical methods used and the main findings are provided for each study in Tables 2.1–2.4, while Table 2.5 provides an overview of the RS variables commonly used in SSA, and Table 2.6 provides the characteristics of the satellites/sensors used in the selected articles.

2.4.1 East Africa

In East Africa, most studies were country-specific as follows: Kenya, four studies, Eritrea, two studies. Other study locations were Ethiopia, Burundi and Tanzania. Three cross-national studies were identified. One study used data that cuts across Kenya, Ethiopia and Uganda [38], while two other studies used national data from Kenya, Uganda and Tanzania [39, 40] (Table 2.1). The East African countries identified in Table 2.1 are currently in the control phase of the World Health Organisation (WHO) malaria elimination continuum [41]. Studies conducted in the region mainly used National Oceanic and Atmospheric Administration (NOAA) Advanced Very High Resolution Radiometer (AVHRR) imagery as a source of proxy climatic and environmental variables for modelling malaria transmission both at the country and cross-national level. NDVI was observed to be the most assessed RS-derived variable and also the most statistically-significant malaria transmission predictor variable across East Africa. In the province of Karuzi in Burundi, Gomez-Elipe *et al.* [42] used NDVI extracted from NOAA-AVHRR at an 8 km × 8 km spatial resolution, while rainfall and maximum and minimum temperatures were obtained from the metrological stations. After employing the autoregressive integrated moving average (ARIMA) model, NDVI, rainfall and maximum temperature were observed to correlate with malaria cases, and hence, it constituted the best predicting model (R^2_{adj} = 82%, $p < 0.0001$ and 93% predicting accuracy). Ceccato *et al.* [43] used Spearman's and Pearson's rank correlations to assess the relationship between malaria incidence and climate and environmental variables anomalies (to eliminate the similar seasonal pattern possessed by both dependent and predictor variables) in Eritrea. The climatic and environmental variables used by these authors included NDVI from NOAA-AVHRR at 8 km × 8 km spatial resolution, rainfall estimates (RFE) from Climate Prediction Centre Merged Analysis of Precipitation (CMAP) at a 2.5° × 2.5° grid and rainfall data from metrological stations.

Table 2.1 Overview of studies that used RS-derived climatic variables and malaria epidemiological data in East Africa

Reference	Study Area(s)	Malaria Epidemiological Data	Climatic and Environmental Data Obtained via RS Technology		Environmental/ Climatic Data from Other Sources	Statistical Method(s)	Main Findings
			Climatic and Environmental Data	Source(s) of RS Data			
[44]	Kenya: Western Kenya	Monthly inpatient confirmed cases	Multivariate El Nino Southern Oscillation Index (MENSOI)	NOAA	Monthly rainfall, mean monthly temperature	Time-series technique of spectral density analysis	MENSOI did not influence teleconnection with monthly malaria incidence.
[39]	Kenya Uganda Tanzania	Historic malaria distribution maps	NDVI, MIR, LST CCD, altitude	NOAA-AVH R, Meteosat, USGS-DEM	-	Temporal Fourier analysis (TFA), discriminant analysis	LST was noted to be the best predictor variable of malaria transmission intensity. NDVI and CCD were identified as secondary predictor variables of transmission intensity. Altitude significantly improved the predictions.
[45]	Kenya: Kisii Central, Gucha, Nandi, and Kericho	Malaria cases (outpatients)	RFE	USGS	Seasonal climate forecast	WHO quartile, Cullen and cumulative sum (C-SUM) epidemic detection methods	Rainfall was able to forecast an epidemic one month in advance, but the outcome of seasonal climate forecast was erroneous and unreliable.
[24]	Kenya: Kisii Central, Gucha, Nandi, and Kericho	Malaria cases (outpatients)	RFE	USGS	Seasonal climate forecast	WHO quartile, Cullen and C-SUM epidemic detection methods	Seasonal climate forecasts did not predict the heavy rainfall. Rainfall estimates gave timely and reliable early warning, but monthly surveillance of malaria cases gave no effective warning.
[38]	Kenya Ethiopia Uganda	Malaria cases (outpatients)	Maximum temperature, minimum temperature and monthly rainfall	National Climate data Centre, NOAA	-	*t*-test, WHO Cullen epidemic detection methods, forward stepwise regression	Malaria incidence was significantly associated with monthly rainfall and maximum and minimum temperature at a time lag of 1–2 and 2–5 months, respectively.
[40]	Kenya Uganda Tanzania	Malariometric data from Mapping Malaria Risk in Africa (MARA/ARMA) (children between 0 and 15 years)	NDVI, MIR, LST CCD, altitude, land cover	NOAA-AVHRR, Meteosat, USGS-DEM, Landsat TM	-	TFA, discriminant analysis	NDVI, CCD and water body area were associated with malaria in the dry Ecozone 1. In Ecozone 2 where it was assumed that water was not generally limiting, LST and MIR were most abundant among the predictor variables selected.
[43]	Eritrea	Monthly clinical malaria cases	RFE, NDVI	CMAP, NOAA-AVHRR	Interpolated rainfall gauge data	Spearman and Pearson rank correlations, principal component analysis, non-hierarchical clustering analysis.	NDVI anomalies were highly correlated with malaria incidence anomalies, particularly in the semi-arid north of the country and along the northern Red Sea coast, which is a highly epidemic-prone area. CMAP rainfall correlated with malaria incidence anomalies, with a lead time of 2–3 months; while weather station rainfall correlated with malaria anomalies with a lag of 2 months.
[42]	Burundi: Karuzi	Monthly inpatient confirmed and unconfirmed cases	NDVI	AVHRR-NOAA	Rainfall, minimum and maximum temperature	ARIMA	NDVI, rainfall, mean maximum temperature and number of cases constituted the formation of the best predicting model (R^2_{adj} = 82%, $p < 0.0001$ and 93% forecasting accuracy in the range ±4 cases per 100 inhabitants). NDVI, rainfall and maximum temperature were noted to correlate with malaria cases.

Table 2.1 (continued)

Reference	Study Area(s)	Malaria Epidemiological Data	Climatic and Environmental Data Obtained via RS Technology		Environmental/ Climatic Data from Other Sources	Statistical Method(s)	Main Findings
			Climatic and Environmental Data	Source(s) of RS Data			
[25]	Eritrea	Monthly clinical malaria cases	RFE, NDVI	CMAP NOAA-AVHRR	-	Regression analysis	The Poisson regression analysis showed that CMAP rainfall estimates were significantly associated with malaria with a lead time of 2–3 months in Gash Barka. NDVI showed a similar relationship in Anseba.
[46]	Somalia	Survey of *P. falciparum* parasite rate (PfPR)	EVI	MODIS	Precipitation, temperature, distance to permanent water bodies	Logistic regression models, kriging, Bayesian binomial generalised linear geostatistical models	The non-spatial bivariate logistic regression analysis showed that EVI, precipitation, maximum and minimum temperature and distance to water were highly significantly associated with PfPR. After employing the above covariates in the multivariate Bayesian geostatistical model, only temperature and precipitation remained significant (odds 95% Bayesian confidence interval (BCI)) at the southern part of Somalia.
[47]	Kenya: Nandi and Kisii	Confirmed and unconfirmed, monthly inpatient and outpatient cases	Dipole mode index (DMI), El Nino-Southern Oscillation (ENSO) index Nino 3 region (NINO3)	NOAA	Rainfall	Time series regression, Poisson generalised linear model (GLM), Pearson's correlation	No strong association was found between NINO3 and the number of malaria cases after adjusting for the effect of DMI. Malaria cases increased by 3.4%–17.9% for each 0.1 increase above a DMI threshold value lagged at 3–4 months. Malaria cases increased by 1.4%–10.7% for each 10-mm increase in monthly rainfall lagged at 1–3 months.
[48]	Tanzania	Survey of confirmed malaria cases among children less than 5 years old	LST, NDVI, altitude	MODIS DEM-USGS	Rainfall, permanent water bodies	Multivariate logistic regression, Bayesian kriging	The bivariate analyses showed that altitude was negatively associated with malaria risk at the 5% significance level, indicating that children at above 1500 m had a lower risk of malaria. Rainfall, NDVI, day and night LST were positively associated with parasitemia risk.
[20]	Ethiopia: Amhara region	Monthly confirmed outpatients cases	LST, NDVI, enhanced vegetation index (EVI), actual evapotranspiration (ETa), RFE	MODIS TRMM, NASA, and the Japan Aerospace Exploration Agency (JAXA)	-	Seasonal autoregressive integrated moving average (SARIMA)	RFE, EVI, LST and ETa served as suitable malaria predictor as they improved the model fit, and they revealed a lagged positive association with malaria cases. ETa, which was utilised in malaria epidemiological study for the first time, showed a significant positive correlation with malaria at lags from 1–3 months in 3 of the 12 sites studied. EVI had a 3-month lag at 3 sites, while rainfall lagged by 1–3 months at 5 sites. LST exhibited a positive association lagged by 1–6 month at 6 sites.
[49]	Somalia	Survey of PfPR data among children of 2 to less than 10 years	EVI	MODIS	Annual mean precipitation, temperature suitability index (TSI), distance to larva breeding sites.	Linear regression, Space-time model-based geostatistical (MBG) method	The inclusion of 1 km^2 MODIS EVI (odds ratio (OR) = 0.81, 95% BCI = 0.19–1.44, *p*-value = 0.011) and other covariates (precipitation, floodplains, distance to main water bodies) in the analysis served as the best predictor for PfPR.

Table 2.2 Overview of studies that used RS-derived climatic variables and malaria epidemiological data in Southern Africa

Reference	Study Area(s)	Malaria Epidemiological Data	Climatic and Environmental Data Obtained via RS Technology		Environmental/ Climatic Data from Other Sources	Statistical Method(s)	Main Findings
			Climatic and Environmental Data	Source(s) of RS Data			
[50]	Zimbabwe	Monthly confirmed and unconfirmed cases (children less than 5 years old)	NDVI	NOAA-AVHRR	Rainfall, maximum temperature, minimum temperature, vapour pressure	Bayesian Poisson model	Vapour pressure, rainfall, mean monthly (28–32°C) and maximum temperature (24–28°C), showed a significant positive correlation with malaria incidence, while NDVI, high monthly maximum and minimum temperatures showed a negative association.
[51]	Botswana	Confirmed malaria incidence data	RFE, sea surface temperature (SST)	CMAP	-	Stepwise regression, Spearman's rank order, Pearson's product moment correlation, quadratic test, logistic regression, Mann–Whitney U-tests	Negative anomalies of December–January SSTs were significantly associated with December–January rainfall estimates (Pearson's $R = -0.55$ (-0.76 to -0.22) and Spearman's $R = -0.59$ (-0.81 to -0.18)), as well as with the standardised malaria incidence anomalies and accounted for nearly 25% of the inter-annual variance in malaria incidence.
[52]	Zimbabwe	Annual confirmed and unconfirmed malaria case (children less than 5 years old)	NDVI	NOAA-AVHRR (NASA)	Rainfall, vapour pressure, mean temperature, maximum temperature, minimum temperature	Markham's seasonality index, Negative binomial regression analysis, Bayesian negative binomial models	In the bivariate analysis NDVI, vapour pressure, rainfall, average monthly (28°C–32°C) and maximum (24°C–29°C) temperature range revealed a significant positive correlation ($p < 0.001$) with malaria incidence. After employing the spatiotemporal model, NDVI became insignificant.
[53]	Botswana	Confirmed malaria incidence data	RFE	CMAP	SST	Probabilistic prediction, Kolmogorov–Smirnov test, quadratic test	Higher than expected malaria years were associated with above-average rainfall, while the lowest malaria years were associated with below average rainfall.
[54]	Botswana	Malaria prevalence data (children between 1 and 14 years age)	NDVI, RFE	NOAA-AVHRR, CMAP	Elevation, surface water land cover, temperature vapour pressure	Univariate logistic regression analysis, stepwise bootstrap method	RFE (OR = 2.01, 95% BCI = 1.47–2.70), annual mean temperature (OR = 5.75, 95% BCI = 4.14–8.08) and elevation (OR = 1.82, 95% BCI = 1.49–2.22) were significantly associated with malaria prevalence after allowing for spatial correlation.
[28]	Angola	Survey of confirmed malaria cases (children less than 5 years old)	Day LST, night LST, NDVI, altitude	MODIS, USGS-DEM	Rainfall	Bayesian logistic regression, Bayesian kriging	NDVI (95% BCI = 6.28, 17.94; OR = 10.62) and rainfall (95% BCI = 6.00, 19.43; OR = 10.80) showed a significantly positive relationship with malaria incidence after carrying out a bivariate analysis.
[55]	Zambia	Survey of confirmed malaria cases among children less than 5 years old	Day LST, night LST, NDVI, land cover, altitude	MODIS, USGS-DEM	RFE, water bodies (lakes, rivers and wetlands)	Lag time analysis, bivariate and multiple geostatistical logistic regression analysis, Bayesian kriging	NDVI, night LST at 1-km^2 spatial resolution and rainfall within the last 2.7 months showed positive significant association, while day LST reflected a significant negative relationship.

Table 2.2 (continued)

Reference	Study Area(s)	Malaria Epidemiological Data	Climatic and Environmental Data Obtained via RS Technology		Environmental/ Climatic Data from Other Sources	Statistical Method(s)	Main Findings
			Climatic and Environmental Data	Source(s) of RS Data			
[56]	Namibia: Northern Namibia	Monthly confirmed malaria cases	EVI, precipitation	MODIS, TRMM-NASA and JAXA	TSI	Non-spatial Poisson regression, Bayesian spatio-temporal zero-inflated conditional autoregressive (CAR) model, zero-Inflated Poisson (ZIP) model	Initially, the univariate non-spatial regression analysis indicated that the EVI (coefficient of regression, 95% BCI: 6.55, 4.25–8.87, $p < 0.001$), the temperature suitability index acquired from the Malaria Atlas project (7.57, 5.34–9.96, $p < 0.001$) and precipitation (0.02, 0.01–0.03, $p = 0.002$) were significant predictors. However, after employing the best performing predictive model (the multivariate model), only EVI (coefficient of regression, 95% BCI: 14.29, 9.24–19.42, $p < 0.001$) was positively correlated.
[57]	Swaziland	Monthly confirmed malaria cases (imported and locally-acquired)	NDVI, NDWI, elevation, TWI	Landsat-7 ETM+, SRTM	Temperature, rainfall, distance to nearest water body	Satterthwaite *t*-tests, logistic regression mixed model, random forest	Case households during the high transmission season tended to be located in areas of lower elevation, closer to bodies of water, in more sparsely-populated areas, with lower rainfall and warmer temperatures and closer to imported cases than random background points (all $p < 0.001$). In relation to model accuracy, NDWI was the most important RS-derived variable followed by NDVI and, lastly, TWI.
[58]	Malawi	Monthly confirmed and unconfirmed cases	Precipitation, altitude	NOAA Climate Prediction Centre SRTM	Temperature	Negative binomial GLM, generalised linear mixed model (GLMM), Kernel density	The negative binomial with only fixed effects was used to determine the best time lags between climatic variables and malaria. It showed that at the 0.05 significance level, precipitation and temperature were statistically significant at Lag 1–3. The maximum relative malaria risk is observed to be the maximum temperature of 28°C and precipitation of 6.24 $mm \cdot day^{-1}$.
[26]	Zambia: Southern Province	Weekly confirmed malaria cases	Rainfall, NDVI, DWP, LST, elevation	TAMSAT, MODIS, ASTER	-	Kruskal-Wallis tests, Ljung–Box Q statistics, Kriging, ARIMAX	NDVI, DWP and night LST were the highly significant predictor variables at the high and low malaria transmission malaria zones partitioned in the study area.

Table 2.3 Overview of studies that used RS-derived climatic variables and malaria epidemiological data in West Africa

Reference	Study Area(s)	Malaria Epidemiological Data	Climatic and Environmental Data Obtained via RS Technology		Environmental/ Climatic Data from Other Sources	Statistical Method(s)	Main Findings
			Climatic and Environmental Data	Source(s) of RS Data			
[59]	Mali	Malaria prevalence data extracted from the MARA/ARMA database	NDVI	NOAA-AVHRR	Rainfall, average maximum temperature, average minimum temperature, distance to the nearest water body	Logistic regression analysis, kriging	Mean NDVI from June–November (wet season), mean maximum temperature from March–May, months with more than 60 mm of rainfall and distance to water bodies were the significant predictor variables for predicting malaria prevalence.
[60]	Mali	Malaria prevalence data extracted from the MARA/ARMA database	NDVI	NOAA/NASA-AVHRR	Temperature, duration of rainy season, distance to water	Garki mode, Bayesian models and kriging	In the raining season, NDVI and temperature had no statistical relationship with entomological inoculation rate (EIR). Distance to water was significantly related to transmission intensity, indicating high transmission in the areas within 4 km of the water source.
[27]	Mali	Malaria prevalence data from the MARA/ARMA database (children between 1 and 10 years old)	NDVI	NASA-AVHRR	Temperature, rainfall, water bodies, season length	Bayesian logistic regression, Bayesian non-stationary model, Bayesian kriging	The non-stationary model showed that NDVI and minimum temperature had a positive statistical relationship with malaria risk, awhile rainfall had a negative statistical relationship.
[61]	Côte d'Ivoire: Man	Confirmed P. falciparum survey in children between 6 and 16 years	NDVI, LST, RFE	MODIS-USGS Meteosat 7	Distance to the nearest river	Bivariate logistic regression models	In bivariate non-spatial models, NDVI, RFE and distance to rivers, were significantly associated with a *P. falciparum* infection. However, after employing the spatial correlation, NDVI showed only a 'borderline' significance with *P. falciparum* prevalence.
[23]	Mali: Bancoumana	Confirmed P. falciparum survey in children between 0 and 12 years	NDVI	NOAA-AVHRR	-	ARIMA	The seasonal analytical approach revealed that the seasonality of *P. falciparum* incidence was significantly explained by NDVI with s 15-day lag ($p = 0.001$). The NDVI threshold was 0.361 ($p = 0.007$).
[6]	West Africa	MARA/ARMA Malaria prevalence data among children between 1 and 10 years	NDVI, land use	NOAA-AVHRR USGS	Temperature, rainfall, soil water storage index (SWS), water bodies, agro-ecological zones	Logistic regression model, non-parametric regression models	NDVI was not associated with malaria in any of the four defined agro-ecological zones (Equatorial forest, Guinea savannah, Sahel region, Sudanese savannah).
[62]	Côte d'Ivoire: Man	Survey of confirmed malaria cases among school children of Grades 3–5	NDVI, LST, RFE DEM	MODIS-USGS Meteosat 7 SRTM	-	Bayesian negative binomial regression models, Bayesian kriging	The bivariate non-spatial analysis identified NDVI, RFE, LST and close proximity to standing water (rivers, swamps and irrigated fields) as significant risk malaria factors. After employing the spatial analyses, only mean RFE remained significant over the malaria transmission season (June–August).
[63]	Senegal	Confirmed malaria cases among children less than 5 years old	Day LST, night LST, NDVI, altitude	MODIS USGS-DEM	Rainfall, permanent rivers and lakes	Bayesian geostatistical zero-inflated binomial (ZIB), Bayesian kriging	Night LST (OR 1.16; 95% BCI (0.66, 1.86)) and NDVI (OR 1.48; 95% BCI (0.88, 2.48)) were noted to have a positive association with malaria parasitemia.
[64]	Côte d'Ivoire	Malaria prevalence data for children aged less than 16 years old	LST, NDVI Elevation	MODIS, USGS-DEM	Rainfall, distance to the nearest water body	Binomial regression models, Bayesian non-spatial and geo-statistical logistic regression models, Bayesian kriging	In the non-stationary spatial model (the best model), the covariates rainfall (OR = 0.76; BCI = 0.70, 0.83) and maximum LST (OR = 0.72; BCI = 0.64, 0.79) were significantly negatively associated with Plasmodium prevalence.

Table 2.4 Overview of studies that used RS-derived climatic variables and malaria epidemiological data covering Central and Western Africa

Reference	Study Area(s)	Malaria Epidemiological Data	Climatic and Environmental Data Obtained via RS Technology		Environmental/Climatic Data from Other Sources	Statistical Method(s)	Main Findings
			Climatic and Environmental Data	Source(s) of RS Data			
[65]	West Africa and Central Africa	Malaria prevalence data extracted from the MARA/ ARMA database	NDVI, land use	NASA-AVHRR USGS-NASA	Temperature, rainfall, soil water storage index, water bodies, agro-ecological zones, transmission seasonality	Multivariate analysis, Garki model, Bayesian linear geostatistical model, Bayesian kriging	NDVI, distance from water, length of season, rainfall and maximum temperature correlated significantly with malaria transmission intensity and were included in the best fitting model. NDVI had a significant positive association with malaria transmission, except for areas distant from water bodies. This negative association between malaria transmission and distance to water was observed in regions with NDVI values greater than 0.6.

Table 2.5 Commonly-used RS variables for malaria modelling and mapping in SSA

RS Variables	Description	Sources
NDVI	This is an indicator of the greenness of the biomass and varies between −1 and +1. It is calculated as [66, 67]: $\frac{(\mathrm{NIR}-\mathrm{Red})}{(\mathrm{NIR}+\mathrm{Red})}$	MODIS, NOAA-AVHRR
LST (day and night)	This can be estimated from thermal infrared sensors. It is sensitive to the thermal characteristics of the ground and atmospheric effects of spectral radiation [68].	MODIS, NOAA-AVHRR
RFE/CCD	This provides indirect estimates of rainfall based on the detection of precipitation particles or the duration a cloud top is below a threshold temperature [69].	TRMM, CMAP, Meteosat
EVI	EVI provides an alternative to NDVI because it improves sensitivity over areas of denser vegetation. It is calculated as [66]: $G\ \frac{(\mathrm{NIR}-\mathrm{Red})}{(\mathrm{NIR}+C1\times\mathrm{Red}-C2\times\mathrm{Blue}+L)}$, where G is a gain factor, $C1$ and $C2$ are aerosol resistance coefficients and L is the canopy background adjustment that addresses nonlinear, differential NIR and red radiant transfer through a canopy.	MODIS
Elevation/altitude	This correlates negatively with temperature and positively with precipitation and can be applied as a surrogate indicator [69].	USGS-DEM, ASTER, SRTM
Land use and land cover	This is related to the natural and physical environment and the human activities on the landscape [66].	MODIS, Landsat TM, USGS-NASA

Table 2.6 Overview of the RS satellites and sensors used in the malaria epidemiology studies in SSA.

Satellite/Sensors	Spectral Range	Spatial Resolution	Revisit Time	Swath Width	Radiometric Resolution
NOAA/NASA-AVHRR	0.58–12.50 µm	1.1 km	12 h	2900 km	10 bit
MODIS	0.40–14.50 µm	250 m, 500 m, 1 km	1–2 days	2330 km	12 bit
Landsat TM [1]	0.45–12.5 µm	30 m, 120 m	16 days	185 km	8 bit
Landsat-7 ETM+ [2]	0.45–12.5 µm	15 m, 30 m, 60 m	16 days	185 km	9 bit (8 bit transmitted)
Meteosat 1–7	0.50–12.5 µm	2.5 km, 5 km	30 min	-	8 bit
Meteosat 8–10	0.40–14.40 µm	1 km, 3 km	15 min		10 bit
TRMM	VIRS [3]: 0.63 µm, 1.60 µm, 3.75 µm, 10.7 µm, and 12 µm	VIRS: 2 km TMI [4]: 5–45 km PR [5]: 4.3 km	3 hourly, daily, monthly	VIRS: 720 km TMI: 780 km PR: 215 km	-
SRTM	-	30 m	16 times per day	C-radar: 225 km X-radar: 50 km	C-radar: 8 bit X-radar: 6 bit
ASTER	VNIR [6]: 0.52–0.86 µm SWIR [7]: 1.60–2.43 µm TIR [8]: 8.125–11.65 µm	VNIR: 15 m SWIR: 30 m TIR: 90 m	5 days 16 days 16 days	60 km 60 km 60 km	VNIR: 8 bit SWIR: 8 bit TIR: 12 bit
CMAP	-	0.25° × 0.25°	5 days, monthly	-	-

[1] Thematic Mapper; [2] Enhanced Thematic Mapper plus; [3] Visible Infrared Scanner; [4] TRMM Microwave Imager; [5] Precipitation Radar; [6] Visible Near Infrared; [7] Shortwave Infrared; [8] Thermal Infrared.

NDVI anomalies were highly correlated with malaria incidence anomalies, particularly in the semi-arid north of the country and along the northern Red Sea coast, which is a highly epidemic-prone area. CMAP rainfall correlated with malaria incidence anomalies, with a lead time of 2–3 months, while weather station rainfall correlated with malaria anomalies with a delay of two months. Generally, the correlation coefficients were between 0.6 and 0.8. Similarly, Graves *et al.* [25] analysed the effects of impregnated nets, larval control, malathion and DDT on malaria cases, while analysing the effects of RS-derived climate variables, such as NOAA-AVHRR NDVI (8 km × 8 km spatial resolution) and CMAP RFE (2.5° × 2.5° grid) in Eritrea at the district level. The Poisson regression analysis employed showed that the relation between the climatic variables and malaria cases varied by zones. The increase in malaria cases was significantly associated with RFE with a lead time of 2–3 months (0.0007711, $p < 0.001$) in the Gash Barka zone and NDVI anomalies in the current and previous months (1.820668, $p < 0.0001$). NDVI also exhibited the same relationship in the Anseba zone, but with a greater coefficient (11.22517, $p < 0.001$). Gosoniu *et al.* [48] employed Bayesian geostatistical models to analyse the effects of parasitemia risk with age, socio-economic status (wealth index and residence), malaria intervention (bed nets) and climatic and environmental factors (Moderate Resolution Imaging Spectrometer (MODIS) land surface temperature (LST), MODIS NDVI, altitude from the United States Geological Service (USGS) digital elevation model (DEM), RFE data from the Meteosat-7 satellite obtained from the Africa Data Dissemination Service (ADDS) and distance to nearest water body obtained from Health Mapper) in Tanzania. Apart from the variable rainfall with a spatial resolution of 8 km × 8 km, the climatic and environmental factors were retrieved at a spatial resolution of 1 km × 1 km. Altitude was negatively associated with malaria risk at the 5% significance level, indicating that children living above 1500 m had a lower risk of malaria, while rainfall, NDVI and day and night LST were positively associated with parasitemia risk. In a study by Omumbo *et al.* [39], malariometric data and RS-derived variables (NDVI, mid-infrared (MIR) reflectance, cold cloud duration (CCD), land surface and air temperature indices and altitude) from Kenya, Uganda and Tanzania were used to update the spatial resolution of their malaria transmission risk map. These authors pre-processed the RS-derived data using the temporal Fourier analysis, and the discriminant analysis that was employed subsequently revealed that NOAA-AVHRR LST was the best predictor of malaria transmission intensity, while NOAA-AVHRR NDVI and CCD derived from the Meteosat satellite were identified as secondary predictors of transmission

intensity. The forecast was significantly improved by altitude derived from USGS-DEM. Areas with moderated malaria were under-forecasted (false negative rate = 27.7%), while malaria-free areas were over-forecasted (false positive rate = 26.3%). In a similar study that used data covering Kenya, Uganda and Tanzania, Omumbo *et al.* [40] discovered that NDVI, CCD and water body areas were associated with malaria in the "dry" Ecozone 1 (arid and highland, with a climate that favors few months of mosquito proliferation). In Ecozone 2 (diverse, with a climate that supports the propagation of mosquitoes for longer transmission seasons), temperature variables were identified as the most abundant variables in the prediction model. The addition of ecological zoning improved the overall model accuracy by 6.1%, and kappa values increased from 0.397–0.477.

2.4.2 Southern Africa

Studies in Southern Africa in which the relationship between RS-derived climatic variables and malaria incidence and/or prevalence that were identified are summarised in Table 2.2. Study sites in Southern Africa mainly included Botswana (three studies), Zimbabwe (two studies) and Zambia (two studies). Other locations included Angola, Namibia, Swaziland and Malawi. Some Southern African countries, including Mozambique, Angola, Zimbabwe and Malawi are in the malaria control stage of the malaria elimination continuum [41]. South Africa, Zambia, Botswana and Namibia are in the pre-elimination stage [41, 70], while Swaziland is in the elimination stage [71]. RS-extracted climatic and environmental variables used in this region were obtained mainly from NOAA-AVHRR and MODIS satellite sensors. NDVI was observed to be the major RS-derived variable linked to malaria transmission in the region followed by RFE. Gosoniu *et al.* [28] fitted Bayesian geostatistical models to assess the effects of malaria intervention (insecticide-treated nets) among children less than 5 years old in Angola between 2006 and 2007 after adjusting for socio-economic status, climatic and environmental factors (MODIS LST, MODIS NDVI, altitude derived from USGS DEM, distance to nearest water body from Health Mapper and rainfall from ADDS). These authors examined the association between malaria incidence and climatic/environmental factors, and found that NDVI (95% Bayesian credible interval (BCI) = 6.28, 17.94; odd ration (OR) = 10.62) and rainfall (95% BCI = 6.00, 19.43; OR = 10.80) had a significantly positive relationship with malaria incidence. Similarly, Riedel *et al.* [55] investigated the relationship that existed between malaria interventions and malaria risk after adjusting for selected RS (MODIS LST, MODIS NDVI,

MODIS land cover, ADDS RFE, altitude from USGS DEM, water bodies) and socio-economic variables in Zambia. A spatially independent and Bayesian geostatistical model was generated that used malaria cases from the Zambia Malaria Indicator Survey conducted in 2006. NDVI, night LST and rainfall (the last 2.7 months) were identified as positive significant predictors of malaria and were fitted in the model. Cohen *et al.* [57] conducted a study aimed at generating a case-based risk map for Swaziland using the 2011 malaria case data obtained from the Swaziland National Malaria Control Programme. Ecological variables, such as rainfall obtained from weather stations, temperature obtained from worldClim database, NDVI, normalised difference water index (NDWI), elevation, topographic wetness index (TWI) and water bodies obtained from the Food and Agriculture Organisation of the United Nations were assessed for their relevance in the formulation of a high spatial and temporal resolution malaria risk map. NDVI and NDWI data were calculated from a high spatial resolution imagery (30m) from the Landsat 7 Enhanced Thematic Mapper plus (ETM+) sensor. The Landsat 7 ETM+ is an improvement of the previous Landsat satellite series that provides medium-resolution multispectral imagery of the Earth's surface [33]. Elevation and TWI were obtained from the Shuttle Radar Topography Mission (SRTM) at 90-m spatial resolution. These authors suggested that during the high transmission season, malaria cases tend to cluster in areas of lower elevation, closer to water bodies, in less populated areas, with lower rainfall and lower temperatures (all $p < 0.001$). In relation to the model accuracy, NDWI was the most important RS-derived predictor followed by NDVI and TWI. Finally, models formulated from the random forest classification were used to produce predicted probability case-based maps. Nygren *et al.* [26] explored the relationship between RS-derived environmental malaria transmission and forecasted malaria cases in the Southern Province of Zambia. The RS-derived variables included MODIS NDVI, MODIS nocturnal dew point (DWP), MODIS LST, rainfall and elevation. The rainfall data were obtained from the Tropical Application of Meteorology using satellite data and ground-based observations (TAMSAT). TAMSAT rainfall data are rainfall estimates obtained from Meteosat (from the thermal infrared channels) and calibrated against rainfall data from rain gauges [72, 73]. In addition, elevation data at 30-m spatial resolution was derived from the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER), which is a sensor on-board the Terra satellite for DEM creation [33]. NDVI, DWP and night LST were the highly significant predictors in the high and low malaria transmission

areas, and the NDVI and DWP improved the ARIMAX models in all areas significantly. The mean average error of the forecast models was between 0.7% and 33.5%.

2.4.3 West Africa

In West Africa, country-specific studies took place mainly in Mali (four studies) and Côte d'Ivoire (three studies). Others studies were conducted in Gambia and Senegal. One regional study used malariometric data obtained from Mapping Malaria Risk in Africa (MARA/ARMA), which covered West Africa, but excluded Cape Verde [6]. All of the West African countries are in the control stage of the WHO malaria elimination continuum, except Cape Verde, which is in the pre-elimination stage [41]. A summary of studies in West Africa that used RS climatic and environmental variables to identify climatic and environmental predictors of malaria is given in Table 2.3. In the region, the most frequently-utilised RS climatic and environmental variables were from NOAA-AVHRR and MODIS sensors with NDVI identified as the major RS climatic predictor of malaria transmission. Giardina *et al.* [63] used malaria prevalence data from Senegal's Malaria Indicator Survey to determine spatially-explicit climatic and environmental variables associated with malaria in Senegal by incorporating Bayesian variable selection methods within a geostatistical framework. The formulated model included night LST (OR = 1.16; 95% BCI (0.66, 1.86)), NDVI (OR = 1.48; 95% BCI (0.88, 2.48)), urban area (OR = 0.19; 95% BCI (0.07, 0.45)) and rural area (OR = 1), and they were noted to have a positive association with malaria parasitemia risk. Similarly, Gosoniu *et al.* [27] estimated the burden of malaria in Mali by using a Bayesian non-stationary model. Malaria prevalence data were extracted from the MARA/ARMA, 1998 database, NDVI from NASA-AVHRR, temperature and rainfall obtained from Hutchinson *et al.* [74], water bodies from World Resources Institute [75] and season length from Gemperli *et al.* [65]. The best sets of variables included in the non-stationary model were NDVI and minimum temperature, which had a positive significant relationship with malaria risk. Contrarily, rainfall had a negative significant relationship. The authors further suggest that stationarity assumptions are vital due to their influence on the significance of environmental parameters and parasitemia risk map. Gaudart *et al.* [23] incorporated RS-derived variables into a temporal model to predict malaria transmission in the locality of Bancoumana, Mali, characterised by Sudanese savannah. Confirmed *P. falciparum* data obtained from a field study of children aged 0-12 years and 15-day composites of NDVI data derived from NOAA-AVHRR between 1981 and 2006 were incorporated in the ARIMA

time series. The analysis revealed that the seasonality of *P. falciparum* incidence was significantly explained by NDVI with a 15-day lag ($p = 0.001$), and the threshold was 0.361 ($p = 0.007$). The deterministic malaria transmission model, with stochastic environmental variables, forecasted an endemoepidemic pattern of malaria, and the value of the adjusted R^2 was 89%. Similarly, in a study conducted by Kleinschmidt *et al.* [59], malaria risk was determined on a large scale by identifying important ecological parameters, and subsequently, a malaria risk map was produced for Mali. These authors used an automatic stepwise variable selection procedure to identify the most reliable predictors of malaria prevalence for the multiple logistic regression model. NDVI from June–November (wet season), mean maximum temperature from March–May, months with more than a 60-mm rainfall and distance to water bodies were the significant predictor variables for predicting malaria prevalence and were incorporated into the final multiple logistic regression model; and finally, a map of malaria risk was formulated. On the other hand, Silue *et al.* [61] used the Bayesian model to produce spatially-explicit risk maps of malaria transmission in Man, Côte d'Ivoire. Initially, these authors analysed the relationship of malaria prevalence data with possible malaria transmission risk factors, including age, use of bed nets, socio-economic status, distance of health facilities, NDVI, rainfall, LST and distance to rivers. NDVI and LST were extracted from MODIS at a 1 × 1 km spatial resolution, while RFE from the Meteosat-7 satellite was obtained from the ADDS at an 8 × 8 km spatial resolution. In bivariate non-spatial models, NDVI, RFE and distance to rivers were significantly associated with a *P. falciparum* infection. However, after employing the spatial correlation analysis, only age was noted to be a significant risk factor for malaria prevalence, while NDVI showed a "borderline" significance.

2.4.4 Central Africa

In the Central African region, the only study identified that examined the association of malaria with RS climatic and environmental characteristics is given in Table 2.4. This is a cross-regional study that used Malaria prevalence data obtained from the MARA/ARMA database and numerous malaria transmission factors, including population density, NDVI, land use, temperature, rainfall, water bodies, soil water storage index, agro-ecological zone and transmission seasonality covering Central and West Africa [65]. The authors discovered that NDVI extracted from the NASA-AVHRR sensor at an 8 × 8-km spatial resolution had a high relationship with malaria across the region, except in areas far away from water bodies.

Furthermore, a negative association was recorded between malaria transmission and distance to water, and this was observed in regions with NDVI values greater than 0.6. The spatial and non-spatial variations were 0.398 and 41.98, respectively. With reference to the WHO malaria elimination continuum, all of the Central African countries are still in the control phase [41], excluding São Tomé and Príncipe, which are currently in the pre-elimination stage.

2.4.5 Commonly-used RS variables and features of satellites and sensors used by the authors in the articles reviewed

Table 2.5 provides an overview of the RS variables commonly used in SSA, while Table 2.6 presents the satellite sensors used in the various selected studies (these sensors have different spatial, spectral, temporal, radiometric and swath width properties). NOAA-AVHRR and MODIS were the most frequently-utilised sources of RS-derived indices, such as NDVI, EVI, LST, ETa and DWP across SSA. In addition, Meteosat [26, 39, 40, 61, 62, 76] and the Climate Prediction Centre Merged Analysis of Precipitation (CMAP) [25, 53, 58, 77, 78] were also used extensively to extract RFE and precipitation data.

2.5 Discussion

This review highlights the contribution of RS technology in modelling malaria transmission and risk in SSA after taking account of potential climatic and environmental variables that can be used to predict malaria transmission. Malaria disease exhibits seasonal and spatial heterogeneity across localities, districts, provinces, countries and also in sub-continental regions. This can be attributed to the complex nature of malaria resulting from the diverse climatic, environmental, social and natural elements supporting the disease. The combination of these factors plays an important role in the endemicity and epidemicity of an area. RS serves as a means of obtaining potential climatic and environmental malaria variables and opens an avenue to better understand and model the environmental and climatic processes fundamentally responsible for the temporal and spatial heterogeneity of malaria disease.

RS has proven to be a vital tool in malaria modelling and prediction. It can contribute to malaria intervention planning and control programs at both local and broad scales and at different malaria risk stratifications. The scarcity of reliable meteorological data, national health policies and priorities, institutional research capacity, availability and the cost of high resolution RS data

for research and public health purposes determines the use of RS in malaria modelling [79]. This notwithstanding, RS-derived variables are gaining widespread acceptance and application in malaria risk modelling in SSA, because the nature and characteristics of a variable of interest can reflect the ecological relevance and contribution to malaria transmission. For instance, RFE provides indirect estimates of rainfall based on the detection of precipitation particles or the duration a cloud top is below a threshold temperature [69]. LST is used as a proxy for temperature, and its values are obtained from land surface emissivity or surface reflectance in relation to their wavelengths and spectral characteristics [68], while NDVI serves as a surrogate for rainfall, temperature, land use and land cover, near-surface humidity and surface water [20, 80]. Thus, RS-derived variables have the potential to provide information that directly exhibits the state of the vector habitat and the potential role that ecology can play in malaria transmission [29, 68].

The robust utilization of RS-derived variables across SSA has shown that malaria predictors and models are peculiar and subject to the influence of the reference data, scale of observation and environmental condition of the study area. For example, in sub-continental East Africa, the investigator observed that NDVI extracted from NOAA-AVHRR at 8 km × 8 km and MODIS at 1 km × 1 km spatial resolutions is an important predictor of malaria transmission at the country level in Kenya, Tanzania [48], Burundi [42] and Eritrea [25, 43]. However, at the local level, in the rich herbaceous and cropland vegetation of the Amhara region, which constitutes the Ethiopian Highlands, NDVI obtained from MODIS at a 1 km × 1 km spatial resolution was not significantly related to malaria. Instead, ETa (which was only recently assessed for its relevance in malaria risk profiling), EVI and LST variables extracted from MODIS at a 1 km × 1 km spatial resolution were observed to be the suitable malaria predictors in Amhara, Ethiopia [20]. This can be explained by the fact that NDVI loses sensitivity in areas of higher vegetation density and at higher EVI values. The vegetation index EVI can be used as a substitute for NDVI, because it preserves more sensitivity over heavier vegetation; hence, good account of the variation and the change in a rich canopy can be recorded [81, 82]. However, it was observed that the application of EVI in malaria risk profiling and modelling was used sparingly in the East African sub-region and other SSA regions.

The pronounced climatic diversity in relation to malaria suitability at the country level in Southern Africa may have contributed to the diverse RS variables identified as significant malaria predictors in the sub-region. However, NDVI extracted from MODIS at 1 km × 1 km [26, 28] and 0.25 km × 0.25 km [55] and Landsat 7 ETM+ at 30 m × 30 m spatial resolutions [57] can be used to explain the geographical spread of malaria in greater parts of the malaria endemic areas when compared to other RS-derived variables identified as significant predictors of transmission. On the other hand, in the malaria endemic region of Northern Namibia, Alegana *et al.* [56] found that MODIS EVI at a 1 km × 1 km spatial resolution and precipitation derived from TRMM at a 0.25° × 0.25° spatial resolution, which was re-sampled to a 1 km × 1 km spatial resolution, were the best malaria predictors. It must be noted that NDVI was not considered in the study, which would have presented a good comparison with EVI.

NDVI continued to exhibit its dominance in usage and significance pertaining to malaria risk determination across the SSA regions. NDVI extracted from either NOAA-AVHRR at an 8 km × 8 km spatial resolution or MODIS at a 1 km × 1 km spatial resolution, respectively, was identified as a suitable malaria predictor in Western Africa, especially in settings characterised by the Sahelian or Sudanian climate at the local level (Bancoumana, Mali) [23] and the country level (Mali and Senegal) [27, 59, 63]. However, in areas characterised by persistent moisture and heavy vegetation, different outcomes were observed. In the Man region of Côte d'Ivoire, Raso *et al.* [62] identified RFE data obtained from the Meteosat 7 satellite at an 8 km × 8 km spatial resolution as the predictor for malaria prevalence. Furthermore, a significant negative association between *Plasmodium* prevalence and MODIS LST at a 1 km × 1 km spatial resolution was recorded in a study that used data covering Cote d'Ivoire [64].

In the Central African region characterised by climatic suitability for malaria proliferation and heavy vegetation, the only study identified and reviewed indicated that NDVI calculated from the NASA-AVHRR sensor at an 8 × 8 km spatial resolution returned a better result for modelling malaria transmission [65]. EVI, which has been suggested to be an alternative predictor over denser vegetation than NDVI, may have been identified as a better malaria predictor in the Central African region, but it was not included in the study. Furthermore, other climatic and environmental factors used in the study may have been identified as suitable malaria predictors, but the authors did not consider the differences that might exist in the

climate-malaria relationship across the study area. In addition, they did not take into account the non-stationary characteristics of malaria data covering large areas. Disregarding this characteristic could result in the wrong specification of the spatial correlation and, therefore, erroneous values of the standard error of the predictors and prediction. In a somewhat similar study conducted by Gosoniu *et al.* [6], the authors used data covering West Africa and addressed the above-mentioned issues by partitioning the study area into four agro-ecological zones and then employing a different non-parametric model in each zone. There is a possibility that more studies in the Central African region may be available, but could not be identified, as the region is dominated by French-speaking countries. Hence, researchers from the region may have their articles published in French.

The tremendous improvements in the RS sensors, better turnaround time and availability of some RS low and medium resolution datasets at no cost [83, 84] may have also contributed to the considerable utilisation of various RS datasets across SSA. The freely available RS datasets obtainable via MODIS and AVHRR satellites can be used to explain the frequent usage of these satellites as compared to RS datasets from other RS sources [33, 85]. Furthermore, RS datasets from MODIS has made it possible to evaluate new and previously unidentified environmental-related malaria predictors. For example, contemporary studies have shown that MODIS DWP at a 5 km × 5 km spatial resolution and MODIS ETa at a 1 km × 1 km spatial resolution can be used to explain and define malaria transmission risk and malaria incidence in the Southern Province of Zambia and the Amhara region of Ethiopia, respectively.

Countries in SSA are at different stages in their fight towards malaria elimination, and this has to be taken into account in line with the characteristics of the RS imagery intended to be used. Low and medium spatial resolution RS data can be useful in studies conducted at national and regional levels in the malaria endemic countries that are still at their malaria control stage, and to derive generalised spatio-temporal models and malaria risk map for robust application of intervention resources. However, in Angola, Botswana, Cape Verde, Namibia, Swaziland and South Africa with significantly low malaria cases [41], high spatial resolution RS data at a local level would be essential to carry out the cluster analysis and detection of foci and hotspots of malaria transmission. This would support adequate monitoring of the disease and delivery of interventions to specific location(s) and/or seasons, ultimately leading to malaria reduction.

New generation satellites, such as Landsat-8, Copernicus: Sentinel-2, the Global Precipitation Measurement (GPM) mission, the Soil Moisture Active/Passive (SMAP) mission, SPOT 6 and SPOT 7 [33], with improved spatial and radiometric resolutions, have potential for malaria transmission and risk modelling, especially in regions where malaria cases are low (Table 2.7). Furthermore, future satellite mission, like Copernicus: Sentinel-3, which would introduce data reliability for long-term monitoring, could also be vital in modelling spatial and temporal malaria transmission and research [83, 84]. The cost of obtaining high spatial resolution datasets remains a challenge in SSA. Hence, lessons can be drawn from the collaborative venture that exists between China and Brazil, which allows their researchers to obtain high spatial resolution data (2.5 m) freely [86].

Table 2.7 Overview of new generation RS satellites and sensors with improved characteristics for malaria modelling.

Satellite/ Sensors	Spectral Range	Spatial Resolution	Revisit Time	Swath Width	Radiometric Resolution
Landsat-8	0.43–12.5 µm	15 m, 30 m, 100 m	16 days	185 km	12 bit
Copernicus: Sentinel-2	0.43–2.28 µm	10 m, 20 m, 60 m	5 days	290 km	12 bit
GPM	-	250 m, 500 m	3 h	120 km, 245 km, 885 km	-
SMAP	-	3 km, 10 km, 40 km	2 days, 3 days	1000 km	-
SPOT 6 and SPOT 7	0.45–0.89 µm	1.5 m, 2 m, 6 m, 8 m	1–5 days	60 km	12 bit

In a bid to identify relevant and potential risk factor(s) or malaria transmission predictors at local, national or regional levels that can be further incorporated into forecast models/early warning systems and malaria risk maps, the statistical methodology employed should accommodate procedures that suit a particular context and setting. According to Tables 2.1–2.4, some classes of generalised linear models (linear regression, logistic regression and Poisson regression) were used frequently in Eastern, Southern and Western Africa as compared to other analytical approaches, but to a varying degree. This can be attributed to the simplistic, flexible and intuitive way this approach accommodates predictors [87]. On the other hand, in the only study identified that was conducted in Central Africa, a multivariate analysis was carried out. This further illustrates the relevance of these approaches in evaluating the relationships between georeferenced environmental variables and prevalence data, identifying potential risk factor(s) and predictor variable(s), explaining the observed variable(s) and forecasting prevalence at

unsampled locations. The reliability of predictive geostatistical models formulated from a multivariate regression analysis is important in malaria mapping, and it depends on the selected variables fitted in the model. Researchers intending to employ either of the above-mentioned approaches should bear in mind that they do not intrinsically consider correlation in the errors [88]. Erroneous serial autocorrelation is likely to result in underestimated standard errors, and in addition, the evaluation of the impact of predictors would be biased. To account for the impact of autocorrelation on estimates, de Jong *et al.* [89] exhibited the relevance of applying heteroscedasticity and autocorrelation consistent estimators. A variety of statistical approaches have been applied across SSA to varying degrees and settings. The exploration and comparison of different statistical approaches and models for a particular setting would be useful in identifying and evaluating prediction accuracy. It will also be useful in identifying approaches that would provide accurate and reliable predictors for either short, long or intermediate prediction [90].

Overall, the quantitative models employed across localities and countries in SSA consistently revealed variability in the relationship between malaria and climatic and environmental variables. However, NDVI was observed to be the most significant predictor of malaria transmission followed by LST and RFE, and thus, they constituted the RS variable(s) that provided the best-fit model. To improve the overall predictive power and model robustness, the investigator recommended the following: (1) Large datasets should be used over longer periods. For example, Nygren *et al.* [26] generated predictive models employing 126 weeks of data. Therefore, it will be difficult to know if the identified RS predictors of malaria transmission and the relationships they found will be sustained over time. (2) The incremental validity approach, which involves incorporating variables as supplemental to an identified predictor, should be practiced, as it can improve the predictive power [91]. For example, the study conducted by Ceccato *et al.* [43] revealed that NDVI predicted about 1%–20% of the variance in the southern and southeastern areas of Eritrea. This means that other RS-derived variables can explain 80%–99% of the variance. However, the addition of other RS-derived variables would be dependent on whether they improve the predictive validity of what the identified predictor predicts. (3) Linear models have been widely used across SSA. However, these models can result in inappropriate static regression and impose unrealistic or general assumptions. Thus, Bayesian

models, which provide extensions of generalised linear models and are formulated to overcome some of the setbacks of linear models, should be employed.

The reviewed studies have shown that RS technology can contribute to the understanding of the complex nature of malaria across SSA. It can provide the potential climatic and environmental variables needed to identify significant spatially-explicit variable(s) associated with malaria risk and transmission. In areas like the Horn of Africa and Kenya, where malaria is highly seasonal, unstable and epidemic, the process of deciding climatic monitoring targets should be handled with caution to avoid the generation of unreliable malaria transmission models. Therefore, the investigator is in support of regular capacity building and multidisciplinary collaboration between relevant departments, e.g., ecology, geography, biological science, epidemiology, entomology, information technology, statisticians, mathematical modellers, public health decision makers and stakeholders, in generating reliable prediction models. Furthermore, although this study can serve as an informative tool for public and environmental health workers, as well as researchers aiming to model potential climatic factors related to malaria and to delineate climate monitoring targets in SSA, some limitations should be noted. Firstly, relevant reports published in languages other than English and/or unpublished reports were excluded from this review. Secondly, studies that used only entomological data were also excluded.

2.6 Conclusions

The investigator conclude that RS technology is a vital tool in determining malaria risk predictors at regional, national and local scales in diverse regions of SSA. This review suggests that the utilization of RS in determining reliable malaria transmission predictors and developing environmental monitoring would require a tailored approach that takes into account the geographical and climatic setting, the stage of the malaria elimination continuum, the characteristics of the RS variables and the analytical approach, which in turn, would support the channelling of intervention resources sustainably. The improvement of this technology has encouraged the acquisition and evaluation of a wide array of historical climatic and environmental variables at different spatial and temporal resolutions depending on the setting and intended usage. This therefore makes RS a relevant tool for identifying reliable climate-related malaria predictors that can be incorporated into an integrated malaria early-warning

system or prediction model. Previously unidentified remotely-sensed variables, such as ETa and DWP, were found to be malaria transmission predictors, and EVI was also noted to be a suitable substitute for NDVI in denser vegetation, which needs to be further explored extensively across relevant localities and regions of SSA. Furthermore, the assessment of different statistical methods and models for a particular location would be useful in identifying and evaluating prediction accuracy depending on the length of prediction. The application of this technology can be further harnessed in generating reliable prediction models by devising means by which relevant skills and training and the easy acquisition of relevant RS-derived variables can be achieved. Therefore, relevant multidisciplinary collaborations, symposiums and capacity development are encouraged.

2.7 Acknowledgments

This study received financial support from the College of Agriculture, Engineering and Science of the University of KwaZulu-Natal through the doctoral research bursary awarded to Osadolor Ebhuoma. We also thank Emmanuella Ebowe for comments on the manuscript.

2.8 Author contributions

Osadolor Ebhuoma and Michael Gebreslasie conceived of and designed the manuscript. Osadolor Ebhuoma performed the literature search. Osadolor Ebhuoma and Michael Gebreslasie planned and performed the study. Osadolor Ebhuoma wrote the paper. All authors read and approved the manuscript.

2.9 Conflicts of interest

The authors declare no conflict of interest.

2.10 References

1. World Health Organisation. *WHO Factsheet: Vector borne diseases*. 2014 Accessed 08/05/2015; Available from: http://www.who.int/kobe_centre/mediacentre/vbdfactsheet.pdf.

2. World Health Organisation. *Malaria*. 2015 Accessed 02/08/2015; Available from: http://www.who.int/mediacentre/factsheets/fs094/en/.

3. Snow, R., S. Hay, and K. Marsh, *Malaria in Africa: sources, risks, drivers and disease burden 2005-2030. Summary paper prepared for the UK Government's Foresight Project on the Detection and Identification of Infectious Diseases (T5. 8).* Report T5, 2006. **8**.

4. Palaniyandi, M., *The role of Remote Sensing and GIS for spatial prediction of vector-borne diseases transmission: A systematic review.* Journal of Vector Borne Diseases, 2012. **49**(4): p. 197.

5. Bai, L., L.C. Morton, and Q. Liu, *Climate change and mosquito-borne diseases in China: a review.* Globalisation and Health, 2013. **9**(1): p. 1.

6. Gosoniu, L., P. Vounatsou, N. Sogoba, N. Maire, and T. Smith, *Mapping malaria risk in West Africa using a Bayesian nonparametric non-stationary model.* Computational Statistics & Data Analysis, 2009. **53**(9): p. 3358-3371.

7. Yoo, E.-H., *Exploring space-time models for West Nile virus mosquito abundance data.* Applied Geography, 2013. **45**: p. 203-210.

8. Grover-Kopec, E., M. Kawano, R.W. Klaver, B. Blumenthal, P. Ceccato, and S.J. Connor, *An online operational rainfall-monitoring resource for epidemic malaria early warning systems in Africa.* Malaria Journal, 2005. **4**.

9. Teklehaimanot, H.D., M. Lipsitch, A. Teklehaimanot, and J. Schwartz, *Weather-based prediction of Plasmodium falciparum malaria in epidemic-prone regions of Ethiopia I. Patterns of lagged weather effects reflect biological mechanisms.* Malaria Journal, 2004. **3**: p. 41.

10. Teklehaimanot, H.D., J. Schwartz, A. Teklehaimanot, and M. Lipsitch, *Weather-based prediction of Plasmodium falciparum malaria in epidemic-prone regions of Ethiopia II. Weather-based prediction systems perform comparably to early detection systems in identifying times for interventions.* Malaria Journal, 2004. **3**.

11. Arab, A., M.C. Jackson, and C. Kongoli, *Modelling the effects of weather and climate on malaria distributions in West Africa.* Malaria Journal, 2014. **13**: p. 126.

12. Palaniyandi, M. *GIS mapping of vector breeding habitats*. Geospatial World, 2013. **14**.

13. Githeko, A.K., *Malaria And Climate Change.* Commonwealth Health Minister's Update 2009, 2009: p. 40-43.

14. Centres for Disease Control and Prevention. *Ecology of Malaria*. 2015 Accessed 03/04/2016; Available from: http://www.cdc.gov/malaria/about/biology/ecology.html.

15. Beck-Johnson, L.M., W.A. Nelson, K.P. Paaijmans, A.F. Read, M.B. Thomas, and O.N. Bjørnstad, *The effect of temperature on Anopheles mosquito population dynamics and the potential for malaria transmission.* PLOS one, 2013. **8**(11): p. e79276.

16. Lindsay, S.W. and W.J.M. Martens, *Malaria in the African highlands: past, present and future.* Bulletin of the World Health Organisation, 1998. **76**(1): p. 33-45.

17. Yamana, T.K. and E.A.B. Eltahir, *Incorporating the effects of humidity in a mechanistic model of Anopheles gambiae mosquito population dynamics in the Sahel region of Africa.* Parasites & Vectors, 2013. **6**(1): p. 1-10.

18. Zacarias, O.P. and M. Andersson, *Spatial and temporal patterns of malaria incidence in Mozambique.* Malaria Journal, 2011. **10**: p. 189.

19. Mills, A., Y. Lubell, and K. Hanson, *Malaria eradication: the economic, financial and institutional challenge.* Malaria Journal, 2008. **7**.

20. Midekisa, A., G. Senay, G.M. Henebry, P. Semuniguse, and M.C. Wimberly, *Remote sensing-based time series models for malaria early warning in the highlands of Ethiopia.* Malaria Journal, 2012. **11**.

21. Kuhn, K., D. Campbell-Lendrum, A. Haines, J. Cox, C. Corvalán, and M. Anker, *Using climate to predict infectious disease epidemics.* Geneva: WHO, 2005.

22. Dinku, T., P. Ceccato, E. Grover-Kopec, M. Lemma, S.J. Connor, and C.F. Ropelewski, *Validation of satellite rainfall products over East Africa's complex topography.* International Journal of Remote Sensing, 2007. **28**(7-8): p. 1503-1526.

23. Gaudart, J., O. Toure, N. Dessay, A.L. Dicko, S. Ranque, L. Forest, J. Demongeot, and O.K. Doumbo, *Modelling malaria incidence with environmental dependency in a locality of Sudanese savannah area, Mali.* Malaria Journal, 2009. **8**.

24. Hay, S., M. Renshaw, S.A. Ochola, A.M. Noor, and R.W. Snow, *Performance of forecasting, warning and detection of malaria epidemics in the highlands of western Kenya.* Trends in Parasitology, 2003. **19**(9): p. 394-9.

25. Graves, P.M., D.E. Osgood, M.C. Thomson, K. Sereke, A. Araia, M. Zerom, P. Ceccato, M. Bell, J. Del Corral, S. Ghebreselassie, et al., *Effectiveness of malaria control during changing climate conditions in Eritrea, 1998-2003.* Tropical Medicine & International Health, 2008. **13**(2): p. 218-28.

26. Nygren, D., C. Stoyanov, C. Lewold, F. Mansson, J. Miller, A. Kamanga, and C.J. Shiff, *Remotely-sensed, nocturnal, dew point correlates with malaria transmission in Southern Province, Zambia: a time-series study.* Malaria Journal, 2014. **13**: p. 231.

27. Gosoniu, L., P. Vounatsou, N. Sogoba, and T. Smith, *Bayesian modelling of geostatistical malaria risk data.* Geospatial Health, 2006. **1**(1): p. 127-39.

28. Gosoniu, L., A.M. Veta, and P. Vounatsou, *Bayesian geostatistical modeling of Malaria Indicator Survey data in Angola.* PLOS One, 2010. **5**(3): p. e9322.

29. Palaniyandi, M., P. Anand, and R. Maniyosai, *Spatial cognition: a geospatial analysis of vector borne disease transmission and the environment, using remote sensing and GIS.* International Journal of Mosquito Research, 2014. **1**(3): p. 39-54.

30. Onyiri, N., *Estimating malaria burden in Nigeria: a geostatistical modelling approach.* Geospatial Health, 2015. **10**(2): p. 163-170.

31. Arksey, H. and L. and O'Malley, *Scoping studies: towards a methodological framework.* International journal of social research methodology, 2005. **8**(1): p. 19-32.

32. Levac, D., H. Colquhoun, and K.K. O'Brien, *Scoping studies: advancing the methodology.* Implement Sci, 2010. **5**(1): p. 1-9.

33. European Space Agency. *eoPortal Directory*. 2000-2005 Accessed 09/06/15; Available from: https://directory.eoportal.org/web/eoportal/satellite-missions.

34. MARA/ARMA. 2009; Available from: http://www.mara-database.org.

35. Hellmuth, M.E., A. Moorhead, M.C. Thomson, and J. Williams, *Cliamte Risk Managment in Africa: Learning from Practice.* 2007: International Research Institute for Climate and Society (IRI), Columbia University, New York, USA.

36. Chenje, M. and P. Johnson, *State of the Environment in Southern Africa*. 1996, Johannesburg: Southern African Research and Documentation Centre, IUCN (The World Conservation Union), and Southern African Development Community.

37. United Nations Environmental Programme. *Africa Environment Outlook 2: Our Evironment, Our Wealth*. 2006 Accessed 14/02/2015; Available from: http://www.unep.org/dewa/Africa/publications/AEO-2/content/index.htm.

38. Zhou, G., N. Minakawa, A.K. Githeko, and G. Yan, *Association between climate variability and malaria epidemics in the East African highlands.* Proceedings of the National Academy of Sciences, 2004. **101**(8): p. 2375-80.

39. Omumbo, J.A., S.I. Hay, S.J. Goetz, R.W. Snow, and D.J. Rogers, *Updating Historical Maps of Malaria Transmission Intensity in East Africa Using Remote Sensing.* Photogrammetric Engineering and Remote Sensing, 2002. **68**(2): p. 161-166.

40. Omumbo, J.A., S.I. Hay, R.W. Snow, A.J. Tatem, and D.J. Rogers, *Modelling malaria risk in East Africa at high-spatial resolution.* Tropical Medicine & International Health, 2005. **10**(6): p. 557-66.

41. World Health Organisation. *World Malaria Report 2014*. 2014 Accessed 14/05/2015; Available from: http://www.who.int/malaria/publications/world_malaria_report_2014/wmr-2014-profiles.pdf?ua=1.

42. Gomez-Elipe, A., A. Otero, M. van Herp, and A. Aguirre-Jaime, *Forecasting malaria incidence based on monthly case reports and environmental factors in Karuzi, Burundi, 1997-2003.* Malaria Journal, 2007. **6**.

43. Ceccato, P., T. Ghebremeskel, M. Jaiteh, P.M. Graves, M. Levy, S. Ghebreselassie, A. Ogbamariam, A.G. Barnston, M. Bell, J. del Corral, et al., *Malaria stratification, climate, and epidemic early warning in Eritrea.* The American Journal of Tropical Medicine and Hygiene, 2007. **77**(6 Suppl): p. 61-8.

44. Hay, S.I., M.F. Myers, D.S. Burke, D.W. Vaughn, T. Endy, N. Ananda, G.D. Shanks, R.W. Snow, and D.J. Rogers, *Etiology of interepidemic periods of mosquito-borne disease.* Proceedings of the National Academy of Sciences, 2000. **97**(16): p. 9335-9.

45. Hay, S.I., E.C. Were, M. Renshaw, A.M. Noor, S.A. Ochola, L. Olusanmi, N. Alipui, and R.W. Snow, *Forecasting, warning, and detection of malaria epidemics: a case study.* Lancet, 2003. **361**(9370): p. 1705-1706.

46. Noor, A.M., A.C. Clements, P.W. Gething, G. Moloney, M. Borle, T. Shewchuk, S.I. Hay, and R.W. Snow, *Spatial prediction of Plasmodium falciparum prevalence in Somalia.* Malaria Journal, 2008. **7**: p. 159.

47. Hashizume, M., T. Terao, and N. Minakawa, *The Indian Ocean Dipole and malaria risk in the highlands of western Kenya.* Proceedings of the National Academy of Sciences, 2009. **106**(6): p. 1857-62.

48. Gosoniu, L., A. Msengwa, C. Lengeler, and P. Vounatsou, *Spatially explicit burden estimates of malaria in Tanzania: bayesian geostatistical modeling of the malaria indicator survey data.* PLOS One, 2012. **7**(5): p. e23966.

49. Noor, A.M., V.A. Alegana, A.P. Patil, G. Moloney, M. Borle, F. Yusuf, J. Amran, and R.W. Snow, *Mapping the receptivity of malaria risk to plan the future of control in Somalia.* BMJ Open, 2012. **2**(4).

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

51. Thomson, M.C., S.J. Mason, T. Phindela, and S.J. Connor, *Use of rainfall and sea surface temperature monitoring for malaria early warning in Botswana.* American Journal of Tropical Medicine and Hygiene, 2005. **73**(1): p. 214-221.

52. Mabaso, M.L., P. Vounatsou, S. Midzi, J. Da Silva, and T. Smith, *Spatio-temporal analysis of the role of climate in inter-annual variation of malaria incidence in Zimbabwe.* Int J Health Geogr, 2006. **5**: p. 20.

53. Thomson, M.C., F.J. Doblas-Reyes, S.J. Mason, R. Hagedorn, S.J. Connor, T. Phindela, A.P. Morse, and T.N. Palmer, *Malaria early warnings based on seasonal climate forecasts from multi-model ensembles.* Nature, 2006. **439**(7076): p. 576-579.

54. Craig, M.H., B.L. Sharp, M.L. Mabaso, and I. Kleinschmidt, *Developing a spatial-statistical model and map of historical malaria prevalence in Botswana using a staged variable selection procedure.* International Journal of Health Geographics, 2007. **6**: p. 44.

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

56. Alegana, V.A., P.M. Atkinson, J.A. Wright, R. Kamwi, P. Uusiku, S. Katokele, R.W. Snow, and A.M. Noor, *Estimation of malaria incidence in northern Namibia in 2009 using Bayesian conditional-autoregressive spatial-temporal models.* Spatial and Spatio-Temporal Epidemiology, 2013. **7**: p. 25-36.

57. Cohen, J.M., S. Dlamini, J.M. Novotny, D. Kandula, S. Kunene, and A.J. Tatem, *Rapid case-based mapping of seasonal malaria transmission risk for strategic elimination planning in Swaziland.* Malaria Journal, 2013. **12**: p. 61.

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

59. Kleinschmidt, I., M. Bagayoko, G.P. Clarke, M. Craig, and D. Le Sueur, *A spatial statistical approach to malaria mapping.* IInternational Journal of Epidemiology, 2000. **29**(2): p. 355-61.

60. Gemperli, A., P. Vounatsou, N. Sogoba, and T. Smith, *Malaria mapping using transmission models: application to survey data from Mali.* American Journal of Epidemiology, 2006. **163**(3): p. 289-297.

61. Silue, K.D., G. Raso, A. Yapi, P. Vounatsou, M. Tanner, K. N'Goran E, and J. Utzinger, *Spatially-explicit risk profiling of Plasmodium falciparum infections at a small scale: a geostatistical modelling approach.* Malaria Journal, 2008. **7**: p. 111.

62. Raso, G., K.D. Silue, P. Vounatsou, B.H. Singer, A. Yapi, M. Tanner, J. Utzinger, and E.K. N'Goran, *Spatial risk profiling of Plasmodium falciparum parasitaemia in a high endemicity area in Cote d'Ivoire.* Malaria Journal, 2009. **8**: p. 252.

63. Giardina, F., L. Gosoniu, L. Konate, M.B. Diouf, R. Perry, O. Gaye, O. Faye, and P. Vounatsou, *Estimating the burden of malaria in Senegal: Bayesian zero-inflated binomial geostatistical modeling of the MIS 2008 data.* PLOS One, 2012. **7**(3): p. e32625.

64. Raso, G., N. Schur, J. Utzinger, B.G. Koudou, E.S. Tchicaya, F. Rohner, K. N'Goran E, K.D. Silue, B. Matthys, S. Assi, et al., *Mapping malaria risk among children in Cote d'Ivoire using Bayesian geo-statistical models.* Malaria Journal, 2012. **11**: p. 160.

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

66. Wardlow, B.D., S.L. Egbert, and J.H. Kastens, *Analysis of time-series MODIS 250 m vegetation index data for crop classification in the US Central Great Plains.* Remote Sensing of Environment, 2007. **108**(3): p. 290-310.

67. Jiang, Z., A.R. Huete, K. Didan, and T. Miura, *Development of a two-band enhanced vegetation index without a blue band.* Remote Sensing of Environment, 2008. **112**(10): p. 3833-3845.

68. Walz, Y., M. Wegmann, S. Dech, G. Raso, and J. Utzinger, *Risk profiling of schistosomiasis using remote sensing: approaches, challenges and outlook.* Parasites & Vectors, 2015. **8**.

69. Machault, V., C. Vignolles, F. Borchi, P. Vounatsou, F. Pages, S. Briolant, J.P. Lacaux, and C. Rogier, *The use of remotely sensed environmental data in the study of malaria.* Geospatial Health, 2011. **5**(2): p. 151-68.

70. United States Agency International Development, *President's Malaria Initiative Zambia: Malaria Operational Plan FY 2018*. 2017, United States Agency International Development.

71. Roll Back Malaria. *Roll Back Malaria Progress and Impact Series: Focus on Swaziland*. 2014 Accessed 06/02/2015; Available from: http://www.rollbackmalaria.org/microsites/wmd2014/report13.html.

72. Maidment, R.I., D. Grimes, R.P. Allan, E. Tarnavsky, M. Stringer, T. Hewison, R. Roebeling, and E. Black, *The 30 year TAMSAT African rainfall climatology and time series (TARCAT) data set.* Journal of Geophysical Research: Atmospheres, 2014. **119**(18).

73. Tarnavsky, E., D. Grimes, R. Maidment, E. Black, R.P. Allan, M. Stringer, R. Chadwick, and F. Kayitakire, *Extension of the TAMSAT satellite-based rainfall monitoring over Africa and from 1983 to present.* Journal of Applied Meteorology and Climatology, 2014. **53**(12): p. 2805-2822.

74. Hutchinson, M., H. Nix, J. McMahon, and K. Ord, *Africa-A topographic and climate database (CD-ROM).* The Australian National University Canberra, ACT, 1996. **200**.

75. World Resources Institute, *African Data Sampler (CD-ROM) Edition 1.* 1995.

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

77. Thomson, M.C., S.J. Mason, T. Phindela, and S.J. Connor, *Use of rainfall and sea surface temperature monitoring for malaria early warning in Botswana.* Am J Trop Med Hyg, 2005. **73**(1): p. 214-21.

78. Ceccato, P., T. Ghebremeskel, M. Jaiteh, P.M. Graves, M. Levy, S. Ghebreselassie, A. Ogbamariam, A.G. Barnston, M. Bell, J. del Corral, et al., *Malaria stratification, climate, and epidemic early warning in Eritrea.* American Journal of Tropical Medicine and Hygiene, 2007. **77**(6): p. 61-68.

79. Stefani, A., I. Dusfour, M.C.B. Cruz, N. Dessay, A.K.R. Galardo, C.D. Galardo, R. Girod, M.S.M. Gomes, H. Gurgel, A.C.F. Lima, et al., *Land cover, land use and malaria in the Amazon: a systematic literature review of studies using remotely sensed data.* Malaria Journal, 2013. **12**.

80. Britch, S.C., K.J. Linthicum, A. Anyamba, C.J. Tucker, E.W. Pak, F.A. Maloney, Jr., K. Cobb, E. Stanwix, J. Humphries, A. Spring, et al., *Satellite vegetation index data as a tool to forecast population dynamics of medically important mosquitoes at military installations in the continental United States.* Military Medicine, 2008. **173**(7): p. 677-683.

81. Matsushita, B., W. Yang, J. Chen, Y. Onda, and G. Qiu, *Sensitivity of the Enhanced Vegetation Index (EVI) and Normalized Difference Vegetation Index (NDVI) to topographic effects: A case study in high-density cypress forest.* Sensors, 2007. **7**(11): p. 2636-2651.

82. Huete, A., K. Didan, T. Miura, E.P. Rodriguez, X. Gao, and L.G. Ferreira, *Overview of the radiometric and biophysical performance of the MODIS vegetation indices.* Remote Sensing of Environment, 2002. **83**(1-2): p. 195-213.

83. Gebreslasie, M. and I. Naidoo, *Earth Observation in Malaria Vector Control and Management*. 2011, South African Medical Research Council (SAMRC).

84. Franke, J., M. Gebreslasie, I. Bauwens, J. Deleu, and F. Siegert, *Earth observation in support of malaria control and epidemiology: MALAREO monitoring approaches.* Geospatial Health, 2015. **10**(1): p. 117-131.

85. National Aeronautics and Space Administration. *Reverb Echo*. n.d Accessed 10/12/14; Available from: http://reverb.echo.nasa.gov/reverb/#utf8=%E2%9C%93&spatial_map=satellite&spatial_type=rectangle.

86. Zhang, Z., M. Ward, J. Gao, Z. Wang, B. Yao, T. Zhang, and Q. Jiang, *Remote sensing and disease control in China: past, present and future.* Parasites & Vectors, 2013. **6**.

87. McCulloch, C.E., S.R. Searle, and J.M. Neuhaus, *Generalized linear mixed models*. 2nd ed. 2008, New York: Wiley.

88. Chatfield, C., *The Analysis Of Time Series: An Introduction*. 6th ed. 2003, Florida: CRC press.

89. de Jong, R.M. and J. Davidson, *Consistency of kernel estimators of heteroscedastic and autocorrelated covariance matrices.* Econometrica, 2000. **68**(2): p. 407-423.

90. Zinszer, K., A.D. Verma, K. Charland, T.F. Brewer, J.S. Brownstein, Z. Sun, and D.L. Buckeridge, *A scoping review of malaria forecasting: past work and future directions.* BMJ Open, 2012. **2**(6).

91. Sternberg, R.J. and E.L. Grigorenko, *The General Factor Of Intelligence: How General Is It?* 2002: Psychology Press.

CHAPTER 3: CLIMATIC AND ENVIRONMENTAL DETERMINANTS FOR MODELLING MALARIA DISEASE RISK IN A PROVINCE WITH LOW MALARIA TRANSMISSION USING BAYESIAN ZERO-INFLATED MODELS IN INLA

This chapter is based on:

Ebhuoma O, Gebreslasie M, Arab A (Ready for submission). Climatic and environmental determinants for modelling malaria disease risk in a province with low malaria transmission using Bayesian zero-inflated models in INLA.

3.1 Abstract

The malaria control and intervention strategy in KwaZulu-Natal (KZN), South Africa (SA) significantly reduced the malaria risk and resulted in zero malaria cases reported in some local municipalities. Therefore, to sustain this and meet the malaria elimination target for the entire province, an evaluation of the different climatic and environmental variables that influence the spatiotemporal malaria transmission is important. This study models the influence of climatic and environmental variables on the spatiotemporal malaria transmission in KZN, SA by assessing different Bayesian statistical models that can handle excess zeros. Considering spatiotemporal dependencies, the investigator employed different Bayesian zero-inflated models in INLA to clinically confirmed monthly malaria cases and the following remotely sensed climatic and environmental variables: precipitation, day and night land surface temperature, normalized difference vegetation index (NDVI), enhanced vegetation index and elevation from the malarious local municipalities in KZN during the period 2005-2014. The Bayesian spatiotemporal zero inflated Poisson (ZIP) was identified as the best model fit based on the deviance information criterion. The spatiotemporal ZIP analysis results indicate that at 95% Bayesian credible interval (BCI) NDVI (0.91; 95% BCI = 0.71, -1.12), precipitation (0.11; 95% BCI = 0.08, 0.14), elevation (0.05; 95% BCI = 0.03, 0.07) and night temperature (0.04; 95% BCI = 0.03, 0.04) are significantly related to malaria transmission in KZN, SA. The area with the highest risk of malaria morbidity in KZN was identified as the north-eastern part of the province. The modelling approach employed in this study presents a valuable tool for understanding and monitoring the influence of climatic and environmental variables on the spatial heterogeneity of malaria in KZN. It will therefore equip the relevant policy makers with information required to channel malaria intervention resources sustainably to vulnerable receptive areas. Also, this study reveals the need to strengthen the already existing cross-border collaborations to fortify KZN's malaria elimination goals.

Keywords: *Malaria transmission, climatic variables, environmental variables, KwaZulu-Natal, excess zeros, zero inflated Poisson.*

3.2 Introduction

Malaria is endemic to the northern part of KwaZulu-Natal (KZN) province, South Africa (SA) [1]. The malaria parasite responsible for over 90% of malaria cases in the province is *Plasmodium falciparum*, while *Anopheles arabiensis* of the *Anopheles gambiae* complex is the major malaria vector species [1]. Low number of malaria cases and excess zero malaria cases have been recorded in the malarious local municipalities arising from the influence of the efficient malaria control and intervention strategy currently in place in KZN [2-4]. Currently, malaria incidence in KZN is between 0.01 to 0.10 case per 1000 persons [1]. This, positions the province within the malaria elimination epidemiological trigger in the WHO elimination continuum [5], as such she strives to achieve malaria elimination by 2020 [1]. To actualise the elimination target, relevant measures and tools for malaria surveillance, prevention and control are necessary. In this regard, the spatiotemporal delineation and prediction of malaria transmission using relevant climatic and environmental variables is important to derive empirical maps of malaria risk and transmission.

Climate is a very important determinant of the spatiotemporal heterogeneity of malaria risk and transmission [6-12]. Rainfall is the main climatic component that contributes to malaria although its effect on malaria vector proliferation and malaria transmission intensity notably varies with rainfall amount variations [12-14]. Temperature plays an integral rote via complex interactions on malaria vector population dynamics on one hand and for parasite development within the vector, on the other hand [15]. Other climatic and environmental variables that have been widely reported in literature for their influential role in malaria transmission include elevation/altitude, relative humidity, land use and land cover [12, 16-21]. Furthermore, urbanisation, human population movement, socio-economic and demographic variables, and malaria intervention have also been reported for their impacts on malaria transmission [2, 22-26].

The progress made in spatial technology and Bayesian geo-statistical modelling have opened new opportunities for exploring the climatic and environmental suitability of malaria using proxy variables from remote sensing, and formulation of spatiotemporal empirical malaria risk maps at different levels [2, 18, 27-30]. This, in turn, has presented a platform for sound inference, prediction and mapping of the variable(s) of interest. For instance, Clements *et al.*

[31] formulated a spatiotemporal malaria risk model in Yunnan province, China based on the relationship between monthly malaria cases and selected RS climatic variables (rainfall and temperature) obtained from WORLDCLIM database. The Bayesian spatiotemporal Poisson regression analysis based on the *Plasmodium vivax* model revealed that at 95% Bayesian credible interval (BCI) rainfall (1.045, 95% BCI = 1.044-1.046) and maximum temperature (1.047, 95% BCI = 1.045-1.050) are significant malaria predictor variables in the province. Also, the *Plasmodium falciparum* based model identified rainfall (1.037, 95% BCI = 1.034-1.040) and maximum temperature (1.053, 95% BCI = 1.047, 1.060) as the significant malaria predictors.

Across sub-Saharan Africa, diverse climatic and environmental variables have also been reported to influence the spatiotemporal heterogeneity of malaria disease markedly. At provincial level in Zambia, the influence of maximum and minimum land surface temperature (LST) derived from moderate resolution imaging spectroradiometer (MODIS) and CHIRPS obtained precipitation data on malaria occurrence was evaluated employing the Bayesian semiparametric Poisson regression analysis [32]. The analysis revealed all the variables studied (precipitation (40.24, 97.5% BCI = 13.70-97.86), minimum LST (1.06, 97.5% BCI = 0.31-2.46) and maximum LST (0.67, 97.5% BCI = 0.36-1.14)) have a significant relationship with malaria incidence. Furthermore, after considering the spatial dependencies, Luapula and North-western province were identified as areas with highest risk of malaria transmission as compared to Lusaka (95% lower) and Western Province (68% lower). In Angola, the malaria parasitaemia risk and prediction maps were delineated after considering the impact of climatic and environmental variables (altitude from United States Geological Service (USGS) EROS, rainfall from Africa data dissemination service (ADDS), rivers and lakes, MODIS derived day and night LST and normalized difference vegetation index (NDVI)), socio-economic status and malaria intervention [33]. The results from the Bayesian geostatistical models, suggested that NDVI value $\geq$ 0.60 (4.34, 95% BCI = 0.53 -36.74) and rainfall between 112-135mm (2.81, 95% BCI = 0.43- 18.57) were the most significant malaria predictors. Other significant climatic and environmental predictors are day and night LST, altitude and proximity to water bodies. The predicted malaria risk map delineated afterwards suggested that the northern and the central regions of Angola are the highest risk malaria areas. Contrarily, the south and south-east regions are the lowest risk malaria areas. In another study, the Bayesian geostatistical logistic regression

approach employed various climatic and environmental variables (such as Advanced Very High Resolution Radiometer (AVHRR) derived NDVI, land use, rainfall, temperature, soil water storage index and water bodies obtained from USGS) to identify malaria predictors and model the spatiotemporal variations of malaria vectors across Mali [34]. The analysis revealed that soil water storage index, NDVI, maximum temperature and the distance to water bodies have a significant positive relationship with the malaria vector dominant in the area (*Anopheles arabiensis)*. While minimum temperature and rainfall were negatively associated. Finally, the Bayesian Kriging approach was used to develop a predictive map of the spatial distribution of the vectors. On a larger scale, important malaria predictors across ten West African countries were identified employing the Bayesian spatiotemporal Hierarchical modelling approach [27]. The modelling was done using malaria cases and various climatic variables (obtained from the National Oceanic and Atmospheric Administration's (NOAA) National Climate Data Centre (NCDC)). The outcome of the analysis revealed that across West Africa total annual precipitation (-0.1055, 97.5% BCI = -0.1808, -0.0325) and mean annual temperature (-0.2034, 97.5% BCI = -0.2528, -0.1545) were the most important variables.

The above studies reviewed highlight the relevance of employing the relevant Bayesian approach in modelling malaria incidence/occurrence and predicting the spatiotemporal distribution of the disease. They also indicate the relevant count data models and analyses were carefully selected in-line with the characteristics of malaria data distribution (which comprised of malaria count distributions without zero-value observations or infrequent zero-values). However, in settings where the number of zeros is greater than expected under a standard count distribution with a fixed mean, a flexible model that can handle the over dispersion resulting from the excess zeros values and still take account of the non-zero values in the model is valuable [35]. In a disease like malaria known for its spatiotemporal heterogeneity, and as more zero cases are observed in areas progressing in their malaria elimination campaign, it is important that the zero values are considered in the spatial and spatiotemporal modelling. They often present valuable information relating to the disease such as the detection rate of the disease, and the occurrence and knowledge of the disease by the population [36, 37]. Varieties of zero-adjusted mixed models are available in that regard. These include the zero-inflated negative binomial (ZINB) model, zero-inflated Poisson (ZIP) model, Poisson hurdle model, and the negative binomial hurdle model [35-40]. For instance, Alegana *et al.* [41] employed a

Bayesian ZIP approach to model the malaria incidence risk in Northern Namibia considering tropical rainfall measuring mission (TRMM) rainfall data, temperature suitability index and MODIS enhanced vegetation index (EVI) data. The multivariate analysis revealed that only EVI (14.29, 95% BCI = 9.24–19.42) was significant and the predicted malaria risk map suggested that areas boarding Angola and Zambia are at the highest risk of malaria transmission. A similar study was conducted in Afghanistan, in which the same Bayesian approach with climatic and environmental variables were used to model the incidence of *Plasmodium vivax* and *Plasmodium falciparum* at district level [42]. The multivariate analysis based on the *Plasmodium vivax* model revealed that only temperature suitability index (0.124, 95% BCI = 0.048-0.202) was significant. But none of the climatic and environmental variables were significant based on the *Plasmodium falciparum* model. The predicted malaria risk map suggested that the eastern and south-eastern Afghanistan areas boarding Pakistan are at the highest risk of malaria incidence. In Senegal, Bayesian geostatistical zero-inflated binomial (ZIB) climatic model formulated by Giardina *et al.* [43] suggests a significant relationship with NDVI (1.48, 95% BCI = 0.88-2.48), night LST (1.16, 95% BCI = 0.66-1.86) and malaria. A study by Kasasa *et al.* [44], showed how two different zero-adjusted models were used to understand the malaria transmission patterns in a small area in Northern Ghana. The Bayesian geostatistical ZIB and ZINB approaches were used to evaluate the sporozoite rate and mosquito densities, respectively. The study revealed a significant spatiotemporal heterogeneity of entomological inoculation rate estimates and malaria transmission intensity existed in the small area. Other studies compared different zero-inflated models and the best-fit model was identified based on a relevant comparative measure(s). For instance, Neelon *et al.* [35] compared Poisson model, Poisson hurdle model, ZIP and Zero altered Poisson model. Based on the lowest deviance information criteria (DIC) and the negative cross-validatory log likelihood measures, the ZIP model produced the best fit model. While, Arab [37] compared the Poisson Hurdle model, ZIP, Poisson Hurdle with probability model, Negative binomial Hurdle model, ZINB model and Negative Binomial Hurdle with probability model. ZIP was also reported to have the lowest DIC value.

Thus, in KZN, the investigator employed various Bayesian spatiotemporal models that can handle zero-inflated surveillance data of malaria cases by taking into account the effects of

significant climatic and environmental variables. They are the ZINB, ZIP, Poisson hurdle and negative binomial hurdle models.

3.3 Methodology

3.3.1 Study area

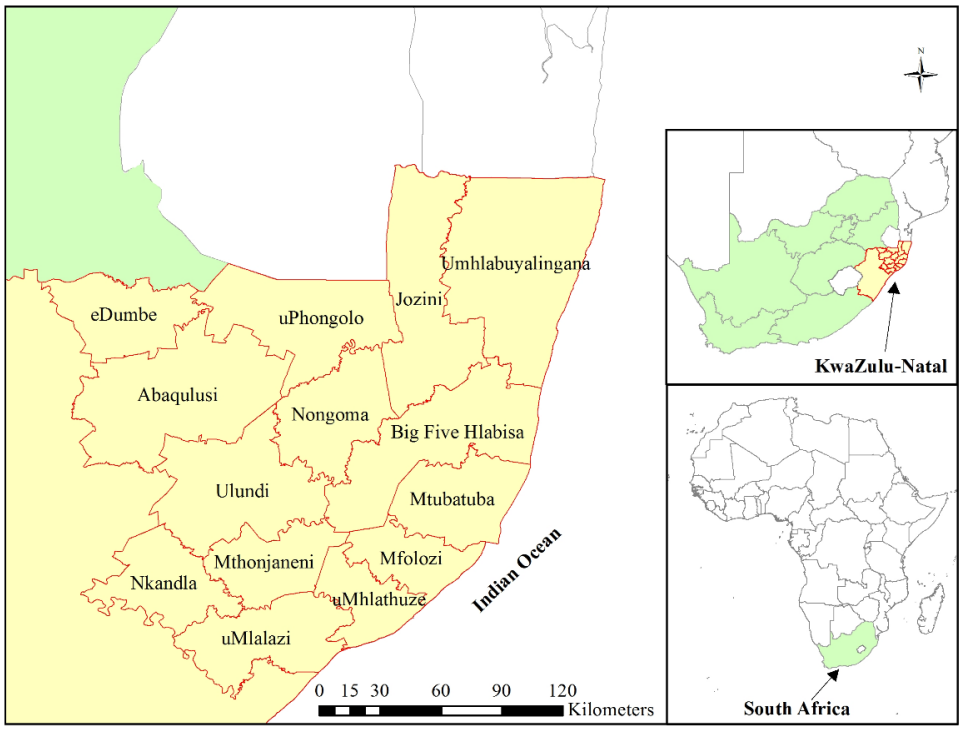


Figure 3.1 Map of the study area showing the malaria endemic areas in KwaZulu-Natal, South Africa.

The study area is located in the north-eastern part of KZN province (Figure 3.1) covering the local municipalities of uMkhanyakude, uThungulu and Zululand district municipalities. It is bordered internationally by The Kingdom of Swaziland and The People's Republic of Mozambique to its north. It has a long shoreline along the Indian Ocean to its east and stretching down south-eastwards. The region possesses a sub-tropical climate with the majority of malaria incidents observed during October to May (the rainy months), with a seasonal peak usually in January and March [1, 45]. The average annual rainfall ranges from 500mm to 2000mm. Along the coastal areas, the summer temperatures is between 24°C to 32°C, and mean winter temperature

is about 20°C. The Midlands generally possesses a mild climate with relatively high summer rainfall and dry winters. The elevation measure of the region varies from sea level to over 3000m. The vegetation of the study area comprises of coastal forest and thornveld along the coast. Towards the inlands, lowveld, highland sourveld, Natal sour sandveld, valley bushveld and tall grassveld vegetation are found. Lowveld and thornveld characterises the low-lying hot and dry regions of Northern KZN [46].

3.3.2 Data

3.3.2.1 Malaria case data

Malaria cases from January 2005 to December 2014 were obtained from the malaria control program of KZN, SA. In SA, when a suspected malaria case is presented, the blood smear of the suspected case is tested for *Plasmodium* using either microscopy or rapid diagnostic test by a certified health officer. If a positive result is obtained, patient details including patient demographics, the health facility the case was reported, symptoms, malaria test results, diagnosis and treatment administered are entered into a malaria case notification form and reported to the relevant provincial malaria control program. At the provincial malaria control program, the details of malaria case(s) are then captured into the malaria information system [1].

The distribution of malaria cases in KZN during the period of the study is characterised by excess zeros or zero-inflated (about 81%) (See Figure 3.2), suggesting the efficiency of the malaria control and intervention strategy currently in place across KZN. [2, 3]. This therefore suggests the appropriateness of employing a zero inflated model.

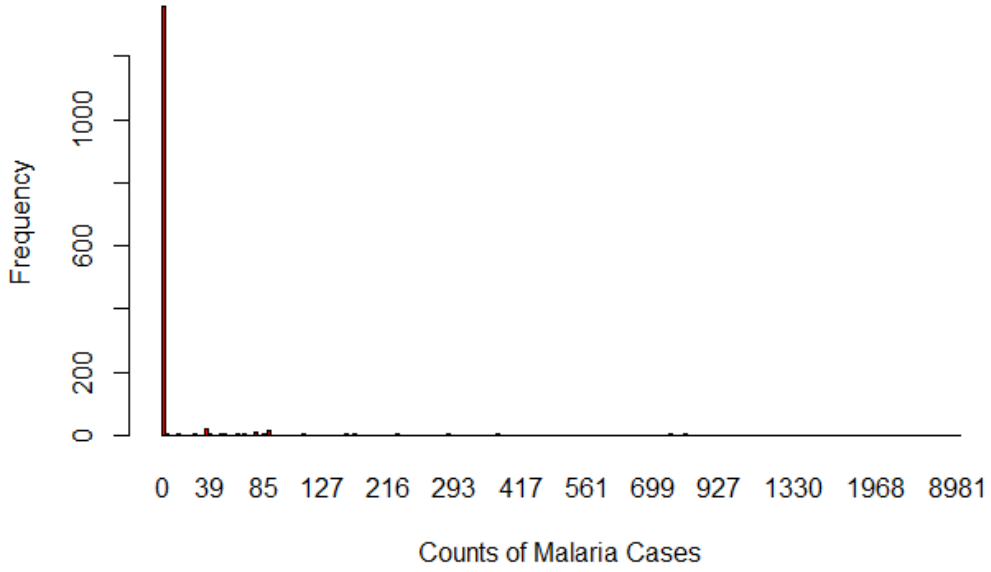


Figure 3.2 The distribution of malaria cases in uMkhanyakude, uThungulu and Zululand District, South Africa (2005-2014)

3.3.2.2 Climatic and environmental variables

In this study, the investigator extracted precipitation data at 2.5 x 2.5 degree spatial resolution from the NOAA National Centres for Environmental Prediction (NCEP) CPC Merged Analysis of Precipitation (CMAP) database. The CMAP precipitation data is a combination of rain gauge observations, numerical model predictions and five different types of satellite estimates (GPI,OPI,SSM/I scattering, SSM/I emission and MSU) [47]. The precipitation data were downloaded as averages over monthly periods.

MODIS derived NDVI, EVI, day and night LST data were downloaded from the USGS Land Processes Distributed Active Archive Centre (LP DAAC) database. Day and night LST data were obtained as averages over an 8-day period at 1km x 1km spatial resolution, while NDVI and EVI were extracted as averages over 16-days period at 250m x 250m spatial resolution. All the MODIS data were aggregated in the form of monthly averages [48].

Elevation data gridded at 1km x 1km were downloaded from the NOAA global Digital Elevation Model (DEM) data from the Global Land One-km Base Elevation (GLOBE) Project [49].

3.3.3 Model development

3.3.3.1 Variable selection for spatiotemporal model

A preliminary analysis (the cross correlation matrix) was carried out to identify the suitable predictor variables to be inputted in the spatiotemporal model to guide against multicollinearity in the models and improve the model fit.

3.3.3.2 Zero-inflated models

Due to the excess zeros or zero-inflated malarial data, the investigator fitted the following zero-inflated models to identify the best model fit.

1) ZINB model

The ZINB model can be described as a mixture of a mass of p for the excess zeros and a mass of (1 - p) for the negative binomial distribution, where $0 \leq p \leq 1$. Thus, the ZINB model is written as [38]:

$$P(Y = k) = \begin{cases} p + (1-p)\left(\frac{\tau}{\tau+\lambda}\right)^{\tau}, & k = 0 \\ (1-p)\frac{\Gamma(\tau+k)}{k!\,\Gamma(\tau)}\left(\frac{\tau}{\tau+\lambda}\right)^{\tau}\left(\frac{\lambda}{\lambda+\tau}\right)^{k} & k = 1,2 \ldots\ldots\ldots\ldots \end{cases} \quad (3.1)$$

The ZINB regression model links p and λ to predictors, i.e.,

$$log(\lambda_i) = x_i\beta \quad (3.2)$$

and

$$logit(p_i) = z_i\gamma \quad (3.3)$$

where $i = 1, 2, \ldots, n$ and xi and zi are d- and q- dimensional vectors of predictors linked to the ith subject, and with β and γ the corresponding vectors of regression parameters, respectively.

2) ZIP model

The ZIP model is a combination of a Poisson distribution part (non-zero component) and a point mass at zero (zero component). The zero data from an observation emerges from both the point mass at zero and the Poisson distribution. In ZIP model, the zero component assumes a probability pi and the Poisson distribution assumes a probability 1 - pi where $i = 0, 1, 2, \ldots, n$. Thus, the ZIP model can be written as [38]:

$$P(Yi = 0) = pi + (1 - pi)exp(-\lambda i) \qquad (3.4)$$

$$P(Yi = k) = (1 - pi)exp(-\lambda i)\,\lambda_i^k/k!, \qquad k = 1,2,3, \ldots\ldots\ldots\ldots \qquad (3.5)$$

The effects of the predictors on the count distribution in a ZIP model can be evaluated by specifying *pi* and λi as a function of predictors.

The probability of excess zeros should be modelled employing a logistic regression model as given below:

$$logit(p_i) = x_i\beta \qquad (3.6)$$

where *xi* is a vector of predictors and β a vector of parameters. On the contrary, the impact of predictors on count data without the excess zeros can be modelled using Poisson regression:

$$log(\lambda_i) = z_i\gamma \qquad (3.7)$$

3) Poisson Hurdle model

The Poisson Hurdle model is a two-part model. The hurdle or logistic regression part models the zero vs. non-zero counts to obtain the zero probabilities. The second part is the zero truncated Poisson or regression part that is used to model the non-zero counts. Thus, the Poisson Hurdle model can be written as [38]:

$$P(Yi = 0) = pi \qquad (3.8)$$

$$P(Yi = k) = (1 - pi)\,\frac{exp(-\lambda i)\,(\lambda i)^k/k!}{1 - exp(-\lambda i)}, \qquad k = 1,2,3, \ldots\ldots \qquad (3.9)$$

pi models all zeros. For this model, the logistic regression should be employed in modeling the probability of zeros.

$$logit(p_i) = x_i\beta \qquad (3.10)$$

while the Poisson regression is the choice model to evaluate the impacts of predictors *zi* on positive count data:

$$log(\lambda_i) = z_i\gamma \qquad (3.11)$$

4) Negative Binomial Hurdle model

The negative binomial hurdle model is similar to the Poisson Hurdle model, but in the second part of the Poisson Hurdle model, the zero truncated Poisson model is replaced with a negative binomial model [38].

$$P(Yi = 0) = pi \qquad (3.12)$$

$$P(Yi = k) = (1 - pi)\frac{\Gamma(k + \tau)}{\Gamma(k+1)\Gamma(\tau)} \times \frac{(1+\tau\lambda)^{-(k+\tau)}\tau^k\lambda^k}{1-(1+\tau\lambda)^{\tau}} \quad k = 1,2,3,\ldots\ldots \qquad (3.13)$$

The logistic regression should be employed in modelling the probability of zeros:

$$logit(p_i) = x_i\beta \qquad (3.14)$$

While the negative binomial regression is the choice model to evaluate the effects of predictors on count data:

$$log(\lambda_i) = x_i\gamma \qquad (3.15)$$

3.3.3.3 Bayesian spatiotemporal model

The Bayesian inference intuitively supports a hierarchical model approach. Implementing the Bayesian hierarchical model approach will allow suitable data sampling variability, parameter uncertainty, and likely spatial and temporal dependences. Therefore, the effects of spatial and temporal dependences were accounted for in the zero-inflated models developed. To this effect, the investigator formulated a hierarchical model for count data Y_i's (for i =1,...,n) and predictor variables $X_i,\ldots,X_p$ following three modelling stages: (1) data model, (2) process model, and (3) parameters models [37].

(1) The data model is written as:

$$Y \sim f(y_i|\theta_i, p), = i = 1, \ldots\ldots, n \qquad (3.16)$$

Let $f(y_i|\theta_i, p)$ represent a zero-inflated or hurdle distribution with parameters θ_i's and mixture probability p.

(2) The process model is written as:

$$S(\theta_i) = \beta_0 + \beta_1 X_{1i} + \ldots + \beta_p X_{pi} + \gamma_i, \quad i = 1, \ldots\ldots, n \qquad (3.17)$$

$S(.)$ represents a function specified based on the conditions on θ_i's. β_i represents the spatial regression coefficients for the specified predictors X_i,...,X_p. γ_i represents noise measurement. Parameters $\gamma = (\gamma_1, ..., \gamma_n)$ represents the noise measurement based on the spatial dependence such that:

$$\gamma \sim N\ (0, \sum) \qquad (3.18)$$

$$\sum\ = \sigma^2\ \mathrm{R}(\varphi) \qquad (3.19)$$

$\sum$ represents the covariance matrix that explains the measure of the relationship each observation has with its neighbours (i.e., the spatial dependence of the data). And the investigator defined $\sum$ based on the geostatistical structure of the data. The spatial correlation is specified based on an exponential covariogram model such that:

$$R(\varphi) = \exp(-\varphi d) \qquad (3.20)$$

where a symmetric spatial correlation is assumed and is based on the Euclidean distance between data points (d) and a spatial range parameter, φ (which is a function of the strength of spatial relationship over spatial locations).

(3) The parameter models

The Bayesian approach regards parameter models as the prior distributions for the set of unknown parameters (e.g., β_i's, τ, and σ^2). This prior distribution and the traditional likelihood are combined to obtain the posterior distribution of the parameter of interest based on the statistical inference using integrated nested Laplace approximation (INLA) via the Gaussian Markov Random Field (GMRF) [50-52]. In this model, a flat non-informative prior distribution with a small mean and large variances were specified to all the unknown parameters. Refer to Ntzoufras [53] and Gelman *et al.* [54] for more reviews on prior determination process.

3.3.3.4 Bayesian spatiotemporal model in R-INLA

INLA is an analytical platform that supports the evaluation of posterior margins in hierarchical models with latent random processes, thus reducing the computation time extensively [50-52]. The investigator, therefore, fitted different Bayesian spatial and spatiotemporal models using R-INLA and the stochastic partial differential equations (SPDE) methods. Based on the DIC

values obtained from the formulated models, the best model fit was selected [55]. Using the INLA/SPDE, the investigator built a mesh made up of triangles across the area of interest (study area) to evaluate the spatial fields (see Figure 3.3).

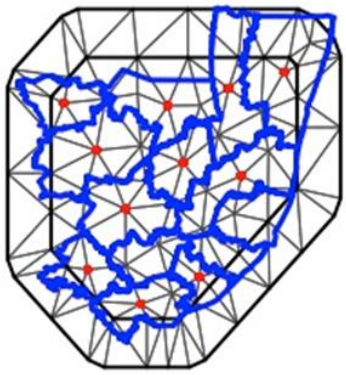


Figure 3.3 INLA/SPDE mesh for the spatial fields

3.4 Results

The cross correlation matrix in Table 3.1 shows high correlation (threshold of correlation coefficient ≥ 0.6) between the following pairs of predictor variables. They are day LST/LogPrecip, day LST/night LST, NDVI/EVI and day LST/EVI. Thus, day LST and EVI were dropped, so that a parsimonious model can be achieved, while precipitation, night LST, NDVI and elevation were selected as the suitable predictor variables and they were subsequently employed in formulation of the spatiotemporal models.

Table 3.1 Correlation matrix of the predictor variables

	LogPrecip	LogElev	Night LST	Day LST	NDVI	EVI
LogPrecip	1.00	-0.03	0.49	0.69	0.50	0.54
LogElev		1.00	-0.09	-0.32	-0.24	-0.29
Night LST			1.00	0.64	-0.02	0.07
Day LST				1.00	0.54	0.60
NDVI					1.00	0.95
EVI						1.00

LogPrecip = Log precipitation; LogElev = Log elevation; LST = land surface temperature; NDVI = Normalised difference vegetation index; EVI= Enhanced vegetation Index.

3.4.1 Models comparison

For comparison, the investigator formulated six models using the selected predictor variables (precipitation, night LST, NDVI and elevation) and the zero inflated malaria case dataset (see Table 3.2).

Table 3.2 Comparison of spatiotemporal models based on their DIC values

Spatiotemporal Model	DIC
ZINB	6819.40
ZIP	4709.88
Negative Binomial Hurdle	7004.70
Poisson Hurdle	4978.05

DIC= Deviance Information Criteria; ZINB = Zero inflated negative binomial; ZIP = Zero inflated Poisson

Based on the smallest DIC values from the models, the zero-inflated Poisson (DIC = 4709.88) was the best fit to zero inflated malaria data compared to the other zero-inflated models. This was followed by the Poisson hurdle model (DIC = 4978.05). The Poisson model had the weakest performance (DIC = 5.990387e+35). The difference between the DIC value of the ZIP model and the Poisson Hurdle is less than 10%. While the DIC value of the ZIP model is substantially different (i.e. more than 10% difference) from that of the ZINB and the negative binomial hurdle models.

Table 3.3 presents the results of different ZIP models. They are; (1) spatial and temporal ZIP model, (2) spatial ZIP model, (3) temporal ZIP model and (4) the ZIP model without spatial and temporal effects. From the results, applying both the spatial and temporal dependences to the ZIP model resulted in the best fit model based on the smallest DIC value. The DIC results of the models also revealed that the temporal effects (DIC = 4874.61) of the spatiotemporal ZIP model contributed substantially in improving the model fit as compared to the spatial aspect (DIC = 967167.90). The investigator, therefore, focused on the spatiotemporal ZIP model in subsequent parts of the results and discussion sections.

Table 3.3 Spatial and spatiotemporal ZIP models

Model	DIC
Spatiotemporal	4709.88
Spatial	967167.90
Temporal	4874.61
No spatial, No temporal	1443548

DIC= Deviance Information Criteria

3.4.2 Posterior inference

The spatiotemporal ZIP analysis results in Table 3.4 indicates that that at 95% BCI all the regression parameters (precipitation, NDVI, night LST and elevation) are significant and they lie within positive values. This implies that all the regression parameters significantly increases the zero-inflation probability i.e. they are more likely to correspond to excess zeros. In other words, higher probability of observing a zero count is associated with higher elevation, precipitation, NDVI, night LST. NDVI (0.68; 95% BCI = 0.47, 0.89) exhibited the strongest relationship with malaria in KZN compared to the other significant variables. This was followed by precipitation (0.07; 95% BCI = 0.04, 0.11).

Table 3.4 Posterior summary statistics for the spatial and temporal zero-inflated models for modelling malaria cases in KwaZulu-Natal, South Africa.

Regression parameter	SD	2.5% BCI	Mean	97.5% BCI
Intercept	0.06	0.38	0.50	0.63
LogPrecip	0.02	0.08	0.11	0.14
NDVI	0.11	0.71	0.91	1.12
Night LST	0.00	0.03	0.04	0.04
LogElev	0.01	0.03	0.05	0.07

LogPrecip = Log precipitation; LogElev = Log elevation; LST = land surface temperature; NDVI = Normalised difference vegetation index; SD = Standard deviation; BCI = Bayesian credible interval

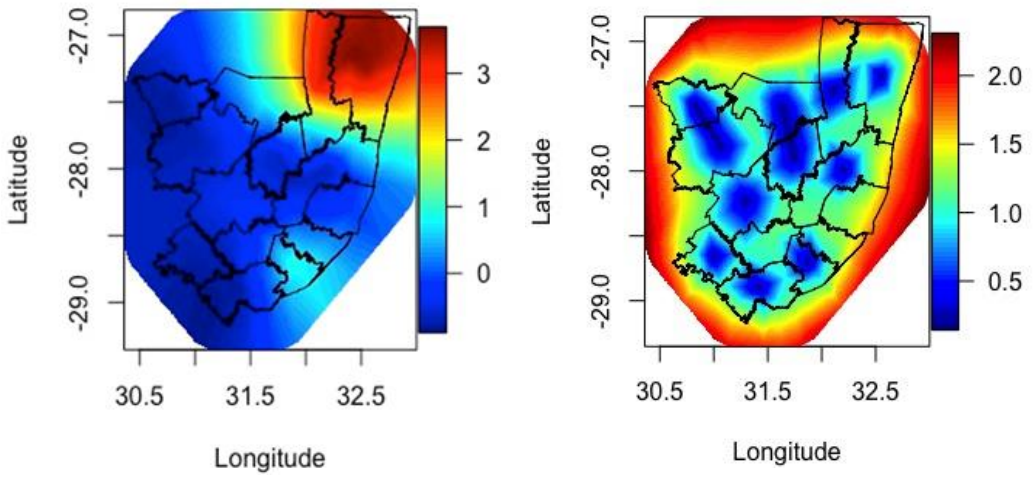


Figure 3.4 Posterior mean (left) and posterior standard deviation (right) of malaria cases in KwaZulu-Natal, South Africa.

The posterior mean map (Figure 3.4) indicates the region with the highest risk of malaria morbidity in KZN province is the Northern Eastern part, and there is a notable declining trend of malaria risk centrally and southwards. The lowest risk area cuts across the south and central parts of the study area with the model showing a slight variation in the south eastern part. The posterior standard deviation map (Figure 3.4) indicate the varying level of uncertainty across the province. The highest posterior errors across the province are at the periphery or borders of the local municipalities.

3.5 Discussion

In this study, the ZIP model was identified as the best model for the over-dispersed, excess zeros and the spatiotemporal dependencies of the malaria case data in the malarious areas of KZN after considering the influence of climatic variables. The results of the posterior statistics from the ZIP model indicates a significant relationship between NDVI, precipitation, elevation, night LST and malaria cases. In addition, the malaria risk map developed from this study showed that the tip of the north-eastern part of KZN province possesses the highest risk of malaria morbidity over the years 2005 to 2014. This is some worth consistent with malaria risk maps developed by the SA department of health in 2007 and 2013 using the geographical distribution of confirmed malaria cases [56]. Thus, improved health management strategies and targeted additional interventions is required to attain significant malaria risk reduction amongst the most vulnerable areas and populations.

To adequately address the issue of over dispersion arising from the excess zero or zero-inflated spatiotemporal malaria data, the spatiotemporal ZIP model is suggested to be a relevant model. This can be attributed to the fact that the ZIP model is a log link linear predictor model and as such, it is more reliable than the other models that employ a logit link function when spatial dependencies are considered [35, 57, 58]. It is worthy to note that some studies with a similar data structure conforms to this assertion. For example, the ZIP model was a suitable model for mapping the malaria incidence data with excess zeros in Afghanistan [42] and Northern Namibia [41]. Similarly, the ZIP model was considered the desirable model for developing a spatiotemporal HIV/TB model in North East SA [57], and an HIV model in New York, USA [40] using mortality data with excess zeros.

A good number of studies have attempted to show the relationship of malaria transmission and diverse climatic variables across different regions. This study indicates that NDVI is a significant variable for malaria transmission in KZN and it is the strongest predictor of malaria disease. The relevance of NDVI in malaria transmission modelling cannot be overemphasised. It is a vegetation index that can be used to assess the level of greenness of a vegetation in question [59, 60], and it can serve as a proxy for precipitation, near-surface humidity and surface water [59, 60]. Also, it was identified as the most important predictor in malaria transmission modelling across SSA [61]. Some studies conducted across SSA have shown that increase in vegetation indices can be used to predict increase in malaria risk [33, 34, 62]. Contrarily, this study shows that increase in NDVI is associated with low malaria risk, and it is consistent with a previous study conducted in Senegal [43]. The relationship between NDVI and malaria in this study can be explained by the fact that NDVI has constantly been reported to be associated with precipitation, near-surface humidity and surface water [59, 60]. However, the strength or form of the relationship is dependent on the structure of the ecosystem. For this reason, the effect high amounts of rainfall have on vector as discussed previously can also be related to high values of NDVI. Although, NDVI can provide information on vegetation intensity, it loses sensitivity over denser vegetation. In light of this characteristic, EVI is suggested to be a reliable substitute [63-65]. However, in this study, EVI was dropped in the preliminary phase of the analysis to guide against multicollinearity.

Another important predictor of malaria in KZN is precipitation. The posterior inference from this study suggests that higher precipitation reduces the risk of malaria transmission in KZN. This type of relationship between rainfall and malaria can be explained by the effect of excessive rainfall flushing the ground, which destroys malaria vector breeding sites and the vector at different aquatic developmental stages [12, 66-68]. Similar suggestions were reported in surveys conducted in Mali [34]. Contrarily, surveys conducted in Angola [33], Zambia [32] and China [31] reported that higher amounts of precipitation increase the risk of malaria transmission. While in some other studies no significant relationship between precipitation and malaria were reported [41, 42].

Several previous studies have reported the relationship between LST and malaria parasitemia risk. For example, in Angola [33] and Tanzania [69] increase in night LST were suggested to be a predictor variable for increased malaria risk. But in this present study, increase in night LST in KZN is related to reducing the risk of malaria. A similar outcome was reported in Senegal [43] and in Ghana [44]. Satellite derived measure of LST takes account of the thermal feature of the land surface, the intervening atmospheric radiation and emissions from a combination of different matters within a location. As an implication, the temporal characteristics or variations of LST may not be closely related to the near surface air temperature [70]. In spite of these limitations, the results show that night LST can be used to capture relative values of temperature spatiotemporally and as a potential pointer of malaria transmission in the malarious region of KZN.

Lowland areas (low elevation) are characterised by favourable temperature for the different developmental stages of malaria vectors and parasites, unlike highland areas that may not favour parasitic and vector development [17, 33, 44, 67]. This study conforms to this assertion. The North-Eastern part of KZN is characterised by a lower elevation of about 107m above sea level compared to the rest of KZN at 3000m above sea level [49], and malaria transmission risk is highest in the North-Eastern part of the province. Thus, the observation about the relationship between elevation and malaria cases is consistent with the study conducted by Gosoniu *et al.* [33] in Angola.

The uncertainties attributed to using climatic and environmental determinants to predict malaria disease makes it challenging for malaria surveillance and interventional purposes [71]. For this reason, in addition to the climatic and environmental inference provided by the spatiotemporal Bayesian ZIP modelling in this study, a map which can further guide the malarial interventional programmes in KZN was produced. The malaria risk map showed that there is a hotspot north-easterly in KZN bordering Swaziland and Mozambique between the years 2005–2014. This map supports similar patterns obtained from previously developed malaria risk maps [72, 73]. Jozini and uMhlabuyalingana local municipalities are the areas with the highest malaria transmission risk. The notable transmission risk in these areas can be explained to some extent by the population movement between neighbouring countries. This type of movement presents the greatest threat to zero local transmission, because their movement patterns are usually from regions of high transmission to regions of low transmission. Thus, this study provides evidence to support the renewed cross-border collaborative efforts, the MOSASWA (Mozambique, South Africa and Swaziland) malaria initiative instituted in 2015 [4]. The initiative aims to boost the progress made by the participating nations towards achieving zero local transmission by further strengthening collaboration between relevant academic institutions, sharing expertise, channelling intervention resources to vulnerable populations in the region (especially the mobile population and border populations) and sourcing for long-term financial support [4].

In addition to the MOSASWA initiative proposed to facilitate KZN's and SA's malaria program transition from pre-elimination to elimination, a modelling approach which takes account of the effects of population movement between the MOSASWA countries and from other malaria endemic countries is important. It will help understand the spatial and temporal implications of mobile population in high transmission areas. It will also serve as a guide for adequate dissemination of chemoprophylaxis message to mobile populations and travellers in malarious and non-malarious areas, and for setting up a quick response strategy to imported cases. Ultimately, it will result in timely channelling of malaria intervention resources to handle the threats that may arise from potential imported cases. Also, the KZN malaria program should be further strengthened and expanded by conducting routine genotyping of vectors, improved insecticide resistance monitoring, close monitoring of intervention resources to ensure adequate implementation, and formulation of malaria elimination commissions to provide technical and managerial guidance to malaria programmes at all levels (district, provincial and national).

3.6 Conclusions

The aim of the malaria programme in KZN is to develop elimination strategies and afterwards eradication strategies. The low and excess zero prevalence recorded in the malarious local municipalities revealed that the Bayesian spatiotemporal zero-inflated models can serve as a suitable tool for the relevant policy makers. Thus, spatiotemporal ZIP Bayesian modelling and the map produced in this study presents a valuable tool for understanding and monitoring the influence of climate variability on the spatial heterogeneity of malaria in KZN. They can play a significant role in the management, prioritising and allocation of intervention resources according to transmission variabilities. Also, this study reveals the importance to strengthen already existing cross-border collaborations for the fortification of KZN's malaria elimination target.

3.7 Competing interests

The authors declare that they have no competing interests.

3.8 Acknowledgments

The authors would like to thank the College of Agriculture, Engineering and Science of the University of KwaZulu-Natal for the doctoral research bursary awarded to Osadolor Ebhuoma. We also thank the malaria control program of KwaZulu-Natal, South Africa for providing the data.

3.9 Ethical approval

Not required.

3.10 References

1. South Africa National Department of Health, *Republic Of South Africa Malaria Elimination Strategy 2011–2018*. 2012: Pretoria, South Africa.

2. Ebhuoma, O., M. Gebreslasie, and L. Magubane, *Modeling malaria control intervention effect in KwaZulu-Natal, South Africa using intervention time series analysis.* Journal of Infection and Public Health, 2017. **10**(3): p. 334-338.

3. Maharaj, R., D. Mthembu, and B. Sharp, *Impact of DDT re-introduction on malaria transmission in KwaZulu-Natal.* South African Medical Journal, 2005. **95**(11): p. 871.

4. Moonasar, D., R. Maharaj, S. Kunene, B. Candrinho, F. Saute, N. Ntshalintshali, and N. Morris, *Towards malaria elimination in the MOSASWA (Mozambique, South Africa and Swaziland) region.* Malaria Journal, 2016. **15**(1): p. 419.

5. World Health Organisation. *World Malaria Report 2014*. 2014 Accessed 14/05/2015; Available from: http://www.who.int/malaria/publications/world_malaria_report_2014/wmr-2014-profiles.pdf?ua=1.

6. Gao, H.-W., L.-P. Wang, S. Liang, Y.-X. Liu, S.-L. Tong, J.-J. Wang, Y.-P. Li, X.-F. Wang, H. Yang, J.-Q. Ma, et al., *Change in Rainfall Drives Malaria Re-Emergence in Anhui Province, China.* PLOS One, 2012. **7**(8): p. e43686.

7. Garske, T., N.M. Ferguson, and A.C. Ghani, *Estimating air temperature and its influence on malaria transmission across Africa.* PLOS One, 2013. **8**(2): p. e56487.

8. Githeko, A.K., *Malaria And Climate Change.* Commonwealth Health Minister's Update 2009, 2009: p. 40-43.

9. Midekisa, A., B. Beyene, A. Mihretie, E. Bayabil, and M.C. Wimberly, *Seasonal associations of climatic drivers and malaria in the highlands of Ethiopia.* Parasites & Vectors, 2015. **8**(1): p. 1-11.

10. Yé, Y., V.R. Louis, S. Simboro, and R. Sauerborn, *Effect of meteorological factors on clinical malaria risk among children: an assessment using village-based meteorological stations and community-based parasitological survey.* BMC Public Health, 2007. **7**(1): p. 101.

11. Zinszer, K., R. Kigozi, K. Charland, G. Dorsey, T.F. Brewer, J.S. Brownstein, M.R. Kamya, and D.L. Buckeridge, *Forecasting malaria in a highly endemic country using environmental and clinical predictors.* Malaria Journal, 2015. **14**(1): p. 1.

12. Zayeri, F., M. Salehi, and H. Pirhosseini, *Geographical mapping and Bayesian spatial modeling of malaria incidence in Sistan and Baluchistan province, Iran.* Asian Pacific Journal of Tropical Medicine, 2011. **4**(12): p. 985-992.

13. Cairns, M., A. Roca-Feltrer, T. Garske, A.L. Wilson, D. Diallo, P.J. Milligan, A.C. Ghani, and B.M. Greenwood, *Estimating the potential public health impact of seasonal malaria chemoprevention in African children.* Nature Communications, 2012. **3**: p. 881.

14. White, M.T., J.T. Griffin, T.S. Churcher, N.M. Ferguson, M.-G. Basáñez, and A.C. Ghani, *Modelling the impact of vector control interventions on Anopheles gambiae population dynamics.* Parasites & Vectors, 2011. **4**(1): p. 153.

15. Craig, M.H., R. Snow, and D. le Sueur, *A climate-based distribution model of malaria transmission in sub-Saharan Africa.* Parasitology Today, 1999. **15**(3): p. 105-111.

16. Stefani, A., I. Dusfour, A.P.S. Corrêa, M.C. Cruz, N. Dessay, A.K. Galardo, C.D. Galardo, R. Girod, M.S. Gomes, H. Gurgel, et al., *Land cover, land use and malaria in the Amazon: a systematic literature review of studies using remotely sensed data.* Malaria Journal, 2013. **12**(1): p. 192.

17. Bodker, R., J. Akida, D. Shayo, W. Kisinza, H.A. Msangeni, E.M. Pedersen, and S.W. Lindsay, *Relationship between altitude and intensity of malaria transmission in the Usambara Mountains, Tanzania.* Journal of Medical Entomology, 2003. **40**(5): p. 706-17.

18. Omumbo, J.A., S.I. Hay, S.J. Goetz, R.W. Snow, and D.J. Rogers, *Updating Historical Maps of Malaria Transmission Intensity in East Africa Using Remote Sensing.* Photogrammetric Engineering and Remote Sensing, 2002. **68**(2): p. 161-166.

19. Cohen, J.M., K.C. Ernst, K.A. Lindblade, J.M. Vulule, C.C. John, and M.L. Wilson, *Topography-derived wetness indices are associated with household-level malaria risk in two communities in the western Kenyan highlands.* Malaria Journal, 2008. **7**: p. 40.

20. Cohen, J.M., K.C. Ernst, K.A. Lindblade, J.M. Vulule, C.C. John, and M.L. Wilson, *Local topographic wetness indices predict household malaria risk better than land-use and land-cover in the western Kenya highlands.* Malaria Journal, 2010. **9**: p. 328.

21. Li, T., Z. Yang, and M. Wang, *Temperature, relative humidity and sunshine may be the effective predictors for occurrence of malaria in Guangzhou, southern China, 2006–2012.* Parasites & Vectors, 2013. **6**(1): p. 155.

22. Tusting, L.S., B. Willey, H. Lucas, J. Thompson, H.T. Kafy, R. Smith, and S.W. Lindsay, *Socioeconomic development as an intervention against malaria: a systematic review and meta-analysis.* The Lancet, 2013. **382**(9896): p. 963-972.

23. Tatem, A.J., P.W. Gething, D.L. Smith, and S.I. Hay, *Urbanisation and the global malaria recession.* Malaria Journal, 2013. **12**(1): p. 133.

24. Ernst, K.C., K.A. Lindblade, D. Koech, P.O. Sumba, D.O. Kuwuor, C.C. John, and M.L. Wilson, *Environmental, socio-demographic and behavioural determinants of malaria risk in the western Kenyan highlands: a case-control study.* Tropical Medicine & International Health, 2009. **14**(10): p. 1258-1265.

25. Tatem, A.J., C.A. Guerra, C.W. Kabaria, A.M. Noor, and S.I. Hay, *Human population, urban settlement patterns and their impact on Plasmodium falciparum malaria endemicity.* Malaria Journal, 2008. **7**.

26. Hay, S.I., C.A. Guerra, A.J. Tatem, P.M. Atkinson, and R.W. Snow, *Urbanisation, malaria transmission and disease burden in Africa.* Nature Reviews Microbiology, 2005. **3**.

27. Arab, A., M.C. Jackson, and C. Kongoli, *Modelling the effects of weather and climate on malaria distributions in West Africa.* Malaria Journal, 2014. **13**: p. 126.

28. Gosoniu, L., P. Vounatsou, N. Sogoba, N. Maire, and T. Smith, *Mapping malaria risk in West Africa using a Bayesian nonparametric non-stationary model.* Computational Statistics & Data Analysis, 2009. **53**(9): p. 3358-3371.

29. Thomson, M., S. Connor, P. Milligan, and S. Flasse, *Mapping malaria risk in Africa: What can satellite data contribute?* Parasitology Today, 1997. **13**(8): p. 313-318.

30. Gebreslasie, M.T., *A review of spatial technologies with applications for malaria transmission modelling and control in Africa.* Geospatial Health, 2015. **10**(2).

31. Clements, A.C., A.G. Barnett, Z.W. Cheng, R.W. Snow, and H.N. Zhou, *Space-time variation of malaria incidence in Yunnan province, China.* Malaria Journal, 2009. **8**(1): p. 180.

32. Shimaponda-Mataa, N.M., E. Tembo-Mwase, M. Gebreslasie, T.N.O. Achia, and S. Mukaratirwa, *Modelling the influence of temperature and rainfall on malaria incidence in four endemic provinces of Zambia using semiparametric Poisson regression.* Acta Tropica, 2017. **166**(Supplement C): p. 81-91.

33. Gosoniu, L., A.M. Veta, and P. Vounatsou, *Bayesian geostatistical modeling of Malaria Indicator Survey data in Angola.* PLOS One, 2010. **5**(3): p. e9322.

34. Sogoba, N., P. Vounatsou, M.M. Bagayoko, S. Doumbia, G. Dolo, L. Gosoniu, S.F. Traore, Y.T. Toure, and T. Smith, *The spatial distribution of Anopheles gambiae sensu stricto and An. arabiensis (Diptera: Culicidae) in Mali.* Geospatial Health, 2007. **1**(2): p. 213-22.

35. Neelon, B.H., A.J. O'Malley, and S.-L.T. Normand, *A Bayesian model for repeated measures zero-inflated count data with application to outpatient psychiatric service use.* Statistical Modelling, 2010. **10**(4): p. 421-439.

36. Arab, A., M.L. Wildhaber, C.K. Wikle, and C.N. Gentry, *Zero-inflated modeling of fish catch per unit area resulting from multiple gears: application to channel catfish and shovelnose sturgeon in the Missouri River.* North American Journal of Fisheries Management, 2008. **28**(4): p. 1044-1058.

37. Arab, A., *Spatial and spatio-temporal models for modeling epidemiological data with excess zeros.* International Journal of Environmental Research and Public Health, 2015. **12**(9): p. 10536-10548.

38. Chipeta, M.G., B.M. Ngwira, C. Simoonga, and L.N. Kazembe, *Zero adjusted models with applications to analysing helminths count data.* BMC Research Notes, 2014. **7**(1): p. 856.

39. Ghosh, S.K., P. Mukhopadhyay, and J.-C.J. Lu, *Bayesian analysis of zero-inflated regression models.* Journal of Statistical Planning and Inference, 2006. **136**(4): p. 1360-1375.

40. Musal, M. and T. Aktekin, *Bayesian spatial modeling of HIV mortality via zero-inflated Poisson models.* Statistics in Medicine, 2013. **32**(2): p. 267-281.

41. Alegana, V.A., P.M. Atkinson, J.A. Wright, R. Kamwi, P. Uusiku, S. Katokele, R.W. Snow, and A.M. Noor, *Estimation of malaria incidence in northern Namibia in 2009 using Bayesian conditional-autoregressive spatial-temporal models.* Spatial and Spatio-Temporal Epidemiology, 2013. **7**: p. 25-36.

42. Alegana, V., J. Wright, S. Nahzat, W. Butt, A. Sediqi, and N. Habib, *Modelling the incidence of Plasmodium vivax and Plasmodium falciparum malaria in Afghanistan 2006–2009.* PLOS One, 2014. **9**.

43. Giardina, F., L. Gosoniu, L. Konate, M.B. Diouf, R. Perry, O. Gaye, O. Faye, and P. Vounatsou, *Estimating the burden of malaria in Senegal: Bayesian zero-inflated binomial geostatistical modeling of the MIS 2008 data.* PLOS One, 2012. **7**(3): p. e32625.

44. Kasasa, S., V. Asoala, L. Gosoniu, F. Anto, M. Adjuik, C. Tindana, T. Smith, S. Owusu-Agyei, and P. Vounatsou, *Spatio-temporal malaria transmission patterns in Navrongo demographic surveillance site, northern Ghana.* Malaria Journal, 2013. **12**: p. 63.

45. Moonasar, D., T. Nuthulaganti, P.S. Kruger, A. Mabuza, E.S. Rasiswi, F.G. Benson, and R. Maharaj, *Malaria control in South Africa 2000–2010: beyond MDG6.* Malaria Journal, 2012. **11**(294): p. 1475-2875.

46. Camp, K.G.T., *A Bio-Resource Classification For KwaZulu-Natal, South Africa*, in *School of Applied Environmental Sciences*. 1999, University of KwaZulu-Natal, South Africa: Pietermaritzburg, South Africa.

47. IRI/LDEO Climate data library. *NOAA NCEP CPC: Climate Prediction Center*. 2016 Accessed 04/09/2016; Available from: https://iridl.ldeo.columbia.edu/SOURCES/.NOAA/.NCEP/.CPC/?Set-Language=en.

48. IRI/LDEO Climate data library. *USGS: United States Geological Survey*. 2016 Accessed 04/09/2016; Available from: https://iridl.ldeo.columbia.edu/SOURCES/.USGS/?Set-Language=en.

49. IRI/LDEO Climate data library. *NOAA NGDC: National Geophysical Data Center*. 2017 Accessed 23/01/2017; Available from: https://iridl.ldeo.columbia.edu/SOURCES/.NOAA/.NGDC/?Set-Language=en.

50. Rue, H., S. Martino, and N. Chopin, *Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations.* Journal of the Royal Statistical Society: Series B (Statistical Methodology), 2009. **71**(2): p. 319-392.

51. Blangiardo, M. and M. Cameletti, *Spatial and spatio-temporal Bayesian models with R-INLA*. 2015: John Wiley & Sons.

52. Held, L., B. Schrödle, and H. Rue, *Posterior and cross-validatory predictive checks: a comparison of MCMC and INLA.* Statistical Modelling and Regression Structures, 2010: p. 91-110.

53. Ntzoufras, I., *Bayesian Modeling Using WinBUGS*. Vol. 698. 2011: John Wiley & Sons.

54. Gelman, A. and J. Hill, *Data analysis using regression and multilevel/hierarchical models*. 2006: Cambridge university press.

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

56. Morris, N., J. Frean, L. Baker, I.S. Ukpe, K.I. Barnes, and P. Kruger, *Re-defining the extent of malaria transmission in South Africa: implications for chemoprophylaxis.* South African Medical Journal, 2013. **103**.

57. Musenge, E., T.F. Chirwa, K. Kahn, and P. Vounatsou, *Bayesian analysis of zero inflated spatiotemporal HIV/TB child mortality data through the INLA and SPDE approaches: applied to data observed between 1992 and 2010 in rural North East South Africa.* International Journal of Applied Earth Observation and Geoinformation, 2013. **22**: p. 86-98.

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

59. Midekisa, A., G. Senay, G.M. Henebry, P. Semuniguse, and M.C. Wimberly, *Remote sensing-based time series models for malaria early warning in the highlands of Ethiopia.* Malaria Journal, 2012. **11**.

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

61. Ebhuoma, O. and M. Gebreslasie, *Remote Sensing-Driven Climatic/Environmental Variables for Modelling Malaria Transmission in Sub-Saharan Africa.* International Journal of Environmental Research and Public Health, 2016. **13**(6): p. 584.

62. Nygren, D., C. Stoyanov, C. Lewold, F. Mansson, J. Miller, A. Kamanga, and C.J. Shiff, *Remotely-sensed, nocturnal, dew point correlates with malaria transmission in Southern Province, Zambia: a time-series study.* Malaria Journal, 2014. **13**: p. 231.

63. Matsushita, B., W. Yang, J. Chen, Y. Onda, and G. Qiu, *Sensitivity of the Enhanced Vegetation Index (EVI) and Normalized Difference Vegetation Index (NDVI) to topographic effects: A case study in high-density cypress forest.* Sensors, 2007. **7**(11): p. 2636-2651.

64. Vina, A., G.M. Henebry, and A.A. Gitelson, *Satellite monitoring of vegetation dynamics: Sensitivity enhancement by the wide dynamic range vegetation index.* Geophysical Research Letters, 2004. **31**(4).

65. Huete, A., K. Didan, T. Miura, E.P. Rodriguez, X. Gao, and L.G. Ferreira, *Overview of the radiometric and biophysical performance of the MODIS vegetation indices.* Remote Sensing of Environment, 2002. **83**(1-2): p. 195-213.

66. Salehi, M., K. Mohammad, M.M. Farahani, H. Zeraati, K. Nourijelyani, and F. Zayeri, *Spatial modeling of malaria incidence rates in Sistan and Baluchistan province, Islamic Republic of Iran.* Saudi Medical Journal, 2008. **29**(12): p. 1791-6.

67. Noor, A.M., P.W. Gething, V.A. Alegana, A.P. Patil, S.I. Hay, E. Muchiri, E. Juma, and R.W. Snow, *The risks of malaria infection in Kenya in 2009.* BMC Infectious Diseases, 2009. **9**: p. 180.

68. Singh, N. and V. Sharma, *Patterns of rainfall and malaria in Madhya Pradesh, central India.* Annals of Tropical Medicine & Parasitology, 2002. **96**(4): p. 349-359.

69. Gosoniu, L., A. Msengwa, C. Lengeler, and P. Vounatsou, *Spatially explicit burden estimates of malaria in Tanzania: bayesian geostatistical modeling of the malaria indicator survey data.* PLOS One, 2012. **7**(5): p. e23966.

70. Vancutsem, C., P. Ceccato, T. Dinku, and S.J. Connor, *Evaluation of MODIS land surface temperature data to estimate air temperature in different ecosystems over Africa.* Remote Sensing of Environment, 2010. **114**(2): p. 449-465.

71. Adu-Prah, S. and E.K. Tetteh, *Spatiotemporal analysis of climate variability impacts on malaria prevalence in Ghana.* Applied Geography, 2015. **60**: p. 266-273.

72. Morris, N., J. Frean, L. Baker, I.S. Ukpe, K.I. Barnes, P. Kruger, A. Mabuza, E. Raswiswi, R. Maharaj, L. Blumberg, et al., *Re-defining the extent of malaria transmission in South Africa: Implications for chemoprophylaxis*. 2013. Vol. 103. 2013.

73. South Africa Medical Research Council, *Malaria Risk Map For South Africa*. 2013, Durban: Malaria Research Unit, South Africa Medical Research Council.

CHAPTER 4: SOCIO-ECONOMIC DETERMINANTS OF MALARIA TRANSMISSION RISK IN KWAZULU-NATAL, SOUTH AFRICA: A BAYESIAN APPROACH

This chapter is based on:

Ebhuoma O, Gebreslasie M, Ropo O (Submitted). Socio-economic determinants of malaria transmission risk in KwaZulu-Natal, South Africa: a Bayesian inference approach.

4.1 Abstract

Socio-economic status (SES) has been suggested to sustain malaria transmission, which in turn can propel the cycle of poverty. Thus, a deep understanding of the SES that influences malaria risk is vital because it will guide towards creating polices and strategies that will concurrently help combat malaria transmission, improve socio-economic conditions and strengthen the malaria elimination campaign in KwaZulu-Natal (KZN), South Africa (SA). The main purpose of this study is to investigate the existed relationship between SES and malaria incidence in KZN, SA, using the Bayesian approach. Database of demography and socio-economic information, and clinically confirmed malaria case data aggregated at the local municipality level for 2011 were obtained from Statistics SA (Census 2011 Municipal report- KZN, SA) and the malaria control program of KZN, SA respectively. The investigator used the 2011 dataset (SES and malaria incidence) because it completely covered the study area and census was conducted in 2011 in SA. The association between SES and malaria incidence was evaluated by employing the Bayesian multiple regression model to obtain the posterior samples via a Markov chain Monte Carlo (MCMC) methodology. The obtained posterior samples shows that, all the SES variables employed are significant and positive determinants of malaria disease at 95% Bayesian credible interval (BCI). From the variables that represent low SES used in this study population, lack of toilet facilities (OR =12.54; 95% BCI = 0.63, 24.38) exhibited the strongest association with malaria and highest risk of malaria disease. This was followed by no education (OR =11.83; 95% BCI = 0.54, 24.27) and lack of electricity supply (OR =10.56; 95% BCI = 0.43, 23.92), respectively. Low SES can potentially sustains malaria transmission and burden. As an implication, poverty alleviation and malaria intervention resources should be incorporated side by side into the socio-economic framework to attain zero malaria transmission. Therefore, the relevant policy makers and departments should stimulate additional sustainable developmental approaches that combines both improved malaria intervention resources and socio-economic conditions, which in turn, will help strengthen the malaria elimination goals in KZN, SA.

Keywords*: Malaria, Socio-economic status, Bayesian, KwaZulu-Natal, South Africa.*

4.2 Introduction

Malaria is endemic in the north eastern part of KwaZulu-Natal (KZN) province, South Africa (SA). The malaria control strategy in the province focuses more on indoor residual spraying, insecticide treated bed nets and case detection alongside treatment with antimalarials [1-5]. A number of studies have focused on the implications of above mentioned interventions on malaria disease transmission and also have suggested ways that they can be utilised optimally [2, 6-9]. Less attention has been paid to understand the influence socio-economic status (SES) and factors have on malaria transmission. Although, the KZN malaria control and intervention strategy has led to commendable malaria control transmission and zero transmission mainly in urban areas [3, 5], transmission is still prevalent in semi urban and rural areas [5]. At a time when the KZN health department is putting more effort to eliminate malaria, it is important to investigate and identify relevant factors of malaria transmission in the province [5]. Literature revealed that the spatial and temporal heterogeneity of malaria transmission is influenced by ecological/environmental and climatic factors such as rainfall, temperature, humidity, altitude and human dwelling close to water bodies [10], human factors such as migration and urbanisation [11], water and environmental management strategies such as drains, dams, water reservoirs and irrigations constructed proximal to human habitation [12], access to quality health care [13] and SES [14]. Thus, the focus of this chapter is to evaluate the implications of SES on malaria transmission in KZN. Addressing SES for malaria control will equip the relevant authorities and policy makers with the necessary information to improve socio-economic conditions on one hand. On the other hand, adopt appropriate malaria intervention strategies in addition to the already existing ones.

For a region where most of these factors are under control and malaria transmission is greatly reduced and control efforts are at elimination stage, it is necessary to focus on factors that remain unexhausted such as SES. An understanding of the SES of a population is important, because it can quantify the small population susceptible to malaria which are normally difficult to identity, favourable conditions for mosquito proliferation and malaria transmission. Socio-economic factors associated with rural settings predispose the rural people to a higher risk of contracting malaria as compared to those in the urban areas. This is because they possess characteristics, and reside in environs with features that support malaria transmission [15]. For instance, literacy level which is generally lower in rural and semi urban settings compared to

urban settings, can affect the knowledge of malaria prevention and control, thus sustaining malaria endemicity. In Bihar and Jharkhand rural areas in India, majority of the people possessed low educational qualification, and they showed inferior knowledge and low acceptance for intervention resources as compared to those with higher education [16]. This outcome is supported by previous cross sectional studies in Manipur, India [17] and in Zambia [18] in which the rural people exhibited low level of knowledge of malaria parasitaemia and epidemiology [19, 20].

Poor housing and socio-economic conditions like lack of sanitary facilities and electricity have implications for malaria transmission and epidemiology. In a cross-sectional study in three high malaria transmission regions in Ethiopia, Ayele *et al.* [21] reported that households without toilet facilities were more likely to contract malaria disease by fitting the generalised linear regression model. However, Monteiro *et al.* [22] found that in a low malaria transmission area in the Brazilian Amazon, malaria had no relationship with basic sanitation. Lack of electricity may result in heightened risk of malaria, because it prevents households from using air conditioners or electric fan which wards off mosquito as suggested by Hewitt *et al.* [23]. However, in a case-control study conducted in a low endemicity urban area of Peru, Rosas-Aguirre *et al.* [24] argued that electricity was not a significant determinant of malaria parasitaemia in the univariate and multivariate analysis.

Although being in employment can facilitate means of transportation to health facility and purchase of medicines, the influence of employment on malaria depends mainly on the type of employment. As such, field labourers are at higher risk of contracting malaria because they are predominantly exposed to malaria vectors in the fields, and this in turn, predisposes their households to higher risk of malaria disease [25, 26]. For example, in the rural areas of central India, labourers showed higher risk of contracting malaria as compared to office workers/businessmen, but lack of employment had no significant impact on malaria occurrence in households in the multivariate analysis [27]. Also, unemployment was not a significant determinant of malaria in Panama after employing the multivariate analysis in a case-control study [28]. While in a case-control study in a low transmission urban area of Peru, employment was not a significant risk factor of malaria in the univariate and multivariate logistic regression analysis [24].

Various reports indicate that the gender at greater risk of malaria varies in different settings, due to occupation. Mosha *et al.* [29] and Yadav *et al.* [30] suggested that males are at higher risk of malaria infection in Tanzania and India, respectively. Although, in Ethiopia [21] and Panama [28] females are more susceptible to malaria. Even though malaria affects males and females of all age groups, children and elderly are more susceptible to malaria, due to low immunity and social factors [31, 32]. Gosoniu *et al.* [33] suggested that children below 5 years old in Tanzania exhibited a significant association with malaria transmission after employing the Bayesian geostatistical models. Furthermore, Abe *et al.* [34] reported that children between 3-5 years in Vietnam were associated with malaria in the univariate and multivariate analysis, while Fayehun *et al.* [31] suggested that people aged 60 years and above are vulnerable to malaria transmission in Nigeria. This is supported by a previous report, which suggests that malaria transmission and mortality is highest amongst people above 65 years [35].

The above studies conducted in different malaria endemic settings suggested that low SES, socio-economic deprivation and poverty sustains and exacerbates malaria transmission at different levels. Most of the studies employed models based on the classical theory to identify relevant socio-economic determinants of malaria transmission. However, the Bayesian modelling approach has been applied by different authors in diverse settings [33, 36-39], and suggested that the approach is a good alternative to classical models [10, 40, 41]. The Bayesian approach regards parameter(s) as random variables that are specified by a prior distribution or information. This prior distribution or information is incorporated with the traditional likelihood to derive the posterior distribution of the parameter(s) of interest on which the statistical inference is based using the Markov chain Monte Carlo (MCMC) methods such as the Metropolis-Hastings algorithm and Gibbs sampler [40]. Accordingly, this chapter aims to provide knowledge on the SES risk factors for malaria transmission in KZN, SA, employing the Bayesian method.

4.3 Methodology

A Bayesian multiple regression approach was employed to assess the association between selected determinants of SES (no education, no electricity, unemployment, no toilet facilities) and demography (gender and vulnerable groups at great risk of contracting malaria- that is, children aged less than 5 years and elderly aged older than 65 years) with malaria incidence in

the local municipalities endemic to malaria in KZN. This would identify important SES that can influence malaria transmission in the endemic areas. In addition, a Bayesian analysis of the posterior distribution of the parameters is specified.

4.3.1 Study area

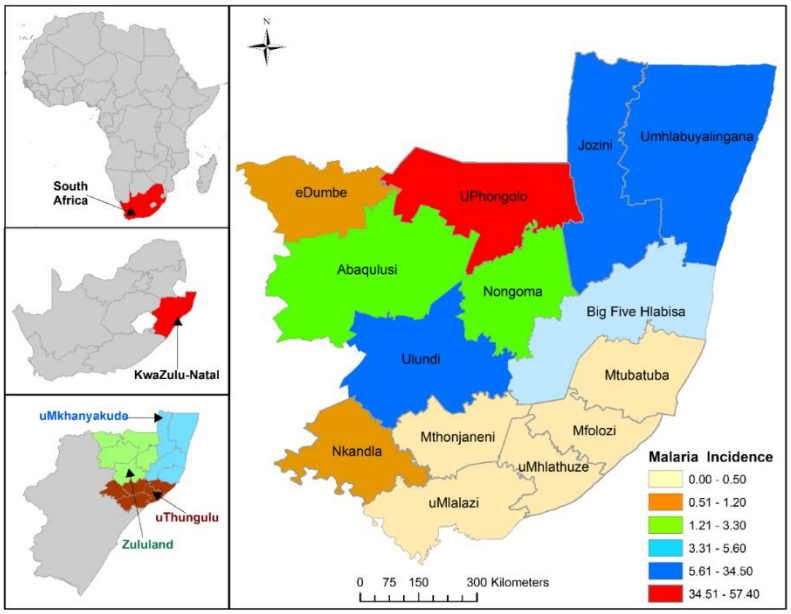


Figure 4.1 The study area and malaria incidence in uMkhanyakude, Zululand and uThungulu District in KwaZulu-Natal, South Africa (2011)

The study area is located in the north-eastern part of KZN province (Figure 4.1) covering the local municipalities of uMkhanyakude, uThungulu and Zululand district municipalities. It is bordered internationally by The Kingdom of Swaziland and The People's Republic of Mozambique to its north. It has a long shoreline along the Indian Ocean to its east and stretching down south-eastwards. The region possesses a sub-tropical climate with the majority of malaria incidents observed during October to May (the rainy months), with a seasonal peak usually in January and March [4, 5]. The average annual rainfall ranges from 500mm to 2000mm. Along the coastal areas, the summer temperatures is between 24°C to 32°C, and mean winter temperature is about 20°C. The Midlands generally possesses a mild climate with relatively high summer rainfall and dry winters. The elevation measure of the region varies from sea level to over 3000m. The vegetation of the study area comprises of coastal forest and thornveld along the coast. Towards the inlands, lowveld, highland sourveld, Natal sour sandveld, valley bushveld and tall grassveld

vegetation are found. Lowveld and thornveld characterises the low-lying hot and dry regions of Northern KZN [42].

4.3.2 Data

The investigator obtained and used data from two separate sources. The database of daily records of clinically confirmed malaria cases for 2011 were obtained from the malaria control program of KZN, SA and aggregated at the local municipality level for uMkhanyakude, uThungulu and Zululand district municipalities. Figure 4.1 shows the distribution of malaria incidence in the respective local municipalities in 2011. The investigator obtained the database of socio-economic and demographic information for the study area from Statistics SA (Census 2011 Municipal report – KZN, SA) [43]. This database is freely available to the general public although, it is protected by scientific and ethical clearance and authorisation. The demographic and socio-economic variables used were: gender, children (less than 5 years old), elderly (above 65 years old), no education, no electricity, no toilet facilities, unemployment.

4.3.3 Model development

To evaluate the association between the predictor variables (demographic and socio-economic variables) with the dependent variable (malaria incidence by year) at the local municipality level, a Bayesian multiple regression of the specified study variables was conducted using the WinBUGS software.

4.3.3.1 Bayesian multiple regression model formulation

A regression model comprise of: (1) dependent variable(s), which represents the stochastic part whose effect is uncertain before the analysis. (2) The predictor variable(s), which represents the non-stochastic or fixed parts and (3) a parameter that links the two set of variables. The model can be expressed as [40]

$$Y|X_1,X_2,\ldots\ldots,X_p \sim \alpha(\boldsymbol{\theta}) \qquad (4.1)$$

where Y is the dependent variable, $X_1,X_2,\ldots\ldots,X_p$ are the predictor variables and $\alpha(\boldsymbol{\theta})$ is a distribution with parameter vector $\boldsymbol{\theta}$.

The simplest and commonly used distribution for the regression model is a normal distribution and it can be written as

$$Y|X_1,X_2,\ldots\ldots, X_p \sim normal(\mu,\sigma^2) \tag{4.2}$$

where $Normal(\mu,\sigma^2)$ is the normal distribution, μ is mean and σ^2 is variance.

While a multiple regression model with a single dependent variable (univariate), having a normal distribution with mean and variance can be summarised and rewritten as [40]

$$Y|X_1,X_2,\ldots\ldots, X_p \sim normal\left(\mu,\left(\beta, X_1,X_2,\ldots\ldots, X_p\right)\sigma^2\right) \tag{4.3}$$

with

$$\mu\left(\beta, X_1,X_2,\ldots\ldots, X_p\right) = \beta_0 + \beta_1 X_1 + \ldots\ldots. + \beta_p X_p = \beta_0 + \sum_{i=1}^{p} \beta_i X_i \tag{4.4}$$

where σ^2 and $\beta = (\beta_0,\beta_1,\ldots\ldots, \beta_p)^T$ are the set of regression parameters under estimation.

To specify the model in WinBUGS, the likelihood function for the observed sample and the prior information (or prior distribution) for the parameters are required. Thus, the likelihood function (extracted from equations (4.3) and (4.4)) is expressed as [40, 44].

$$Y_i \sim normal\left(\mu_i,\sigma^2\right) \tag{4.5}$$

$$\mu_i = \beta_0 + \beta_1 x_{i1} + \ldots\ldots. + \beta_p x_{ip} \quad \text{for } i = 1, \ldots\ldots.., n \tag{4.6}$$

While the prior distribution for all the parameters are assumed to have the structure

$$S(\beta,\tau) = \prod_{j=0}^{p} S(\beta_i)\, S(\tau)\,, \tag{4.7}$$

$$\beta_i \sim normal\left(\mu\beta_i, R_j^2\right) \text{ for } j = 0,1, 2\ldots\ldots\ldots,p \text{ and}$$

$$\tau \sim gamma\ (a,b) \tag{4.8}$$

From the above, the normal distribution is selected to exhibit the prior information for β and the gamma distribution for the precision parameter, $\tau = \sigma^{-2}$.

However, because the investigator considered $\boldsymbol{\theta}$ to be a continuous parameter, a uniform prior distribution for the regression model was set up. Therefore, the formulated uniform prior distribution employed in this analysis is given as

$$\beta_i \sim uniform\ (0.0, 25) \text{ for } j = 0,1, 2\text{…….,p and}$$

$$\tau \sim \text{gamma } (0.001,0.001) \tag{4.9}$$

The likelihood coupled with the priors were used to obtain the posterior distributions and results of the model parameters via the MCMC approach known as the Gibbs sampling technique. The MCMC approach adopted in this chapter is described in details elsewhere [40].

4.3.3.2 MCMC implementation and convergence

The WinBUGS software was used to obtain the posterior samples of model parameters for the multiple regression model via the MCMC approach known as the Gibbs sampling. The investigator assumed flat but proper priors as expressed in equation (4.8) for the model. The investigator ran three parallel chains (with different starting points) for 500,000 iterations of the MCMC with a burn-in period of 200,000 iterations and a thinning interval of 5. Thus, to ascertain if convergence was reached in the model, four different diagnostic tests were employed. These are the Monte Carlo (MC) errors calculation of all the regression and precision parameters, assessing the autocorrelation plot, the trace plot and the Gelman-Rubin convergence diagnostic test. The Gelman- Rubin convergence diagnostic test was employed, because more than one chain (three parallel chains) were generated simultaneously.

For notational and modelling convenience in WinBUGS, the Greek symbols β and τ were written as beta and tau, respectively. In the model developed, beta1 was used for male, beta2 for female, beta3 for children aged less than 5 years, beta4 for adult aged greater than 65 years, beta5 for no education, beta6 for no electricity, beta7 for no toilet facilities, beta8 for unemployment, and tau for the precision parameter.

4.4 Results

4.4.1 Spatial distribution of malaria incidence

The spatial distribution of malaria data of all the local municipalities obtained from the malaria control program of KZN, SA showed that malaria incidence ranged from 0 to 57.40 during 2011 (Figure 4.1). About 72% of the total malaria incidence were reported from uPhongolo, Jozini and uMhlabuyalingana local municaplities reported in 2011 were observed to be clustered in the areas neighbouring Mozambique and Swaziland. uPhongolo local municipality which recorded the highest incidence of malaria (about 37%) shared boarders with Swaziland and Mpumalanga province northwards (Figure 4.1).

4.4.2 Posterior distribution of determinates associated with increased risk of malaria disease.

In interpreting the posterior statistics, it is vital to note that a regression parameter with a positive posterior mean, exhibited a significant positive relationship with malaria incidence at 95% Bayesian credible interval (BCI) if the interval does not include zero (0) (Table 4.1). While a regression parameter with negative posterior mean and interval containing zero (0), suggested an inverse relationship. But, from the results in Table 4.1 regression parameters with negative posterior mean were not observed.

The posterior summary statistics (Table 4.1) obtained from the Bayesian multiple regression model analysis suggests that all the variables employed in the model are significant and positive determinants of malaria at 95% BCI. However, in terms of gender, males (odds ratio (OR) =8.78; 95% BCI =0.29, 23.18) were slightly at a higher risk of contracting malaria disease as compared to females (OR =8.36; 95% BCI = 0.28, 22.83) in the study population. Considering the population groups vulnerable to malaria parasitaemia, children less than 5 years old (OR =11.43; 95% BCI = 0.50, 24.18) are less likely to contract malaria disease compared to adults over 65 years old (OR =12.17; 95% BCI = 0.58, 24.33).

From the variables or determinants that represents low SES used in this study population, lack of toilet facilities (OR =12.54; 95% BCI = 0.63, 24.38) exhibited the strongest association with malaria and highest risk of malaria disease. This was followed by no education (OR =11.83; 95% BCI = 0.54, 24.27) and lack of electricity supply (OR =10.56; 95% BCI = 0.43, 23.92)

respectively, while unemployment was identified as the weakest significant variable as compared to the other variables that represent low SES.

Table 4.1 Posterior summary statistics for the multiple regression model.

Regression parameter	SD	2.5% BCI	Estimate	97.5% BCI	MC error
beta1	6.55	0.29	8.78	23.18	0.0158
beta2	6.39	0.28	8.36	22.83	0.0148
beta3	7.16	0.50	11.43	24.18	0.0174
beta4	7.21	0.58	12.17	24.33	0.0170
beta5	7.19	0.54	11.83	24.27	0.0172
beta6	6.99	0.43	10.56	23.92	0.0166
beta7	7.22	0.63	12.54	24.38	0.0165
beta8	6.81	0.35	9.67	23.59	0.0162
Precision parameter					
tau	0.001026	4.11E-04	0.001807	0.004326	2.35E-06

SD= Standard deviation; MC error= Monte Carlo error; BCI = Bayesian credible interval beta1=Male; beta2=Female; beta3=Children (< 5 years); beta4=Adult (>65 years); beta5=no education; beta6= no electricity; beta7= no toilet facilities; beta8 = Unemployment.

4.4.3 MCMC output and convergence diagnostics

The MCMC output results indicated that in obtaining the posterior samples of all the monitored parameters (regression coefficients and precision parameter), no convergence issue was experienced. The visual inspection of the trace plots or simulation plot for selected regression parameters presented in Figure 4.2 reveals that all the generated values were within a parallel zone and notable patterns were not observed. Thus, the simulation is uniform throughout the plot.

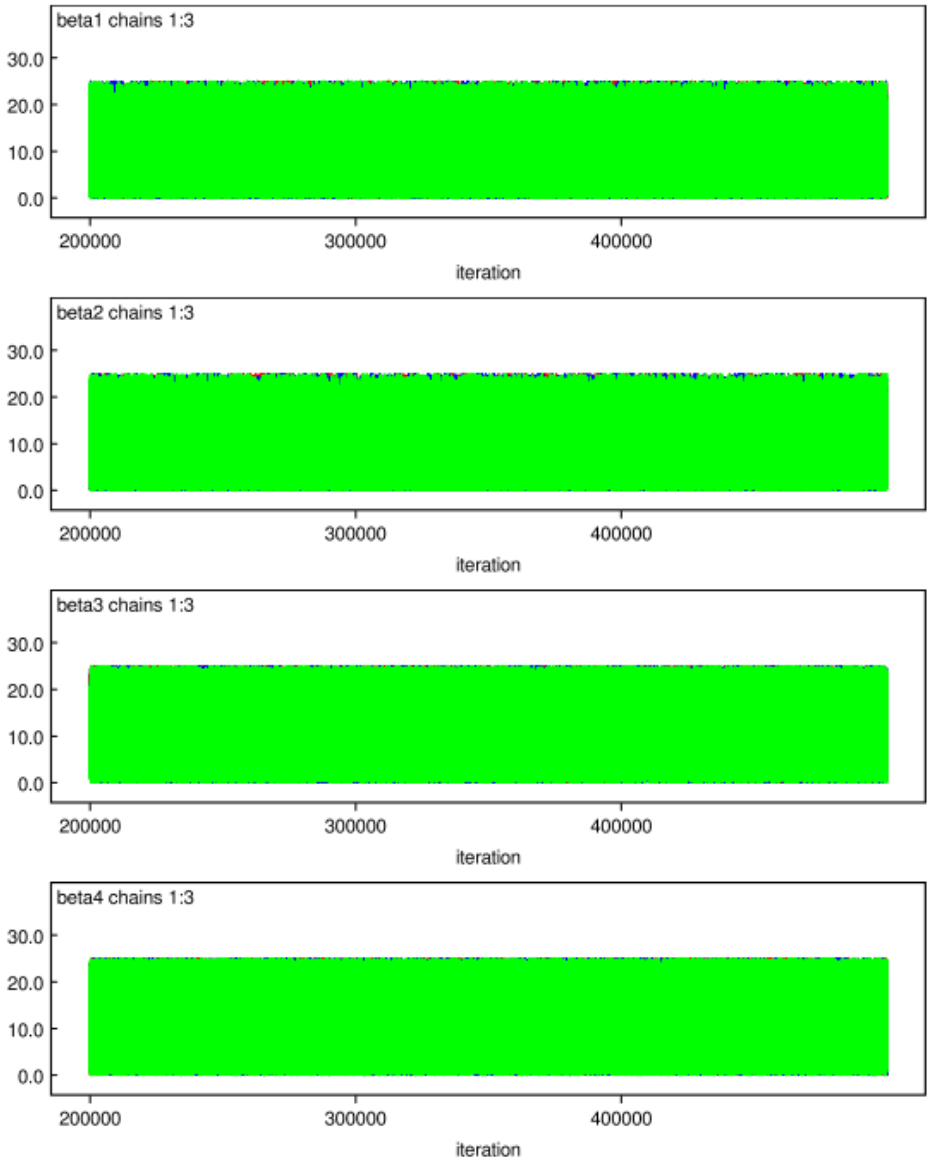


Figure 4.2 Trace plots for some selected predictor parameters

The autocorrelations functions for all the monitored parameters decayed sharply and were low (Figure 4.3). This indicates quick mixing of the chains and fast convergence to the posterior distributions. At this point, it was assumed convergence was reached. However, the investigator continued with the diagnostics and observed that the value of the MC errors calculated for the regression and precision parameters are very small compared to their corresponding estimated posterior standard deviations (SD). In other words, the MC error is less than 5% of the corresponding estimated posterior SD (See Table 4.1). This indicates the posterior mean was

estimated with high precision and convergence was also assumed. Finally, a formal convergence diagnostic was implemented by using the Gelman and Rubin plots and the shrink factor (also known as the scale reduction point estimate) by using three different chains with three different initial points.

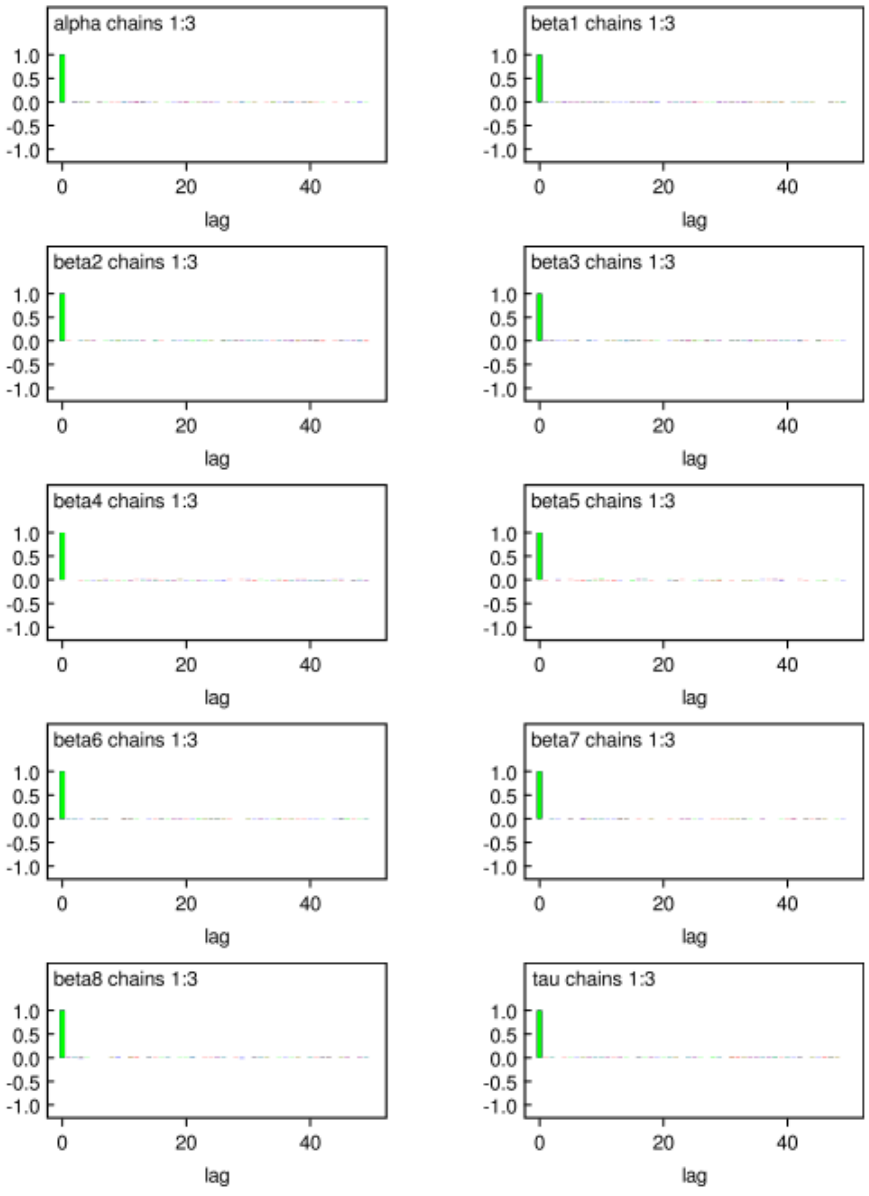
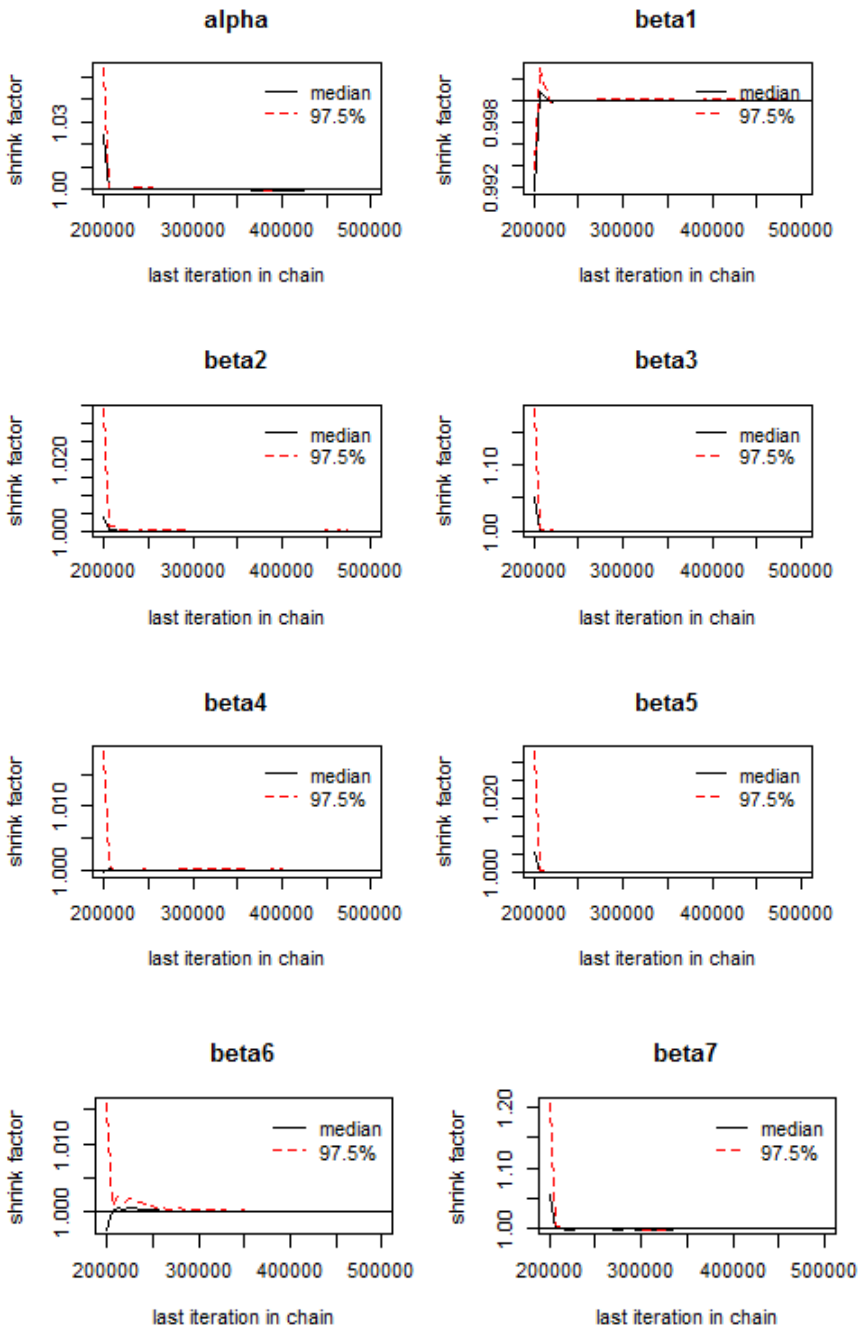


Figure 4.3 Autocorrelation function plot for the regression and precision parameters

From the Gelman and Rubin shrink factor plots illustrated in Figure 4.4, the shrink factor for all the monitored parameter got to 1 abruptly as the number of iterations increases. At this point,

the investigator can confidently say convergence was reached and also sampling is from the right posterior means.



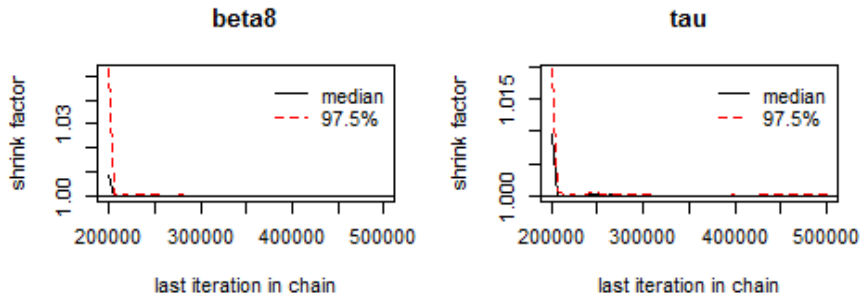


Figure 4.4 Gelman and Rubin shrink factor plots for the regression and precision parameters

4.5 Discussion

This chapter evaluates the relationship between selected SES with malaria incidence risk in north eastern part of KZN employing a Bayesian inference approach. The investigator formulated a Bayesian regression model after adequate burn-ins and successful convergence.

It is a well-known fact that malaria disease does not thrive on its own [14, 21, 22, 28, 45]. Considering the nature of SES on malaria and its dual influence, deprived communities and households maybe faced with a lingering and reinforcing cycle of malaria disease. Population characterised by low SES possesses elements that exposes and sustains high malaria transmission. On the contrary, high SES communities or households possesses measures to sustain low or none malaria transmission, that may support their high SES. Hence, the investigator suggests that the following determinants of low SES are associated with malaria risk at the local municipalities endemic to malaria in KZN in order of highest malaria risk transmission. They are lack of toilet facilities, lack of education, lack of electricity, and unemployment. The contributory mechanism between these low SES and malaria transmission runs side by side, which in turn, can result in a vicious cycle of poverty. The findings corresponds with some previous studies conducted in India, South America and sub-Saharan Africa in which households without toilet facilities, unemployment, illiteracy and lack of electricity were at higher risk of contracting malaria compared to those exposed to medium and high SES [21, 27, 28, 46]. However, these findings are contrary to the study conducted by Somi *et al.* [47] in Tanzania who found no relationship between malaria and SES. Also, in the study

conducted by Obaldia [28], lack of electricity and unemployed did not exhibit an association with malaria disease.

Taking into account the epidemiological nature of malaria, and its inclination with the environment, it should not be surprising that lack of toilet facilities has the highest risk of being exposed to malaria as compared to other determinants that represent low SES of malaria assessed in this chapter. This demonstrates the heightened risk and exposure to malaria disease in households without decent and safe toilet facilities, and as such, open defecation is practised. Consequently, increased malaria transmission risk is certain due to numerous and daily exposure to exophagic mosquito vectors from outdoor defecation. This is consistent with the findings by Ayele *et al.* [21] and Ayele *et al.* [46], who suggested that households without toilet facilities in Amhara, Oromiya and Southern Nation Nationalities and People regions of Ethiopia were more likely to test positive for malaria disease. On the contrary, no association was reported in Panama [28] and in the Brazilian Amazon [22].

Another important determinant of malaria identified in this chapter is education. Notable risks of malaria transmission risk among those without education can be attributed to weak knowledge and understanding of malaria transmission and prevention [45, 48]. They may also not understand the resting behaviour and breeding of mosquitoes. Thus, the provision of quality education and setting up communication activities to equip the uneducated people and also those with low level of education in KZN about malaria may result in better application of intervention resources, and reduce the risk of malaria transmission.

The investigator found out that lack of electricity was another risk factor that can support malaria transmission. Lack of electricity may result in households sleeping outdoors in malaria transmission seasons, due to lack of power to operate electric fan or air conditioner that can serve as a form of malaria intervention to ward off mosquitoes. This was revealed in a cohort study conducted in Pakistan aimed at assessing various malaria vector intervention techniques. It was reported that electric fan significantly reduced the total number of *Anopheles stephensi and culicine* mosquitoes entering the huts, and the amount of blood-fed mosquitoes caught [23].

Unemployment was identified as a significant factor of malaria transmission, but not as strong as the other socio economic variables evaluated in this chapter. This can be explained by the fact that health care is free in SA [49], thus eliminating the financial burden among those that are unemployed and in pursuance of medical attention. In addition, free routine indoor residual spraying is carried out in endemic areas by the KZN malaria program [50]**.** Nevertheless, its significance can be attributed to the direct relationship between unemployment and the other variables that represent low SES, and in turn, they present a complex interrelationship with malaria disease.

Apart from socio-economic factors, other demographic factors exhibited significant effects on the risk of malaria in KZN. Regarding gender, both male and female exhibited significant effects on the risk of malaria, but males showed a higher risk of contracting malaria disease in the study area. This corresponds to a previous study conducted in Tanzania and India that suggests that males are at higher risk of malaria [29, 30]. Contrarily, females were discovered to be more susceptible to malaria in Ethiopia [21] and Panama [28]. Female susceptibility can be linked to domestic activities and endophagic mosquitoes, while males susceptibility can be attributed to exophagic mosquitoes and outdoor activities like fishing, agriculture and hunting activities [28]. In terms of age group susceptible to malaria in KZN, the findings suggest elderly are more vulnerable to malaria compared to children less than 5 years old. This complements previous surveys that suggested elderly people are more prone to contracting malaria disease [35, 51]. Even though the mechanism is not clear, it is assumed that waning immunity with age or low immunity among elderly people may be responsible. It is worthy of note that the susceptibility and disparity in malaria transmission pertaining to gender and age can be attributed to behavioural/ life style and level of immunity respectively.

Considering the links between low SES and malaria transmission, conceptualizing interventions aimed at improving the living conditions by SA's department of social development, and the investment of malaria intervention resources (prevention and treatment) by both SA department of health and social development is vital for sustained poverty alleviation in malaria-endemic communities. This means the provision of malaria intervention resources should be considered as a means of both health intervention and poverty alleviation. This proposed double barrelled approach and collaborations between both departments can possibly result in sustained malaria

elimination. Countries like Greece, Italy, Spain and USA proved that the double barrelled approach lead to sustained malaria elimination [14, 52, 53]. They incorporated rigorous anti-malaria intervention, socio-economic improvement and adequate environmental management strategies. In other words, they modified human dwellings or behaviour to limit human-vector interaction-e.g., improved housing conditions, mosquito proofing of houses, usage of dichlorodiphenyltrichloroethane (DDT) in their indoor residual spaying regime, installation and maintenance of drains and draining of swampy areas to permanently destroy breeding sites [14, 52, 53]. In the same vein, a randomised controlled trial conducted by Kirby *et al.* [54] in The Gambia revealed that better housing conditions can significantly reduce malaria burden. In the randomised controlled trial, house screening (full screening of windows, doors and closing of eaves) revealed that the risk of children contracting malaria disease reduced by 50%.

In summary, this chapter suggests that the risk of malaria disease exhibits a significant effect in areas or population in a lower socio-economic bracket. As an implication, poverty alleviation and malaria intervention resources should be incorporated side by side into the socio-economic framework to attain zero malaria transmission. Therefore, the relevant policy makers and departments should invest more in sustainable developmental approaches that combine both improved malaria intervention resources and socio-economic conditions. This can ultimately help strengthen the malaria elimination goals in KZN.

4.6 Conclusions

Malaria can potentially sustain poverty in areas of low SES. The low malaria transmission and burden of malaria in KZN presents an avenue to study the relationship between determinants of low SES and malaria at the local municipality level to help guide provision of relevant intervention resources, policy and legislation. This chapter suggests that the following determinants of low SES play a significant role in malaria transmission and burden: illiteracy, lack of electricity, lack of toilet facilities and unemployment. Other factors that had an effect on the risk of malaria are gender, children less than 5 years old and adult above 65 years old. This means low SES, socio-economic deprivation and poverty can sustain and exacerbate malaria transmission in KZN. Eradicating and paying close attention to these risk factors, can contribute and accelerate the attainment of malaria elimination status alongside poverty alleviation. These outcomes suggest that a collaborative venture between SA's department of

social development and the department of health can provide improved and sustainable socio-economic conditions and, allocation of malaria intervention resources, and in turn, result in control and elimination of malaria in KZN, SA.

4.7 Competing interests

None declared.

4.8 Acknowledgments

The authors would like to thank the College of Agriculture, Engineering and Science of the University of KwaZulu-Natal for the doctoral research bursary awarded to Osadolor Ebhuoma. We also thank the malaria control program of KwaZulu-Natal, South Africa, and Statistics South Africa for providing the data.

4.9 Ethical approval

Not required.

4.10 Reference

1. Coetzee, M., P. Kruger, R. Hunt, D. Durrheim, J. Urbach, and C. Hansford, *Malaria in South Africa: 110 years of learning to control the disease.* South African Medical Journal, 2013. **103**(10): p. 770-778.

2. Maharaj, R., J. Raman, N. Morris, D. Moonasar, D.N. Durrheim, I. Seocharan, P. Kruger, B. Shandukani, and I. Kleinschmidt, *Epidemiology of malaria in South Africa: From control to elimination.* South African Medical Journal, 2013. **103**(10): p. 779-783.

3. Partnership Roll Back Malaria, *Progress and impact series: focus on South Africa*. 2013: Geneva: Country Reports.

4. Moonasar, D., T. Nuthulaganti, P.S. Kruger, A. Mabuza, E.S. Rasiswi, F.G. Benson, and R. Maharaj, *Malaria control in South Africa 2000–2010: beyond MDG6.* Malaria Journal, 2012. **11**(294): p. 1475-2875.

5. South Africa National Department of Health, *Republic Of South Africa Malaria Elimination Strategy 2011–2018*. 2012: Pretoria, South Africa.

6. Barnes, K.I., D.N. Durrheim, F. Little, A. Jackson, U. Mehta, E. Allen, S.S. Dlamini, J. Tsoka, B. Bredenkamp, and D.J. Mthembu, *Effect of artemether-lumefantrine policy*

and improved vector control on malaria burden in KwaZulu–Natal, South Africa. PLOS Med, 2005. **2**(11): p. e330.

7. Ebhuoma, O., M. Gebreslasie, and L. Magubane, *Modeling malaria control intervention effect in KwaZulu-Natal, South Africa using intervention time series analysis.* Journal of Infection and Public Health.

8. Maharaj, R., D. Mthembu, and B. Sharp, *Impact of DDT re-introduction on malaria transmission in KwaZulu-Natal.* South African Medical Journal, 2005. **95**(11): p. 871.

9. Maharaj, R., N. Morris, I. Seocharan, P. Kruger, D. Moonasar, A. Mabuza, E. Raswiswi, and J. Raman, *The feasibility of malaria elimination in South Africa.* Malaria Journal, 2012. **11**: p. 423-423.

10. Ebhuoma, O. and M. Gebreslasie, *Remote Sensing-Driven Climatic/Environmental Variables for Modelling Malaria Transmission in Sub-Saharan Africa.* International Journal of Environmental Research and Public Health, 2016. **13**(6): p. 584.

11. Tatem, A.J., P.W. Gething, D.L. Smith, and S.I. Hay, *Urbanisation and the global malaria recession.* Malaria Journal, 2013. **12**(1): p. 133.

12. Matthys, B., E.K. N'Goran, M. Kone, B.G. Koudou, P. Vounatsou, G. Cisse, A.B. Tschannen, M. Tanner, and J. Utzinger, *Urban agricultural land use and characterisation of mosquito larval habitats in a medium-sized town of Cote d'Ivoire.* Journal of Vector Ecology, 2006. **31**(2): p. 319-33.

13. Snow, R., M. Craig, C. Newton, and R. Steketee, *The Public Health Burden of Plasmodium Falciparum Malaria in Africa: Deriving the Numbers (The Disease Control Priorities Project (DCPP) Working Paper Number 11, Washington DC., 2003).* 2008, Working Paper.

14. Tusting, L.S., B. Willey, H. Lucas, J. Thompson, H.T. Kafy, R. Smith, and S.W. Lindsay, *Socioeconomic development as an intervention against malaria: a systematic review and meta-analysis.* Lancet, 2013. **382**(9896): p. 963-972.

15. Kelly-Hope, L.A. and F.E. McKenzie, *The multiplicity of malaria transmission: a review of entomological inoculation rate measurements and methods across sub-Saharan Africa.* Malaria Journal, 2009. **8**: p. 19.

16. Singh, R., S. Haq, and R. Dhiman, *Studies on knowledge, attitude and practices in malaria endemic tribal areas of Bihar and Jharkhand, India.* Journal of Tropical Diseases, 2013.

17. Singh, T.G., R.K. Singh, and E.Y. Singh, *A study of knowledge about malaria and treatment seeking behaviour in two tribal communities of Manipur.* Indian Journal of Public Health, 2003. **47**(2): p. 61-5.

18. Kaona, F., M.T. Siajunza, C. Manyando, S. Khondowe, and G.K. Ngoma, *Utilisation of malarial drugs at a household level: results from a KAP study in Choma, southern province and Mporokoso, northern province of Zambia.* Central African Journal of Medicine, 2000. **46**(10): p. 268-70.

19. Dunn, C.E., A. Le Mare, and C. Makungu, *Malaria risk behaviours, socio-cultural practices and rural livelihoods in southern Tanzania: implications for bednet usage.* Social Science & Medicine, 2011. **72(**3): p. 408-17.

20. Appiah-Darkwah, I. and S.K. Badu-Nyarko, *Knowledge of malaria prevention and control in a sub-urban community in Accra, Ghana.* International Journal of Tropical Medicine, 2011. **6**(3): p. 61-9.

21. Ayele, D.G., T.T. Zewotir, and H.G. Mwambi, *Prevalence and risk factors of malaria in Ethiopia.* Malaria Journal, 2012. **11**.

22. Monteiro, T.H.A., T.d.S.S. Chaves, H.J.d. Matos, N.F.d.L. Sofffiatti, R.J.d.P.S.e. Guimarães, L.H.R. Guimarães, A.M.R. Ventura, and R.L.D. Machado, *Basic sanitation, socioeconomic conditions, and degree of risk for the presence and maintenance of malaria in a low-transmission area in the Brazilian Amazon.* Revista da Sociedade Brasileira de Medicina Tropical, 2015. **48**: p. 573-579.

23. Hewitt, S.E., M. Farhan, H. Urhaman, N. Muhammad, M. Kamal, and M.W. Rowland, *Self-protection from malaria vectors in Pakistan: an evaluation of popular existing methods and appropriate new techniques in Afghan refugee communities.* Annals of Tropical Medicine & Parasitology, 1996. **90**(3): p. 337-44.

24. Rosas-Aguirre, A., O.J. Ponce, G. Carrasco-Escobar, N. Speybroeck, J. Contreras-Mancilla, D. Gamboa, E. Pozo, S. Herrera, and A. Llanos-Cuentas, *Plasmodium vivax malaria at households: spatial clustering and risk factors in a low endemicity urban area of the northwestern Peruvian coast.* Malaria Journal, 2015. **14**: p. 176.

25. Singh, N., A.K. Mishra, M.M. Shukla, and S.K. Chand, *Forest malaria in Chhindwara, Madhya Pradesh, central India: a case study in a tribal community.* American Journal of Tropical Medicine and Hygiene, 2003. **68**(5): p. 602-7.

26. Dysoley, L., A. Kaneko, H. Eto, T. Mita, D. Socheat, A. Borkman, and T. Kobayakawa, *Changing patterns of forest malaria among the mobile adult male population in Chumkiri District, Cambodia.* Acta Tropica, 2008. **106**(3): p. 207-12.

27. Sharma, R.K., M.P. Singh, K.B. Saha, P.K. Bharti, V. Jain, P. Singh, N. Silawat, R. Patel, M. Hussain, and S. Chand, *Socio-economic & household risk factors of malaria in tribal areas of Madhya Pradesh, central India.* Indian Journal of Medical Research, 2015. **141**(5): p. 567.

28. Obaldia, N., *Determinants of low socio-economic status and risk of Plasmodium vivax malaria infection in Panama (2009–2012): a case–control study.* Malaria Journal, 2015. **14**(1): p. 14.

29. Mosha, J.F., H.J. Sturrock, J.M. Brown, R. Hashim, G. Kibiki, D. Chandramohan, and R.D. Gosling, *The independent effect of living in malaria hotspots on future malaria infection: an observational study from Misungwi, Tanzania.* Malaria Journal, 2014. **13**: p. 445.

30. Yadav, K., S. Dhiman, B. Rabha, P. Saikia, and V. Veer, *Socio-economic determinants for malaria transmission risk in an endemic primary health centre in Assam, India.* Infectious Diseases of Poverty, 2014. **3**(1): p. 19.

31. Fayehun, O.A. and K.K. Salami, *Older persons and malaria treatment in Nigeria.* African Population Studies, 2014. **27**(2): p. 10.

32. Schwartz, E., S. Sadetzki, H. Murad, and D. Raveh, *Age as a risk factor for severe Plasmodium falciparum malaria in nonimmune patients.* Clinical Infectious Diseases, 2001. **33**(10): p. 1774-1777.

33. Gosoniu, L., A. Msengwa, C. Lengeler, and P. Vounatsou, *Spatially explicit burden estimates of malaria in Tanzania: bayesian geostatistical modeling of the malaria indicator survey data.* PLOS One, 2012. **7**(5): p. e23966.

34. Abe, T., S. Honda, S. Nakazawa, T.D. Tuong, N.Q. Thieu, X. Hung le, K. Thuan le, K. Moji, M. Takagi, and T. Yamamoto, *Risk factors for malaria infection among ethnic minorities in Binh Phuoc, Vietnam.* Southeast Asian Journal of Tropical Medicine and Public Health, 2009. **40**(1): p. 18-29.

35. Checkley, A.M., A. Smith, V. Smith, M. Blaze, D. Bradley, P.L. Chiodini, and C.J.M. Whitty, *Risk factors for mortality from imported falciparum malaria in the United Kingdom over 20 years: an observational study.* BMJ, 2012. **344**.

36. Arab, A., *Spatial and spatio-temporal models for modeling epidemiological data with excess zeros.* International Journal of Environmental Research and Public Health, 2015. **12**(9): p. 10536-10548.

37. Giardina, F., L. Gosoniu, L. Konate, M.B. Diouf, R. Perry, O. Gaye, O. Faye, and P. Vounatsou, *Estimating the burden of malaria in Senegal: Bayesian zero-inflated binomial geostatistical modeling of the MIS 2008 data.* PLOS One, 2012. **7**(3): p. e32625.

38. Cancre, N., A. Tall, C. Rogier, J. Faye, O. Sarr, J.F. Trape, A. Spiegel, and F. Bois, *Bayesian analysis of an epidemiologic model of Plasmodium falciparum malaria infection in Ndiop, Senegal.* American Journal of Epidemiology, 2000. **152**(8): p. 760-70.

39. Musal, M. and T. Aktekin, *Bayesian spatial modeling of HIV mortality via zero-inflated Poisson models.* Statistics in Medicine, 2013. **32**(2): p. 267-281.

40. Ntzoufras, I., *Bayesian Modeling Using WinBUGS*. Vol. 698. 2011: John Wiley & Sons.

41. Ghosh, S.K., P. Mukhopadhyay, and J.-C.J. Lu, *Bayesian analysis of zero-inflated regression models.* Journal of Statistical Planning and Inference, 2006. **136**(4): p. 1360-1375.

42. Camp, K.G.T., *A Bio-Resource Classification For KwaZulu-Natal, South Africa*, in *School of Applied Environmental Sciences*. 1999, University of KwaZulu-Natal, South Africa: Pietermaritzburg, South Africa.

43. Statistics South Africa, *Census 2011 Municipal Report – KwaZulu-Natal*. 2012: Pretoria, South Africa.

44. Lykou, A. and I. Ntzoufras, *WinBUGS: a tutorial.* Wiley Interdisciplinary Reviews: Computational Statistics, 2011. **3**(5): p. 385-396.

45. Amoran, O.E., *Impact of health education intervention on malaria prevention practices among nursing mothers in rural communities in Nigeria.* Nigerian Medical Journal, 2013. **54**(2): p. 115-122.

46. Ayele, D.G., T.T. Zewotir, and H.G. Mwambi, *The risk factor indicators of malaria in Ethiopia.* International Journal of Medicine and Medical Sciences, 2013. **5**(7): p. 335-347.

47. Somi, M.F., J.R. Butler, F. Vahid, J.D. Njau, S.P. Kachur, and S. Abdulla, *Use of proxy measures in estimating socioeconomic inequalities in malaria prevalence.* Tropical Medicine & International Health, 2008. **13**(3): p. 354-364.

48. Kroeger, A., R. Meyer, M. Mancheno, and M. Gonzalez, *Health education for community-based malaria control: an intervention study in Ecuador, Colombia and Nicaragua.* Tropical Medicine & International Health, 1996. **1**(6): p. 836-46.

49. Coovadia, H., R. Jewkes, P. Barron, D. Sanders, and D. McIntyre, *The health and health system of South Africa: historical roots of current public health challenges.* Lancet, 2009. **374**(9692): p. 817-34.

50. Brooke, B., L. Koekemoer, P. Kruger, J. Urbach, E. Misiani, and M. Coetzee, *Malaria vector control in South Africa*. 2013. Vol. 103. 2013.

51. Gjorup, I.E. and A. Ronn, *Malaria in elderly nonimmune travelers.* Journal of Travel Medicine, 2002. **9**(2): p. 91-3.

52. Sachs, J. and P. Malaney, *The economic and social burden of malaria.* Nature, 2002. **415**(6872): p. 680-685.

53. Keiser, J., B.H. Singer, and J. Utzinger, *Reducing the burden of malaria in different eco-epidemiological settings with environmental management: a systematic review.* Lancet Infectious Diseases, 2005. **5**(11): p. 695-708.

54. Kirby, M.J., D. Ameh, C. Bottomley, C. Green, M. Jawara, P.J. Milligan, P.C. Snell, D.J. Conway, and S.W. Lindsay, *Effect of two different house screening interventions on exposure to malaria vectors and on anaemia in children in The Gambia: a randomised controlled trial.* Lancet, 2009. **374**(9694): p. 998-1009.

CHAPTER 5: MODELLING MALARIA CONTROL INTERVENTION EFFECT IN KWAZULU-NATAL, SOUTH AFRICA USING INTERVENTION TIME SERIES ANALYSIS

This chapter is based on:

Ebhuoma O, Gebreslasie M, Magubane L (2017). Modelling malaria control intervention effect in KwaZulu-Natal, South Africa using intervention time series analysis. Journal of Infection and Public Health 10(3): 334-338.

5.1 Abstract

The change of the malaria control intervention policy in South Africa (SA), re-introduction of dichlorodiphenyltrichloroethane (DDT), may be responsible for the low and sustained malaria transmission in KwaZulu-Natal (KZN). The investigator evaluated the effect of the re-introduction of DDT on malaria in KZN and suggested practical ways the province can strengthen their already existing malaria control and elimination efforts, to achieve zero malaria transmission. The investigator obtained confirmed monthly malaria cases in KZN from the malaria control program of KZN from 1998 to 2014. The seasonal autoregressive integrated moving average (SARIMA) intervention time series analysis (ITSA) was employed to model the effect of the re-introduction of DDT on confirmed monthly malaria cases. The result is an abrupt and permanent decline of monthly malaria cases (w_0= -1174.781, *p*-value = 0.003) following the implementation of the intervention policy. The sustained low malaria cases observed over a long period suggests that the continued usage of DDT did not result in insecticide resistance as earlier anticipated. It may be due to exophagic malaria vectors, which renders the indoor residual spraying not totally effective. Therefore, the feasibility of reducing malaria transmission to zero in KZN requires other reliable and complementary intervention resources to optimise the existing ones.

Keywords: *Intervention Time Series Analysis (ITSA), Malaria, Dichlorodiphenyltrichloroethane (DDT), Seasonal Autoregressive Integrated Moving Average (SARIMA).*

5.2 Introduction

South Africa (SA)'s malaria vector control (i.e., intervention) depends mainly on indoor residual spraying with dichlorodiphenyltrichloroethane (DDT). As an implication, SA in general and KwaZulu-Natal (KZN) in particular have made significant progress over the past two decades in reducing malaria disease caused by *Plasmodium falciparum* [1, 2]. This decline can be associated with the change of SA's malaria vector control policy during the peak of the 1999/2000 malaria epidemic. In March 2000, DDT was re-introduced for malaria vector control purposes after it was discontinued in 1995. The re-introduction and continued use of DDT have been possible because the national government, with the help from international scientists and an independent advocacy group successfully obtained an exemption in the Stockholm Convention on Persistent Organic Pollutants in 2000 [1]. After DDT was introduced in March 2000, an abrupt decline in malaria cases was observed in the time series data [3]. While the level of malaria control achieved in KZN is encouraging, local transmission has not reached zero [1, 2], thus, it remains a cause for concern for SA as she targets malaria elimination in 2020 [2].

The impact of the re-introduction of DDT (the known intervention) in KZN is assumed to be associated with the notable alteration of the malaria time series or change of the mean function [3] and can be evaluated employing an intervention time series analysis (ITSA) [4, 5]. The ITSA is a thorough and reliable analytical method that allows the effect of an intervention to be separated from the general trends and serial dependencies in time, thereby allowing sound statistical inference to be made if the intervention had an effect on the time series[4, 5]. This method gives analysts the opportunity to draw inferences from the impact assessment and confirms the substantive notion of a particular area or region. In otherwords, the ITSA model can be the best form of impact assessment from a statistical point of view as long as the inference(s) and conclusion(s) drawn from the analysis is/are reconciled with the prevailing theory in a substantive sense [4]. Substantial ITSA studies applied the univariate approach due to its simplistic application in modeling and data availability [6-10]. Similarly, the impact of the re-introduction of DDT in KZN can be reliably evaluated employing the same approach.

This paper thus attempts to determine if the long-term use of DDT significantly lead to the decline and ultimately will lead to zero malaria transmission in KZN. The outcome of this

chapter will serve as a validation of the substantive significance of DDT on malaria in KZN. This is vital to the province's malaria control and elimination efforts, because it will bring to light the necessity of identifying other practical ways that can be used to upscale the existing malaria vector control strategies to achieve zero malaria transmission. Thus, this chapter also seeks to suggest other reliable and complementary intervention resources that can support and optimise the already existing interventions. It is for these reasons that the historical series of malarial cases in KZN will be utilised as the dependent series, in the univariate seasonal autoregressive integrated moving average (SARIMA) model with known intervention, and its characteristics will allow the behaviour of malaria over time, to be evaluated.

5.3 Methodology

5.3.1 Study area

The three district municipalities (uMkhanyakude, uThungulu, and Zululand) in KZN province endemic to malaria were included in this chapter (Figure 5.1). uMkhanyakude is situated in the northern region of KZN province with a population of 625,846 [11]. uThungulu and Zululand district municipalities are located in the north-eastern part of KZN province with a population of 907,519 and 803,575 respectively [11]. The study areas are bordered by Swaziland and Mozambique to the north and the Indian Ocean stretching from the east down to the southeast. The province is characterised by the sub-tropical climate with most of the malaria cases occurring during the rainy months (October to May), with a seasonal peak usually in January and March [2, 12]. The average annual rainfall ranges from 500mm to 2000mm. Along the coastal areas, the summer temperatures is between 24°C to 32°C, and mean winter temperature is about 20°C. The Midlands generally possesses a mild climate with relatively high summer rainfall and dry winters. The elevation measure of the region varies from sea level to over 3000m. The vegetation of the study area comprises of coastal forest and thornveld along the coast. Towards the inlands, lowveld, highland sourveld, Natal sour sandveld, valley bushveld and tall grassveld vegetation are found. Lowveld and thornveld characterises the low-lying hot and dry regions of Northern KZN [13].

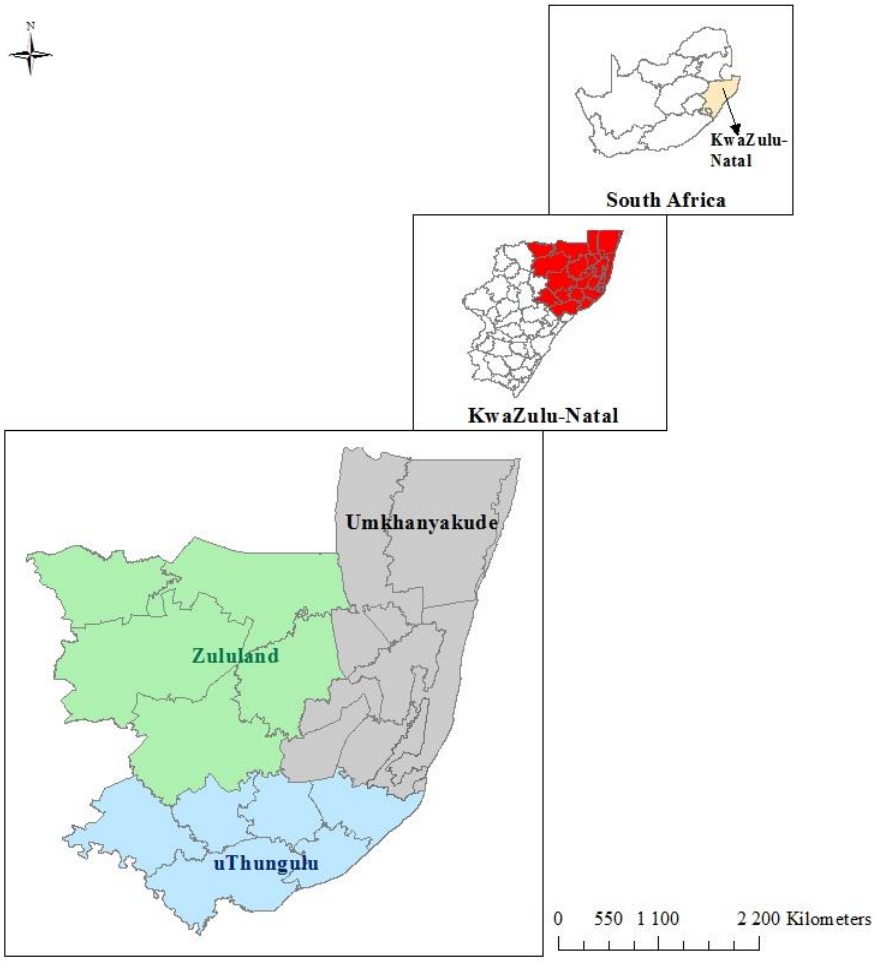


Figure 5.1 Map of the study area

5.3.2 Data source

Monthly aggregates of clinically confirmed malaria cases from January 1998 to December 2014 were collected from the malaria control program of KZN, SA. A malaria case is a person whose blood smear tested positive to *Plasmodium* after undergoing a rapid diagnostic test or slide microscopy at a health facility [14]. Since 1956, it became a legal requirement to notify malaria cases to the relevant health authorities in SA [15]. Confirmed malaria cases at health facilities are reported to the relevant district health office which is subsequently reported to the provincial malaria control program. At the provincial malaria control program, the malaria control worker collects and inputs information relating to the malaria case into the malaria information system. The information includes patient's personal details, the health facility the case was reported, symptoms, malaria test results, diagnosis and type of treatment administered [2, 16].

5.3.3 Analytical method

The Box and Tiao approach of the ITSA [17] was employed to examine the effect of the malaria intervention. It involved a two-step analytical process. The first step is the identification of the most suitable SARIMA model, referred to as the noise component of the model, using the dataset not impacted by the intervention or before the intervention. This part of the ITSA is bounded by Box and Jenkin's SARIMA model approach [18], and it involves the following steps: model identification, parameter estimation, and diagnostic checking. Exhaustive presentations of these procedures are found elsewhere [4, 18, 19]. The second step involves re-estimating the identified model using the full dataset to test the effects of the intervention on the behaviour of the time series, and it is known as the intervention component. Hence, by comparing the level of pre-intervention and post-intervention time series, the statistical significance of the intervention was evaluated.

The ITSA model is written as [17]:

$$Y_t = f(I_t) + N_t \qquad (5.1)$$

Where Y_t denotes the dependent variable for a certain time or is an observed time-series, the function $f(I_t)$ denotes a "function of the variable I_t", the intervention component (also referred to a transfer function), N_t denotes the noise component determined by an univariate SARIMA *(p,d,q)(P,D,Q)* structure and *t* denotes the discrete time.

The impact of an intervention on the time series can be either abrupt or gradual in onset and either permanent or temporary in duration. Therefore, the shape of the time series after the onset of an intervention determines which transfer function (i.e., zero order transfer function, first order transfer function or pulse function) will be used to model the impact [4]. The time series data for this study shows that the response of the malaria cases after the onset of the known intervention (i.e. re-introduction of DDT) had an abrupt and permanent shift in the process. The investigator, therefore, used a "zero order transfer function" to determine the effect of the re-introduction of DDT in KZN.

The zero order transfer function is written as [4]:

$$f(I_t) = w_0 I_t \qquad (5.2)$$

Where w_0 denotes the parameter estimate of a transfer function. The variable I_t is defined as a step variable or step variable such that,

I_t= 0 before the intervention

and I_t= 1 at and after the intervention.

Therefore the impact assessment model is:

$$Y_t = w_0 I_t + N_t \qquad (5.3)$$

From the time series, $I_t^{03/2000} = \begin{cases} 0 \ \ if \ t \ < 03/2000 \ (before \ the \ intervention) \\ 1 \ \ if \ t \ \geq 03/2000 \ (at \ and \ after \ the \ intervention) \end{cases}$

The modeling and analysis were performed using the R statistical software version 3.2.3. The map showing the study area was developed using the ArcGIS 10.2 software (ESRI, Redlands, CA, USA).

5.4 Results

5.4.1 Exploratory data analysis

The time series of confirmed monthly malaria cases from January 1998 to December 2014 in KZN is shown in Figure 5.2. Before the re-introduction of DDT, the time series shows markedly increase in malaria cases. The re-introduction of DDT into the SA malaria control program in March 2000, coincided with the beginning of the abrupt decline of malaria cases in KZN, which continued until June 2001. Afterward, relatively low and sustained cases were recorded with a noticeable spike between December 2003 to November 2004. The quick decline in malaria cases after November 2004 and sustained low malaria cases onwards can be attributed to the continous application of the intervention. Overall, this is a good example of a time series that the impact of the known intervention is abrupt in onset and permanent in duration. Thus, an empirical assessment of the impact of the known intervention which combines a noise and intervention models is conducted.

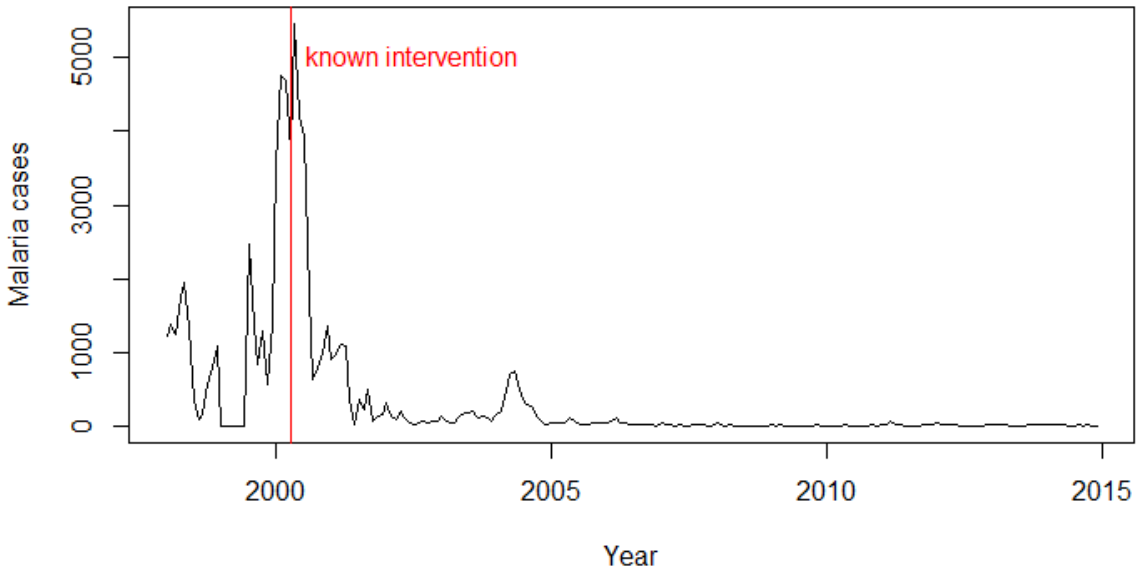


Figure 5.2 Time series plot of confirmed monthly malaria cases in KZN from 1998-2014

5.4.2 SARIMA or noise model

After the SARIMA model was fitted to the time series data that was not impacted by the intervention (January 1998 to February 2000), the autocorrelation functions (ACF) and partial autocorrelation functions (PACF) of the non-seasonal and seasonal differenced series lead to the identification of two plausible models. They are SARIMA $(0,1,0)\ (4,1,1)_{12}$ and SARIMA $(0,1,0)\ (0,1,1)_{12}$. Details of how the plausible models were identified are found elsewhere [4, 18, 19]. Based on the goodness of fit statistics (Table 5.1) and parameter estimation (Table 5.2), the investigator selected the SARIMA $(0,1,0)\ (0,1,1)_{12}$ model as the most suitable model, and it was subsequently used to develop the intervention model. The SARIMA $(0,1,0)\ (0,1,1)_{12}$ model was selected as the best model fit because it possesses a lower Bayesian information criterion (BIC) value (which is based on the likelihood function and the Akaike information criterion (AIC), and a *p*-value less than 0.05 . Also, all its parameter estimates are significant.

Table 5.1 Goodness of fit statistics of the plausible SARIMA models

Statistic	SARIMA $(0,1,0)(4,1,1)_{12}$	SARIMA $(0,1,0)(0,1,1)_{12}$
AIC	2860.88	2864.85
BIC	2880.39	2871.36
LL	-1424.44	-1430.43

AIC=Akaike information criterion; BIC=Bayesian information criterion; LL=Log likelihood

Table 5.2 Parameter estimation of the plausible SARIMA models

	SARIMA $(0,1,0)(4,1,1)_{12}$			SARIMA $(0,1,0)(0,1,1)_{12}$		
Type	**Coef.**	**S.E. of Coef.**	***p*-values**	**Coef.**	**S.E. of Coef.**	***p*-values**
SAR1	-0.5272	0.5238	0.16033	-	-	-
SAR2	-0.2297	0.5516	0.33994	-	-	-
SAR3	-0.1002	0.4133	0.40508	-	-	-
SAR4	0.0351	0.2427	0.44291	-	-	-
SMA1	-0.5398	0.5159	0.15127	-0.8999	0.0693	< 0.001

SAR= Seasonal autoregressive; SMA = Seasonal moving average; Coef.= coefficient; S.E. = Standard error.

5.4.3 Intervention model

Table 5.3 presents the results of zero-order transfer model assessment, while Table 5.4 presents the diagnostic test for the full intervention model. The parameter estimate for SMA1 and the intervention event (w_0) presented in Table 5.3 are statistically significant. This means the intervention model for malaria cases showed that the re-introduction of DDT resulted in an abrupt and permanent decline in monthly cases in KZN (w_0= -1174.781, p-value = 0.003).

The final diagnostics indicates the estimated model is statistically adequate as given in Table 5.4. The Ljung-Box test of randomness of the residuals from the intervention model revealed that there is no autocorrelation at the 0.05 level (X^2 = 58.386; p-value = 0.145). This implies that the full intervention model for the malaria cases in KZN is adequate.

Table 5.3 Impact of DDT on malaria cases in KZN province/Parameter estimates for the tentative intervention model

Parameter	Coef.	S.E. of Coef.	*p*-value
SMA1	-0.729	0.068	< 0.001
T1-MA0 (w_0)	-1174.781	401.897	0.003

SMA = Seasonal moving average; w_0= parameter estimate of a transfer function; Coef.= coefficient; S.E. = Standard error

Table 5.4 Ljung-Box test for the intervention model

X^2	*df*	*p*-value
58.386	48	0.145

5.5 Discussion

The reliability of the analysis lies in the fact that the analysis was conducted in line with the recommendations specified by Gilmour *et al.* [8] to ensure issues associated with ITSA were avoided. The investigator choose the right indicator data (i.e., malaria case data) linked to the intervention to avoid biased outcomes and erroneous inferences. The investigator also correctly specified the onset of the known intervention (re-introduction of DDT) through reliable reports [1, 3]. Nevertheless, in instances where the onset of the known intervention(s) is/are not clear, several onset dates from the time series may be modelled by obtaining the onset dates from related studies, experts and directly from individuals who witnessed the event. Furthermore, the nature of the dataset clearly portrayed the long term change that occurred from the onset of the intervention. As such, the investigator choose the right transfer function for the model, the zero-order transfer function because it fits an intervention that is abrupt in onset and permanent in duration. Other transfer functions are first-order transfer function and pulse function, and they are explained in details elsewhere [4, 6]. It is worthy of note that the flexibility of determining and using transfer functions peculiar to a time series makes the ITSA preferable over linear models [4].

The ITSA model developed presents a reliable confirmation of the importance the malaria intervention DDT has on the malaria cases in KZN, by estimating the impact of its robust use over a long period. The findings strongly suggest that the re-introduction of DDT in March 2000

led to an immediate and permanent drop in malaria in KZN. Afterward, low and sustained cases were recorded, but malaria transmission did not reach zero. In addition to the outcome of this chapter, there are strong reasons to agree that the substantive significance of DDT in combating malaria exists. Firstly, marked increase in malaria cases were noted when the use of DDT was abandoned in the late 1970s in Madagascar [20], and in 1995 in South Africa [1]. Secondly, cross-sectional studies reveal an association between indoor DDT and low malaria cases in the highlands [21] and Western Foothill area [22] of Madagascar, and in KZN, South Africa [3]. The outcome of this chapter is very similar in scope to the previous study by Maharaj *et al.* [3], who revealed that the re-introduction of DDT led to significant decline in malaria cases. However, their study used a shorter post-intervention data to assess the impact of DDT (from 2000 to 2002) by employing inferential statistics. Other studies [21, 22] also suggested an association between the use of indoor DDT and low malaria cases. It suggests that the prolonged and continuous usage of DDT did not result in malaria vector resistance as anticipated earlier by Hargreaves *et al.* [23]. But the lingering low malaria transmission can be attributed to the exophagic and resting behavior of the malaria vectors, which makes them not susceptible to be the insecticide-treated surfaces indoors. Thus, the feasibility of eliminating malaria in KZN would require up-scaled and re-energised technical and operational resources, to reduce local transmission to zero and sustain the elimination in the face of repeated reintroduction from the seasonal transmission and imported cases.

The re-introduction of DDT bolstered the malaria control efforts and campaigns in KZN markedly. Other reliable and complementary interventions resources will be needed to further strengthen and optimise the already existing malaria vector control resources in KZN in particular and SA at large to achieve malaria elimination in 2020. Such as (1) the introduction of pre-erythrocytic malaria vaccines, RTS,S/AS01 for residence living in malaria endemic areas and visitors, (2) the introduction of genetically modified sterile male *Anopheles arabiensis* mosquitoes and (3) the practical application of reactive case detection to determine the optimal screening radius peculiar to each of the malaria transmission settings or the focal screening and treatment depending on the level of malaria endemicity. Furthermore, renewed attention and campaigns on malaria disease are essential. Malaria disease has suffered limited attention in the last 5 years due to more focus on other diseases like HIV/AIDS, tuberculosis and diabetes. This is reflected in the limited number of malaria-related studies in the region geared towards malaria

elimination. For example, studies on the imported cases of malaria, studies aimed at identifying reliable climatic, vegetative and socio-economic determinants of malaria in KZN, studies on the hot spot of malaria transmission are lacking. Such studies can serve as invaluable tools for monitoring the progress made in the province towards the malaria elimination efforts, and can help identify areas that need reinforcement of intervention resources. Furthermore, such studies can also serve as a guide for the preparation against malaria re-introduction once zero transmission is achieved.

5.6 Conclusions

The ITSA is a well-suited approach that provides a rigorous test of the association between the effect of an intervention policy and a public health issue at a population level. The intervention time series analysis modelled the effect of known intervention (re-introduction of DDT in March 2000) on data of counts of events (malaria cases) regularly collected in time (from 1998 to 2014). It resulted in an abrupt and permanent decline of confirmed malaria cases in KZN, which confirms that DDT played a significant role in the low and sustained malaria case in the province. The sustained low malaria cases recorded over the last two decades suggests that the continued usage of DDT did not result in insecticide resistance as earlier anticipated. However, for the province to achieve zero malaria transmission, renewed attention should be given to malaria in KZN in terms of more financial commitment and multidisciplinary research efforts. This should comprise of location specific studies and the introduction of other reliable and complementary intervention resources to support and optimise the already existing malaria intervention.

5.7 Acknowledgements

The authors would like to thank the College of Agriculture, Engineering and Science of the University of KwaZulu-Natal for the doctoral research bursary awarded to Osadolor Ebhuoma. We also thank the malaria control program of KwaZulu-Natal, South Africa for providing the data.

5.8 Competing interests

None declared.

5.9 Ethical approval

Not required.

5.10 Reference

1. Coetzee, M., P. Kruger, R. Hunt, D. Durrheim, J. Urbach, and C. Hansford, *Malaria in South Africa: 110 years of learning to control the disease.* South African Medical Journal, 2013. **103**(10): p. 770-778.

2. South Africa National Department of Health, *Republic of South Africa malaria elimination strategy 2011–2018*. 2011: Pretoria, South Africa.

3. Maharaj, R., D. Mthembu, and B. Sharp, *Impact of DDT re-introduction on malaria transmission in KwaZulu-Natal.* South African Medical Journal, 2005. **95**(11): p. 871.

4. McCleary, R., R.A. Hay, E.E. Meidinger, and D. McDowall, *Applied Time Series Analysis For The Social Sciences*. 1980: Sage Publications Beverly Hills, CA.

5. Harvey, A.C., *Forecasting, structural time series models and the Kalman filter*. 1992: Cambridge university press.

6. Pridemore, W.A. and A.J. Snowden, *Reduction in suicide mortality following a new national alcohol policy in Slovenia: An interrupted time-series analysis.* American Journal of Public Health, 2009. **99**(5): p. 915-920.

7. Darkwah, K., G. Okyere, and A. Boakye, *Intervention analysis of serious crimes in the eastern region of Ghana.* International Journal of Business and Social Research, 2012. **2**(7): p. 132-138.

8. Gilmour, S., L. Degenhardt, W. Hall, and C. Day, *Using intervention time series analyses to assess the effects of imperfectly identifiable natural events: a general method and example.* BMC Medical Research Methodology, 2006. **6**(1): p. 16.

9. Girard, D.Z., *Intervention times series analysis of pertussis vaccination in England and Wales.* Health Policy, 2000. **54**(1): p. 13-25.

10. Pridemore, W.A., M.B. Chamlin, and E. Andreev, *Reduction in male suicide mortality following the 2006 Russian alcohol policy: an interrupted time series analysis.* American Journal of Public Health, 2013. **103**(11): p. 2021-2026.

11. Statistics South Africa, *Census 2011 Municipal Report – KwaZulu-Natal*. 2012: Pretoria, South Africa.

12. Moonasar, D., T. Nuthulaganti, P.S. Kruger, A. Mabuza, E.S. Rasiswi, F.G. Benson, and R. Maharaj, *Malaria control in South Africa 2000–2010: beyond MDG6.* Malaria Journal, 2012. **11**(294): p. 1475-2875.

13. Camp, K.G.T., *A Bio-Resource Classification For KwaZulu-Natal, South Africa*, in *School of Applied Environmental Sciences*. 1999, University of KwaZulu-Natal, South Africa: Pietermaritzburg, South Africa.

14. South African National Department of Health, *Guidelines for the Treatment of Malaria in South Africa*. 2010: Pretoria, South Africa.

15. South Africa National Department of Health, *Notification Of Diseases*. 1956: Pretoria, South Africa

16. Khosa, E., L.R. Kuonza, P. Kruger, and E. Maimela, *Towards the elimination of malaria in South Africa: a review of surveillance data in Mutale Municipality, Limpopo Province, 2005 to 2010.* Malaria Journal, 2013. **12**(7).

17. Box, G.E.P. and G.C. Tiao, *Intervention Analysis with Applications to Economic and Environmental Problems.* Journal of the American Statistical Association, 1975. **70**(349): p. 70-79.

18. Box, G.E., G.M. Jenkins, G.C. Reinsel, and G.M. Ljung, *Time Series Analysis: Forecasting And Control*. 2015: John Wiley & Sons.

19. Harvey, A.C., *Forecasting, Structural Time Series Models, And The Kalman Filter*. 1990, Cambridge; New York: Cambridge University Press.

20. Mouchet, J., S. Laventure, S. Blanchy, R. Fioramonti, A. Rakotonjanabelo, P. Rabarison, J. Sircoulon, and J. Roux, *The reconquest of the highlands of Madagascar by malaria.* Bulletin De La Societe De Pathologie Exotique, 1997. **90**(3): p. 162-168.

21. Romi, R., M. Razaiarimanga, R. Raharimanga, E. Rakotondraibe, L. Ranaivo, V. Pietra, A. Raveloson, and G. Majori, *Impact of the malaria control campaign (1993-1998) in the highlands of Madagascar: parasitological and entomological data.* The American Journal of Tropical Medicine and Hygiene, 2002. **66**(1): p. 2-6.

22. Ratovonjato, J., M. Randrianarivelojosia, M.E. Rakotondrainibe, V. Raharimanga, L. Andrianaivolambo, G. Le Goff, C. Rogier, F. Ariey, S. Boyer, and V. Robert, *Entomological and parasitological impacts of indoor residual spraying with DDT, alphacypermethrin and deltamethrin in the western foothill area of Madagascar.* Malaria Journal, 2014. **13**(1): p. 1.

23. Hargreaves, K., R. Hunt, B. Brooke, J. Mthembu, M. Weeto, T. Awolola, and M. Coetzee, *Anopheles arabiensis and An. quadriannulatus resistance to DDT in South Africa.* Medical and Veterinary Entomology, 2003. **17**(4): p. 417-422.

CHAPTER 6: A SEASONAL AUTOREGRESSIVE INTEGRATED MOVING AVERAGE (SARIMA) FORECASTING MODEL TO PREDICT MONTHLY MALARIA CASES IN KWAZULU-NATAL, SOUTH AFRICA

This chapter is based on:

Ebhuoma O, Gebreslasie M, Magubane L (In press). A seasonal autoregressive integrated moving average (SARIMA) forecasting model to predict monthly malaria cases in KwaZulu-Natal, South Africa. South African Medical Journal.

6.1 Abstract

South Africa in general and KwaZulu-Natal (KZN) province in particular, stepped up its effort to eliminate malaria such that it has been reporting consistently low cases since 2000. To strengthen the malaria control that leads to elimination efforts in KZN, this chapter is aimed at developing a forecasting model of malaria cases in KZN by using Seasonal Autoregressive Integrate Moving Average (SARIMA) time series model. This was carried out retrospectively using the monthly reported malaria case data from January 2005 to December 2014. The dataset was acquired from the Malaria control programme of KZN and it was split into two. The first dataset from January 2005 to December 2013 was used to construct a SARIMA model by adopting the approach propounded by Box-Jenkins as follows: model identification, parameter estimation, diagnostic checking, and forecasting after performing the Yeo-Johnson transformation to manage the zeros and close to zero values. The second dataset from January to December 2014 was used to validate the forecast generated from the best fit model. Three plausible models were identified, and based on the goodness of fit statistics and parameter estimation, the SARIMA $(0,1,1)(0,1,1)_{12}$ model was selected as the best fit model. The SARIMA $(0,1,1)(0,1,1)_{12}$ model was used to forecast malaria cases during 2014, and it was observed to fit closely with the reported malaria cases during January to December 2014. The outcome of this chapter suggests that the SARIMA $(0,1,1)(0,1,1)_{12}$ model can serve as a vital tool for modelling and forecasting monthly malaria cases in KZN, SA. It can, therefore, play a key role in shaping the KZN malaria control and elimination effort so that intervention resources can be channelled sustainably and efficiently.

Keywords: *SARIMA, Time series, Malaria, Elimination, Forecast, KwaZulu-Natal, South Africa*

6.2 Introduction

Malaria transmission in South Africa (SA) is restricted to the north-eastern parts of KwaZulu-Natal (KZN), Limpopo and Mpumalanga provinces. Across these malaria endemic provinces, satisfactory progress in malaria disease burden have been recorded and currently malaria incidence is low. Limpopo presents the highest burden of malaria in SA, where malaria incidence ranges from 1.7 to 2.4 cases per 1000 persons at risk, while KZN has the lowest burden of malaria disease (0.01 to 0.10 cases per 1000 persons at risk) [1-3] . Accordingly, SA aims to eliminate malaria by 2020 and prevent the resurgence of malaria transmission in subsequent years [2]. On this endeavour, there is a pressing need to develop robust and reliable predictive models which can strengthen the public health service in decision making for effective targeted malaria transmission combating and elimination strategies.

The development of predictive models is a vital part of malaria surveillance essential to policy makers and public health workers to project the future occurrence of the disease and act proactively [4]. One approach to develop a malaria predictive model is to use historical malaria case data and employ analytical predictive models such as mathematical modelling, machine-learning approach (artificial neural networks) and statistical methods (generalised linear models and seasonal autoregressive intergrated moving average (SARIMA) models). An understanding of the assumptions underlying a predictive model, the advantage(s) and the disadvantage(s) are vital when developing a forecast model [5]. The SARIMA approach can exhibit temporal trends like seasonality and autocorrelation (which is a correlation of a time series with its own past and future values) [6] that is actualised by eliminating high-frequency noise in the data. Furthermore, due to the model's ability to perform automated model determination over a time series, predictions can be said to be reliable if longer time series data are employed. The formulated models are easy to interpret in a retrospective study [5]. Nevertheless, the formulation of the models requires general mathematics and statistics skills, and an understanding of a relevant statistical package/software for the execution of analysis. The required mathematical and statistical skills are not limited to trigonometry, complex numbers, calculus, linear regression (multiple regression and weighted least square) and basic probability [7], while the analysis can be implemented using either an open source (free) statistical packages such as R statistics or python, or a licensed package (STATA, MATLAB, SAS, MiniTab or SPSS).

Based on the need for KZN to enhance their malaria control and elimination efforts, and the epidemiological potential of the SARIMA time series model to that effect, this chapter was designed to develop a SARIMA temporal model using long-term historical malaria case data from KZN. Modelling and predicting malaria cases can potentially assist the KZN malaria control programme and other relevant authorities in making an informed decision about future malaria cases, and provide a basis for malaria prevention and control. It will help optimise the allocation of intervention resources. It can also allow some time to prepare health facilities (procure and re-distribute anti-malarials and diagnostics equipment), engage in targeted and or scaled up vector control activities, and raise awareness amongst locals and visitors/travellers. Furthermore, to the best of the investigator's knowledge, no study has attempted to model and forecast malaria cases in KZN by employing the SARIMA time series analysis. Therefore in this chapter, the investigator seeks to address this gap, by fitting SARIMA models and predict malaria cases in KZN.

6.3 Methodology

6.3.1 Study area

Three district municipalities namely uMkhanyakude, uThungulu and Zululand are malarious areas in KZN province, SA, and were included in this chapter. The study areas are bordered internationally by Swaziland and Mozambique to the north, and the Indian Ocean stretching from the east down to the southeast (Fig. 6.1). The province is characterised by the sub-tropical climate with most of the malaria cases occurring during the rainy months from October to May, with a seasonal peak usually in January and March [8, 9]. The average annual rainfall ranges from 500mm to 2000mm. Along the coastal areas, the summer temperatures is between 24°C to 32°C, and mean winter temperature is about 20°C. The Midlands generally possesses a mild climate with relatively high summer rainfall and dry winters. The elevation measure of the region varies from sea level to over 3000m. The vegetation of the study area comprises of coastal forest and thornveld along the coast. Towards the inlands, lowveld, highland sourveld, Natal sour sandveld, valley bushveld and tall grassveld vegetation are found. Lowveld and thornveld characterises the low-lying hot and dry regions of Northern KZN [10].

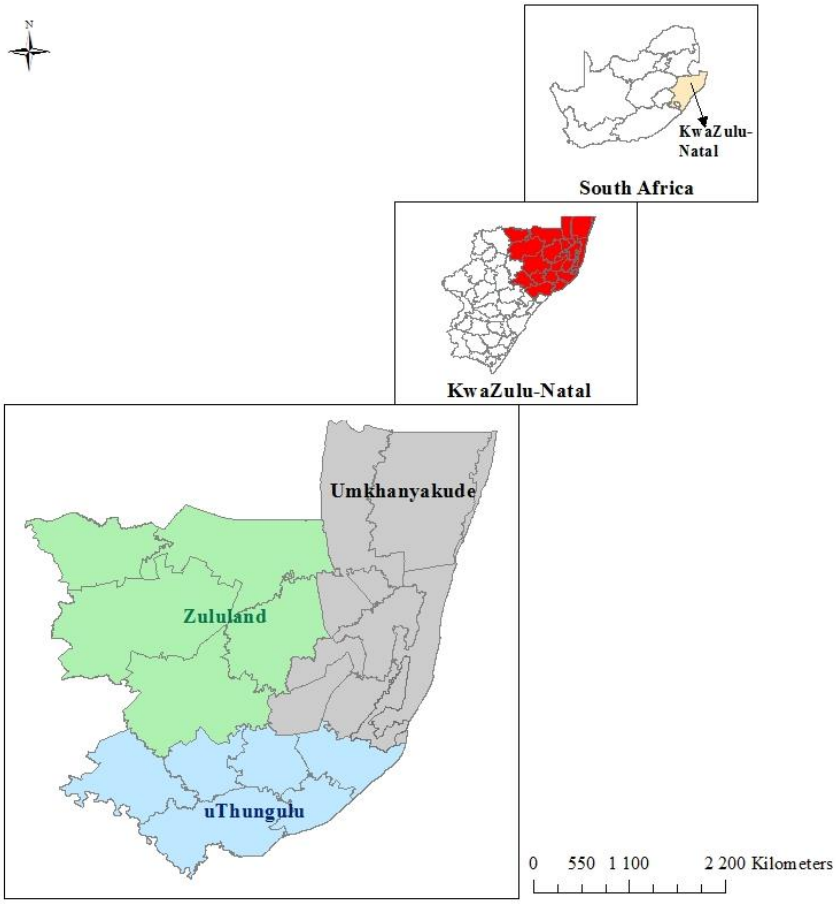


Figure 6.1 Map of the study area

6.3.2. Malaria data

The investigator used confirmed monthly malaria cases including all age groups from January 2005 to December 2014 obtained from the Malaria Control Programme KZN. A malaria case is a person whose blood smear tested positive to *Plasmodium* after undergoing a rapid diagnostic test or slide microscopy at a health facility [11]. Since 1956, it became a legal requirement to notify malaria cases to the relevant health authorities in SA [12]. Confirmed malaria cases at health facilities across the malarious provinces in SA are reported by telephone to the relevant district health office and are subsequently reported to the provincial malaria control programme. At the provincial malaria control programme, the malaria control worker collects and inputs information relating to the malaria case data into the malaria information system. The information includes patient demographics, the health facility the case was reported, symptoms, malaria test results, diagnosis and treatment administered [8, 13].

6.3.3 Statistical analysis

The analytical approach to this chapter is bounded by the Box-Jerkin's SARIMA model. The SARIMA model combines a non-seasonal and seasonal components, and can be specified as SARIMA (p, d, q) x $(P,D,Q)_s$ where, p, d, and q refers to the orders of the non-seasonal autoregressive (AR), non-seasonal differencing, and non-seasonal moving average (MA) parts of the model. P, D, and Q refers to the orders of the seasonal AR, seasonal differencing, and seasonal MA parts of the model and s is the length of the seasonal period. The AR, process accounts for previously observed value up to a specified maximum lag, plus an error term. The process of differencing is referred to as the integration part that accounts for stabilisation of the data by removing seasonality or trend. While the MA process accounts for previous error terms making forecasting easier. The algebraic form of the SARIMA model is given as [14]:

$$\Phi\ (B^S)\ \phi\ (B)\ \Delta^d\ \Delta_S^D\ X_t = \theta_0 + \Theta\ (B^S)\ \theta\ (B)\ a_t \qquad (6.1)$$

The non-seasonal factors are given as:

AR: $\phi(B) = 1 - \phi_1 B - \ldots - \phi_p B^p$ (6.2)

MA: $\theta(B) = 1 + \theta_1 B + \ldots + \theta_q B^q$ (6.3)

The seasonal factors are given as:

Seasonal AR: $\Phi(B^S) = 1 - \Phi_1 B^S - \ldots - \Phi_P B^{PS}$ (6.4)

Seasonal MA: $\Theta(B^S) = 1 + \Theta_1 B^S + \ldots + \Theta_Q B^{QS}$ (6.5)

where, X_t : data series, a_t: random error (with mean zero and variance σ^2), B: backward shift operator, ϕ: coefficient non-seasonal autoregressive, θ: coefficient non-seasonal moving average, Φ: coefficient seasonal autoregressive, Θ: coefficient seasonal moving average, Δ^d: difference operator and d is order of differencing, Δ_S^D: seasonal difference operator and D is seasonal order of differencing, s: length of the seasonal period.

Therefore, a SARIMA $(p,d,q)(P,D,Q)_{12}$ model was constructed using monthly malaria case data from January 2005 to December 2013 and a forecast of malaria cases from January 2014 to December 2014 following the steps below:

Step 1: Transformation of time series data and model identification.

The power transformation known as the Yeo-Johnson transformation was employed on the time series to stabilise the variance. While the SARIMA non-seasonal and seasonal differencing were conducted to achieve stationarity of the time series by eliminating the trend and seasonality. From the non-seasonal and seasonal differenced data, the non-seasonal and seasonal components of the model were formulated by examining their autocorrelation function (ACF) and partial autocorrelation function (PACF). The ACF and PACF were used to determine the degree of differencing and appropriate autoregressive and moving average terms.

Step 2: Parameter estimation.

Parameters of the model in step 1 were estimated to verify if all the parameters in the plausible model were significant.

Step 3: Model validation.

To test for the adequacy of the selected SARIMA model, he investigator used the residuals of the fitted model to find ACF plot of the residuals and the Box-Ljung test. Q-Q plot and Shapiro-Wilk test were used to test for normality of the residuals. If all the diagnostics tests are well meant, then the SARIMA model in step 2 is appropriate.

Step 4: Forecasting

The selected SARIMA model in step 3 was used to forecast malaria cases from January 2014-December 2014. The reported malaria cases for 2014 were used to validate the forecast.

6.4 Results

6.4.1 Model identification

The time series data covers 120 months, starting from January 2005 to December 2015 and it depicts notable seasonality and downward trend of malaria cases as shown in Figure 6.2. The Yeo-Johnson transformation method and differencing were employed to stabilise the variance and eliminate the seasonal trend respectively. The Yeo-Johnson transformation suppressed the fluctuations, which in turn, enhanced the normality of the data (Figure 6.3). This type of transformation suited the dataset because it contains zero values and values close to zero.

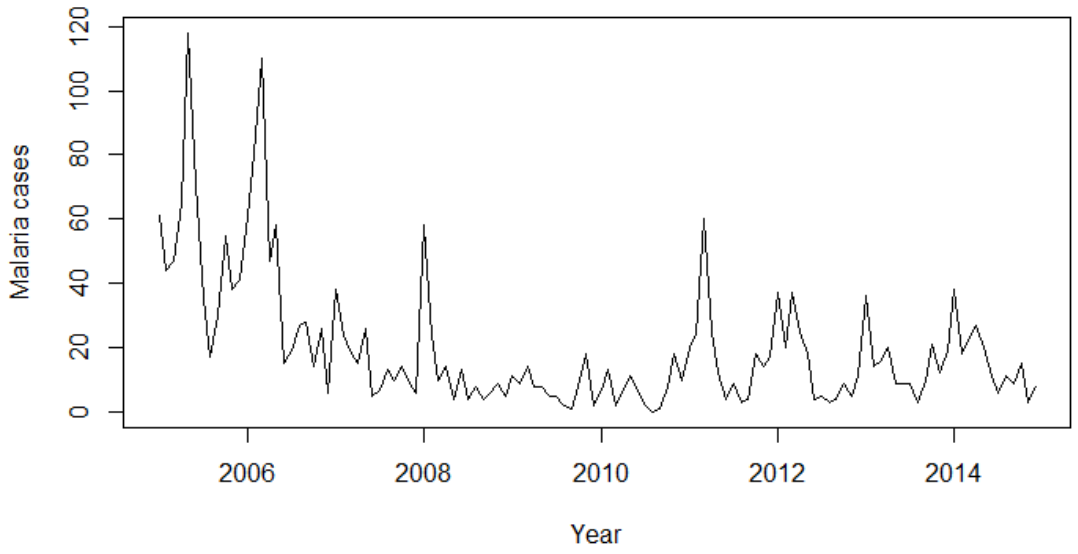


Figure 6.2 Time series plot of monthly malaria cases in KwaZulu-Natal, 2005-2014

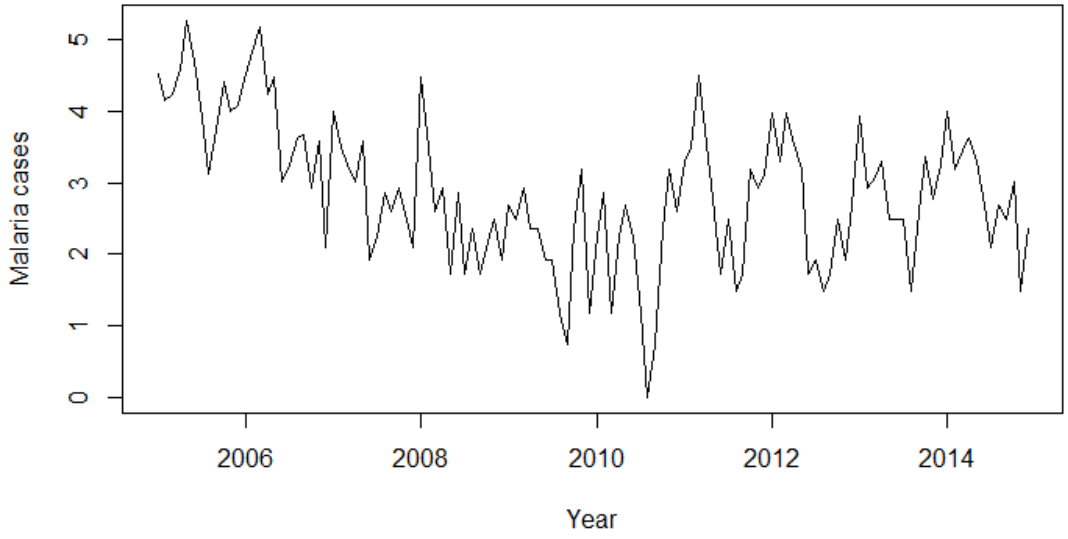


Figure 6.3 Yeo-Johnson's transformation of time series plot of monthly malaria cases in KwaZulu-Natal, 2005-2014

The ACF plot of the transformed malaria case data in Figure 6.4 depicts seasonality, which dies down slightly. While the PACF plot of the malaria case data in Figure 6.5 tails off after lag 1, and decays in sine wave fashion. In an initial attempt to remedy the non-stationarity of the time

series (depicted in Figure 6.3), and eliminate the trend and seasonality (indicated in the ACF plot in Figure 6.4), the non-seasonal differencing was employed.

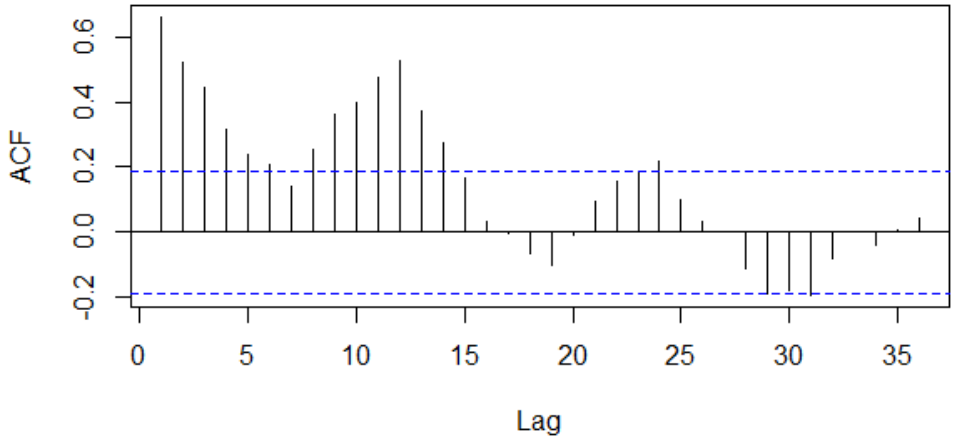


Figure 6.4 ACF plot of the transformed monthly malaria case data

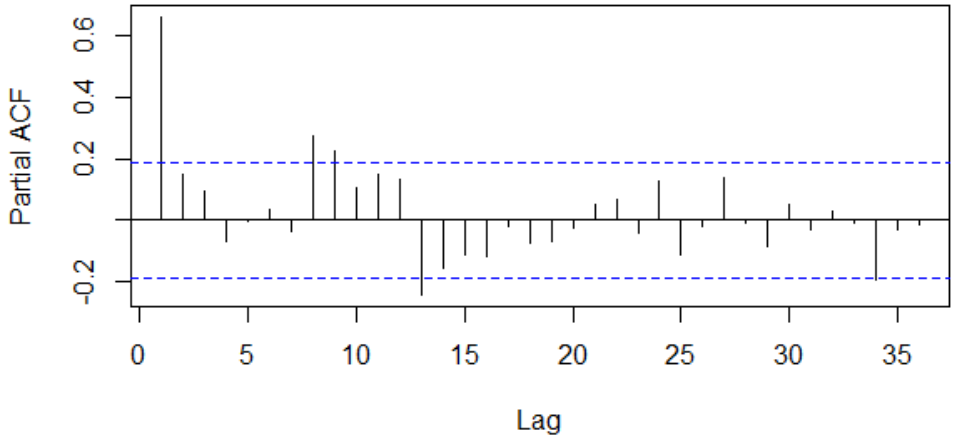


Figure 6.5 PACF plot of the transformed monthly malaria case data

Figures 6.6, 6.7 and 6.8 presents the output of the non-seasonal differenced malaria cases. ACF plot (Figure 6.7) indicates seasonality is still evident (lags 7, 19 and 31). Therefore, the investigator differenced the data a second time (seasonal differencing) to eliminate the effect of seasonality in the model and to seek for a better model fit.

The non-seasonal component of the model was identified by examining the ACF plot (Figure 6.7) and the PACF plot (Figure 6.8) of the non-seasonal differenced malaria cases. The ACF values in Figure 6.7 declines steadily after 1 lag and the PACF (Figure 6.8) decays exponentially in a sine wave fashion. This suggests a moving average of order 1, resulting in an autoregressive moving average $(0,1,1)_{12}$ model (i.e. p = 0, d = 1 and q = 1).

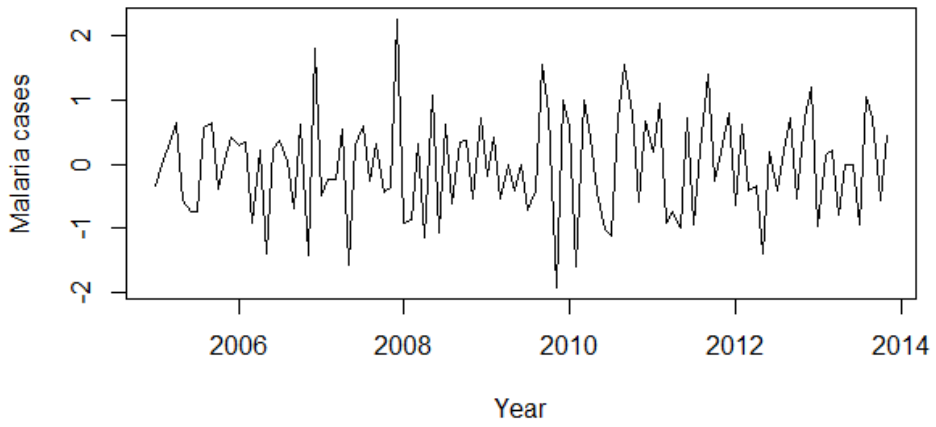


Figure 6.6 Time series plot of the transformed non-seasonal differenced monthly malaria cases exhibiting stationarity

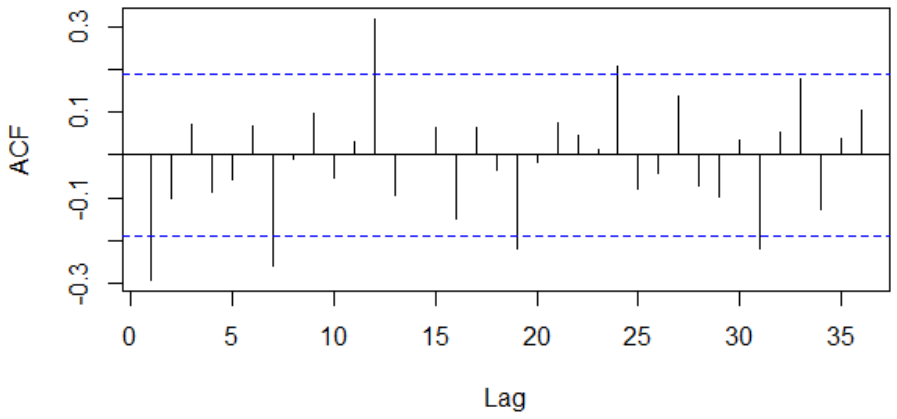


Figure 6.7 ACF plot of the transformed non-seasonal differenced monthly malaria cases

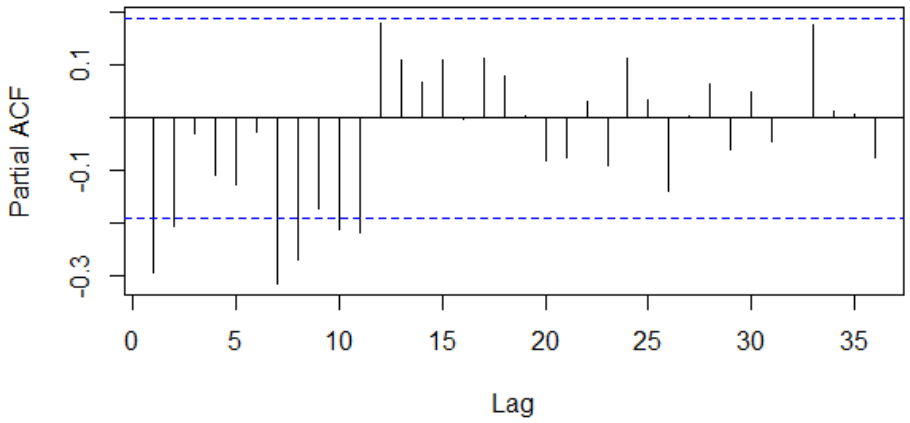


Figure 6.8 PACF plot of the transformed non-seasonal differenced monthly malaria cases

Figures 6.9 and 6.10 shows the ACF and PACF plots of seasonal differenced malaria cases. The ACF of the seasonal differenced malaria cases cuts off after 1 lag which suggest a seasonal moving average MA (1) model, while the PACF plot declines after 3 lags which suggests a seasonal autoregressive AR (3) model. Therefore, based on the non-seasonal differencing and seasonal differencing, seasonality was eliminated from the time series data and, a stationary mean (i.e. D=0) was achieved. This resulted in the identification of three plausible SARIMA models. They are SARIMA $(0,1,1)(3,1,1)_{12}$, SARIMA $(0,1,1)(0,1,1)_{12}$, and SARIMA $(0,1,1)(3,1,0)_{12}$. The investigator, therefore, proceed to the next phase of the Box-Jenkins SARIMA analytical approach, known as parameter estimation or fitting stage of the suggested models.

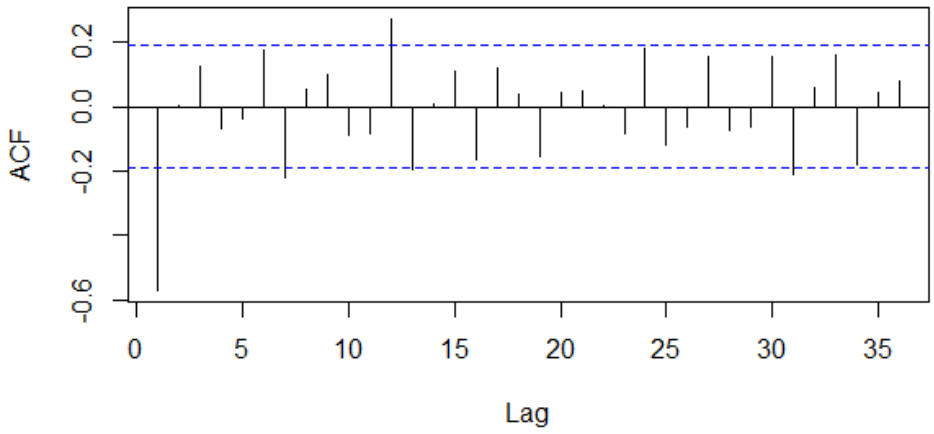


Figure 6.9 ACF plot of the transformed seasonal differenced monthly malaria cases

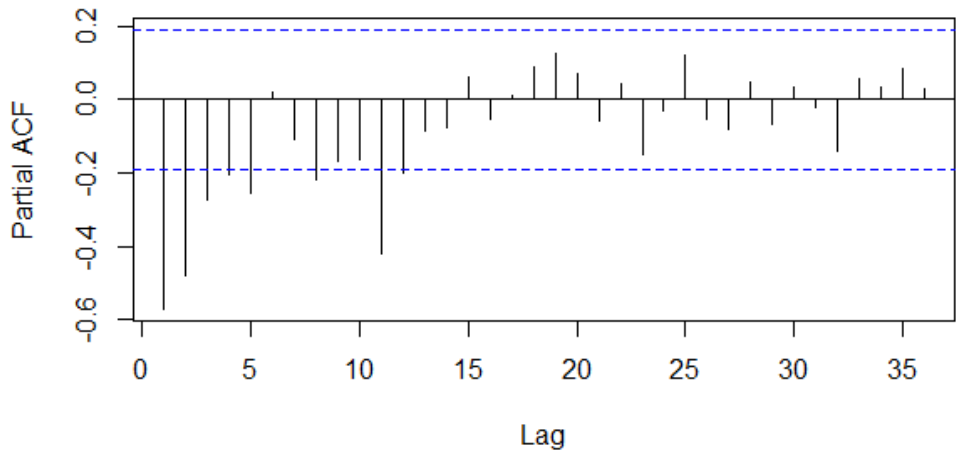


Figure 6.10 PACF plot of the transformed seasonal differenced monthly malaria cases

6.4.2 Model testing and parameter estimation

The goodness of fit statistics employed were the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), log-likelihood and the standard error. The model with the lowest BIC value and with a *p*-value less than 0.05 was selected as the best model fit. The BIC values are based on the likelihood function and the AIC. The SARIMA $(0,1,1)(0,1,1)_{12}$ model has the smallest BIC (Table 6.1) and all the estimates provided in Table 2 are significant. Therefore, based on the goodness of fit statistics (Table 6.1) and parameter estimation (Table 6.2), the investigator identified the SARIMA $(0,1,1)(0,1,1)_{12}$ model as the most suitable model for forecasting. To further evaluate the suitability of the forecast model, the investigator proceed to the next phase of the Box-Jenkins methodology, which is the validation of the model.

Table 6.1 Goodness of fit statistics of the plausible SARIMA models

Statistic	SARIMA $(0,1,1)(3,1,1)_{12}$	SARIMA $(0,1,1)(0,1,1)_{12}$	SARIMA $(0,1,1)(3,1,0)_{12}$
AIC	202.77	199.6	208.28
BIC	218.10	207.26	221.02
LL	-95.39	-96.80	-99.12

AIC=Akaike information criterion; BIC=Bayesian information criterion; LL=Log likelihood

Table 6.2 Parameter estimation of the plausible SARIMA models

	SARIMA $(0,1,1)(3,1,1)_{12}$			SARIMA $(0,1,1)(0,1,1)_{12}$			SARIMA$(0,1,1)(3,1,0)_{12}$		
Type	**Coef**	**S.E of coef**	***P*-values**	**Coef**	**S.E of coef.**	***P*-values**	**Coef**	**S.E of coef**	***P*-values**
MA1	-0.7029	0.0803	0.0000	-0.7156	0.0785	0.0000	-0.7274	0.0767	0.0000
SAR1	0.2003	0.1189	0.0460	-	-	-	-0.5261	0.1083	0.0000
SAR2	-0.0002	0.1197	0.4995	-	-	-	-0.3182	0.1186	0.0037
SAR3	-0.0862	0.1328	0.2582	-	-	-	-0.1561	0.1265	0.1087
SMA1	-1.0000	0.3738	0.0037	-0.7272	0.1316	0.0000	-	-	-

MA=Non-seasonal moving average; SAR= Seasonal autoregressive; SMA = Seasonal moving average S.E = Standard error; Coef.= coefficient

6.4.3 Model validation

Even though the choosen SARIMA $(0,1,1)(0,1,1)_{12}$ model seems most suitable, validation of the model is vital to ascertain if the model possesses systematic patterns that can be removed to enhance the functionality of the model and, this is done by examining the residuals of the model. The model verification was tested by verifying (1) the ACF of the residuals to check for autocorrelation, and (2) the normal probability plot of the residuals.

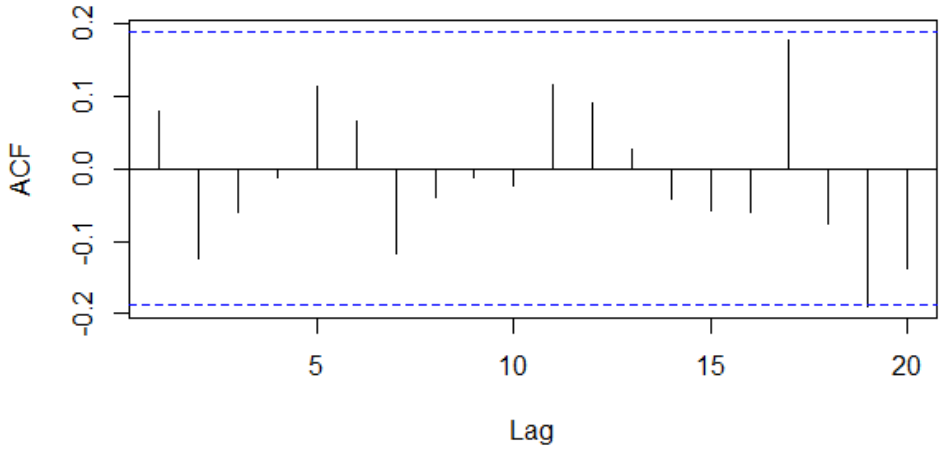


Figure 6.11 ACF plot of residual for SARIMA $(0,1,1)(0,1,1)_{12}$ model

The ACF plot of residuals (Figure 6.11) suggests that the residuals have a constant variance, and the autocorrelations were modelled out leaving only one significant value as indicated by the spike in lag 19. Also, the Box-Ljung test results (x-squared = 60.499, df=48, p-value = 0.1064) revealed that the p-value exceeded 5% and this implies the model is adequate (i.e. there

is no autocorrelation). The Shapiro-Wilk test results for normality possesses a test statistic of W=0.98811 and p-value =0.4595 and, the Q-Q plot (Figure 6.12) depicts some outliers on the tails which suggests that the normality of the residuals is not rejected. Therefore, the investigator proceed to use the SARIMA $(0,1,1)(0,1,1)_{12}$ model for forcasting since it provides a reasonable fit to the highly seasonal and non-seasonal time series data.

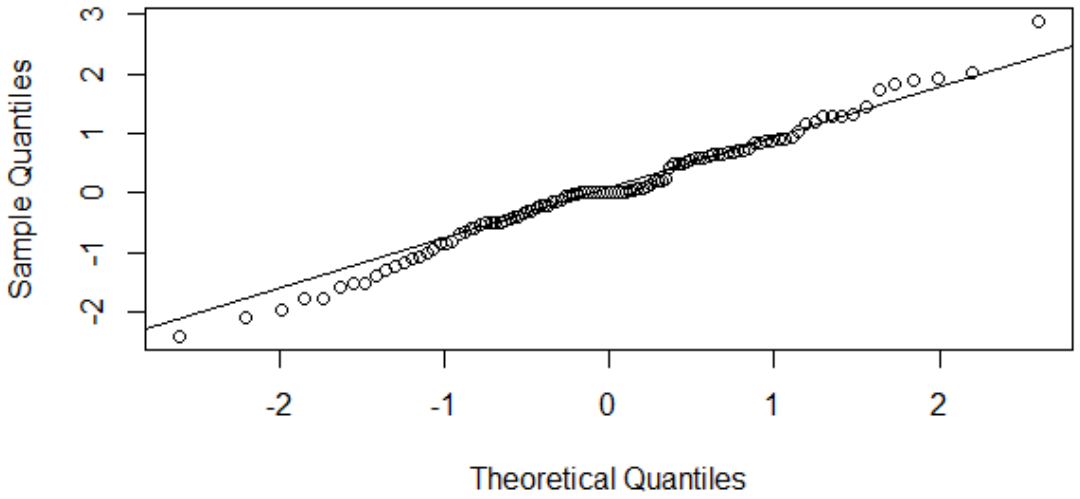


Figure 6.12 The Q-Q plot of residuals of the selected SARIMA $(0,1,1)(0,1,1)_{12}$ model

6.4.4 Forecasting

The selected SARIMA $(0,1,1)(0,1,1)_{12}$ model was used to forecast monthly malaria cases from January 2014 to December 2014 as shown in figure 6.13. The predicted estimates of monthly malaria cases are represented by the blue line in the figure, while the 95% and 85% confidence bounds are shaded in ligher and darker grey respectively.

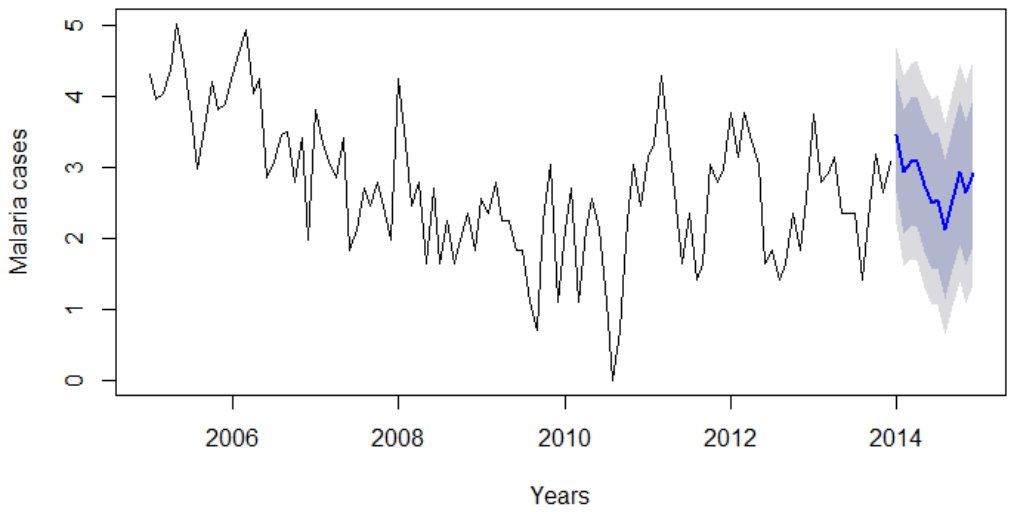


Figure 6.13 Observed monthly malaria cases from January 2005-December 2013 and predicted monthly malaria cases from January 2014 to December 2014

Table 6.3 presents the values of predicted malaria cases at 95% confidence level and, the model overpredicted malaria cases by 8%.

Table 6.3 Predicted monthly malaria cases for KwaZulu-Natal, South Africa from January 2014 to December 2014 using the SARIMA $(0,1,1)(0,1,1)_{12}$ model

Month	Predicted case	95% CL (Lower)	95% CL (Upper)
January	3.805152	2.5555447	5.054760
February	3.407794	2.1086728	4.706916
March	3.525228	2.1784115	4.872044
April	3.420779	2.0279002	4.813658
May	3.065883	1.6284168	4.503349
June	2.470701	0.9899892	3.951412
July	2.435048	0.9123190	3.957777
August	1.825197	0.2615790	3.388815
September	2.209246	0.6057816	3.812710
October	3.086150	1.4438058	4.728495
November	2.951564	1.2712388	4.631889
December	2.926445	1.2089793	4.643912

CL=confidence level

The plot of the observed monthly malaria cases and predicted cases for 2014 (Figure 6.14) shows that the values for monthly predicted cases tends to follow the reported values quite closely except in August, November and December where pronounced differences were observed.

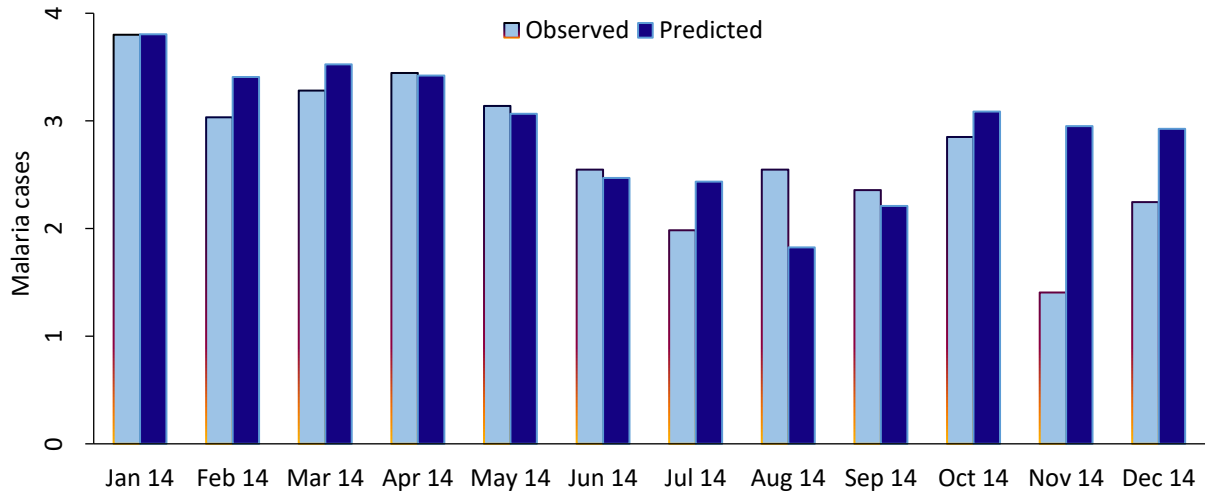


Figure 6.14 Observed and predicted monthly malaria cases from January 2014 to December 2014

6.5 Discussion

Time series predictions are generated by models based on changes over time in previously observed values or historical datasets [15]. The SARIMA forecast model can serve as a vital tool for public health workers and epidemiologists. It can be applied as a malaria early warning system and, can provide vital information for the relevant authority to act proactively [15, 16]. This chapter presents an example how the SARIMA model which is particularly relevant for a disease that exhibits seasonality [6] was employed in modelling and predicting malaria cases in a relatively low malaria-transmission region, where targeted interventions are extremely vital to strengthen KZN malaria control and elimination efforts. The model can provide information to support policy makers and public health efforts so that intervention resources can be provided and chanelled in a sustainable and effective way. It can also serve as a tool for providing relevant information to locals and visitors prior to high malaria transmission months. This, in turn, will be pivotal in transforming SA's current malaria programme to elimination in 2020.

The epidemiological potential and functionality (epidemiological studies, disease surveillance, and forecasting) of the SARIMA time series, have also been explored by different authors in different capacities [17-21]. These authors ensured the time series processes attained stationarity in the homogenous sense (stationary in its level) and variance, which are indispensable conditions of a SARIMA model. This was done by carrying out the first differencing, and the seasonal differencing, which results in a stationary time series by removing trends and seasonal effects. However, in instances where the variance of a time series trends downwards (or increases) as the level of the series decreases (or increases), the time series must be transformed before the analysis or differencing [15]. This will lead to a time series stationary in the homogenous sense and variance, and in turn, improves and leads to the formulation of a better model fit [15]. Some studies employed the log-transformation to achieve a stationary variance [22-24], and it is the most commonly used transformation approach. Other studies employed the Box-Cox power transformation [25, 26] which is valid for datasets containing positive variables. In this chapter, the investigator employed a seldom-used power transformation known as the Yeo-Johnson transformation because the time series systematically trended downwards and, had zero values and values close to zero [27].

Even though malaria transmission in KZN is limited as a result of effective malaria control measures [9, 28-30], the SA National Department of Health still regards malaria a significance disease due to its propensity to cause an epidemic [8, 11]. In SA, specific population groups are at higher risk of contracting malaria. They are infants and children below five years old living in localities of stable malaria transmission, elderly (above 65 years old), people living with HIV/AIDS, non-immune pregnant women, semi-immune pregnant women living in high malaria transmission localities, semi-immune HIV-infected pregnant women living in localities of stable transmission, non-immune travellers and migrants [11]. Nevertheless, the entire population is vulnerable to malaria epidemic due to little or no immunity [11]. To avert an epidemic, SA has in place an outbreak threshold of confirmed cases at districts and provinces endemic to malaria, and health facilities located in these areas. When the threshold is reached or exceeded, reactive measures are taken by the relevant malaria divisions [8] The malaria control and elimination efforts needed will, therefore, require scaling up and revising of the epidemic preparedness and response strategy.

In addition to the SARIMA model, further studies should be conducted utilising either epidemiological or entomological data or a combination of both with environmental and socio-economic malaria triggers at a fine scale to delineate the needed intervention resources. Furthermore, conducting epidemiological research, and studies in a particular setting employing different parameters, mechanisms or possible confounder can inform on the deficiencies in the knowledge of the malaria disease about that study area. This could result in identifying unanswered questions and loopholes crucial to model reliability, which can be bridged by either incorporating more variables or eliminating certain variables. Nevertheless, without practical applications of developed models, we will not be informed about the reliability of the models and how they can be improved. It is also important that relevant stakeholders (researchers, statistical analyst, the South African Weather Services, physicians, public health workers, epidemiologist, entomologists and policy makers) are brought together to reshape the malaria elimination strategy so that reliable and operational malaria case predictions models can be generated. Other importance of such multi-disciplinary collaboration are the identification of reliable malaria predictors and accurate hotspots of malaria transmission [31]. In addition, a reliable and direct means of accessing and sharing information among the relevant stakeholders is of utmost necessity.

The weakness of this chapter is that it attempted to develop a single model for the entire malaria area of KZN. Separate models for each of the district municipalities might provide an in-depth assessment of the malaria trends across the municipality district, which in turn may help identify possible differences in the implementation of prevention measures, patients' seeking behaviours and migration of people. The structuring of the data, and the mode of the forecasting using monthly data and forecasting maybe responsible for the overestimated and underestimated monthly forecast observed. Conducting daily data analysis could result in improved model fit and daily forecasts, which can then be aggregated into weekly and monthly forecasts. The univariate analysis approach employed in this chapter could also be another reason for the overestimated and underestimated forecasts. The incorporation of predictors into the SARIMA model (multivariate SARIMA model) over a longer time frame could improve the model fit and the forecast if the exogenous factors responsible for trend, seasonality and outliers are incorporated into the model.

6.6 Conclusions

The SARIMA forecast model is a vital tool that has the potential for malaria early warning, and can provide information to the relevant authority to act proactively. In this chapter, time series SARIMA models guided by the Box-Jenkins approach were constructed using historical monthly malaria case data of KZN from 2005 to 2014. The best fit SARIMA $(0,1,1)(0,1,1)_{12}$ model was used to predict 2014 monthly malaria cases. The predicted values were validated with the reported cases and, it was observed to fit closely with the actual malaria cases having tolerable error values. The outcome of the validation indicates that the model could be used for predicting malaria cases in KZN, SA. The practical application of the generated model is encouraged. Furthermore, studies that employ daily data and incorporate possible malaria transmission risk factors, and confounders in multivariate time-series models are recommended.

6.7 Conflicts of interests

The authors declare that they have no competing interests.

6.8 Acknowledgments

The authors would like to thank the College of Agriculture, Engineering and Science of the University of KwaZulu-Natal for the doctoral research bursary awarded to Osadolor Ebhuoma. We also thank the malaria control program of KwaZulu-Natal, South Africa for providing the data.

6.9 Ethical approval

Not required.

6.10 References

1. World Health Organisation. *Malaria country profile 2016*. 2016. Accessed 11/09/2017; Available from: http://www.who.int/malaria/publications/country-profiles/en/.

2. Elimination 8. *Annual report 2016*. 2016. Accessed 11/09/2017; Available from: https://malariaelimination8.org/wp-content/uploads/2017/05/e8-annual-report-2016.pdf.

3. Raman, J., N. Morris, J. Frean, B. Brooke, L. Blumberg, P. Kruger, A. Mabusa, E. Raswiswi, B. Shandukani, E. Misani, et al., *Reviewing South Africa's malaria*

elimination strategy (2012–2018): progress, challenges and priorities. Malaria Journal, 2016. **15**(1): p. 438.

4. Cunha, G.B.d., J.F. Luitgards-Moura, E.L.M. Naves, A.O. Andrade, A.A. Pereira, and S.T. Milagre, *Use of an artificial neural network to predict the incidence of malaria in the city of Canta, state of Roraima.* Revista da Sociedade Brasileira de Medicina Tropical, 2010. **43**(5): p. 567-570.

5. Zinszer, K., A.D. Verma, K. Charland, T.F. Brewer, J.S. Brownstein, Z. Sun, and D.L. Buckeridge, *A scoping review of malaria forecasting: past work and future directions.* BMJ Open, 2012. **2**(6).

6. Chatfield, C., *The Analysis Of Time Series: An Introduction*. 6th ed. 2003, Florida: CRC press.

7. Shumway, R. and D. Stoffer, *Time Series Analysis Using the R Statistical Package*. 2017: Free dog publishing.

8. South Africa National Department of Health, *Republic Of South Africa Malaria Elimination Strategy 2011–2018*. 2012: Pretoria, South Africa.

9. Moonasar, D., T. Nuthulaganti, P.S. Kruger, A. Mabuza, E.S. Rasiswi, F.G. Benson, and R. Maharaj, *Malaria control in South Africa 2000–2010: beyond MDG6.* Malaria Journal, 2012. **11**(294): p. 1475-2875.

10. Camp, K.G.T., *A Bio-Resource Classification For KwaZulu-Natal, South Africa*, in *School of Applied Environmental Sciences*. 1999, University of KwaZulu-Natal, South Africa: Pietermaritzburg, South Africa.

11. South Africa National Department of Health, *Guidelines For The Treatment Of Malaria In South Africa-2016*. 2016: Pretoria, South Africa.

12. South Africa National Department of Health, *Notification Of Diseases*. 1956: Pretoria, South Africa.

13. Khosa, E., L.R. Kuonza, P. Kruger, and E. Maimela, *Towards the elimination of malaria in South Africa: a review of surveillance data in Mutale Municipality, Limpopo Province, 2005 to 2010.* Malaria Journal, 2013. **12**(7).

14. Brockwell, P.J. and R.A. Davis, *Time Series: Theory And Methods*. 1991: Springer Science & Business Media.

15. McCleary, R., R.A. Hay, E.E. Meidinger, and D. McDowall, *Applied Time Series Analysis For The Social Sciences*. 1980: Sage Publications Beverly Hills, CA.

16. Midekisa, A., G. Senay, G.M. Henebry, P. Semuniguse, and M.C. Wimberly, *Remote sensing-based time series models for malaria early warning in the highlands of Ethiopia.* Malaria Journal, 2012. **11**.

17. Ekezie, D., O. Jude, and O. Idochi, *Modelling and forecasting malaria mortality rate using Sarima models (a case study of Aboh Mbaise general hospital, Imo State Nigeria).* Science Journal of Applied Mathematics and Statistics, 2014. **2**(1): p. 31-41.

18. Kumar, V., A. Mangal, S. Panesar, G. Yadav, R. Talwar, and D. Raut, *Forecasting malaria cases using climatic factors in Delhi, India: a time series analysis.* Malaria Research and Treatment, 2014. **1**.

19. Permanasari, A.E., I. Hidayah, and I.A. Bustoni. *SARIMA (Seasonal ARIMA) implementation on time series to forecast the number of Malaria incidence*. in *Information Technology and Electrical Engineering (ICITEE), 2013 International Conference on*. 2013. IEEE.

20. Briët, O.J., P. Vounatsou, D.M. Gunawardena, G.N. Galappaththy, and P.H. Amerasinghe, *Models for short term malaria prediction in Sri Lanka.* Malaria Journal, 2008. **7**(1): p. 76.

21. Wangdi, K., P. Singhasivanon, T. Silawan, S. Lawpoolsri, N. White, and J. Kaewkungwal, *Development of temporal modelling for forecasting and prediction of malaria infections using time-series and ARIMAX analyses: a case study in endemic districts of Bhutan.* Malaria Journal, 2010. **9**.

22. Zinszer, K., R. Kigozi, K. Charland, G. Dorsey, T.F. Brewer, J.S. Brownstein, M.R. Kamya, and D.L. Buckeridge, *Forecasting malaria in a highly endemic country using environmental and clinical predictors.* Malaria Journal, 2015. **14**(1): p. 1.

23. Homan, T., N. Maire, A. Hiscox, A. Pasquale, I. Kiche, K. Onoka, C. Mweresa, W.R. Mukabana, A. Ross, and T.A. Smith, *Spatially variable risk factors for malaria in a geographically heterogeneous landscape, western Kenya: an explorative study.* Malaria Journal, 2016. **15**(1): p. 1.

24. Krefis, A.C., G.N. Schwarz, A. Krüger, J. Fobil, B. Nkrumah, and S. Acquah, *Modeling the relationship between precipitation and malaria incidence in children from a holoendemic area in Ghana.* American Journal of Tropical Medicine and Hygiene, 2011. **84**.

25. Trebbia, G., M. Lacombe, C. Fermanian, L. Falaize, M. Lejaille, A. Louis, C. Devaux, J.C. Raphaël, and F. Lofaso, *Cough determinants in patients with neuromuscular disease.* Respiratory Physiology & Neurobiology, 2005. **146**(2): p. 291-300.

26. Ren, Z., D. Wang, J. Hwang, A. Bennett, H.J. Sturrock, A. Ma, J. Huang, Z. Xia, X. Feng, and J. Wang, *Spatial-Temporal Variation and Primary Ecological Drivers of*

Anopheles sinensis Human Biting Rates in Malaria Epidemic-Prone Regions of China. PLOS one, 2015. **10**(1): p. e0116932.

27. Yeo, I.K. and R.A. Johnson, *A new family of power transformations to improve normality or symmetry.* Biometrika, 2000. **87**(4): p. 954-959.

28. Ebhuoma, O., M. Gebreslasie, and L. Magubane, *Modeling malaria control intervention effect in KwaZulu-Natal, South Africa using intervention time series analysis.* Journal of Infection and Public Health, 2017. **10**(3): p. 334-338.

29. Maharaj, R., D. Mthembu, and B. Sharp, *Impact of DDT re-introduction on malaria transmission in KwaZulu-Natal.* South African Medical Journal, 2005. **95**(11): p. 871.

30. Maharaj, R., J. Raman, N. Morris, D. Moonasar, D.N. Durrheim, and I. Seocharan, *Epidemiology of malaria in South Africa: from control to elimination.* South African Medical Journal, 2013. **103**.

31. Ebhuoma, O. and M. Gebreslasie, *Remote Sensing-Driven Climatic/Environmental Variables for Modelling Malaria Transmission in Sub-Saharan Africa.* International Journal of Environmental Research and Public Health, 2016. **13**(6): p. 584.

CHAPTER 7: GENERAL CONCLUSIONS AND RECOMMENTATIONS

The aim of this study was to contribute to malaria epidemiology by exploring the factors that influence malaria transmission in the malarious regions of KZN, SA by employing different spatial and temporal models. To achieve this aim, the specific objectives were: 1) To systematically appraise the existing body of literature on RS-derived climatic and environmental determinants of malaria transmission in SSA by identifying determinants peculiar to regions in SSA, appraise modelling approaches, and current research shortfalls; 2) to identify the climatic and environmental determinants of malaria transmission in the malarious regions in KZN, SA, and develop a malaria risk map; 3) to determine the socio-economic factors that influence malaria transmission at local municipality level in the malarious regions in KZN, SA; 4) to evaluate the malaria control intervention (the use of DDT) on malaria transmission in the malarious regions in KZN, SA; and 5) to develop a forecasting model to predict malaria in the malarious regions in KZN, SA.

In the second chapter of this study titled "Remote sensing-derived climatic/environmental variables for modelling malaria transmission in sub-Saharan Africa", the investigator presented a systematic collation and summary of relevant studies that utilised RS-derived climatic and environmental variables for modelling malaria transmission in SSA. The investigator identified thirty-five peer-reviewed articles that met the final selection criteria. Across the SSA sub-region, NDVI was the most frequently returned as a statistically-significant variable to model both spatial and temporal malaria transmission. In terms of analytical approaches, linear models were widely used across SSA. Based on the finding, the investigator suggested the following: 1) the utilization of RS in determining reliable malaria transmission predictors and developing environmental monitoring, a tailored approach is required that takes into account the geographical/climatic setting, the stage of the malaria elimination continuum, the characteristics of the RS variables and the analytical approach; 2) Bayesian modelling approach can be used as reliable substitute for linear models. They provide extensions of generalised linear models and are formulated to overcome some of the setbacks of linear models.

The third chapter titled, "Modelling of malaria transmission using climatic and environmental variables in KwaZulu-Natal province with low malaria transmission via Bayesian zero-inflated

models in INLA", identifies different climatic and environmental variables that influence malaria transmission and develop malaria risk map for the malarious regions of KZN. The results of the posterior statistics from the ZIP model revealed that malaria cases are significantly related to NDVI, precipitation, elevation and night temperature. Also, the malaria risk map developed showed that the North Eastern part of KZN province possesses the highest risk of malaria morbidity. The modelling approach employed in this chapter presents a valuable tool for understanding and monitoring the influence of climatic and environmental variables on the spatial heterogeneity of malaria in KZN. It will therefore equip the relevant policy makers with information required to channel malaria intervention resources sustainably to vulnerable receptive areas. Also, this chapter reveals the need to strengthen the already existing cross-border initiatives to boost KZN's malaria elimination goals. However, a modelling approach which takes account of the effects of population movement between the E8, MOSASWA countries and from other malaria endemic countries was recommended.

The fourth chapter titled, "Socio-economic determinants of malaria transmission risk in KwaZulu-Natal, South Africa: a Bayesian approach", identified relevant SES that influenced malaria transmission. This chapter suggests that the following determinants of low SES play a significant role in malaria transmission and burden: illiteracy, lack of electricity, lack of toilet facilities and unemployment. Other factors that had an effect on the risk of malaria are gender, children less than 5 years old and adult above 65 years old. This means low SES, socio-economic deprivation and poverty can maintain and exacerbate malaria transmission in KZN. As an implication, poverty alleviation and malaria intervention resources should be incorporated side by side into the socio-economic framework. Therefore, the relevant policy makers and departments should invest more on sustainable developmental approach that combines both improved malaria intervention resources and socio-economic conditions. This can ultimately help strengthen the malaria elimination goals in KZN.

The fifth chapter titled, "Modelling malaria control intervention effect in KwaZulu-Natal, South Africa using intervention time series analysis", evaluated the effect of the re-introduction of DDT to combat malaria transmission in KZN and suggested practical ways the province can strengthen the already existing malaria control and elimination efforts, to achieve zero malaria transmission. The ITSA revealed an abrupt and permanent decline of confirmed malaria cases

in KZN. This confirms that DDT contributed significantly to the low and sustained malaria case in the province. The sustained low malaria cases recorded over the last two decades suggests that the continued usage of DDT did not result in insecticide resistance as earlier anticipated. Thus, for the province to eliminate malaria transmission completely, renewed attention should be given to malaria in KZN in terms of more financial commitment and multidisciplinary research efforts. This should comprise of location specific studies and the introduction of other reliable and complementary intervention resources to support and optimise the already existing malaria intervention. Such as (1) the introduction of pre-erythrocytic malaria vaccines, RTS,S/AS01 for residence living in malaria endemic areas and visitors, (2) the introduction of genetically modified sterile male *Anopheles arabiensis* mosquitoes and (3) the practical application of reactive case detection to determine the optimal screening radius peculiar to each of the malaria transmission settings or the focal screening and treatment depending on the level of malaria endemicity.

The sixth chapter titled, "A Seasonal autoregressive integrated moving average (SARIMA) forecasting model to predict monthly malaria cases in KwaZulu-Natal, South Africa", develops a forecast model of monthly malaria cases in KZN by using the Seasonal Autoregressive Integrate Moving Average (SARIMA) time series approach. Three plausible models were identified, and the SARIMA (0,1,1)(0,1,1)12 model was selected as the best fit model. The SARIMA (0,1,1)(0,1,1)12 model was used to forecast malaria cases during 2014, and it was observed to fit closely with malaria cases reported in 2014. The outcome of this chapter suggests that the SARIMA (0,1,1)(0,1,1)12 model can serve as a vital tool for modelling and forecasting monthly malaria cases in KZN, SA. Nevertheless, the practical application of the generated model is suggested. Furthermore, studies that employ daily data and incorporate possible malaria transmission risk factors, and confounders in multivariate time-series models were recommended.

The different modelling or analytical approaches employed in this study illustrates the importance of conduction various spatial and temporal studies in a setting with low malaria endemicity seeking to achieve malaria elimination. It also presents a road map for malaria elimination in KZN. It was able to identify the high malaria risk areas in KZN that requires re-enforcement of malaria intervention resources and the contributory factors (SES, climatic and

environmental predictors) of malaria transmission, evaluate to relevance of applying DDT for malaria intervention purposes, and predict malaria case. These information can serve as a vital tool for the KZN malaria control program, policy makers, relevant research institutions, statisticians, the community health workers, SA's DOH, SA's department of social development, the E8 and the MOSASWA communities towards a holistic and integrated approach to achieve malaria elimination in the province in 2020 and in SA at large. Considering the multifaceted nature and interrelated links of the various elements that play significant roles in the proliferation of malaria, a solid cooperation with the relevant stakeholders cannot be overemphasised. This approach will provide measures for sustainable poverty alleviation in the malarious communities, and provide intervention resources that can be implemented sustainably. Thus, the recommendations that emerged from the various studies should be practically applied by the relevant authorities and more studies should be conducted in accordance with the identified research gap. In addition to the specified research gap, relevant entomological studies should be conducted and the suggested epidemiological studies should be conducted at much smaller scale. Due to that dynamic nature of malaria in KZN, the recommendations provided in this study should be treated as a road map for malaria elimination in KZN. As the KZN epidemiological environment evolves, elimination approaches should also evolve through adaptive and innovative approaches, and constant reviews and evaluation of the approaches until elimination is finally attained.

The models generated from the various studies demonstrated the need for the KZN malaria program, relevant policy makers and stakeholders to further strengthen the KZN malaria elimination efforts. The required malaria elimination fortification are not limited to the implementation of additional sustainable developmental approach that combines both improved malaria intervention resources and socio-economic conditions, strengthening of existing community health workers, strengthening the relationship with the relevant stakeholders and strengthening of the already existing cross-border collaborations.