Malnutrition related child morbidity and mortality : A space-time based analysis using Kilifi County Hospital Data 2002 to 2015

Student: KM Wambui

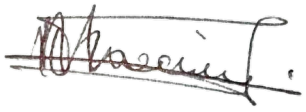
Supervisor: Associate Prof Eustasius Musenge

A RESEARCH REPORT SUBMITTED TO THE SCHOOL OF PUBLIC HEALTH, UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG, IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE IN BIOSTATISTICS

November, 2017

DECLARATION

I, KM Wambui, student No: 1504769 declare that this research report is my own work. It is being submitted for the degree of Master of Science in Epidemiology in the field of Biostatistics at the University of Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination at this or any other University.



KM Wambui November 2017

DEDICATION

This research is dedicated to my loving mum Wambui and Mwai's family at large for the uncontinued prayers and support during the research period , ***Kũgũrũ nĩ ĩrata thĩ***, (Perseverance prevails).

To **Jehova Ngai** (THE MOST HIGH) , all is unto you.

ABSTRACT

Background:

Globally malnutrition is an underlying cause of death and accounts for over 45% of under-5 mortality mainly resulting from diarrhoea and pneumonia. The post-2015 era has seen, more than 25% of Kenya's population being food insecure, with considerable geographic and temporal disparities. Our primary aim was to understand the determinants of malnutrition related morbidity and mortality in the rural Kilifi HDSS, with a special focus on children admitted in Kilifi County Hospital (KCH) during 2002-2015.

Methodology:

Our study participants were all the children between the ages of 6 months to 15 years who were admitted two times or more at the KCH. The outcomes were derived from malnutrition-related admissions based on wasting (WHZ<-2) and oedema and the discharge outcome whether alive or died. There were 3114 children with a total of 7620 admissions for children with more than one admission.

In the exploratory data analysis, temporality and seasonality were determined using SARIMA time series models. Morans I index was used to investigate for the presence of spatial autocorrelation. SatScan was used to identify the spatial clusters of malnutrition related admissions and mortality. To understand mortality patterns, geo-additive logistic models were fitted to the KCH data. Mixed effects negative binomial models with separate space and temporal random effects were fit using the Maximum Likelihood and Bayesian Estimation procedures. The Bayesian methods were used to estimate the spatial parameters using Markov Chain Monte Carlo (MCMC) assisted with either Metropolis Hastings or Integrated Nested Laplace Approximations (INLA).

Results:

There were 17,740 children observed over the period of study and 4.01% of those died. A total of 23,347 admission events were observed of which 7,128 were malnutrition related. Out of the 17,740 children admitted, 3,114 had one or more admission event. A seasonal hike in the May to July month was identified for malnutrition admission. Children with more than one admission, (7620 admissions) ~24%(n=1858) had a malnutrition event and 6.24% of them died. Spatial hotspots clusters were identified in the North and South of the creek and areas near Kilifi Town was identified as cold spots. Children with two or more severe diseases are more likely to have a malnutrition admission event and females are less likely to be admitted with malnutrition. There was a protective effect as the children grew older and also as their body weights increased.The males had a higher risk of death compared to the females and a year increase in age reduced the risk of death by 15%.

Conclusion:

A better understanding of the factors that contribute to malnutrition attributable admission and mortality can be used to advocate for and develop earlier and more appropriate responses. Additionally, this can provide an indication of future trends and the potential impact of interventions.Importantly, including spatial and temporal random effects biostatistical modelling can help reduce bias reporting and help understand better the patterns of morbidity and mortality. Campaigns providing food and/or vitamin or other supplements can contribute to reducing morbidity and ultimately deaths in Kenyan children and building more health facilities to reduce the distance of travel to care is highly recommendable.

ACKNOWLEDGEMENTS

The achievement and realization of this research project has been through the mentorship of my supervisor **Associate Prof Eustasius Musenge**. I am very grateful for the professional and personal advice he accorded me. As you continue mentoring many more, may blessings come your way.

Secondly I thank Sub-Saharan African Consortium for Advanced Biostatistics (SSACAB) training programme for funding me to advance my career in Bio-statistics. All the lectures, administrators and staff at the School of Public Health my gratitude for the support during the course period. My classmates; especially Biostatistics cohort Karimi, Chidumwa and Kaunda thank you for making the course fun and enjoyable.

Special thanks to KEMRI-Wellcome Trust for providing me with data for my masters especially David Amadi , Prof. James Berkley and Prof. Greg Fegan.

Special gratitude to my mum Wambui Mwai for the support and prayers during my course work and my career path.

Finally without you God, I could not have made it. *Siyabonga*

Ũgĩ wa mũndũ ũmwe ndũrĩmaga (A single mans ability cannot till all fields)

TABLE OF CONTENTS

Contents

List of Figures

List of Tables

NOMENCLATURE

- **KHDSS** - Kilifi Health and Demographic Surveillance System
- **KCH** - Kilifi County Hospital
- **MySQL** -My Structured Query Language
- **WHZ** -Weight for Height Z Scores
- **WHO** -World Health Organisation
- **LRTI** -Lower Respiratory Tract Infection
- **CSF** -Cerebrospinal fluid
- **UNICEF** -United Nations Children's Fund
- **MODIS** -Moderate Resolution Imaging Spectroradiometer
- **EVI** -Enhanced Vegetation Index
- **SD** -Standard Deviation
- **ARIMA** -AutoRegressive Integrated Moving Average
- **SARIMA** -Seasonal AutoRegressive Integrated Moving Average
- **ACF** -Autocorrelation Function
- **PACF** -Partial autocorrelation function
- **AR1** -Auto Regressive order 1
- **MLE** -Maximum Likelihood Estimation
- **MCMC** -Markov chain Monte Carlo

- **INLA** -Integrated Nested Laplace Approximations
- **DIC** -Deviance Information Criterion
- **GLM** -Generalised Linear Models

CHAPTER 1: INTRODUCTION AND BACKGROUND

1.1 Background

1.1.1 Burden of Malnutrition

Malnutrition is a vital underlying cause of death and accounts for 45% of childhood mortality and morbidity by increasing susceptibility to major infections like diarrhoea and pneumonia (1,2). Among hospitalised children with malnutrition, in the absence of suitable treatment, the case-fatality rates range from 30% to 40% (1,3,4). Majorly, it affects childhood development and increases the risk of other diseases in adulthood. These malnutrition related outcomes can significantly affect the socio-economic productivity in their adult life (5). Sub-Saharan Africa and Asia remain the areas with the highest prevalence of malnutrition. Specifically, in Sub-Saharan Africa, malnutrition has been reported to be among the leading risk factors in populations health (1,2,6).

In East and Southern Africa, research estimates have shown that approximately 40%, 30% and 6% of children under five are stunted, underweight and wasted respectively (7). In patients with multiple infections and malnutrition have been shown to complicate the management

of patients, therefore leading to a higher case-fatality in hospitals (3,8). Malnutrition remains a major contributor to inpatient morbidity and mortality among children in rural areas in Kenya, despite the incentives to overcome malnutrition. Malnutrition has also been associated with increased inpatient costs and increased risk of inpatient mortality (3,9,10).

In Kenya, considerable geographic disparities of the food insecure population have been reported (11). Food inavailability has been shown to increase the risk of malnutrition potentially. Nonetheless, malnutrition is the result of a much broader range of risk factors. The UNICEF malnutrition conceptual framework (Appendix 7) shows a wide spectrum of determinants that worsen malnutrition leading to a vicious cycle (2).

1.1.2 Spatial-temporal analysis in sub-Saharan Africa

Spatial-temporal Bayesian models have been applied in different health research studies especially in malaria and other infectious diseases in sub-Saharan Africa (SSA). Spatial models in SSA have importantly contributed to understanding the spread of malaria and a potential role for alerts on epidemics. According to a review done by Gebreslasie et al., he describes the importance of spatial models in the study of epidemiology and the transmission risk of malaria in Africa (12).

Multivariate and Semi-Parametric Bayesian spatial models have also been applied to understand the risk and spatial variation of HIV in Africa (13,14). This modelling approach of spatial multivariate and semi-parametric models shows the adaptability and viability of spatial models in understanding HIV in sub-Saharan Africa (SSA) (13,14).

Malnutrition has been shown to be heterogeneous spatially although very little of this work has been done in sub-Saharan Africa. Lack of detailed spatial data at administrative and homestead levels has been a major limitation for residence centred investigation of malnutrition. According to a review of SSA spatial analysis on malnutrition, most of them applied spatial regression models at a meso level, but few studies used the Bayesian spatial modelling approach (15).In Somalia, Kinyoki et al. observed a clear seasonal variation in wasting in under 5s due to variations in climate, food security and diseases. Bayesian hierarchical space–time models using stochastic partial differential equation (SPDE) was used in Somalia (16).

Malnutrition being a compounded phenomenon, using joint spatial-temporal models can give a better understanding of this issue. Spatial analysis allows health research to integrate health, environmental and population data. This allows them to investigate the relationship between these

factors at different scales (17).

1.1.3 Advantages of Bayesian spatial models

The main advantage of Bayesian modelling approach is the ability to incorporate prior information on the parameters used in the models. This method offers a better way over the classical regression models by allowing concurrent modelling of spatial-temporal autocorrelation while still estimating the usual fixed effects. In spatial models, Bayesian models over the frequentist modelling approach make it possible to include the intrinsic Gaussian CAR prior distribution to represent the shared spatial components. Modelling of shared spatial components using the frequentist approach can also be cumbersome and computer intensive (13,18,19).

1.2 Problem statement

Malnutrition remains a serious challenge in public health and has been linked to an increased number of deaths and morbidity of children. This accounts for approximately 50% of global deaths in under-fives, which enhances the risk of infectious diseases. Conversely, infections aggravate malnutrition by different reasons leading to increased number of admissions

and a longer length of hospital stay (1,6,20,21). Malnutrition has also been reported to interact with various components of the environment potentially increasing the risk of malnutrition and low birth weight (22,23).

Shortage of data on detailed geographic, temporal, household and homestead level is a major setback in research on malnutrition related morbidity and association with potential risk factors (15).

1.3 Justification of the study

According to a review done by Marx et al., little research has been done on spatial-temporal malnutrition attributable mortality and morbidity. Some studies have associated this with a lack of detailed geographic household and homestead level data (15,24). Non-bayesian based methods are the most commonly used approaches to the analysis of malnutrition attributable morbidity and mortality, and they are limited in the manner they handle large data set, temporal and spatial random effects (25,26).

It is now feasible to map malnutrition attributable mortality and morbidity status at high spatial resolutions using Bayesian approach due to the availability of georeferenced data. Additionally, recent developments and accessibility of modern statistical tools together with advances in

computing speeds offer opportunities for fitting spatial-temporal models that converge quickly without loss of predictive precision (27). This will provide a better understanding of space-time variation of morbidity and mortality associated with malnutrition to improve the prevention, early detection and outcome of diseases in malnourished children.

1.4 Literature review

1.4.1 Child malnutrition morbidity and mortality studies

Currently, Asia Latin America and sub-Saharan Africa hosts two-thirds of the world's malnourished children, and it is responsible for over 50% of under-five mortality (1,28). In a study done in a rural area in Kenya, malnutrition accounted for a higher rate of inpatient morbidity and mortality. Bejon et al. concluded that despite the great effort to fight malnutrition, the malnutrition attributable fraction for hospital mortality was over 50% and not a significant change noted in East Africa since 1980's. Middle Upper Arm Circumference (MUAC) was a better marker for malnutrition compared to the other conventional clinical definitions; wasting and stunting in this study (3).

In a longitudinal study conducted in East rural Ethiopia among under-5

years, wasting was observed to occur during the dry season though seasonality was not a significant covariate. This could have been a limitation of using a fixed effect model as stated in their paper to avoid variable omission bias (28). Enhanced vegetation index (EVI) was highly associated with all the malnutrition indicators according to a study done in Somalia. EVI unit increase in Somalia accounted for over 30% reduction in all the indicators (29). This study concluded that infections and seasonal variations are key drivers of malnutrition, but a spatial-temporal study similar to what we propose would help understand this better.

In Malawi, spatial patterns of childhood morbidity determined by Bayesian geo-additive models similar to those we propose revealed some urban agglomerations to be associated with a higher childhood morbidity, especially diarrhoea. Other significant factors were environmental, climatic factors and remoteness in some rural areas. Modelling was also controlled for geographic factors (30).

1.4.2 Spatial-temporal modelling utilisation in epidemiology

In a few studies done in Africa and other studies elsewhere, the available evidence suggests that malnutrition is spatially heterogeneous (30,31). A Malawi study suggested more emphasis be placed upon childhood

morbidity association with remoteness and geographic variation (30). In Brazil, malnutrition was associated with primary diagnosis at admission and was shown to be higher in different regions (20). Non-random clustering of infant mortality due to malnutrition and diarrhoea deduced from verbal autopsy (VA) was observed in rural South Africa (32).

Geographical variability has also been seen in hospital admissions using non-separable spatial-temporal models(33). Since malnutrition amplify common life-threatening infections that lead to admissions, understanding the spatial and temporal patterns of the childhood malnutrition attributable mortality and morbidity can be of great importance (15,31). Developing models that will help better estimate the impacts of geographical variation on population changes in health has been recommended by various authors (23,34,35). Bayesian space–time geostatistical models and structured additive logistic regression models have been used to predict and estimate the risk factors of malaria in sub-Saharan Africa respectively (36,37).

Noor et al. applied the Bayesian spatial-temporal models to show how a reduction in transmission of malaria has been achieved and also predicted the malaria burden in SSA for the period between 2000 and 2010 (36). This

is complimented by Chirombo et al. using structured additive models which showed similar results in Malawi with second order random walk priors for their continuous variables (37).

Bayesian Poisson zero-inflated models using Integrated Nested Laplace Approximation (INLA) approach for inference were used to identify protective factors for HIV/TB child mortality (38). This was applied on the mortality outcome variable, which had a Poisson distribution similar to the morbidity outcome in this study. Inverse probability weighting combined with prediction models were used to obtain the statistics and spatial epidemiology of the prevalence of HIV in Malawi (13). This approach applied Bayesian multivariate models to jointly model and map the HIV risks (13). In Kenya, a spatial variation of HIV infections was observed using Bayesian Spatial Semi-Parametric model showing the adaptability and viability of spatial models in understanding HIV in SSA (14).

In Somalia, a Bayesian approach was applied to understand the spatial-temporal distribution of wasted children. The study showed a seasonality patterns in the prevalence of wasting recommending consideration of seasons implementation of nutrition programs (16). Ricardo et al., identified clusters of severe anaemia and malnutrition were

identified as one of the factors contributing to the spatial heterogeneity of anaemia risk (39).

1.4.3 Child malnutrition morbidity and mortality determinants

Agricultural production, food prices, food availability and food security remain important determinants of malnutrition in sub-Saharan Africa. However, while malnutrition exacerbates common life-threatening infections, these infections themselves cause growth faltering and without a sufficient recovery period can worsen malnutrition (1,3,6,8,23). Thus, diseases, especially diarrhoea, are a key driver of malnutrition. Recently, it has also become clear that chronic intestinal inflammation, known as environmental enteric dysfunction (EED), is widespread in the developing countries and is an important mediator of stunting (1,40).

1.4.4 Hospital based models

Kazembe et. al. used semiparametric regression model with data from Zomba district hospital to show the risk of dying in hospital is lower in the dry season but high for those reffered to the hospital. Significant difference was observed in both structured and unstructured spatial effects.The health facility effects revealed considerable differences by type of facility

or practice offered (41).In Tanzania, prompter care-seeking behaviour , improved quality of care at health facilities and better adherence to treatment was reccommended after analysing hospital data to understand the malaria attributabke mortality (42). This shows the importance of geo referenced data in informing important epidemiological policies and design of interventions.

Most of the malnutrition studies have used survey or household data.Therefore, we propose to map the mortality attributable to malnutrition and malnutrition related morbidity using the available geo-referenced household data for admitted children in the paediatric ward at county Hospital in Kilifi and their non-admitted counterparts in the Kilifi Health and Demographic Survey System (KDHSS) (43). This should better our understanding of the variation of malnutrition attributable mortality and malnutrition related morbidity.

1.5 Research question

Is there a variation in malnutrition attributable mortality and malnutrition related morbidity space-time patterns among children admitted in Kilifi county Hospital from 2002-2015, Kenya?

1.6 Aim and objectives of the study

1.6.2 Study aim

To investigate the variability in childhood morbidity and mortality attributable to malnutrition in Kilifi county between 2002 and 2015 using space-time models.

1.6.3 Study objectives

a) To describe the spatial-temporal trends of malnutrition related morbidity and mortality in Kilifi county, Kenya between 2002 and 2015.

b) To investigate the space-temporal variations in malnutrition related morbidity in Kilifi county, Kenya between 2002 and 2015.

c) To investigate the spatial variations of impact factors for malnutrition attributable mortality patterns in Kilifi county, Kenya between 2002 and 2015.

CHAPTER 2: METHODOLOGY

Introduction

This chapter describes the study site, design and population of the data used for this research project. Additionally, we describe the data collection and management process for the different statistical approaches used to analyse the data. The analysis was done in three main phases, exploratory data analysis, temporal investigation and spatial-temporal models. Time series models were used for investigating temporality. Moran I and Kulldorff statistics that were used to investigate spatial auto correlation, hotspots and coldspots respectively are introduced for the exploratory analysis. Finally, we present Bayesian spatial-temporal negative binomial model and geo-additive logistic model for the inferential statistics. In this analysis we used WHZ and oedema as our malnutrition defining variables since the children were aged between 6 months to 15yrs and MUAC is a reliable malnutrition marker for children below 60 months (3,44).

2.1 The Kilifi Health and Demographic Surveillance System (KHDSS)

The KHDSS was launched in 2000 in Kilifi County which is located in the Coastal region of Kenya as shown on Figure~2.1. The area is a semi-arid area with subsistence farming being the main economic activity for the population in this region. The study area was mainly set to define rates of mortality, migration and fertility. KHDSS has 15 administrative locations and 40 sub-locations. Special 4-monthly visits are done to the households to record any death, migration and birth event.

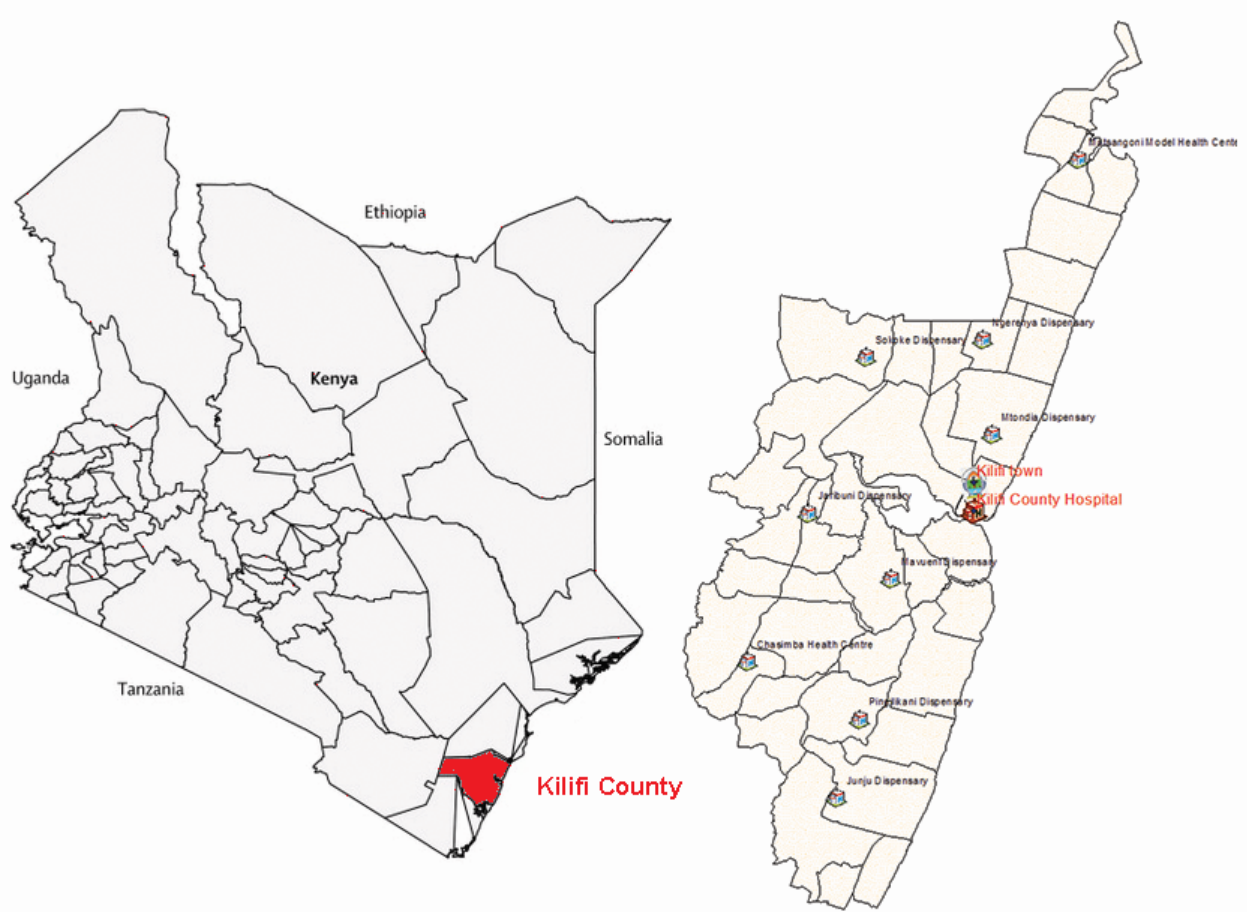


Figure 2.1: The location of the KHDSS and the different dispensaries and the main Hospital

We utilise data from the Kilifi County Hospital (KCH) paediatric ward admissions. KCH is the main referral hospital in this region, and 80% of the admissions come from KHDSS. The hospital is located in the middle of KHDSS which was launched in 2000 to capture the majority of the inpatient details at KCH paediatric ward. The area has approximately 280,000 residents with almost 4000 paediatric admissions per year. The mortality and morbidity events at the hospital are synchronised real time with the population register updated by 4-monthly visits to the households (43).

2.2 Study Design and Data Management

Our study design was a retrospective cohort observational study. To understand malnutrition related morbidity, our outcome was the number of malnutrition related admissions over the period of 2002 to 2015. Our cases for the mortality were children who had a death outcome and malnutrition as the main predictor variable (3).

The mortality and morbidity events captured at the hospital are integrated with the population register updated by 4-monthly visits to the households. The linkage of the surveillance data and admissions data is done real time with the matching of individuals at the point of admissions by the field

workers. The data is then de-identified by assigning a unique identifier and a corresponding person identifier for the KHDSS residents. A MySQL 5.7 database updated by the medical officers at KCH using a web based PHP 5.3.19 application. The medical officers update the medical background of each admission with a unique identifier and a corresponding person identifier for the KHDSS residents (43,45,46).

2.2.1 Data Quality

A team of qualified field workers trained on using the KHDSS surveillance system do the matching of the data on real-time admissions. Each field worker has a unique username to access the web based system and the matching procedures of the individuals is explained elsewhere (43). Qualified medical and clinical officers enter the history and clinical examination of the patients in the system after matching has been done by the field workers. The data quality checks on the clinical measurements are implemented ensure values entered into the system are within the expected range. The database has a daily backup, and each event or record entered into the system has a unique event identifier and an audit trail.

2.3 Study Population

The study utilised data of children admitted in KCH from 2002 to 2015 and aged between 3 months and 13 yrs who live in KHDSS and were health system users.The highest number of admission events was 20, but for the morbidity outcome, we selected up to 11 admission events. This was considered in the modelling approach.

2.3.1 Inclusion Criteria

We included children who lived within the KHDSS and were admitted in Kilifi county hospital between 2002-2015. The age limit for the children was 3 months to 13 years. The household geolocations for the participant was a requirement to be included in the analysis.

2.3.2 Exclusion Criteria

We excluded children with unknown discharge outcome and the malnutrition status was unknown at admission.Trauma admission events were also excluded from the analysis. Children with missing person identifier were excluded from the analysis.

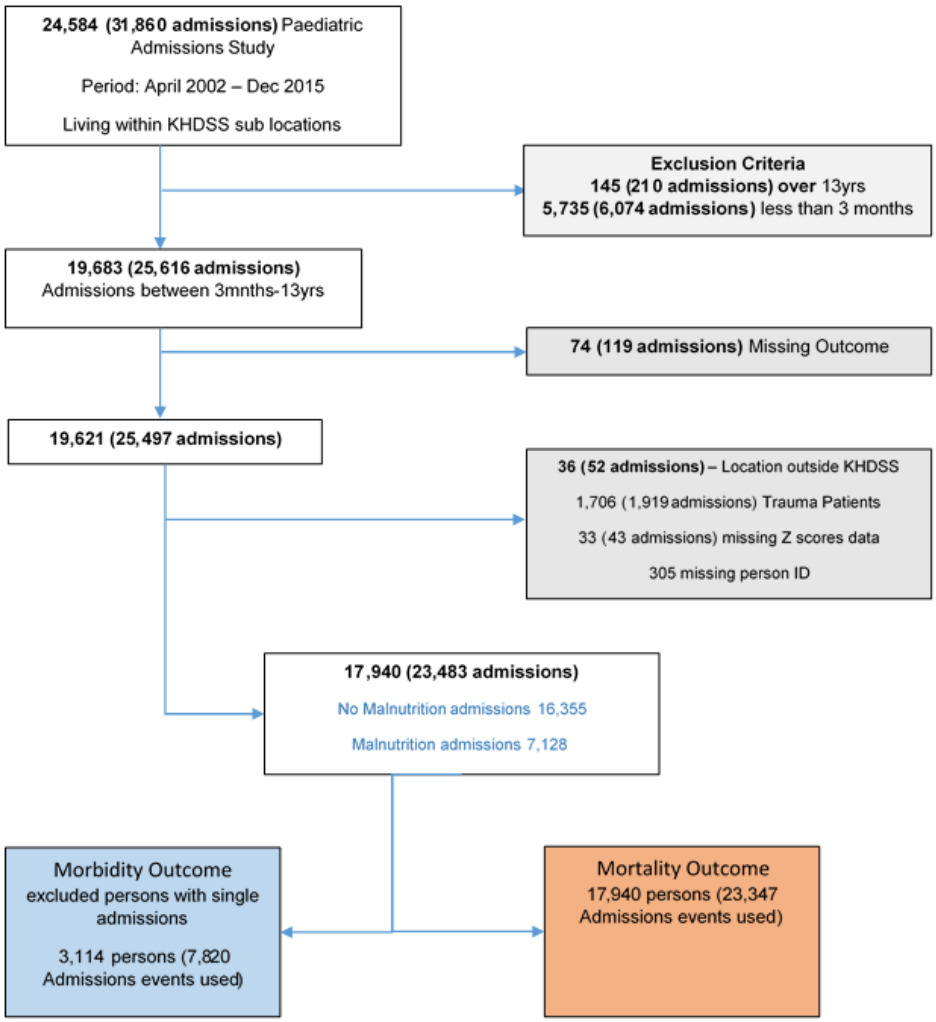


Figure 2.2: Data flow diagram

2.4 Outcome and Explanatory Variables

The outcome variables used to define a malnutrition related admission were Weight for Height Z-score (WHZ) for children under-5 years or BMI for age Z-score for children over 5 years and the presence of nutritional oedema

at admission. The Weight for Height Z-score was calculated using the 2006 WHO child growth standards (47) and the BMI for age Z-score was calculated using WHO growth reference for school-aged children (48). The weight of the children during admission was done using an electronic scale (Weylux; H Fereday and Sons, London, United Kingdom). The heights of the children were measured with a wall-mounted scale except for children with less than two years whose length was measured using a calibrated board. Wasting which is low WHZ (WHZ < -2) or low BMI for age Z-score (BMIZ <-2) is one of the key indicators of malnutrition and is related to illness or food insecurity (49). Oedema is a clinical sign of undernutrition which may include swelling of the feet, skin and hair changes (3). Nutrition oedema is recorded by the medical officers at KCH during admission.

Child severe illness was defined as a child admitted with either gastroenteritis, LRTI, blood and CSF culture positive, malaria and fever or meningitis (3). Other factors affecting malnutrition related morbidity and mortality were selected using the UNICEF malnutrition conceptual framework guidance (2). The conceptual framework defines the immediate causes of undernutrition to be individual related, i.e., diseases and inadequate dietary intake and the underlying causes to be the environmental and household factors as shown in Appendix 7.

Additional covariates that are associated with malnutrition morbidity were identified by a stepwise analysis based on a generalised linear regression model and also those that have a biologically plausible relationship (2,50). The environmental predictor variables were extracted from the Moderate Resolution Imaging Spectroradiometer (MODIS) using the MODISTools and MODIS package in R version 3.3.2 (51–53). To extract values from the Kilifi EVI and Rainfall Raster files, the values interpolated from the values of the four nearest raster cells of each admission coordinate provided (54). The immediate and underlying factors of undernutrition were considered in this analysis.

2.5 Exploratory Data Analysis

The summary statistics, frequency tables of the demographic characteristics were done using the repeated measures tabulation and frequency statistics.The summaries are created from the panel data of the number of admissions and the monthly time variable. Continuous variables are summarised with sample size (n), overall and between means, between, within and overall standard deviations of the panel data. Categorical variables are summarised as counts and percentages of overall, between and within components of the panel data.

2.5.1 Temporal Exploratory Data Analysis

To explore the temporal patterns, the data were defined as a regular time series data using monthly time points of counts of malnutrition related admissions and mortality. A line plot of the time series data was done for the period 2002-2015. Augmented Dickey-Fuller unit-root was used to test for stationarity; the null hypothesis is that the series has no root or there is no trend in the model. The stationary series is then examined on the Auto-Correlation Function (ACF) for the MA lags, and Partial Auto-Correlation Function (PACF) for the AR lags. This was followed by fitting Autoregressive integrated moving average (ARIMA) and seasonal ARIMA models (51,55). The Seasonal ARIMA is useful where a time series data shows a periodic pattern that reoccurs with about the same intensity in a year. This makes it adequate for the admissions data since admissions in a hospital seems to follow a pattern.

The time series orders autoregression, integration and moving-average of 1,0,1 respectively were selected using the table of the autocorrelations with Bartlett's statistic (51). Seasonality was determined using the same approach with pointwise confidence intervals also based on Bartlett's formula and selected the peaks.

In general the seasonal autoregressive moving average model is defined as;

$$\text{ARIMA} \quad \underset{\text{Non seasonal}}{(p,d,q)} \quad \underset{\text{seasonal}}{(P,D,Q)_s}$$

where s is the number period per season.SARIMA is an $ARIMA\,(p,d,q)$ model whose residuals at ℓ_t are $ARIMA\,(P,D,Q)$.

In general SARIMA model with period s is defined as;

$$\underset{\text{NS AR(1)}}{(1-\varphi_p L)}\underset{\text{S AR(1)}}{(1-\theta_1 L^s)}\underset{\text{NS D}}{(1-L)}\underset{\text{SD}}{(1-L^s)}_{yt} = \underset{\text{NS AR(1)}}{(1+\phi_p L)}\underset{\text{S AR(1)}}{(1+\Theta_1 L^s)}_{\ell_t} \qquad \text{(Eqn:1)}$$

NS - Non-seasonal, AR - Auto Regressive, S- Seasonal, D -Difference

By definition $ARIMA\,(p,d,q)\,(P,D,Q)_s$ is fitted as $\Theta\left(L^s\right)\varphi\left(L\right)\left(1-L\right)^d\left(1-L^s\right)$ and ${}^{D}Y_t = \Theta\left(L^s\right)\theta(L)_{\varepsilon_t}$ where L is the lag operator defined as $L^k = \frac{Y_{t-k}}{Y_t}$, $\varphi\left(L\right) = 1-\varphi_1\left(L^1\right)-\varphi_2\left(L^2\right)-$ ..... $-\,\varphi_p\left(L^p\right)$ is the polynomial function of order p with a set number of

vector coefficients $\underset{\sim}{\varphi}$. $\theta(L) = 1 - \theta_1(L^1) + \theta_2(L^2) + ... + \theta_q(L^q)$ is a moving average polynomial of order q with a vector of coefficients $\underset{\sim}{\theta}$. $\Theta\left(L^S\right) = 1 - {}_{\varphi_{S,1}}L^S - {}_{\varphi_{S,2}}L^{2S} - .... - {}_{\varphi_{S,P}}L^{PS}$ and $\Theta\left(L^S\right) = 1 - {}_{\varphi_{S,1}}L^S - {}_{\varphi_{S,2}}L^{2S} - .... - {}_{\varphi_{S,Q}}L^{QS}$ are seasonal parameters of order P and Q respectively that meet the stationarity assumption. d is the number of differencing passes needed to stationarize the series, D is the number of seasonal differences,and ε_t are the error terms with mean zero and variance σ^2 (56,57).

The Portmanteau test is used to confirm the significane of the seasonality as a covariate. The Portmanteau test null hypothesis is there is no serial correlation using the white noise under the $Chi2$ statistic (58). Our model was defined as; SARIMA$(p,d,q)(P,D,Q)_s$, where s is the number of period per season. This was defined as SARIMA(1,0,1) $(1,0,1)_s$ and modelled as $\mu_t = \rho(y_{t-1}) + \theta_{\varepsilon_{t-1}} + \varepsilon_t$ where ρ and θ are the products of the seasonal parameters identified.

2.5.2 Spatial Exploratory Data Analysis

Maps of the prevalence of children admitted with malnutrition and the patients who died were done. The Global Moran I index was used

to calculate the measure of sublocations spatial autocorrelation or randomness or dispersion on malnutrition admission. A distance matrix of the sub-locations centroids was used to calculate the Morans I index (59,60). The Moran's I index was calculated based on the large sample theory distribution of the index, which states that as the sample size increases the indexs tends to a normal distribution.

$$I = \frac{n}{W} \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} w_{i,j} z_i z_j}{\sum_{i=1}^{n} z_i^2} \qquad \text{(Eqn:2)}$$

Where z_j is the deviation of a malnutrition admission from a sub-location $i(z_i = Y_i - \widetilde{Y})$, $w_{i,j}$ is the spatial weight between locations i and j and $W = \sum_{i=1}^{n} \sum_{j \neq i}^{n} w_{i,j}$ is the aggregate of all spatial weights. The Moran I index is interpreted depending on the three outcomes; $I > E(I)$ shows a positive autocorrelation implying clustering and $I \approx 0$ shows no spatial auto correlation and $I < E(I)$ shows a negative autocorrelation implying a dispersed neighbouring value, where $E(I)$ is the normally distributed population mean (38,61).

Local heterogeneity and clustering of malnutrition attributable mortality and morbidity were determined using SaTScan software to observe the hot and cold spots in Kilifi County. SaTScan applies the Kulldorff spatial scan

statistic which imposes a circular window and calculates the likelihood of observing the events inside and outside the study area. The circle with the maximum likelihood is defined as the most likely cluster (62,63).

2.6 Inferential Statistics

2.6.1 Spatial-Temporal Negative Binomial Model

Our outcome was defined as the counts of malnutrition related admissions until the person with the highest number of malnutrition related admissions. Secondly, the variance of the malnutrition related admissions was higher than the mean, so a negative binomial distribution fitted our count data well (25,32,64). The morbidity outcome data followed a negative binomial distribution; which is a generalisation of the Poisson distribution used to provide better epidemiological estimates of factors associated with malnutrition morbidity (25). The negative binomial spatial-temporal model was applied to achieve objective 2; to identify the spatial and temporal pattern of malnutrition admissions in Kilifi (64). We included children with at least two admissions events for the spatial-temporal negative binomial model due to computational requirements(65). The readmissions would also help investigate vulnerability due to malnutrition morbidity.

Model Definition

The definition of the negative binomial spatial model is

$$y_{ijt}|\,\psi_{ijt},\underset{\sim}{\Omega} \sim\ \pi\left(y_{ijt}|\,\psi_{ijt};\underset{\sim}{\Omega}\right) \quad \text{(Eqn:3)}$$

For $i = 1, 2, 3, ..., n_t = 3114$, individuals data , $j = 1, 2, 3, ..., q = 40$, are the sub locations and $t = 1, 2, 3, ..., p = 11$ is the admission visit number. $\underset{\sim}{\Omega}$ (used as Ω thereafter) is a vector of parameters to be estimated and $\psi_{ijt}$is the linear predictor.

Where $\underset{\sim}{y} = [y_{ijt}]_{n\times p}$ are the observed admission data which follow a negative binomial distribution of the form $y_{ijt} \sim\ NegBinomial(p_{ijt}, r_{ijt})$ (66).

$$\text{p(y} = [y_{ijt}]\,|p, r) = \begin{pmatrix} y_{ijt}-1 \\ r-1 \end{pmatrix} p^r(1-p)^{y_{ijt}}; y \geq r, r = 1, 2, 3...; 0 < p < 1 \quad \text{(Eqn:4)}$$

where $r = \frac{1}{\theta}$ and $p_{ijt} = \frac{1}{1+\theta\mu_{ijt}}$.

Using our data and y is the number of admissions from each individual i,r

is the number of successful malnutrition admissions from sub-location j at admission t, and p is the probability of a malnutrition admission.

The conditional mean is $E\left(y_{ijt}|p,r\right)=\frac{r(1-p_{ijt})}{p_{ijt}}$ and conditional variance var$(y_{ijt}|p,r)=\frac{r(1-p_{ijt})}{p_{ijt}^2}$.

Taking the natural logs from Eqn:4 above,

$$=\ln\begin{pmatrix}y_{ijt}-1\\r-1\end{pmatrix}+r\ln\left(p_{ijt}\right)+y_{ijt}\ln(1-p_{ijt}) \qquad \text{(Eqn:5)}$$

And in general the exponential family form of a negative binomial is expressed as a member of the generalized linear model which has a link function and a cumulant as shown below;

$$f\left(y;p,r\right)=\exp\left\{y_{ijt}\underbrace{\ln\left(1-p_{ijt}\right)}_{\text{link function}}+r\underbrace{\ln\left(p_i\right)}_{\text{cumilant}}+\ln\begin{pmatrix}y_{ijt}-1\\r-1\end{pmatrix}\right\} \qquad \text{(Eqn:6)}$$

Since

$$\begin{pmatrix}y_{ijt}-1\\r-1\end{pmatrix}=\frac{\Gamma\left(y_{ijt}\right)}{\Gamma\left(y_{ijt}-r\right)\Gamma\left(r\right)}$$

from $\Gamma(x) = (x-1)!$ and $\begin{pmatrix} x \\ m \end{pmatrix} = \frac{x!}{m!(x-m)!}$ and using $r = \frac{1}{\theta}$ and $p_{ijt} = \frac{1}{1+\theta\mu_{ijt}}$ from Eqn:4 above, and replace in Eqn:5 , this is gives

$$= \ln\left(\frac{\Gamma(y_{ijt})}{\Gamma\left(y_{ijt}-\frac{1}{\theta}\right)\Gamma\left(\frac{1}{\theta}\right)}\right) + \frac{1}{\theta}\ln\left(\frac{1}{1+\theta\mu_{ijt}}\right) + y_{ijt}\ln\left(1-\frac{1}{1+\theta\mu_{ijt}}\right) \quad \text{(Eqn:7)}$$

Then after taking the exponents, this can be expressed as;

$$= \exp\left\{\ln\Gamma(y_{ijt}) + y_{ijt}\ln\left(\frac{\theta\mu_{ijt}}{1+\theta\mu_{ijt}}\right) - \ln\Gamma\left(y_{ijt}-\frac{1}{\theta}\right) + \frac{1}{\theta}\ln\left(\frac{1}{1+\theta\mu_{ijt}}\right) - \ln\Gamma\left(\frac{1}{\theta}\right)\right\} \quad \text{(Eqn:8)}$$

From the marginal of the joint distribution of the Poisson-Gamma, the Negative-Binomial distribution (66) can be expressed as follows;

$$= \exp\left\{c_0(\mu_{ijt},\theta) + \ln\Gamma(y_{ijt}) - \ln\Gamma\left(y_{ijt}-\frac{1}{\theta}\right) - \ln\Gamma\left(\frac{1}{\theta}\right)\right\} \quad \text{(Eqn:9)}$$

Where

$$c_0(\mu_{ijt},\theta) = y_{ijt}\ln\left(\frac{\theta\mu_{ijt}}{1+\theta\mu_{ijt}}\right) + \frac{1}{\theta}\ln\left(\frac{1}{1+\theta\mu_{ijt}}\right)$$

$$c_0\left(\mu_{ijt},\theta\right)=y_{ijt}\ln\left(\frac{\theta\mu_{it}}{1+\theta\mu_{it}}\right)-\frac{1}{\theta}\ln\left(1+\theta\mu_{it}\right)$$

Where $\mu_{ijt}=\exp\left(\psi_{ijt}\right)=\log\left(E_{ijt}\right)+x_{ijt}\underset{\sim}{\beta}+\underset{\sim_j}{\phi}+\underset{\sim_j}{\vartheta}+\underset{\sim_t}{\gamma}+\varepsilon_{ijt}$ with E_{ijt} as the ages of the children and is the exposure time variable, x_{ijt} are the covariates design matrix, $\underset{\sim}{\beta}$ vector of fixed coefficients. And the INLA modelled latent variables for structured $(\underset{\sim_j}{\phi})$ and unstructured $(\underset{\sim_j}{\vartheta})$ space and time $(\underset{\sim_t}{\gamma})$

Our spatial-temporal model was fitted assuming the age of the child as the exposure variable and was fitted as follows

$$\ln(\mu_{ijt})=x_{ijt}\underset{\sim}{\beta}+\underset{\sim_j}{\phi}+\underset{\sim_j}{\vartheta}+\underset{\sim_t}{\gamma}+\varepsilon_{ijt}+\log(E_{ijt}) \qquad \text{(Eqn:10)}$$

$$\mu_{ijt}=\exp\left(x_{ijt}\underset{\sim}{\beta}+\underset{\sim_j}{\phi}+\underset{\sim_j}{\vartheta}+\underset{\sim_t}{\gamma}+\varepsilon_{ijt}\right)\times E_{ijt} \qquad \text{(Eqn:11)}$$

Under maximum likelihood estimation, the negative binomial form was expressed as (32,66)

$$L\left(\mu_{ijt}|y_{ijt},\theta\right)=$$
$$\prod_{i=1}^{n_t}\exp\left\{c_0\left(\mu_{ijt},\theta\right)+\ln\Gamma\left(y_{ijt}-\frac{1}{\theta}\right)-\ln\Gamma\left(y_{ijt}\right)-\ln\Gamma\left(\frac{1}{\theta}\right)\right\} \qquad \text{(Eqn:12)}$$

The log likelihood is obtained by taking the log of the likelihood; Eqn:12

$$\ell\left(\mu_{ijt}|y_{ijt},\theta\right) = \exp\left[\sum_{i=1}^{n_t}\left\{c_0\left(\mu_{ijt},\theta\right)+\ln\Gamma\left(y_{ijt}-\frac{1}{\theta}\right)-\ln\Gamma\left(y_{ijt}\right)-\ln\Gamma\left(\frac{1}{\theta}\right)\right\}\right] \quad \text{(Eqn:13)}$$

Considering values with the parameters , substituting μ_{ijt} with the linear predictor and $c_0\left(\mu_{ijt},\theta\right)$, the negative binomial log-likelihood from Eqn:13 the model coefficients can be expressed as

$$\propto \exp\left[\sum_{i=1}^{n_t}\left\{y_{ijt}\ln\left(\frac{\theta\exp(\psi_{ijt})}{1+\theta\exp(\psi_{ijt})}\right)-\frac{1}{\theta}\ln\left(1+\theta\exp(\psi_{ijt})\right)\right\}\right] \quad \text{(Eqn:14)}$$

Thus our model from Eqn:6 can be implemented using GLM with the link function being (66–68).

$$y_{ijt}\ln\left(\frac{\theta\exp(\psi_{ijt})}{1+\theta\exp(\psi_{ijt})}\right)$$

Since the prior distribution of μ_{ijt} is a multivariate normal with a mean zero and a $\Sigma_{q\times q} = \Sigma_{40\times 40}$ variance matrix, $\mu_{ijt} = MVN\left(0,\Sigma_{40\times 40}\right)$. The likelihood contribution for the j^{th} sub location is obtained by integrating μ_{ijt} out of the joint probability density $f\left(y_{ijt}|\mu_{ijt},\theta\right)$

$$L\left(\Omega,\sum,\theta\right)=\left(2\pi\right)^{-\frac{q}{2}}\left|\sum\right|^{-\frac{1}{2}}\int f\left(y_{ijt}\middle|\mu_{ijt},\theta\right)\exp\left(\mu_{ijt}^{'}\sum^{-1}\frac{\mu_{ijt}}{2}\right)d\mu_{ijt} \quad \text{(Eqn:15)}$$

The above equation (Eqn:15) has no closed form, and thus approximation method is used for Maximum likelihood estimation.

The spatial-temporal model with Bayesian approach estimation was defined as follows

$$\text{Posterior}\left[p(\text{parameters}\middle|\text{data})\right]\propto\text{Likelihood}\times\text{Priors}$$

The posterior likelihood of our malnutrition admission data is defined as;

$$L\left(\Omega\equiv\left\{\underset{\sim}{\beta},\underset{\sim_j}{\phi},\underset{\sim_j}{\vartheta},\underset{\sim_t}{\gamma}\right\}\middle|y_{ijt}\right)=\prod_{i=1}^{n}p\left(y_{ijt}\middle|\Omega\right)\times p\left(\underset{\sim}{\beta},\underset{\sim_j}{\phi},\underset{\sim_j}{\vartheta},\underset{\sim_t}{\gamma}\right) \quad \text{(Eqn:16)}$$

The full conditional for our model is expressed as;

$$p\left(\Omega\middle|y_{ijt}\right)\propto L\left(y_{ijt}\middle|\Omega\right)\times p\left(\underset{\sim_k}{\beta}\right)\times p\left(\underset{\sim_j}{\phi}\middle|\tau_c\right)\times p\left(\underset{\sim_j}{\vartheta}\middle|\tau_h\right)\times p\left(\underset{\sim_t}{\gamma}\middle|\tau_e\right) \quad \text{(Eqn:17)}$$

$$= L\left(y_{ijt}|\Omega\right)\times p\left(\underset{\sim k}{\beta}\right)\times p\left(\underset{\sim j}{\phi}|\tau_c\right)\times p\left(\tau_c\right)\times p\left(\underset{\sim j}{\vartheta}|\tau_h\right)\times p\left(\tau_h\right)\times p\left(\underset{\sim t}{\gamma}|\tau_e\right) \quad \text{(Eqn:18)}$$

The prior for the beta coefficients for $k-1$ prdictors in the model is assumed to be $\underset{\sim k}{\beta} \overset{iid}{\sim} N\left(\mu_\beta, {\sigma_\beta}^2\right)$ therefore

$$p\left(\beta_k\right) = \frac{1}{\sqrt{2\pi\sigma_\beta^2}}\exp\left[-\frac{1}{2}\left(\frac{\beta_k-\mu_\beta}{\sigma_\beta}\right)^2\right] \quad \text{(Eqn:19)}$$

Spatial Components:

The unstructured random effects $p\left(\underset{\sim j}{\vartheta}|\tau_h\right)$, where $\vartheta_j \overset{iid}{\sim} N\left(0, \frac{1}{\tau_h}\right)$. Therefore

$$p\left(\vartheta_j|\tau_h\right) = \frac{1}{\sqrt{\frac{2\pi}{\tau_h}}}\exp\left[-\frac{1}{2}\left(\frac{\vartheta_j-0}{\frac{1}{\sqrt{\tau_h}}}\right)^2\right] \quad \text{(Eqn:20)}$$

with a Gamma function prior used $\tau_h \sim Gamma\left(\alpha_h, \beta_h\right)$ expressed as

$$p(\tau_h) = \frac{\left(\beta_h\right)^{\alpha_h}}{\Gamma\left(\alpha_h\right)}\tau_h^{\alpha_h-1}\exp(-\beta_h\tau_h), \alpha_h > 0; \beta_h > 0 \quad \text{(Eqn:21)}$$

Therefore the unstructured random effects prior in BUGS is expressed as;

$$p\left(\vartheta_j|\tau_h\right) \times p\left(\tau_h\right) \propto \tau_h^{\alpha_h-1} \exp(-\beta_h\tau_h) \times \exp\left[-\frac{1}{2}\left(\frac{\vartheta_j - 0}{\frac{1}{\sqrt{\tau_h}}}\right)^2\right] \quad \text{(Eqn:22)}$$

Similarly in INLA the unstructured random effects $\underset{\sim}{\vartheta}_j = f_{spat\_unst}\left(s_j\right)$ scaled with parameter τ_h.

In BUGS, a CAR prior is used for the structured spatial random effects,a CAR prior is used for the structured spatial random effects (19) $\underset{\sim}{\phi} \sim CAR\left(\tau_c\right)$ and a CAR prior given by prior is given as $\phi_j|\phi_i, j \neq i, \tau_c \sim N\left(\frac{\phi_j}{\bar{\phi_j}}, \frac{1}{\tau_c m_j}\right)$ therefore

$$p\left(\phi_j|\tau_c\right) = \frac{1}{\sqrt{\frac{2\pi}{\tau_c m_j}}} \exp\left[-\frac{1}{2}\left(\frac{\phi_j - \frac{\phi_j}{\bar{\phi_j}}}{\sqrt{\frac{1}{\tau_c m_j}}}\right)^2\right] \quad \text{(Eqn:23)}$$

Hence the likelihood of the neighbouring sub-locations is given as

$$p\left(\phi_j|\tau_c\right) \propto \exp\left\{-\frac{\tau_c}{2}\sum_{i=1}^{n_t} w_{ij}(\phi_j - \phi_i)^2\right\} \quad \text{(Eqn:24)}$$

where w_{ij} denotes the adjacency matrix and ij shows that the sub location j is a neighbour of sub location i and m_j is the number of neigbours for sub

location j. A conjugate hyper prior of $\tau_c \sim\ Gamma\left(\alpha_c, \beta_c\right)$ is assumed

$$p(\tau_c) = \frac{\left(\beta_c\right)^{\alpha_c}}{\Gamma\left(\alpha_c\right)} \tau_c^{\alpha_c - 1} \exp\left(-\beta_c \tau_c\right), \alpha_c > 0; \beta_c > 0 \qquad \text{(Eqn:25)}$$

Therefore;

$$p\left(\phi_j | \tau_c\right) \times p(\tau_c) \propto \exp\left\{-\frac{\tau_c}{2} \sum_{i=1}^{n} w_{ij} (\phi_j - \phi_i)^2\right\} \tau_c^{\alpha_c - 1} \exp\left(-\beta_c \tau_c\right) \qquad \text{(Eqn:26)}$$

In INLA,the structured temporal component $\underset{\sim j}{\phi} = f_{spat}\left(s_j\right)$,a Besag CAR prior was used similar to CAR prior in BUGS, given as

$$s_j | s_i, j \neq i, \tau_c \sim\ N\left(\frac{1}{m_j} \sum_{j \sim\ i} s_j, \frac{1}{\tau_c m_j}\right) \qquad \text{(Eqn:27)}$$

Where $\phi_j\ =\ (\phi_1, \phi_2, \phi_3, ..., \phi_{40})$, m_j is the number of neighbours of sub location j ,$j\ \sim\ i$ indicates that the two sub locations j and i are neighbours.

AR(1) Temporal Component:

The temporal component was only implemented using R-INLA only and a first order autoregressive model with normal first term prior was used.

$\underset{\sim t}{\gamma} = f_{temp}\left(y_t\right) \sim\ AR\left(1\right)$ first order autoregressive model with normal

first term prior $y_t = \rho y_{t-1} + \varepsilon_t$ (69,70) where

$$y_1 = N\left(0, \frac{1}{\tau_e\left(1-\rho^2\right)}\right)$$

$|\rho_t| < 1$ for stationarity and $\varepsilon_t \sim N\left(0, \tau_e^{-1} = \sigma_t^2\right)$ is the white noise process.

The auto regressive parameter is expressed as

$$p\left(\rho|\sigma_t\right) = \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp\left(\frac{-\left(y_t - \rho y_{t-1}\right)^2}{2\sigma_t^2}\right) \qquad \text{(Eqn:28)}$$

Finally combining the likelihood and the prior to obtain our posterior distribution as shown in Eqn:43

$$p\left(\Omega|y_{ijt}\right) \propto \exp\left[\sum_{i=1}^{n_t}\left\{y_{ijt}\ln\left(\frac{\theta\exp(\psi_{ijt})}{1+\theta\exp(\psi_{ijt})}\right) - \frac{1}{\theta}\ln\left(1+\theta\exp(\psi_{ijt})\right)\right\}\right] \times$$
$$\frac{1}{\sqrt{2\pi\sigma_\beta^2}} \exp\left[-\frac{1}{2}\left(\frac{\beta_k - \mu_\beta}{\sigma_\beta}\right)^2\right] \times \tau_h^{\alpha_h - 1}\exp(-\beta_h\tau_h) \times \exp\left[-\frac{1}{2}\left(\frac{\vartheta_j - 0}{\frac{1}{\sqrt{\tau_h}}}\right)^2\right] \times$$
$$\exp\left\{-\frac{\tau_c}{2}\sum_{i=1}^{n} w_{ij}(\phi_j - \phi_i)^2\right\} \tau_c^{\alpha_c - 1}\exp\left(-\beta_c\tau_c\right) \times$$
$$\frac{1}{\sqrt{2\pi\sigma_t^2}} \exp\left(\frac{-\left(y_t - \rho y_{t-1}\right)^2}{2\sigma_t^2}\right) \qquad \text{(Eqn:29)}$$

Where

$$\exp\left(\psi_{ijt}\right) = \exp\left(x_{ijt}\underset{\sim}{\beta} + \underset{\sim}{\phi}_j + \underset{\sim}{\vartheta}_j + \underset{\sim}{\gamma}_t + \varepsilon_{ijt}\right)$$

This combination has no closed form; estimation is used to estimate the parameters. To solve for the parameters, we use the adapted Stochastic Partial Differential Equations (SPDE) in INLA, and the MCMC with Metropolis- Hastings algorithms approaches in WinBUGS (25,32,71).

In our implementation, combination of the unstructured $\underset{\sim}{\vartheta}_j = f_{spat\_unst}(s_j)$ and the structured $\underset{\sim}{\phi}_j = f_{spat}(s_j)$ priors was modelled as a convolution-prior resulting to a Besag-York-Mollie model. The model is a union of a Besag model Eqn:27 and iid model Eqn:22 . This BYM model allows to get the posterior marginal of the sum of structured Besag and unstructured iid spatial components (72).

A negative binomial Bayesian hierarchical space-time model was implemented through an adapted Stochastic Partial Differential Equations (SPDE) approach to produce continuous maps of malnutrition attributable morbidity at level 5 spatial resolution. This approach was selected over the Markov Chain Monte Carlo (MCMC) due to its suitability for Gaussian Markov Random Fields (GMRF) and its computing capability of temporal models (26,27,69,73,74).

2.6.2 Spatial Logistic Model

Our outcome for objective 3 was the discharge outcome of the patient, whether discharged alive or dead. A Bayesian geo-additive spatial logit model with the death as the outcome with a logit link function is fitted to determine the malnutrition attributable mortality. This was considered over the count model since the mortality outcome was fitted as a binomial outcome to allow adjusting for the individual admission effects.

Model Definition

Since our outcome is binary the general distribution is;

$$y_i = \begin{cases} 1 \text{ if there was a death} \\ 0 \qquad otherwise \end{cases}$$

Then the likelihood of death is a binomial distribution (a sequence of Bernoulli trials) in general defined as;

$$L\left(x_{ij}, y_{ij}|p_{ij}\right) = L\left\{Y_{ij} = y_{ij}\right\} = \begin{pmatrix} r_{ij} \\ y_{ij} \end{pmatrix} p_{ij}^{y_{ij}} (1 - p_{ij})^{r_{ij} - y_{ij}} \quad \text{(Eqn:30)}$$

Where p_{ij} represents the probability of the subject i from sub-location j who has a covariate vector x_{ij} and $y_{ij} = 1$ indicates death or alive if it takes value zero.

The r_{ij} are total admissions with a mean $E\left(y_{ij}\right)=\mu_{ij}=r_{ij}p_{ij}$ and the variance is $\operatorname{var}\left(y_{ij}\right)=\delta_{ij}=r_{ij}p_{ij}\left(1-p_{ij}\right)$.

Considering that a logistic regression model can be defined as follows with cluster-level random effects $\underset{\sim}{u}_{ij}$.

$$\text{logit}\left(p_{ij}\left(x_{ij}\right)\right)=x_{ij}\underset{\sim}{\beta}+\underset{\sim}{\phi}_{j}+\underset{\sim}{\vartheta}_{j}+\varepsilon_{ij} \qquad \text{(Eqn:31)}$$

Then making p_{ij} the subject of the formula, then logistic model was modelled as

$$\ln\left(\frac{p_{ij}}{1-p_{ij}}\right)=x_{ij}^{T}\underset{\sim}{\beta}+\underset{\sim}{\phi}_{j}+\underset{\sim}{\vartheta}_{j}+\varepsilon_{ij} \qquad \text{(Eqn:32)}$$

Considering the binomial mixed-effects model, for the sub location clusters, $j=1,2,3,....,40$ the conditional distribution of $\underset{\sim}{y}_{j}=\left(y_{j1},y_{j2},...,y_{jn_{j}}\right)^{T}$ given a set of sub location level random effects $\underset{\sim}{u}_{ij}$; expressed as

$$f\left(y_{ij}|\mu_{ij}\right)=\prod_{i=1}^{n_{j}}\left[\binom{r_{ij}}{y_{ij}}\left\{H(\underset{\sim}{\psi}_{ij})\right\}^{y_{ij}}\left\{1-H(\underset{\sim}{\psi}_{ij})\right\}^{r_{ij}-y_{ij}}\right] \qquad \text{(Eqn:33)}$$

Where $\underset{\sim}{\psi}_{ij}=x_{ij}^{T}\underset{\sim}{\beta}+\underset{\sim}{z}_{ij}\underset{\sim}{u}_{j}$

Then, exponentiation and log transformation of $H(\underset{\sim}{\psi}_{ij})$, this can be written as,

$$H\left(\underset{\sim}{\psi}_{ij}\right) = \left(\frac{\exp\left(x_{ij}\underset{\sim}{\beta} + \underset{\sim}{\phi}_{j} + \underset{\sim}{\vartheta}_{j} + \varepsilon_{ij}\right)}{1 + \exp\left(x_{ij}\underset{\sim}{\beta} + \underset{\sim}{\phi}_{j} + \underset{\sim}{\vartheta}_{j} + \varepsilon_{ij}\right)}\right) = \frac{\exp\left(\underset{\sim}{\psi}_{ij}\right)}{1 + \exp\left(\underset{\sim}{\psi}_{ij}\right)} \quad \text{(Eqn:34)}$$

From Eqn:32

$$= \exp\left\{\ln\begin{pmatrix} r_{ij} \\ y_{ij} \end{pmatrix} + y_{ij}\ln\left(H\left(\underset{\sim}{\psi}_{ij}\right)\right) + (r_{ij} - y_{ij})\ln\left(1 - H\left(\underset{\sim}{\psi}_{ij}\right)\right)\right\} \quad \text{(Eqn:35)}$$

Expanding the brackets inside the exponent and simplifying we obtain

$$= \prod_{i=1}^{n_j} \exp\left\{\ln\begin{pmatrix} r_{ij} \\ y_{ij} \end{pmatrix} + y_{ij}\ln\left(\frac{\exp\left(\underset{\sim}{\psi}_{ij}\right)}{1 + \exp\left(\underset{\sim}{\psi}_{ij}\right)}\right) + (r_{ij} - y_{ij})\ln\left(\frac{1}{1 + \exp\left(\underset{\sim}{\psi}_{ij}\right)}\right)\right\} \quad \text{(Eqn:36)}$$

$$= \prod_{i=1}^{n_j} \exp\left\{\ln\begin{pmatrix} r_{ij} \\ y_{ij} \end{pmatrix} + y_{ij}\ln\left(\exp\left(\underset{\sim}{\psi}_{ij}\right)\right) - r_{ij}\ln\left(1 + \exp\left(\underset{\sim}{\psi}_{ij}\right)\right)\right\} \quad \text{(Eqn:37)}$$

$$= \prod_{i=1}^{n_j} \exp\left\{\ln\begin{pmatrix} r_{ij} \\ y_{ij} \end{pmatrix} + y_{ij}\underset{\sim}{\psi}_{ij} - r_{ij}\ln\left(1 + \exp\left(\underset{\sim}{\psi}_{ij}\right)\right)\right\} \quad \text{(Eqn:38)}$$

$$= \exp\left\{\sum_{i=1}^{n_j}\left[y_{ij}\psi_{ij} - r_{ij}\ln\left(1+\exp\left(\underset{\sim}{\psi}_{ij}\right)\right) + \ln\begin{pmatrix} r_{ij} \\ y_{ij}\end{pmatrix}\right]\right\} \quad \text{(Eqn:39)}$$

Let $\underset{\sim}{r}_j = \left(r_{j1}, r_{j2}, ..., r_{jn_j}\right)^T$, $\underset{\sim}{y}_j = \left(y_{j1}, y_{j2}, ..., y_{jn_j}\right)^T$ and $\underset{\sim}{\psi}_j = \left(\psi_{j1}, \psi_{j2}, ..., \psi_{jn_j}\right)^T$ then

$$c\left(\underset{\sim}{y}_j, \underset{\sim}{r}_j\right) = \sum_{i=1}^{n_j}\ln\begin{pmatrix} r_{ij} \\ y_{ij}\end{pmatrix}$$

which does not depend on the model parameters on Eqn:39 in compact form is thus expressed as

$$f\left(\underset{\sim}{y}_j|\underset{\sim}{u}_j\right) = \exp\left[\sum_{i=1}^{n_j}\left(\underset{\sim}{y}_j'\underset{\sim}{\psi}_j - r_j'\ln\left(1+\exp\left(\underset{\sim}{\psi}_j\right)\right)\right) + c\left(\underset{\sim}{y}_j, \underset{\sim}{r}_j\right)\right] \quad \text{(Eqn:40)}$$

Given that the prior distribution of the $u_j = MVN\left(0, \Sigma\right)$ is a multivariate normal with a mean zero and a $\Sigma_{q\times q} = \Sigma_{40\times 40}$ variance matrix. The likelihood contribution for the j^{th} cluster is obtained by integrating $\underset{\sim}{u}_j$ out of the joint density $f\left(\underset{\sim}{y}_j|\underset{\sim}{u}_j\right)$.

$$L_j\left(\beta, \Sigma\right) = \left(2\pi\right)^{-\frac{q}{2}}|\Sigma|^{-\frac{1}{2}}\int f\left(\underset{\sim}{y}_j|\underset{\sim}{u}_j\right)\exp\left(-u_J'\Sigma^{-1}\frac{u_j}{2}\right)du_j \quad \text{(Eqn:41)}$$

$$= \exp\{c(y_j, r_j)\}(2\pi)^{-\frac{q}{2}}|\Sigma|^{-\frac{1}{2}} \int \exp\left\{h\left(\beta, \Sigma, \underset{\sim}{u}_j\right)\right\} du_j \quad \text{(Eqn:42)}$$

Where

$$h\left(\beta, \Sigma, \underset{\sim}{u}_j\right) = y_j^{'}\underset{\sim}{\psi}_j - r_j^{'}\ln\left(1 + \exp(\underset{\sim}{\psi}_j)\right) - u_j^{'}\Sigma^{-1}\frac{\underset{\sim}{u}_j}{2}$$

And for convenience in the arguments of $h(.)$ we supress the dependence on the observable data $\left(\underset{\sim}{y}_j, \underset{\sim}{r}_j, \underset{\sim}{x}_j, \underset{\sim}{z}_j\right)$

The above equation has no closed form and has to be approximated for the maximum likelihood estimation using mean–variance adaptive Gauss–Hermite quadrature and Laplacian approximation crossed random-effects models (75).

The model fit within STATA does not cater for spatial random effects as such we extended this to the Bayesian framework such that the link function predictor ψ_j is extended such that

$$\psi_j^* = x_{ij}\underset{\sim}{\beta} + z_{ij}\underset{\sim}{u}_j + \varepsilon_{ij}$$

Letting $z_{ij}\underset{\sim}{u}_j = \underset{\sim}{\vartheta}_j + \underset{\sim}{\phi}_j$

The spatial Bayesian model is defined as follows

$$Posterior[p(parameters|data)] \propto Likelihood \times Priors$$

The full conditional for our model can be expressed as

$$p\left(\Omega \equiv \left\{\underset{\sim}{\beta}, \underset{\sim}{\phi}, \underset{\sim}{\vartheta}\right\}|y_{ij}\right) \propto P\left(\Omega|y_{ij}\right) = L\left(y_{ij}, x_{ij}|\Omega\right) \times p\left(\underset{\sim k}{\beta}\right) \times p\left(\underset{\sim j}{\phi}|\tau_c\right) \times p\left(\underset{\sim j}{\vartheta}|\tau_h\right) \quad \text{(Eqn:43)}$$

$$= L\left(y_{ij}|\Omega\right) \times p\left(\underset{\sim k}{\beta}\right) \times p\left(\underset{\sim j}{\phi}|\tau_c\right) \times p\left(\tau_c\right) \times p\left(\underset{\sim j}{\vartheta}|\tau_h\right) \times p\left(\tau_h\right) \quad \text{(Eqn:44)}$$

The prior for the beta coefficients in the model is $\beta_k \overset{iid}{\sim} N\left(\mu_\beta, {\sigma_\beta}^2\right)$, is as defined in Eqn:19

The unstructured random effects $p\left(\underset{\sim j}{\vartheta}|\tau_h\right)$, where $\underset{\sim j}{\vartheta} \overset{iid}{\sim} N\left(0, \frac{1}{\tau_h}\right)$ with a Gamma function prior used $\tau_h \sim Gamma\left(\alpha_h, \beta_h\right)$ was defined as shown in equation Eqn:22 above.

A CAR prior is used for the structured spatial random effects $\phi_j \sim CAR\left(\tau_c\right)$ and a CAR prior is given as $\phi_j|\phi_i, j \neq i, \tau_c \sim N\left(\frac{\phi_j}{\bar{\phi}_j}, \frac{1}{\tau_c m_j}\right)$ with a conjugate hyper prior of $\tau_c \sim Gamma\left(\alpha_c, \beta_c\right)$ assumed. The pdf is as described in equation Eqn:26 and a similar approach with a Besag prior as shown in Eqn:27

Combining the equations for our posterior distribution as shown in Eqn:43

$$p\left(\Omega|y_{ij}\right) \propto \exp\left[\sum_{i=1}^{n_j}\left(\underset{\sim}{y}_j^{'}\underset{\sim}{\psi}_{ij} - r_j^{'}\ln\left(1+\exp\left(\underset{\sim}{\psi}_{ij}\right)\right)\right)\right] \times$$
$$\frac{1}{\sqrt{2\pi\sigma_\beta^2}}\exp\left[-\frac{1}{2}\left(\frac{\beta_k-\mu_\beta}{\sigma_\beta}\right)^2\right] \times \tau_h^{\alpha_h-1}\exp(-\beta_h\tau_h) \times \exp\left[-\frac{1}{2}\left(\frac{\vartheta_j-0}{\frac{1}{\sqrt{\tau_h}}}\right)^2\right] \times$$
$$\exp\left\{-\frac{\tau_c}{2}\sum_{i=1}^{n} w_{ij}(\phi_j-\phi_i)^2\right\}\tau_c^{\alpha_c-1}\exp\left(-\beta_c\tau_c\right)$$

(Eqn:45)

Where

$$\underset{\sim}{\psi}_{ij} = \exp\left(x_{ij}\underset{\sim}{\beta} + \underset{\sim}{\phi}_j + \underset{\sim}{\vartheta}_j + \varepsilon_{ij}\right)$$

The above equation also has no closed form and has to be approximated. To approximate for the parameters we use the adapted Stochastic Partial Differential Equations (SPDE) in INLA and the MCMC with Metropolis-Hastings algorithms approaches (26,27,69,73,74).

The estimated parameters we then used to calculate the mortality attributable fraction. To calculate the confidence intervals for the attributable fractions, the ordinary bootstrap with 1000 iterations was used (3,76).

2.6.3 Model assesment and goodness of fit

We adopted both multilevel models and Bayesian approach for model estimation. The advantage of Bayesian models over the multilevel models is the combination of the prior information with the data through Bayes theorem to obtain posterior distributions as shown in section 2.6.1 and 2.6.2. Convergence of the MCMC and SPDE approach was monitored using the trace plots. For model comparison and best fit, we use the Deviance Information Criterion (DIC), where the small DIC was considered better. DIC is defined as $\bar{D}\ (\theta) + pD$ where $\bar{D}\ (\theta) = E\left[D(\theta)|y\right]$ which is the posterior mean of the deviance, $D(\theta)$. pD is the difference in the posterior mean deviance and the deviance evaluated at the posterior mean of the parameters, $pD = \bar{D}\left(\theta\right) - D\left(E\left(\theta|y\right)\right)$. The posterior mean of the deviance is used to asses the best model fit and pD penalizes for number of parameters in the model, and values between 5-10 indicates the model fits better (71,77).

2.9 Ethical Clearance

The project utilises data from KHDSS which was received in an anonymised format. The study is approved by the KEMRI (Kenya Medical Research

Institute) Scientific Steering Committee in Kilifi (78). Ethical approval under clearance certificate number M1611104 was received from the Human Research Ethics Committee of the University of the Witwatersrand. The certificate is in Appendix 1.

CHAPTER 3: RESULTS

3.1 Introduction

In this chapter, we show the results of the analysis of the KHDSS admissions data for the period between 2002-2015. Temporal and spatial exploratory analysis using time series and hotspot analysis respectively are presented. We adjusted for the potential confounders, spatial-temporal random effects and other associated factors.

Inferential analysis of the spatial-temporal negative binomial model and spatial logistic model for objective 2 and objective 3 respectively are presented, but we discuss the models of best fit only. Firstly we present the exploratory analysis and then the inferential statistics.

3.2 Exploratory Data Analysis

3.2.1 Temporal Exploratory Data Analysis

The temporal analysis was fitted using monthly data for the period from April 2002 to December 2015.The auto correlation function (ACF) and partial auto correlation function(PACF) of the monthly transformed series using data from 2002 to 2015 show the peaks at different periods as shown

in Appendix 3 for morbidity. A significant serial correlation (seasonality) of 7 months was observed in children admitted with malnutrition and 4 months was observed for non-malnutrition admissions. We report the admission pattern that was unique for malnutrition related admissions. This was confirmed with a significant ($p<0.001$) Portmanteau test for white noise. An Autoregressive of order 1 (AR1) was selected to be used in the spatial-temporal model which fitted the observed data well as shown in Appendix 3.

Panels (A) and (B) of Figure~4.3 show graphs of the time series plots for the monthly malnutrition related admissions. A significant seasonality of July was observed for malnutrition related admissions, but in general, the malnutrition admissions decline over the period. The graphs show the prediction using SARIMA model. The Panel B of Figure~4.3 shows the non-malnutrition related admission. The red line on Figure~4.3 shows that the Auto Regressive of order 1 predicts the data well and this was implemented in the spatial-temporal model for malnutrition related admissions. The tables in Appendix 2 shows the monthly counts of malnutrition admissions and mortality for the period 2002 to 2015.

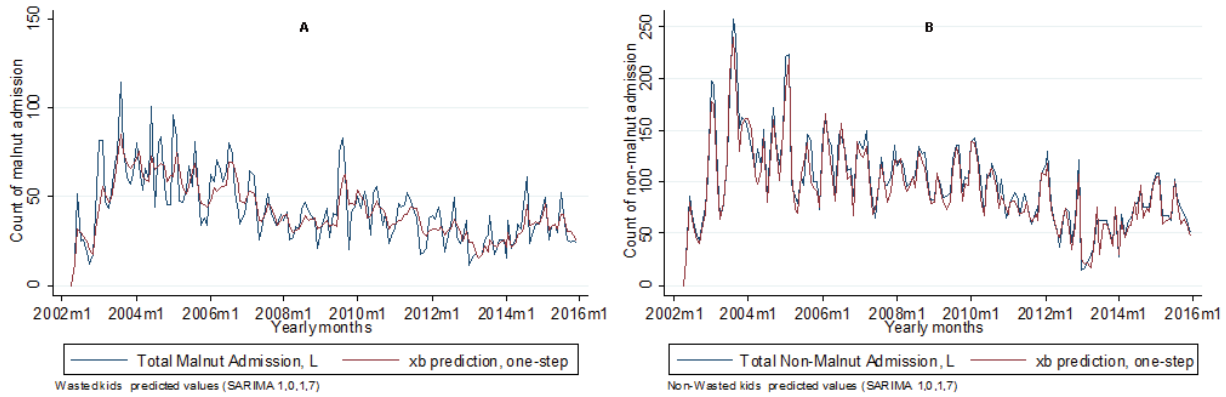


Figure 4.3: A time series plot for monnthly number of admissions from 2002-2015 in KHDSS

3.2.2 Spatial Exploratory Data Analysis

A Global Moran's I index of 0.1 (p-value<0.001, sd=0.029) was observed. Based on the calculated Global Morans I index, we rejected the zero spatial autocorrelation hypothesis. This shows spatial clustering of the admissions from the different sub locations. The Kulldorff spatial scan statistic using SaTScan showed hotspots in the Northern and Southern areas of the Kilifi urban area. Areas closer to the main hospital (KCH) and Kilifi town were observed as cold spots of malnutrition related admissions as shown on the Figure~4.4 panel A. Four of the five hot spots and cold spots were significant (i.e. p<0.05), the cluster (Mwapula, Marere, Magogoni-K, Mdangarani, Vinagoni, Chivara) was the one observed not to be significant for malnutrition related admissions.The temporal clustering of malnutrition related admissions as shown in Appendix 4, kept on shifting

over the periods but apparently similar regions were observed. Three mortality clusters we observed as shown on Figure~4.4 panel. We only did hotspot analysis for mortality, and the clusters close to urban Kilifi town was observed during the period 2002 to 2006, the other mortality hotspots were observed until 2008.

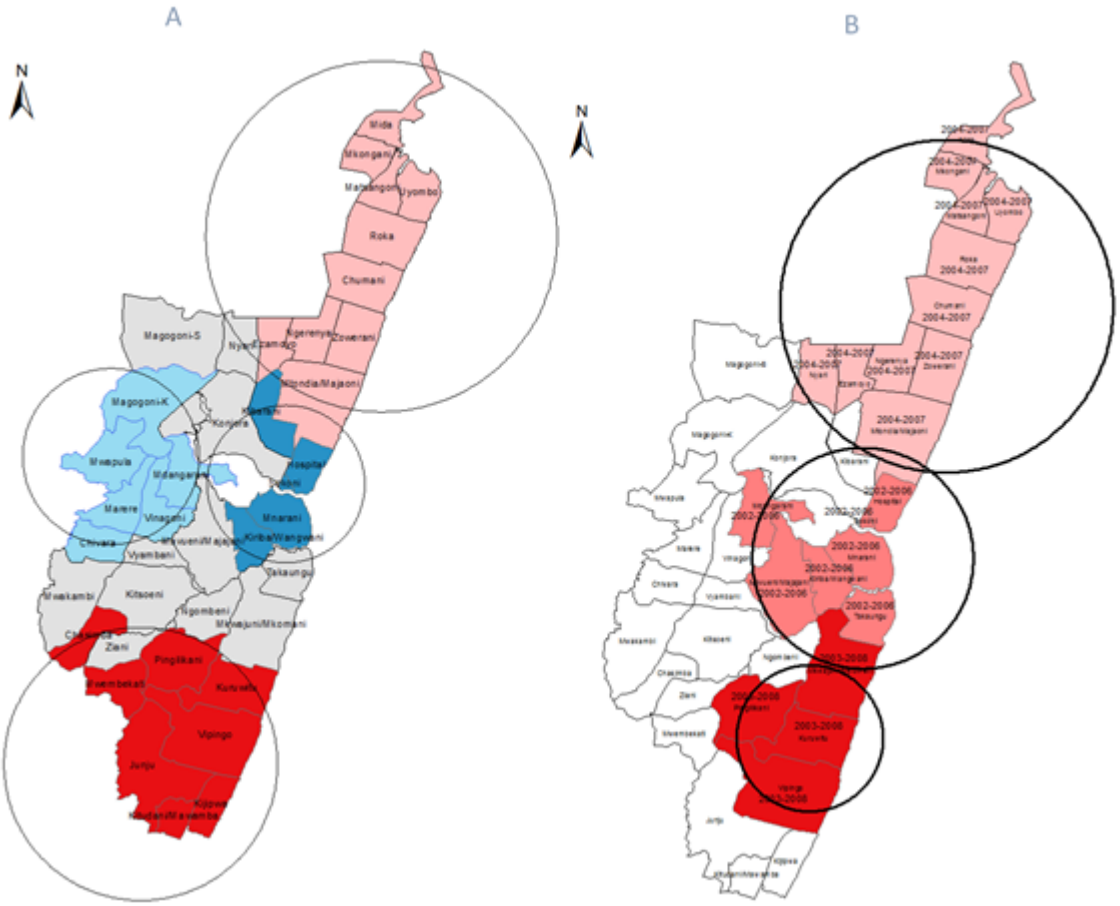


Figure 4.4: Malnutrition related hotspots (red shade) and coldspots (blue shade) in KHDSS 2002-2015, A- morbidity coldspots and hotspots and B- mortality hotsposts

3.3 Bivariate Analysis

The data used was for the children aged 3 months to 15 years admitted in Kilifi County Hospital over the 14 years between 2002 and 2015 whose sub location data was available. This totalled to 23,483 admissions from 17,940 individuals across the 40 sub-locations in the KHDSS. Malnutrition related readmission analysis had 7,820 admission events from 3,114 children as shown on Table~4.1. In objective three we used the full set of data with 16,355 non-malnutrition related admissions and 7,128 malnutrition related admissions. The average age of the patients who died was higher than that of those who were discharged alive but had many numbers of admission days as shown on Table~4.1. The admissions that had a malnutrition was 30.24% of the total inpatient records of 2365 individuals. Out of the 1271 respondents who had a malnutrition admission in at least one of his or her admission, 73% of their admissions are a malnutrition related admission. The between column tells us that (1271/3114) 40.82% of the repeated admissions happened with the children being malnourished.

Males had a higher percentage (56.23% between males) of experiencing two or more admissions compared to the females (43.77% between females), which is also a similar case in the mortality outcome. Malaria

admissions happened 68% of the times in all admissions that occurred. At least 33% of the admissions had a malaria admission with a significant association in the cases and controls. Blood culture and CSF culture positive patients who died had a higher between percentage compared to those discharged alive as shown on Table~4.2 below.

Table 4.1: Characteristics of the study population for the continous variables a bivariate analysis by mortality; SD(O,B,W) is the overall, within and between Standard Deviation

Variables	Overall			Alive			Died			P-value
	N (total admissions)	Mean	SD (O,B,W)	N (total admissions)	Mean	SD (O,B,W)	N (total admissions)	Mean	SD (O,B,W)	
Age (months)	3114 (7820)	35.95	32.56 28.53 14.73	3114 (7621)	35.71	32.35 28.48 14.46	199 (199)	45.16	38.69 38.69 N/A	0.0001
Admission Days	3113 (7780)	5.18	6.61 4.95 4.45	3113 (7591)	5.15	6.53 5.04 4.34	189 (189)	6.51	9.16 9.16 N/A	0.0067
Number of Admissions	3114 (7820)	3.49	2.56 1.59 N/A	3114 (7621)	3.50	2.58 1.59 N/A	199 (199)	2.96	1.75 1.75 N/A	0.0036
Total Admissions	3114 (7820)	3.22	2.27 1.37 N/A	3114 (7621)	3.23	2.29 1.37 N/A	199 (199)	2.67	1.37 1.37 N/A	0.0007
Weight (kgs)	3109 (7736)	10.96	5.17 4.69 2.16	3108 (7541)	10.97	5.13 4.68 2.13	195 (195)	10.53	6.38 6.38 N/A	0.2597
Height (cm)	3092 (7526)	85.17	19.06 17.09 8.53	3088 (7350)	85.12	18.98 17.14 8.40	176 (176)	87.05	22.45 22.45 N/A	0.1939
MUAC (cm)	3113 (7729)	13.84	2.02 1.81 0.93	3112 (7537)	13.88	1.98 1.79 0.91	192 (192)	12.36	2.85 2.85 N/A	<0.00001
HAZ06	2789 (6167)	-1.68	1.52 1.40 0.63	2785 (6040)	-1.65	1.50 1.40 0.62	127 (127)	-3.05	1.61 1.61 N/A	<0.00001
WAZ06	2777 (6128)	-1.87	1.51 1.43 0.47	2776 (6013)	-1.83	1.48 1.42 0.46	115 (115)	-4.01	1.69 1.69 N/A	<0.00001
WHZ06	2746 (5968)	-1.29	1.54 1.39 0.69	2742 (5862)	-1.25	1.51 1.37 0.67	106 (106)	-3.39	1.93 1.93 N/A	<0.00001
EVI Value	3114 (7820)	0.39	0.08 0.06 0.05	3114 (7621)	0.39	0.08 0.07 0.05	199 (199)	0.40	0.08 0.08 N/A	0.1798
Rain (mm)	3070 (7426)	31.92	44.30 32.42 32.84	3068 (7236)	31.78	44.02 32.78 32.30	190 (190)	37.27	53.79 53.79 N/A	0.1002

Table 4.2: Categorical Characteristics of the study population; N-total number of individuals,Adm N - Total Admissions, (%B,%W) is the percentage between and within

Variables		Overall			Alive			Died			
		N	Adm N (%)	Between N (%B;%W)	N	Adm N (%)	Between N (%B;%W)	N	Adm N(%)	Between N (%B;%W)	Pvalue
Gender	Male	3114	4415 (56.46)	1751 (56.23;100)	3114	4298 (56.4)	1751 (56.23;100)	199	117 (58.79)	117 (58.79;100)	0.471
	Female		3405 (43.54)	1363 (43.77;100)		3323 (43.6)	1363 (43.77;100)		82 (41.21)	82 (41.21;100)	
Malnut Kid	NO	3114	5455 (69.76)	2522 (80.99;86.67)	3114	5391 (70.74)	2522 (80.99;87.56)	199	64 (32.16)	64 (32.16;100)	<0.001
	YES		2365 (30.24)	1271 (40.82;73.03)		2230 (29.26)	1225 (39.34;73.94)		135 (67.84)	135 (67.84;100)	
Severe diseases	0	3114	3658 (46.78)	2081 (66.83;66.11)	3114	3563 (46.75)	2045 (65.67;67.21)	199	95 (47.74)	95 (47.74;100)	<0.001
	1		3808 (49.70)	2354 (75.59;67.34)		3720 (48.81)	2322 (74.57;68.34)		88 (44.22)	88 (44.22;100)	
	2		332 (4.25)	305 (9.92.26;46.29)		322 (4.23)	302 (9.70;47.86)		10 (5.03)	10 (5.03;100)	
	3		22 (0.28)	21 (0.67;47.62)		16 (0.21)	16 (0.51;50.00)		6 (3.02)	6 (3.02;100)	
Severe anaemia	NO	3107	6870 (92.02)	3041 (97.88;95.02)	3100	6708 (92.1)	3025 (97.58;95.3)	183	162 (88.52)	162 (88.52;100)	0.982
	YES		596 (7.98)	405 (13.04;53.7)		575 (7.9)	392 (12.65;55.44)		21 (11.48)	21 (11.48;100)	
Hypoglycaemia	NO	2864	6189 (99.1)	2850 (99.51;99.45)	2856	6057 (99.15)	2843 (99.54;99.48)	136	132 (97.06)	132 (97.06;100)	0.303
	YES		56 (0.9)	53 (1.85;55.94)		52 (0.85)	49 (1.72;56.67)		4 (2.94)	4 (2.94;100)	
Malaria	NO	3108	5795 (78.07)	2697 (86.78;88.6)	3099	5637 (77.87)	2677 (86.38;88.96)	184	158 (85.87)	158 (85.87;100)	<0.001
	YES		1628 (21.93)	1052 (33.85;68.3)		1602 (22.13)	1037 (33.46;69.2)		26 (14.13)	26 (14.13;100)	
Diarrhoea	NO	3114	6357 (81.5)	2908 (93.38;85.6)	3114	6221 (81.77)	2890 (92.81;86.29)	192	136 (70.83)	136 (70.83;100)	0.503
	YES		1443 (18.5)	1118 (35.9;55.88)		1387 (18.23)	1089 (34.97;56.96)		56 (29.17)	56 (29.17;100)	
Menegitis	NO	3113	7717 (99.13)	3110 (99.9;99.12)	3113	7536 (99.22)	3107 (99.81;99.27)	190	181 (95.26)	181 (95.26;100)	0.002
	YES		68 (0.87)	64 (2.06;47.28)		59 (0.78)	56 (1.8;51.06)		9 (4.74)	9 (4.74;100)	
LRTI	NO	3113	5736 (73.68)	2775 (89.14;82.21)	3113	5582 (73.5)	2743 (88.11;82.76)	190	154 (81.05)	154 (81.05;100)	<0.001
	YES		2049 (26.32)	1371 (44.04;60.66)		2013 (26.5)	1357 (43.59;62.12)		36 (18.95)	36 (18.95;100)	
Gastroenteritis	NO	3113	6788 (87.19)	2970 (95.41;89.95)	3113	6621 (87.18)	2957 (94.99;90.31)	190	167 (87.89)	167 (87.89;100)	0.004
	YES		997 (12.81)	803 (25.8;54.98)		974 (12.82)	790 (25.38;56.02)		23 (12.11)	23 (12.11;100)	
Transfused	NO	3114	7209 (92.61)	3073 (98.68;94.59)	3114	7053 (92.86)	3066 (98.46;95.03)	189	156 (82.54)	156 (82.54;100)	0.014
	YES		575 (7.39)	408 (13.1;50.78)		542 (7.14)	383 (12.3;52.29)		33 (17.46)	33 (17.46;100)	
Blood Culture	NO	3088	7207 (96.29)	3072 (99.48;96.62)	3086	7064 (96.75)	3066 (99.35;97.14)	184	143 (77.72)	143 (77.72;100)	<0.001
	YES		278 (3.71)	252 (8.16;47.58)		237 (3.25)	217 (7.03;49.61)		41 (22.28)	41 (22.28;100)	
CSF Culture	NO	1078	1375 (98.78)	1068 (99.07;99.73)	1058	1348 (99.12)	1051 (99.34;99.78)	32	27 (84.38)	27 (84.38;100)	<0.001
	YES		17 (1.22)	16 (1.48;80.21)		12 (0.88)	12 (1.13;77.78)		5 (15.63)	5 (15.63;100)	

3.4 Spatial-Temporal Negative Binomial Model Fit and Diagnostics

3.4.1 Model fit results

Multivariable analysis was done adjusting for different random effects non-spatial, spatial and spatial-temporal. The age of the child at admission was used as the exposure duration (offset)variable, and the cumulative count of admissions was used as the temporal variable for more epidemiological intuitive results. The estimated coefficients are reported in Table~4.3 with 95% confidence interval based on the Maximum Likelihood Estimation (MLE) approach and 95% credible intervals for the Bayesian approach. As shown on Table~4.3 four models were fit, a multilevel model (MLE approach) adjusting for the sub-location random effects, a spatial model using Gibbs sampling and the INLA estimation procedure which has a better computational capability (79).

In general, the spatial-temporal model with a lower Deviance Information Criterion compared to non-spatial models was used to assess the best model fit. Bayesian spatial-temporal model using R-INLA was the model with the best fit with a DIC= 13640.20 which is the lowest among the spatial-temporal models but with a higher number of effective parameters $pD = 35.50$ because of the extra temporal parameters.

Table 4.3: Non-spatial and spatial multivariable Negative Binomial models.

	Multi level model Random Effects (STATA)		Bayesian Multilevel Spatial Random Effects Model (WINBUGS)	Bayesian Multilevel Random Effects Model (INLA)	Spatial Structured and Unstructured and Temporal Random Effects Model(INLA)
variables	Coeficients (95% C.I)	p-value	Mean (95% Cr .I)	Mean (95% Cr .I)	Mean (95% Cr .I)
EVI	-0.48 (-0.98;0.02)	0.061	-0.6 (-1.81;-0.03)	-0.49 (-1.02;0.04)	-0.18 (-0.68;0.31)
Rainfall	0.04 (0.01;0.09)	0.059	0.04 (0;0.09)	0.03 (-0.01;0.08)	0.01 (-0.03;0.05)
Gender: Male	-0.2 (-0.28;-0.13)	<0.01	0.2 (0.12;0.27)	-0.15 (-0.22;-0.07)	-0.16 (-0.23;-0.09)
Severe Disease					
1	0.001 (-0.08;0.08)	0.927	0.01 (-0.09;0.07)	-0.08 (-0.16;0)	-0.04 (-0.12;0.03)
2	-0.06 (-0.26;0.14)	0.565	-0.07 (-0.28;0.13)	-0.21 (-0.42;0)	-0.12 (-0.32;0.07)
3	0.32 (-0.27;0.92)	0.286	0.31 (-0.3;0.93)	0.24 (-0.4;0.86)	0.22 (-0.38;0.77)
Total Admission	0.13 (0.11;0.14)	<0.01	0.13 (0.11;0.14)	0.14 (0.13;0.16)	-0.02 (-0.04;0)
Admission Days	0.03 (0.02;0.03)	<0.01	0.03 (0.02;0.03)	0.04 (0.03;0.04)	0.03 (0.03;0.03)
Weight (kg)	-0.2 (-0.21;-0.19)	<0.01	-0.2 (-0.2;-0.19)	-0.08 (-0.09;-0.07)	-0.11 (-0.12;-0.1)
DIC (pD)			13169.33 (66.38)	14498.08 (25.03)	13640.20 (35.50)

The Bayesian spatial-temporal multi variable model showed some explanatory variables as determinants of child malnutrition re-admission which was consistent with the non-spatial models. The results are as shown on Table 4.3. The environmental variables were significantly associated with a malnutrition admission in the multilevel model, but a different observation in the spatial-temporal models. Rainfall increase was associated with a higher risk of malnutrition admission (RR=1.04; Cr.I=(1.01-1.09)) in the spatial model but after adjusting for the temporal component the interval of the level of significance changes (RR= 1.02; Cr.I=(0.97-1.05)) in the spatial-temporal model. Similarly, adjusting for the spatial and temporal model changed the significance of the EVI variable in the model though it still showed a reduction of the risk of malnutrition admission. The weight of a child which is an important marker of malnutrition; we observed that an increase in weight of a child reduced the relative risk of admission by 20% (RR=0.89; Cr.I= (0.88-0.91)). The children with more admission days had a higher risk of malnutrition re-admission.

The Figure~4.5 shows the overall hot spots and cold spots which were consistent with the SaTScan results, but some sub-locations close to the Kilifi urban were identified as hotspots in this model. Areas with darker red

shows that they have a higher risk of admission with malnutrition and blue shades (in the coloured version) indicate less probability of a malnutrition admission. The temporal risk ratios map were generated after running the Bayesian spatial-temporal model and are as shown on the Figure~4.6. The maps were generated to explain differences in the temporal trend of malnutrition related admissions for the different sub-locations. The blue shades symbolise the cold spots; the red shades exhibits the hot spots for malnutrition related readmissions.

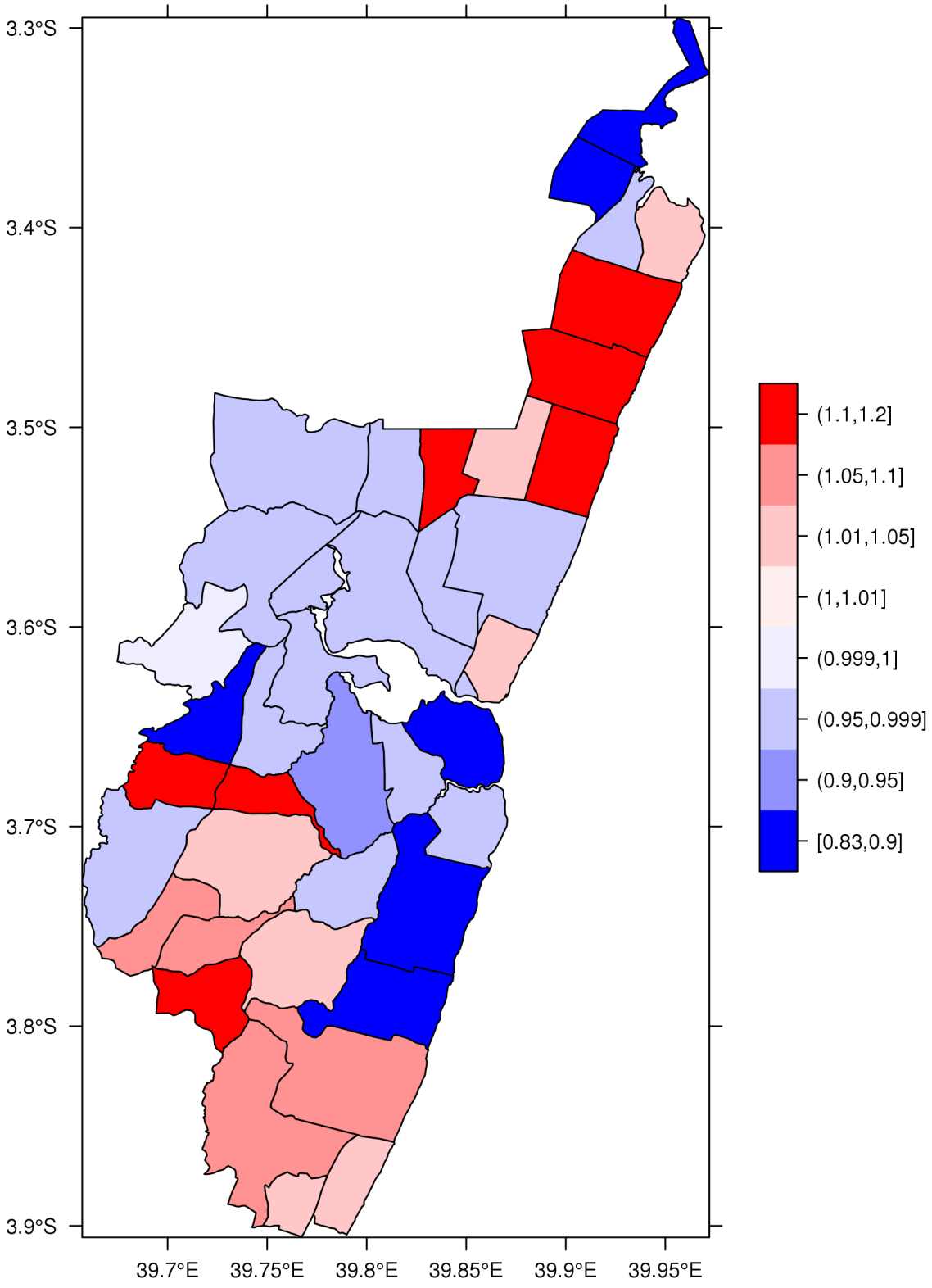


Figure 4.5: Malnutrition related hotspots (red shade) and coldspots (blue shade) in KHDSS 2002-2015, overall results from the Bayesian spatial-temporal model

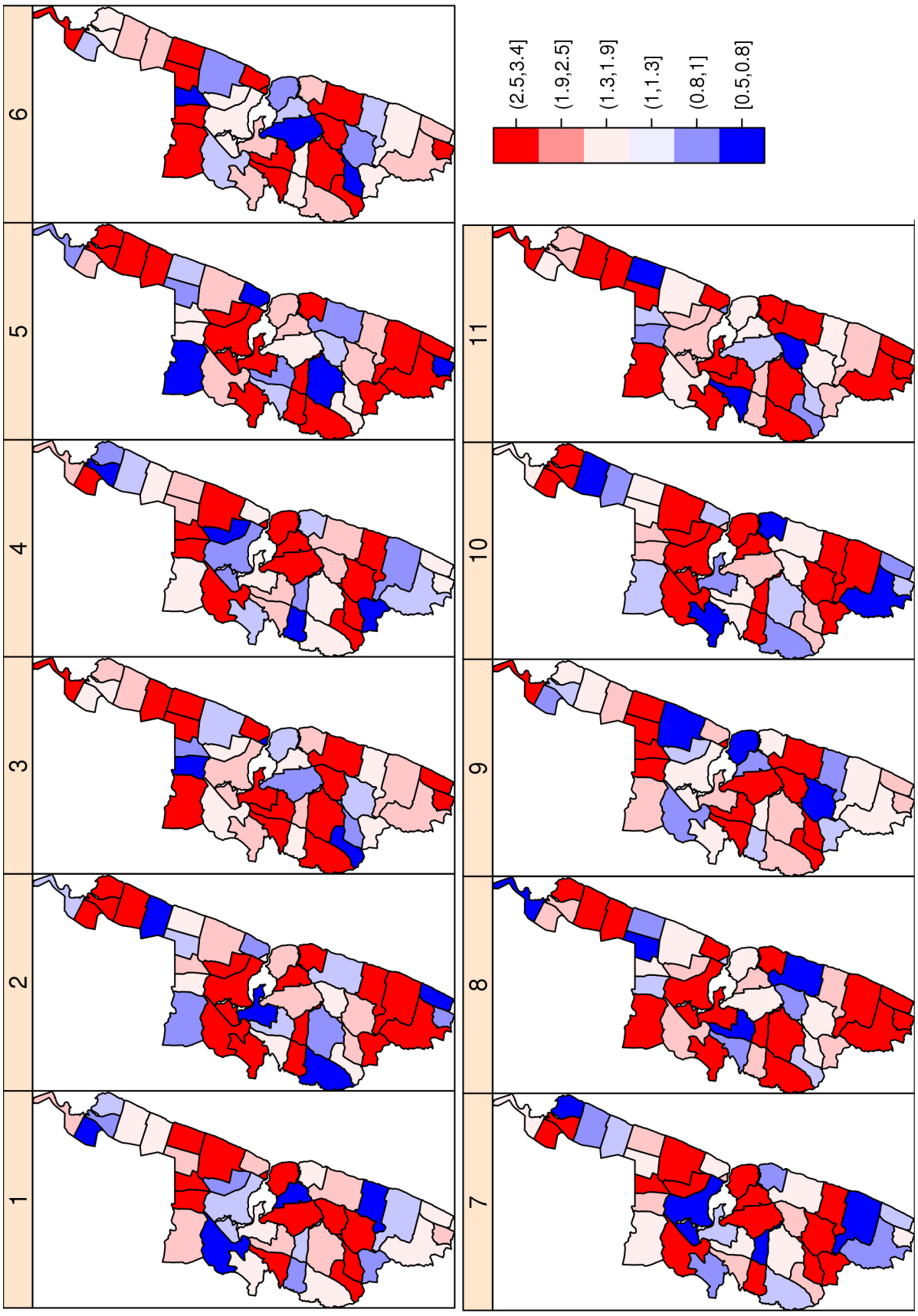


Figure 4.6: Malnutrition related hotspots (red shade) and coldspots (blue shade) in KHDSS 2002-2015, space-time interaction results from the Bayesian spatial-temporal model

3.4.2 Model fit diagnostics

Using the final spatial-temporal model we investigated for the model convergence and correlation. As a way of showing the convergence graphically the model convergence we show posterior mean density plots Figure~4.7 and the WinBUGS model convergence graphs are shown in Appendix 5. The convergence graphs show a perfectly symmetrical pattern observed for the estimated parameters indicating that the model captured the true values very well (80).

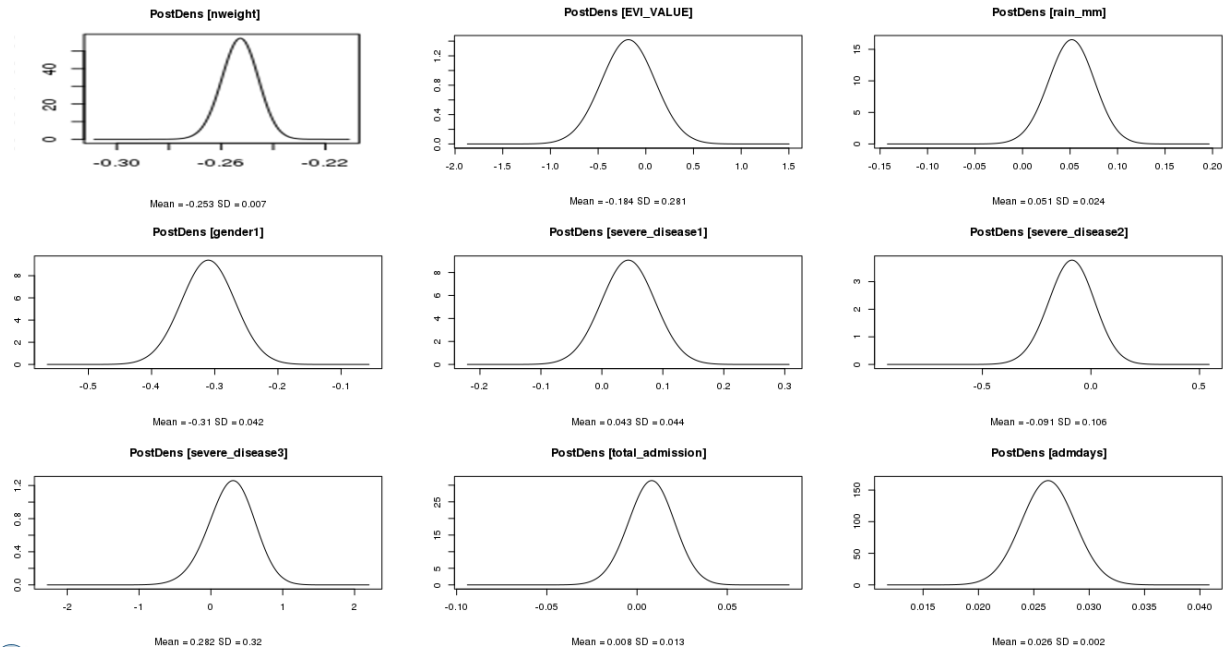


Figure 4.7: Posterior density plots for fixed effects of the NB spatiotemporal model

3.5 Spatial Logistic Model Fit and Diagnostics

3.5.1 Model fit results

Multivariable analysis was done using non-spatial and spatial models for admission specific variables with WHZ scores being the main predictor, environmental variables and sub-location information. Table~4.4 shows the association between different admission predictors and mortality. The model with structured and unstructured random effects had a higher DIC due to the extra spatial parameters. We used the pD that assesses the model complexity to select the most parsimonious model, which is the model with structured and unstructured random effects model as shown on Table~4.4.

Table 4.4: Spatial logisitc model for mortality in Kilifi County Hospital, 2002-2015

	Multi level model		Bayesian Multilevel Random Effects Model (INLA)	Bayesian Spatial Random Effects Model (INLA)
variables	Coeficients (95% C.I)	p-value	Mean (95% Cr .I)	Mean (95% Cr .I)
WHZ	-0.52 (-0.58;-0.46)	<0.001	-0.22 (-0.26;-0.17)	-0.22 (-0.26;-0.17)
Gender:Male	0.08 (-0.1;0.25)	0.353	-0.1 (-0.23;0.03)	-0.1 (-0.23;0.03)
Severe Disease				
1	0.02 (-0.18;0.21)	0.45	-0.25 (-0.4;-0.11)	-0.26 (-0.4;-0.11)
2	0.92 (0.61;1.23)	<0.001	0.54 (0.29;0.78)	0.53 (0.28;0.78)
3	2.68 (2.13;3.23)	<0.001	2.16 (1.69;2.62)	2.16 (1.68;2.61)
Total Admission	-0.28 (-0.37;-0.18)	0.89	-0.32 (-0.4;-0.24)	-0.32 (-0.4;-0.24)
Admission Days	-0.02 (-0.03;0)	0.568	-0.01 (-0.02;0)	-0.01 (-0.02;0)
Time to Readmission	0.02 (0.01;0.02)	<0.001	0.01 (0.001;0.02)	0.01 (0;0.02)
Age in Years	0.02 (-0.01;0.05)	0.188	0.01 (-0.01;0.03)	0.01 (-0.01;0.03)
EVI	-0.18 (-1.36;0.99)	0.008	-0.39 (-1.29;0.5)	-0.45 (-1.37;0.46)
Rainfall	0.09 (0.01;0.18)	0.005	0.06 (-0.02;0.13)	0.06 (-0.01;0.13)
Attributable Fraction	38.30(35.64, 42.82)			24.91(21.04, 30.84)
DIC (pD)			7690.56 (26.37)	7690.93 (22.41)

The model with unstructured and structured random effects model showed several predictors of mortality. WHZ which is one the determinant of malnutrition admission was strongly associated with mortality (p<0.001) and RR = -0.52 CrI= (-0.58;-0.46) in both spatial and non-spatial models respectively. The malnutrition attributable fraction for the non-spatial model was high than in the spatial model 38.30(C.I 35.64,42.82). The attributable fraction was lower in the spatial model, after adjusting for the spatial random effects 24.91(C.I = 21.04;30.84) but this was the most effective model with a $pD = 22.41$.

The weather variables and gender were significantly associated with mortality. The model shows high admission mortality when there is high rainfall, and EVI was significantly associated with a lower risk of admission mortality. The children who had a higher number of admissions and many admission days had a lower probability (~28% and 2% lower odds respectively) of death as shown on Table~4.4. The diagnostic plot for the posterior means are as shown on Figure~4.9

After adjusting for the spatial random effects, the attributable fraction differs by clusters as shown on Figure~4.8. The maps present the clustering of malnutrition attributable mortality after adjusting for the spatial random

effects. Areas on the south of the creek are observed to have a higher (more than 30%) mortality attributable to mortality. Gede location did not have an estimated fraction due to fewer numbers in from the region.

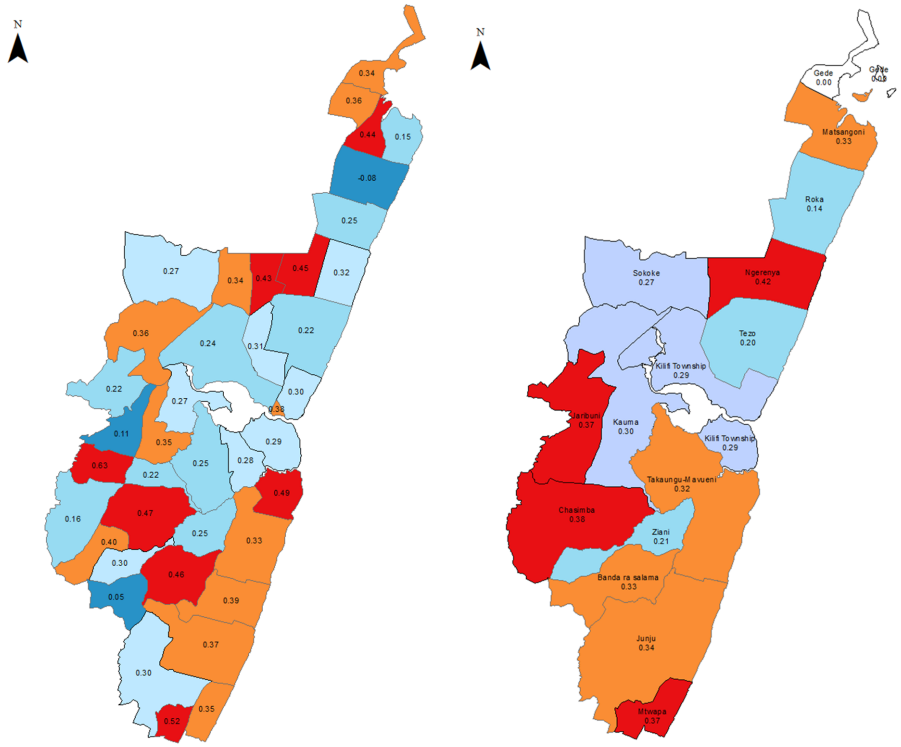


Figure 4.8: Malnutrition attributalble mortality in KHDSS, 2002-2015

3.5.2 Model fit diagnostics

The selected model which had the lowest $pD = 22.41$ was investigated for convergence and correlation. Posterior mean density plots on Figure~4.9 were used to show convergence graphically. The convergence graphs show a perfectly symmetrical pattern observed for the estimated parameters indicating that the model captured the true values very well

(80).

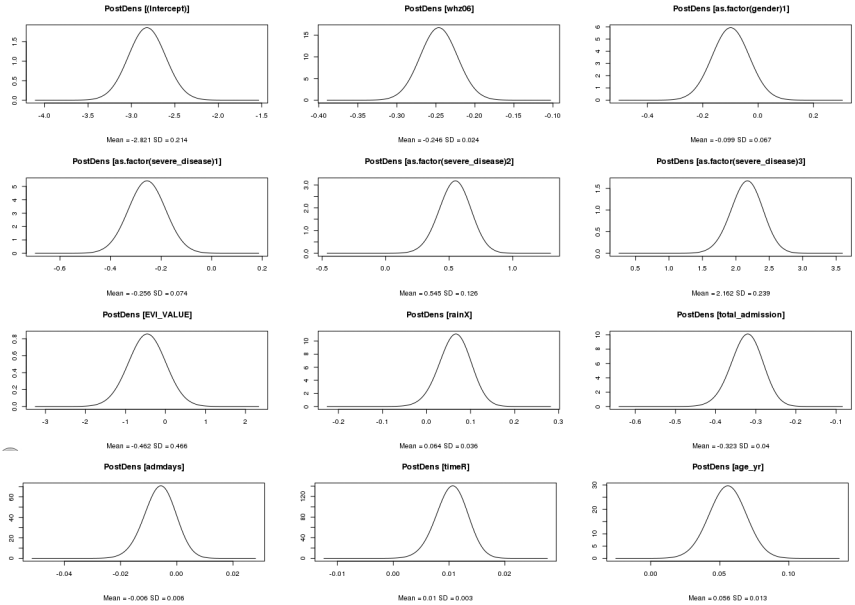


Figure 4.9: Posterior density plots for fixed effects of the logisitc spatial model

CHAPTER 4: DISCUSSION AND CONCLUSION

4.1 Introduction

This chapter discusses the key results from our analysis by comparing these with other studies. We also discuss the strength and limitations of the study. The two main models under discussion are the spatial-temporal model for morbidity and spatial model for mortality.We adjusted for potential confounders and other associated factors, adjusting for spatial and temporal for objective 2 and spatial only for objective 3.

In general, hotspots of malnutrition related admissions were found in North and South areas of the Kilifi town as shown on Figure~4.4. The mortality attributable to malnutrition had a spatial heterogeneity as it was observed in Figure~4.8 and the spatial model in section 3.3.

4.2 Spatial-Temporal Patterns of malnutrition related morbidity

A 19-year analysis done on Kilifi showed that a trend for neonatal admissions and severe diseases. Severe disease was defined as neonatal sepsis, prematurity, neonatal jaundice, neonatal encephalopathy, tetanus and neonatal meningitis.The severe diseases accounted for approximately

75% of the admissions, though seasonality of admissions was not reported (81). Stella et al. reported a seasonality during rainfall peaks for malaria admissions in the Kenya Coast (82). Malaria hotspots and hotspots within hotspots were also observed in the period between 2003 to 2011 (83). From our analysis, a seasonality of malnutrition related admissions was observed in Kilifi County hospital with peaks occurring in July.

The severity of disease was defined as the presence of either gastroenteritis, LRTI, blood and CSF culture positive, malaria and fever or meningitis. The explanatory variables; severe disease, admission weight, days of admission and environmental factors; EVI and rainfall were identified as significant factors from our model. In Somalia, infections and the Enhanced Vegetation Index were observed as key drivers of malnutrition (29). Similarly, in our analysis children with a higher number of severe diseases were associated with an increased risk of a malnutrition admission.

In Malawi, it was reported that remoteness, environmental factors and geographical factors were important drivers of children morbidity (30). Rainfall increased the risk of malnutrition admissions, but EVI was associated with the reduction of malnutrition admission. The malnutrition admissions mostly during the rainy season which was different from a

longitudinal malnutrition study done in Ethiopia where acute malnutrition happened during the dry seasons, though the season was not significant in their study (28). In our analysis, we combined both the spatial and temporal random effects and the environmental variables. In our analysis, rainfall increased the risk of a malnutrition related admission and the EVI index reduced the risk of a malnutrition related admission. This could be possibly be explained by the infections that occur during rainy seasons, i.e., diarrhoea, environmental enteropathy or malaria.

In Somalia, Damaris et al. observed a distinct seasonal variation in wasting of under 5 due to changes in climate, food security and diseases. The peaks of malnutrition were observed during the dry season and were reported to have an elevated effect for the rainy season (16). This can also be used to explain the peaks of malnutrition related re-admissions in Kilifi county hospital during the rainy season.

In a study done in Spain, malnutrition was significantly associated with the number of days a patient stayed in hospital thus increasing the cost of hospital stay (21). In our model, the length of hospital stay reduced the risk of a malnutrition admission. Additionally, higher time to readmission increased the risk of a malnutrition admission. Treatment of malnourished

children is considered an important factor for nutrition specific interventions and programs (1).

4.3 Spatial patterns malnutrition attributable mortality

The report for malnutrition attributable fractions in KHDSS report fractions for the whole study region (3). In our study, we build on the study to report location specific attributable fractions; having recognised the spatial heterogeneity in Kilifi HDSS. The importance of adjusting for spatial location in malnutrition prevalences is advised in a study done DR Congo in 2011 (31).

In our model WHZ was a strong predictor of mortality, a unit increase in WHZ score reduced the risk of death by approximately 20%. In a study done in Kilifi children aged 6 to 60 months malnutrition was a strong predictor of mortality with inpatient deaths attributable mortality reported as 19% using WHZ as the marker and 51% using MUAC (3).

In our spatial-temporal model, the inpatient mortality attributable fraction using WHZ was 21.98% with a confidence band of 18.73%-27.81%. The attributable fractions also showed a spatial heterogeneity as shown on Figure~4.8. Most children are treated in the community, and most of the

mortality occurs without a hospital admission, so the reported attributable fractions might be a lower estimate. Similarly, the maps show spatial variation in attributable fractions emphasising on the role of location in malnutrition related modelling (31).

Children who were older had a lower risk of death adjusting for the location and WHZ which is comparable to what was reported in a Tanzanian study where the prevalence of malnutrition decreased in older children (8). Children who had a higher number of admission days had a lower risk of a malnutrition admission in our model. The environmental factors were not significantly associated with mortality in both the multilevel and spatial-temporal models.

4.4 Computational aspects of Spatial-Temporal Modelling

The data we use for our analysis was imbalanced; some individuals had more admission events than others at different times. A better computing speed for approximation was needed. To handle this we also excluded individuals who had a single admission event for a reduced computing time.

Sabrina et al. in 2014 review recommended the importance of combining spatial and temporal components in understanding the compounded

phenomenon of malnutrition (15). With the different recommendations of spatial models to get an improved understanding of malnutrition; some of them require higher computing resources for an imbalanced and large data sets (15,84). In our models, we utilise a spatial-temporal and spatial model to understand malnutrition related admissions and malnutrition attributable mortality using KHDSS data.

In most spatial-temporal models with imbalanced data, the temporal components are fitted as a random effect or a Bayesian generalised linear mixed model is implemented (71,84,85). We performed the latter in our analysis; a spatial-temporal negative binomial model, which is a family of GLM using a Bayesian approach. In the spatial mortality model, we did not include the temporal component since mortality is not a changing covariate for the individuals.

Due to the complexity of our data set over 23,000 admissions with 17,000 individuals, we utilised the nested Laplace approximation to handle our spatial-temporal model. The R-INLA interface provided an interface to handle the spatial-temporal negative binomial model with a better computing capability. This is similar to what Musenge et al. in 2013 utilised. In the analysis, they used Integrated Nested Laplace

Approximation (INLA) to fit spatial models with Gaussian Markov Random Fields which had better computing time over MCMC (38). This approach has also been applied in other health research methods and has been shown to give faster and accurate results of posterior estimates (27,36,69).

4.5 Strengths and limitations of the study

The strength and limitations of the results and discussion in this research project are discussed in four broad categories;

- Exposures and outcomes classification
- Modelling approach to catering for confounding and effect modification
- sub-Saharan Africa relevance of the results
- Computational requirements

4.5.1 Exposures and outcomes classification

A major drawback in the utlisation of Negative Binomial model over the truncated Negative Binomial model in modelling the morbidity variable due to lack of implementations truncated negative binomial (TNB) in INLA or BUGS packages (86). The main interest for our modelling was to adjust for the spatial and temproral random effects so we considered

the spatial-temporal negative binomial model over the truncated negative binomial. New approaches which would conisder TNB in INLA implemtation would further help in modelling this.

4.5.2 Modelling approach to catering for confounding and effect modification

Detailed spatial-temporal data has been a major drawback in understanding the epidemiology of malnutrition (15). One of the strengths of the data points used in this analysis is that they are consistently corrected, updated, and a proper audit trail is in place. However, in our modelling, selection bias may have occurred since we selected individuals who had two or more admissions and kept the admissions events up to the highest number of a malnutrition admission for objective 2. This was handled by using a negative binomial regression model which caters for this (66). Additionally, we used geo-additive regression models to cater for multiple confounding variables and also catered for time-varying covariates in the spatial-temporal model.

Bayesian models helped in catering for the complex structured data for inpatient admissions in KCH. This approach helped in adjusting for geographical location confounding and including the prior information on the distribution of the data.

4.5.3 sub-Saharan Africa relevance of the results

The results from our analysis show the importance of adjusting for the spatial random effects and spatial covariates. This is comparable to what other studies have shown in sub-Saharan Africa, especially temporality of admissions and spatial heterogeneity (29,31,82).

Environmental variables, EVI and rainfall, have been shown to be key drivers of malnutrition and infectious diseases. In Somalia, Kenya and the Democratic Republic of Congo, EVI was shown to reduce the risk of malnutrition in children. However, temporality was not considered in the different projects (29,31,82). Similarly, this was observed in our spatial-temporal models.

Logistic regression has been applied in different studies to estimate attributable fraction (3,76,87) but few studies adjust and report spatially adjusted attributable fractions. In our spatial model for mortality, we adjusted for spatial random effects, and as shown in section 3.3, heterogeneity of attributable fractions is observed.

4.5.4 Computational requirements

Bayesian inference using MCMC required high computing capabilities compared to the Laplace Approximation in handling imbalanced data. This led to fitting the spatial-temporal Bayesian models in INLA alone. This was due to the complex structure of the data. We treated the spatial component as a random effect without considering the temporality of the locations since children of age 6 months to 15 years rarely migrate.

Computational capability was one of the things required for bulk download of MODIS data for rainfall and EVI. The MODIS reprojection process for EVI was a challenge but using an R Slurm Job was a better way to handle this (53,88). Slurm job is an open source platform that clusters large jobs with a scheduling system in Linux clusters. This helps in performance and also notifications of jobs when done (88).

4.6 Conclusion

A better understanding of the spatial and environmental factors that contribute to malnutrition related re-admissions can be used to advocate for and develop earlier and more appropriate responses and provide an indication of future trends and the potential impact of interventions.

Admission with malnutrition is an important marker of re-admission, understanding the spatial risk distribution can be used to understand mechanisms of post discharge mortality. This would also help in the research of tremendous costs associated with the treatment of morbidity that could be prevented through better child nutrition.

Campaigns providing food and or vitamin or other supplements can help reduce deaths in Kenyan children and building more health facilities to reduce the distance of travel to care is highly recommendable.

The developments of spatial-temporal models account for individual, location and temporal aspects in health-related data analysis. As we have observed, spatial-temporal Bayesian models provide better fitting models compared to non-spatial models. Especially in targeted interventions, maps produced from Bayesian models can be of great importance to policy makers and financing organisations.

Studies considering malnutrition related interventions should consider time and locations in their approach. Available health care significantly reduce the risk of malnutrition as we have observed, so having medical facilities closer and evenly distributed can help reduce the burden of mortality and morbidity.

A more sustainable solution would be to empower and equip smaller facilities. This is drawn from our results that those closer to the Kilifi Hospital were most likely not to die and also not to be readmitted in relation to other health facilities.

REFERENCES

1. Black RE, Victora CG, Walker SP, Bhutta ZA, Christian P, Onis M de, et al. Maternal and child undernutrition and overweight in low-income and middle-income countries. Lancet. 2013;382(9890):427–51.

2. Black RE, Allen LH, Bhutta ZA, Caulfield LE, Onis M de, Ezzati M, et al. Maternal and child undernutrition: global and regional exposures and health consequences. Lancet. 2008;371(9608):243–60.

3. Bejon P, Mohammed S, Mwangi I, Atkinson SH, Osier F, Peshu N, et al. Fraction of all hospital admissions and deaths attributable to malnutrition among children in rural Kenya. Am J Clin Nutr. 2008;88(6):1626–31.

4. Bhutta ZA, Ahmed T, Black RE, Cousens S, Dewey K, Giugliani E, et al. What works? Interventions for maternal and child undernutrition and survival. The Lancet. 2008;371(9610):417–40.

5. Grantham-McGregor S, Baker-Henningham H. Review of the evidence linking protein and energy to mental development. Public health nutrition. 2005;8(7a):1191–201.

6. Forouzanfar MH, Alexander L, Anderson HR, Bachman VF, Biryukov S, Brauer M, et al. Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks in 188 countries, 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. The Lancet. 2015 Dec;386(10010):2287–323.

7. UNICEF, UNICEF, others. WHO-world bank joint child malnutrition estimates. 2012. New York, UNICEF, Geneva, WHO, Washington, World Bank [Internet]. 2012 [cited 2017 May 16]; Available from: http://www.who.int/nutgrowthdb/jme{_}unicef{_}who{_}wb.pdf

8. Sunguya BFP, Koola JI, Atkinson S. Infections associated with severe malnutrition among hospitalised children in East Africa. Tanzania health research bulletin. 2006 Sep;8(3):189–92.

9. Institute I, Haddad LJ, Hawkes C, Emorn U, Achadi E, Ahuja A, et al. Global Nutrition Report 2015: Actions and accountability to advance nutrition and sustainable development. 2015 [cited 2017 Jun 19]; Available from: https://books.google.co.ke/books?id=kdqICgAAQBAJ

10. Burgos R, Sarto B, Elío I, Planas M, Forga M, Cantón A, et al. Prevalencia De Malnutrición Y Sus Factores Etiológicos en Hospitales. Nutricion Hospitalaria. 2012;27(2):469–76.

11. Ngare DK, Muttunga JN. Prevalence of malnutrition in Kenya. East African medical journal. 1999 Jul;76(7):376–80.

12. Gebreslasie MT. A review of spatial technologies with applications for malaria transmission modelling and control

in Africa. Geospatial health. 2015;10(2):328.

13. Manda S, Masenyetse L, Cai B, Meyer R. Mapping HIV prevalence using population and antenatal sentinel-based HIV surveys: a multi-stage approach. Population health metrics. 2015;13:22.

14. Ngesa O, Mwambi H, Achia T, Auvert B, Taljaard D, Lagarde E, et al. Bayesian Spatial Semi-Parametric Modeling of HIV Variation in Kenya. Paraskevis D, editor. PLoS ONE. 2014 Jul;9(7):e103299.

15. Marx S, Phalkey R, Aranda-Jan CB, Profe J, Sauerborn R, Höfle B, et al. Geographic information analysis and web-based geoportals to explore malnutrition in Sub-Saharan Africa: a systematic review of approaches. BMC Public Health. 2014;14(1):1189.

16. Kinyoki DK, Berkley JA, Moloney GM, Odundo EO, Kandala NB, Noor AM. Space-time mapping of wasting among children under the age of five years in Somalia from 2007 to 2010. Spatial and Spatio-temporal Epidemiology. 2016;16:77–87.

17. Fletcher-Lartey SM, Caprarelli G. Application of gis technology in public health: Successes and challenges. Parasitology. 2016;143(4):401–15.

18. Mariella L, Tarantino M. Spatial Temporal Conditional Auto-Regressive Model: A New Autoregressive Matrix. Austrian Journal of Statistics. 2010;39(3):223–44.

19. Lawson AB. Bayesian disease mapping: Hierarchical modeling in spatial epidemiology. CRC press; 2013.

20. Waitzberg DL, Caiaffa WT, Correia MIT. Hospital malnutrition: The brazilian national survey (ibranutri): A study of 4000 patients. Nutrition. 2001;17(7):573–80.

21. Álvarez-Hernández J, Planas Vilá M, León-Sanz M, García de Lorenzo A, Celaya-Pérez S, García-Lorda P, et al. Prevalence and costs of malnutrition in hospitalized patients; the predyces study. Nutrición hospitalaria. 2012;27(4).

22. Campbell-lendrum D, Prüss-üstün A, Diarmid H. Climate change Quantifying the health impact at national and local levels. 2007 [cited 2017 Sep 26];1(14). Available from: http://www.unscn.org/layout/modules/resources/files/Climate_change_quantifying_the_impacts.pdf

23. Lloyd SJ, Kovats RS, Chalabi Z. Climate change, crop yields, and undernutrition: development of a model to quantify the impact of climate scenarios on child undernutrition. Environ Health Perspect. 2011/08/17. 2011;119(12):1817–23.

24. Rytter MJH, Kolte L, Briend A, Friis H, Christensen VB. The immune system in children with malnutrition–a

systematic review. PloS one. 2014;9(8):e105017.

25. Musenge E, Vounatsou P, Kahn K. Space-time confounding adjusted determinants of child HIV/TB mortality for large zero-inflated data in rural South Africa. Spatial and Spatio-temporal Epidemiology. 2011;2(4):205–17.

26. Fahrmeir L, Lang S. Bayesian inference for generalized additive mixed models based on Markov random field priors. Journal of the Royal Statistical Society: Series C (Applied Statistics). 2001;50(2):201–20.

27. Martins TG, Simpson D, Lindgren F, Rue H. Computational Statistics and Data Analysis Bayesian computing with INLA: New features. Computational Statistics and Data Analysis. 2013;67:68–83.

28. Joosten KF, Hulst JM, Worku A, Joosten K, Hulst J, Antwi S, et al. Prevalence of malnutrition in pediatric hospital patients. Current Opinion in Pediatrics. 2008 Oct;20(5):590–6.

29. Kinyoki DK, Berkley JA, Moloney GM, Kandala N-B, Noor AM. Predictors of the risk of malnutrition among children under the age of 5 years in Somalia. Public health nutrition. 2015;18(17):1–9.

30. Kandala N-B, Magadi MA, Madise NJ. An investigation of district spatial variations of childhood diarrhoea and fever morbidity in Malawi. Social Science & Medicine. 2006;62(5):1138–52.

31. Kandala N-B, Madungu TP, Emina JB, Nzita KP, Cappuccio FP, Black R, et al. Malnutrition among children under the age of five in the Democratic Republic of Congo (DRC): does geographic location matter? BMC Public Health. 2011;11(1):261.

32. Sartorius BK, Kahn K, Vounatsou P, Collinson MA, Tollman SM, Hammer G, et al. Young and vulnerable: Spatial-temporal trends and risk factors for infant mortality in rural South Africa (Agincourt), 1992-2007. BMC Public Health. 2010;10(1):645.

33. Mogeni P, Williams TN, Fegan G, Nyundo C, Bauni E, Mwai K, et al. Age, Spatial, and Temporal Variations in Hospital Admissions with Malaria in Kilifi County, Kenya: A 25-Year Longitudinal Observational Study. Grais RF, editor. PLOS Medicine. 2016 Jun;13(6):e1002047.

34. Portier CJ, Thigpen Tart K, Carter SR, Dilworth CH, Grambsch AE GJ et al. A human health perspective on climate change: A report outlining the research needs on the human health effects of climate change. 2010;1–80. Available from: www.niehs.nih.gov/climatereport

35. Wolf J, Armstrong B. The association of season and temperature with adverse pregnancy outcome in two German states, a time-series analysis. PLoS One. 2012;7(7):e40228.

36. Noor AM, Kinyoki DK, Mundia CW, Kabaria CW, Mutua JW, Alegana VA, et al. The changing risk of Plasmodium

falciparum malaria infection in Africa: 2000–10: a spatial and temporal analysis of transmission intensity. The Lancet. 2014;383(9930):1739–47.

37. Speybroeck N, editor. Using Structured Additive Regression Models to Estimate Risk Factors of Malaria: Analysis of 2010 Malawi Malaria Indicator Survey Data. PLoS ONE. 2014 Jul;9(7):e101116.

38. Musenge E, Chirwa TF, Kahn K, Vounatsou P. 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(1):86–98.

39. Soares Magalhães RJ, Langa A, Pedro JM, Sousa-Figueiredo JC, Clements AC, Vaz Nery S. Role of malnutrition and parasite infections in the spatial variation in children's anaemia risk in northern Angola. Geospatial health. 2013 May;7(2):341.

40. Rytter MJH, Kolte L, Briend A, Friis H, Christensen VB, Black R, et al. The Immune System in Children with Malnutrition—A Systematic Review. Akiyama T, editor. PLoS ONE. 2014 Aug;9(8):e105017.

41. Kazembe LN, Chirwa TF, Simbeye JS, Namangale JJ. Applications of bayesian approach in modelling risk of malaria-related hospital mortality. BMC Medical Research Methodology. 2008;8:6.

42. De Savigny D, Mayombana C, Mwageni E, Masanja H, Minhaj A, Mkilindi Y, et al. Care-seeking patterns for fatal malaria in tanzania. Malaria Journal. 2004;3(1):27.

43. Scott JAG, Bauni E, Moisi JC, Ojal J, Gatakaa H, Nyundo C, et al. Profile: The Kilifi health and demographic surveillance system (KHDSS). International Journal of Epidemiology. 2012;41(April):650–7.

44. Berkley J., Mwangi I., Griffiths K., Ahmed I., Mithwani S., English M., Newton C. MK. Assessment of Severe Malnutrition Among Hospitalized Children in Rural Kenya. American Medical Association. 2005;294(5):591–7.

45. Hinz S, DuBois P, Stephens J. MySQL 5.7 reference manual [Internet]. John Wiley & Sons, Inc. 2016 [cited 2017 Jun 6]. Available from: http://dl.acm.org/citation.cfm?id=558011

46. Greenspan J, Bulger B. MySQL/php database applications [Internet]. John Wiley & Sons, Inc. 2001 [cited 2017 Jun 6]. Available from: http://dl.acm.org/citation.cfm?id=558011

47. Leroy J. ZSCORE06: Stata module to calculate anthropometric z-scores using the 2006 WHO child growth standards. 2011 [cited 2017 May 16]; Available from: http://econpapers.repec.org/RePEc:boc:bocode:s457279

48. Onis M de, Onyango AW, Borghi E, Siyam A, Nishida C, Siekmann J. Development of a WHO growth reference

for school-aged children and adolescents. Bulletin of the World Health Organization. 2007 Sep;85(9):660–7.

49. Kinyoki DK, Kandala N-B, Manda SO, Krainski ET, Fuglstad G-A, Moloney GM, et al. Assessing comorbidity and correlates of wasting and stunting among children in Somalia using cross-sectional household surveys: 2007 to 2010. BMJ Open. 2016;6(3):e009854.

50. Miller AJ. Subset selection in regression. Second Edi. Isham V, editor. Florida: Chapman & Hall/CRC; 2002.

51. Tuck SL, Phillips HRP, Hintzen RE, Scharlemann JPW, Purvis A, Hudson LN. MODISTools - downloading and processing MODIS remotely sensed data in R. Ecology and Evolution. 2014;4(24):4658–68.

52. Mattiuzzi M. MODIS: Acquisition and processing of modis products [Internet]. 2017 [cited 2017 May 16]. Available from: https://CRAN.R-project.org/package=MODIS

53. R Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2016.

54. Hijmans RJ. Raster: Geographic data analysis and modeling [Internet]. 2016 [cited 2017 May 16]. Available from: https://CRAN.R-project.org/package=raster

55. Coghlan A. A Little Book of R For Time Series Release 0.2. 2017 [cited 2017 Jul 19]; Available from: https://media.readthedocs.org/pdf/a-little-book-of-r-for-time-series/latest/a-little-book-of-r-for-time-series.pdf

56. Martinez EZ, Silva EASD, Fabbro ALD. A SARIMA forecasting model to predict the number of cases of dengue in Campinas, State of São Paulo, Brazil. Revista da Sociedade Brasileira de Medicina Tropical. 2011;44(4):436–40.

57. Newton HJ, Cox NJ, Garrett JM, Pagano M, Royston JP. STATA September 2000 Technical STB-57 Bulletin. Stata Press [Internet]. 2000 [cited 2017 May 16]; Available from: http://www.stata.com/products/stb/journals/stb57.pdf

58. Yaffee RA. A Review of Stata versions 9 and 10 Time Series and Forecasting capability. JSS Journal of Statistical Software [Internet]. [cited 2017 May 16]; Available from: http://www.jstatsoft.org/

59. Paradis E, Claude J, Strimmer K. APE: Analyses of phylogenetics and evolution in R language. Bioinformatics. 2004;20:289–90.

60. Gittleman JL, Kot M. Adaptation: statistics and a null model for estimating phylogenetic effects. 1990;39(3):227–41.

61. Lai D. A simple test for spatial pattern in regional health data - Letter to the editor. Statistics in Medicine. 1995;14:1393–4.

62. Coleman M, Coleman M, Mabuza AM, Kok G, Coetzee M, Durrheim DN, et al. Using the SaTScan method to

detect local malaria clusters for guiding malaria control programmes. Malaria Journal. 2009;8(1):68.

63. Hjalmars U, Kulldorff M, Gustafsson G, Nagarwalla N. Childhood leukaemia in Sweden: using GIS and a spatial scan statistic for cluster detection. Stat Med. 1996;15(7-9):707–15.

64. Thomsen JL, Parner ET. Methods for analysing recurrent events in health care data. examples from admissions in ebeltoft health promotion project. Family Practice Advance Access. 2006;

65. Ridout M, Hinde J, DeméAtrio CG. A score test for testing a zero-inflated poisson regression model against zero-inflated negative binomial alternatives. Biometrics. 2001;57(1):219–23.

66. Hilbe JM. Negative binomial regression. Cambridge University Press; 2011.

67. Cameron AC, Trivedi PK. Regression analysis of count data. Vol. 53. Cambridge university press; 2013.

68. Guikema SD, Goffelt JP. A flexible count data regression model for risk analysis. Risk Analysis. 2008;28(1):213–23.

69. Blangiardo M, Cameletti M, Baio G, Rue H. Spatial and spatio-temporal models with R-INLA. 2013;

70. Brandt PT, Williams JT. A linear poisson autoregressive model: The poisson ar(p) model. Political Analysis. 2001;9(2):164–84.

71. Neelon B, Ghosh P, Loebs PF. A Spatial Poisson Hurdle Model for Exploring Geographic Variation in Emergency Department Visits. Available from: http://www.people.vcu.edu/{~}dbandyop/pubh8472/PoissonHurdle.pdf

72. Martino S, Rue H. Implementing Approximate Bayesian Inference using Integrated Nested Laplace Approximation: a manual for the inla program. 2009.

73. Lindgren F, Rue H avard. Bayesian Spatial Modelling with R-INLA. Journal of Statistical Software. 2015;63(19):1–26.

74. Geman S, Geman D. Stochastic relaxation, gibbs distributions and the bayesian restoration of images*. Journal of Applied Statistics. 1993;20(5-6):25–62.

75. Vaughn BK. Data analysis using regression and multilevel/hierarchical models, by gelman, a., & hill, j. Journal of Educational Measurement. 2008;45(1):94–7.

76. Mwangi TW, Ross A, Snow RW, Marsh K. Case Definitions of Clinical Malaria under Different Transmission Conditions in Kilifi District, Kenya. 2005;

77. Spiegelhalter DJ, Best NG, Carlin BP, Linde A van der. Bayesian measures of model complexity and fit. Journal

of the Royal Statistical Society: Series B (Statistical Methodology). 2002 Oct;64(4):583–639.

78. Boga M, Davies A, Kamuya D, Kinyanjui SM, Kivaya E, Kombe F, et al. Strengthening the Informed Consent Process in International Health Research through Community Engagement: The KEMRI-Wellcome Trust Research Programme Experience. PLoS Medicine. 2011 Sep;8(9):e1001089.

79. Rue H, Martino S, Chopin N. Approximate bayesian inference for latent gaussian models by using integrated nested laplace approximations.pdf. Journal of the Royal Statistical Society. 2009;Series B(71):319–92.

80. Wilgenbusch J. AWTY: A system for graphical exploration of mcmc convergence in bayesian phylogenetic inference. http://ceb csit fsu edu/awty. 2004;

81. Mwaniki MK, Gatakaa HW, Mturi FN, Chesaro CR, Chuma JM, Peshu NM, et al. An increase in the burden of neonatal admissions to a rural district hospital in Kenya over 19 years. BMC Public Health. 2010/10/12. 2010;10:591.

82. Karuri SW, Snow RW. Forecasting paediatric malaria admissions on the Kenya Coast using rainfall. Global health action. 2016;9:29876.

83. Bejon P, Williams TN, Nyundo C, Hay SI, Benz D, Gething PW, et al. A micro-epidemiological analysis of febrile malaria in Coastal Kenya showing hotspots within hotspots. eLife. 2014;3:e02130.

84. Lesaffre E, Lawson AB. Bayesian biostatistics. John Wiley & Sons; 2012.

85. Cook RJ, Lawless J. The statistical analysis of recurrent events. Springer Science & Business Media; 2007.

86. Blangiardo M, Cameletti M. Spatial and spatio-temporal bayesian models with r-inla. John Wiley & Sons; 2015.

87. Smith T, Schellenberg JA, Hayes R. Attributable fraction estimates and case definitions for malaria in endemic areas. Statistics in medicine. 1994 Nov;13(22):2345–58.

88. Jette MA, Yoo AB, Grondona M. SLURM: Simple linux utility for resource management. In: In lecture notes in computer science: Proceedings of job scheduling strategies for parallel processing (jsspp) 2003. Springer-Verlag; 2002. pp. 44–60.

APPENDICES

Appendix 1 - Plagiarism declaration report

PLAGIARISM DECLARATION TO BE SIGNED BY ALL HIGHER DEGREE STUDENTS

SENATE PLAGIARISM POLICY: APPENDIX ONE

I __KM Wambui_____ (Student number: ___1504769___) am a student registered for the degree of __MSc Biostatistics_____ in the academic year _2016/2017_____.

I hereby declare the following:

- I am aware that plagiarism (the use of someone else's work without their permission and/or without acknowledging the original source) is wrong.
- I confirm that the work submitted for assessment for the above degree is my own unaided work except where I have explicitly indicated otherwise.
- I have followed the required conventions in referencing the thoughts and ideas of others.
- I understand that the University of the Witwatersrand may take disciplinary action against me if there is a belief that this is not my own unaided work or that I have failed to acknowledge the source of the ideas or words in my writing.

Signature: _____ Date: 7 September 2017

Clearance Certificate

R14/49 Mr Wambui Kennedy Mwai

HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)

CLEARANCE CERTIFICATE NO. M1611104

NAME: (Principal Investigator)	Mr Wambui Kennedy Mwai
DEPARTMENT:	School of Public Health Kilifi County Hospital (KCH) Paediatric Ward, Kenya
PROJECT TITLE:	Space-Time Patterns of Child Mortality and Morbidity Attributable to Malnutrition in Kilifi County during 2002-2015, Kenya: A Retrospective Case Control
DATE CONSIDERED:	25/11/2016
DECISION:	Approved
CONDITIONS:	South African Human Research Ethics Committees (HRECs) have no standing outside South Africa. Ethics approval is also required from local HRECs in Kenya
SUPERVISOR:	Eustasius Musenge
APPROVED BY:	Professor P Cleaton-Jones,Chairperson, HREC (Medical)
DATE OF APPROVAL:	30/11/2016

This clearance certificate is valid for 5 years from date of approval. Extension may be applied for.

DECLARATION OF INVESTIGATORS

To be completed in duplicate and **ONE COPY** returned to the Research Office Secretary in Room 301, Third Floor, Faculty of Health Sciences, Phillip Tobias Building, 29 Princess of Wales Terrace, Parktown, 2193, University of the Witwatersrand. I/we fully understand the conditions under which I am/we are authorized to carry out the above-mentioned research and I/we undertake to ensure compliance with these conditions. Should any departure be contemplated, from the research protocol as approved, I/we undertake to resubmit the application to the Committee. **I agree to submit a yearly progress report**. The date for annual re-certification will be one year after the date of convened meeting where the study was initially reviewed. In this case, the study was initially reviewed in November and will therefore be due in the month of November each year. Unreported changes to the application may invalidate the clearance given by the HREC (Medical).

Principal Investigator Signature Date

PLEASE QUOTE THE PROTOCOL NUMBER IN ALL ENQUIRIES

Appendix 2 - Number of malnutrition admission cases by month.

Table 16.5: Number of malnutrition admission cases by month in KCH admissions from KHDSS, 2002-2015

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Total
2002				14	51	25	26	19	12	17	41	82	287
2003	82	48	43	51	69	75	114	77	61	57	63	80	820
2004	67	54	66	60	101	44	80	84	56	45	45	96	798
2005	84	48	47	54	67	56	81	51	34	38	34	63	657
2006	58	70	65	58	64	80	74	52	42	35	40	53	691
2007	64	62	40	26	32	47	51	41	34	35	40	36	508
2008	42	26	27	33	32	43	47	42	39	36	21	31	419
2009	36	43	27	41	39	76	83	65	20	42	44	50	566
2010	43	53	48	28	51	56	47	34	41	23	29	32	485
2011	46	44	45	52	48	43	38	17	18	21	38	39	449
2012	37	44	36	19	28	34	49	27	23	27	36	11	371
2013	16	18	15	17	26	28	39	17	20	26	26	15	263
2014	36	21	27	35	32	45	61	23	30	35	35	44	424
2015	49	26	33	35	29	52	40	26	24	25	24	27	390
Total	660	557	519	523	669	704	830	575	454	462	516	659	7,128

Table 16.6: Number of deaths by month in KCH admissions from KHDSS, 2002-2015

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Total
2002				1	0	2	1	2	1	1	2	4	14
2003	12	7	4	5	11	12	19	15	6	9	8	6	114
2004	9	4	8	8	15	4	10	9	4	11	7	5	94
2005	12	4	9	6	7	6	4	0	6	6	3	6	69
2006	7	8	5	3	12	11	8	11	7	8	6	9	95
2007	4	6	1	1	7	5	4	7	4	8	7	6	60
2008	5	2	2	2	4	5	2	5	4	5	5	2	43
2009	4	4	5	6	13	11	16	8	3	6	8	8	92
2010	10	6	3	5	7	8	7	5	3	7	3	6	70
2011	3	6	11	3	11	4	7	3	4	3	4	8	67
2012	6	10	7	4	2	5	6	5	3	9	8	2	67
2013	7	1	4	3	8	3	3	4	3	3	2	5	46
2014	8	2	3	5	6	5	5	9	6	5	3	12	69
2015	5	3	4	8	5	9	4	2	2	4	2	7	55
Total	92	63	66	60	108	90	96	85	56	85	68	86	955

Appendix 3 - Time series plots

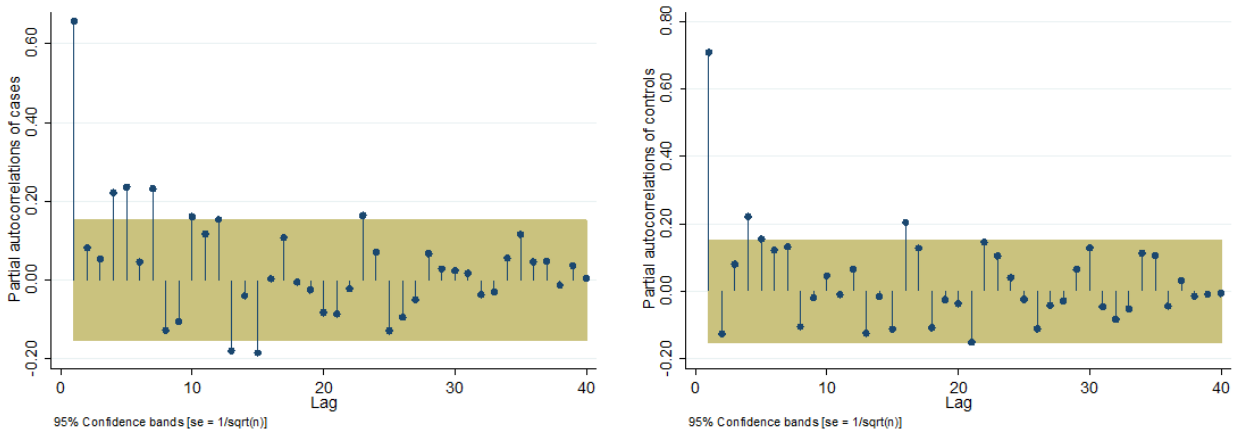


Figure 16.10: A partial autocorrelation function calculated for both cases and controls

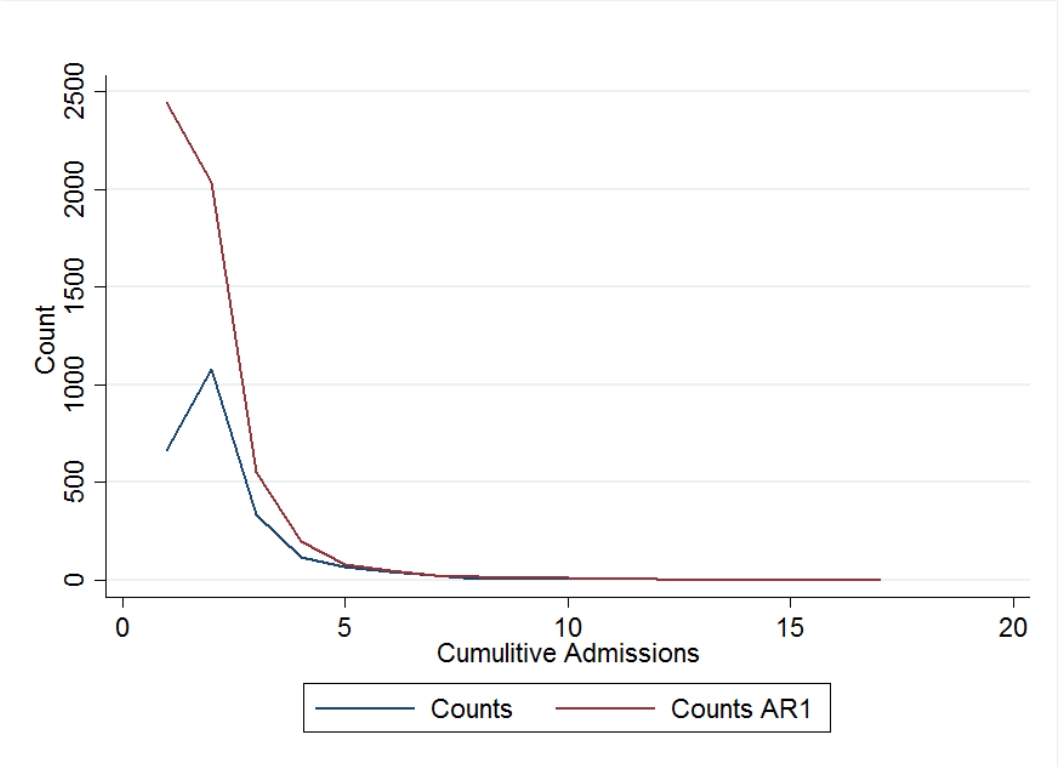
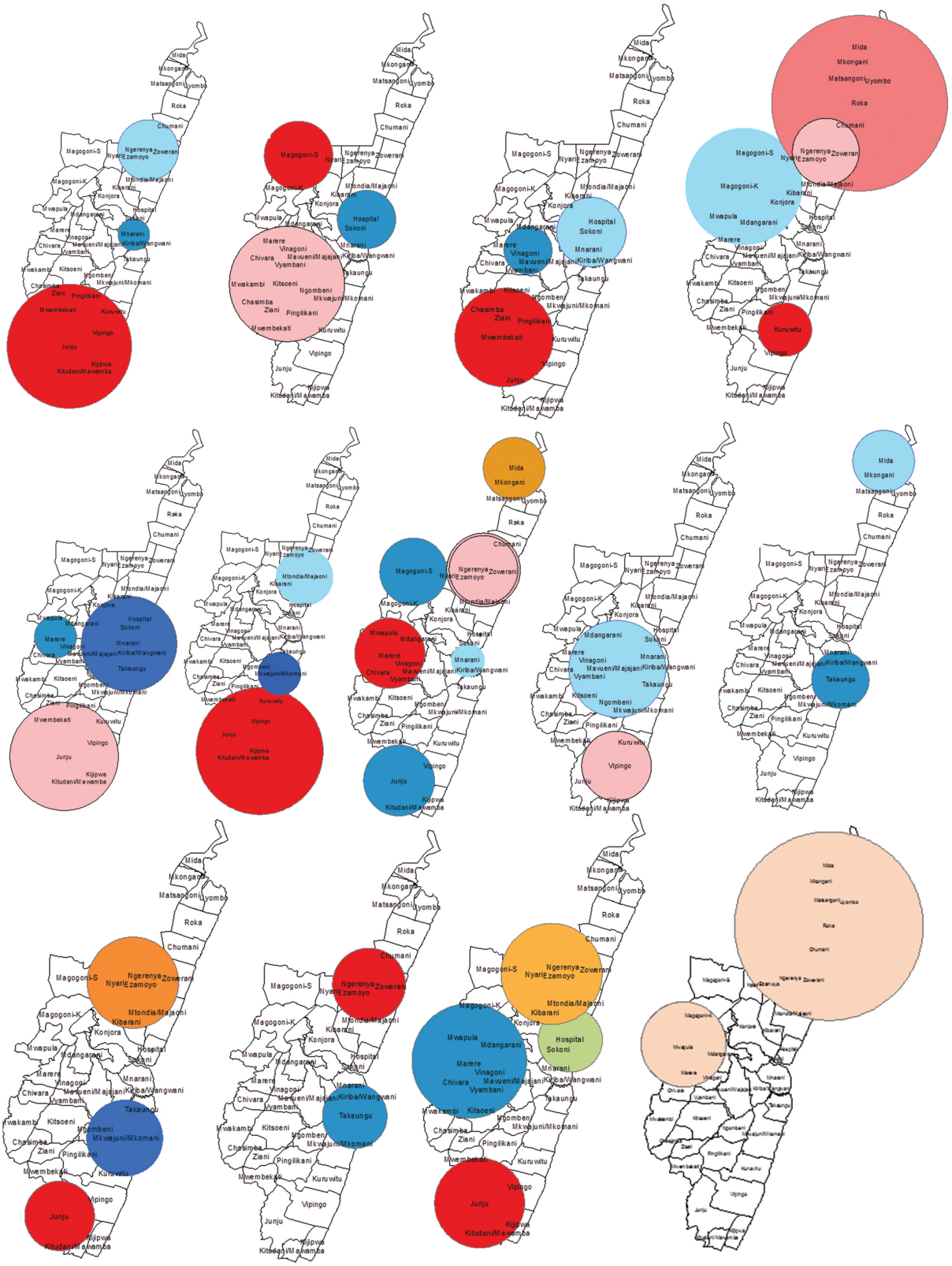


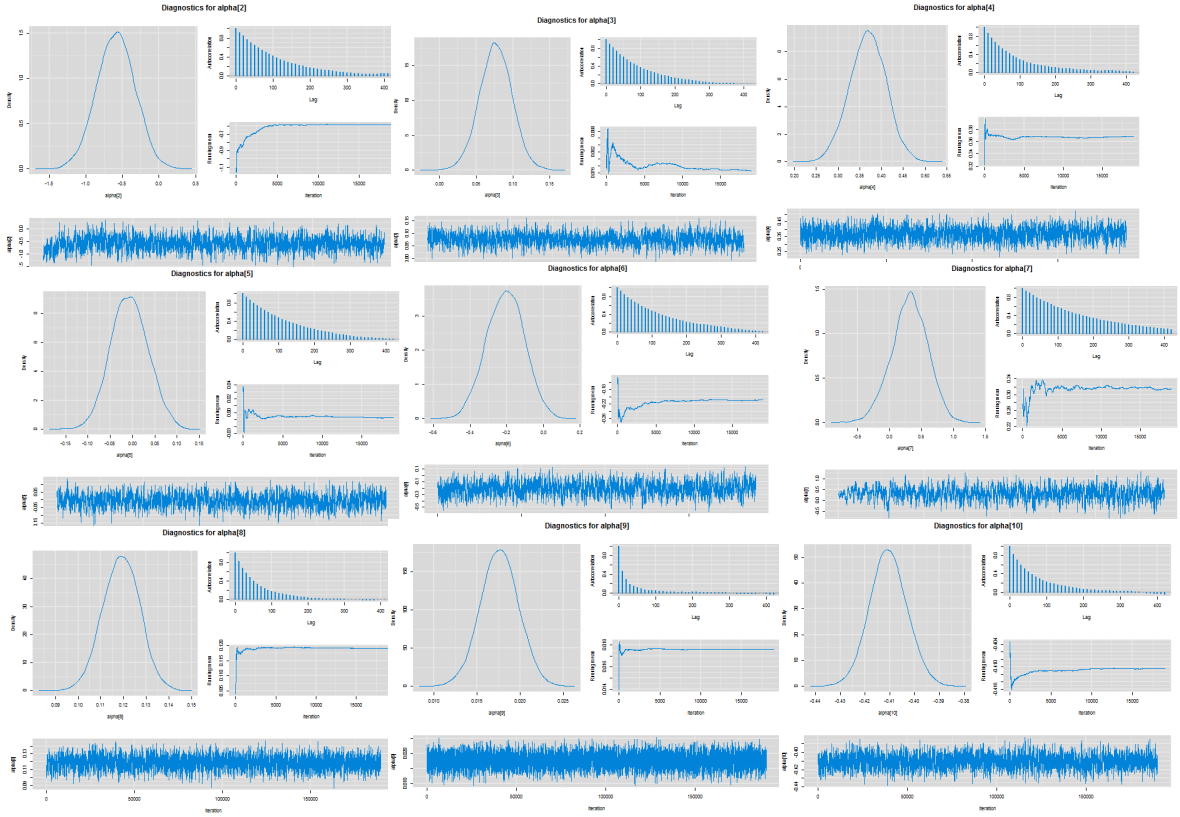
Figure 16.11: A test of AR1 for the temporal variable used in the Spatial - Temporal model.

Appendix 4 - Spatial Temporal Clustering using SaTScan for 2003-2015



Appendix 5 - Winbugs diagnostic plots for convergence.

alpha2 -alpha10 shows our covariates in the model.



Appendix 3 - Analysis Codes

The full codes are available at https://github.com/Keniajin/msc_thesis_2017

C.1 Morbidity Analysis

C.1.1 Stata Multi-level model

```
menbreg cumulitive_count EVI_VALUE  rain_mm  i.nsex i.severe_disease total_admission  ///
admdays nweight ,  exposure(nagem)  nolog || sublocs:
```

C.1.2 WinBUGS Model

```
model {
  for (i in 1:N) {
    # Likelihood
    cumulitive_count[i] ~ dnegbin(p[i], r)
    p[i] <- r / (mu[i] + r)
    log(mu[i]) <-
      log(nagem[i]) + alpha[1] + alpha[2] * EVI_VALUE[i] + alpha[3] * rain_mm[i] + alpha[4] *
      nsex[i] +
      alpha[5] * equals(severe_disease[i], 1) + alpha[6] * equals(severe_disease[i], 2) +
      alpha[7] * equals(severe_disease[i], 3) + alpha[8] * total_adm[i] + alpha[9] *
      admDays[i] + alpha[10] * nweight[i] + Phi[sublocation[i]]
  }#end loop

  ##########
  #   Priors   #
  ##########
  #r
  r ~ dcat(pi[])

  ## 1:11 is the number of succcessful admissions
```

```
  for (i in 1:7) {
    pi[i] <- 1 / 7
  }
  ### Define the priors for the model parameters specification
  # Baseline Covariate Coefficient
  alpha[1] ~ dflat()

  for (j in 2:10) {
    alpha[j] ~ dnorm(0, 0.001)
  }
  # Bivariate CAR Prior for Phi -- Spatial Main Effects
  Phi[1:40] ~ car.normal(adj[], weights[], num[], tau)     # num specifies no. of neighbors
  for (i in 1:sumNumNeigh) {
    weights[i] <- 1
  }
### prior for tau
  tau  ~ dgamma(0.5, 0.0005)
}#end model
```

C.1.3 Bayesian Multilevel Spatial Random Effects Model (INLA)

```
## Mapping
library(rgeos)
library(maptools)
library("ggplot2")
library(INLA)
library(spdep)
## coloring the spplot
library(colorspace)

#'         \\\_Model_1_\\\         #
################ MOdel 1
```

```
############### ###########################################################
############### replicate the final model in STATA spatial unstructured DIC
############### 11229.11
formulaUH0 <- cumulitive_count ~ EVI_VALUE + rain_mm + gender +
    severe_disease + total_admission + admdays + nweight + f(Adj_ID,
    model = "iid", prior = "normal", param = c(0, 0.001), initial = 1)

resultUH0 <- inla(formulaUH0, family = "nbinomial", data = admData2,
    control.compute = list(dic = TRUE, cpo = TRUE), E = log(nagem),
    control.predictor = list(compute = TRUE))
## summary in 3 decimal places
summary(resultUH0)
exp(resultUH0$summary.fixed)
## write the files to a CSV
write.csv(data.frame(resultUH0$summary.fixed), "results1_14504_36.csv")

pdresultUH0 <- resultUH0$dic$p.eff
## summary in 3 decimal places
summary(resultUH0)
```

C.1.4 Spatial Model -INLA

```
## Load the packages in the above code
#'         \\\_Model 2\\\         #
############### MOdel
############### 2###########################################################
############### spatial model structured and unstrustured without
############### to comapare with Winbugs DIC 14498.08 PD 25.03
formulaUHB <- cumulitive_count ~ EVI_VALUE + rain_mm +
    gender + severe_disease + total_admission + admdays +
    nweight + f(Adj_ID, model = "bym", graph = klf.adj,
    scale.model = TRUE, hyper = list(prec.unstruct = list(prior = "loggamma",
```

```
        param = c(0.0111, 0.001)), prec.spatial = list(prior = "loggamma",
        param = c(0.0011, 0.001))))
resultUHB <- inla(formulaUHB, family = "nbinomial",
    data = admData2, control.compute = list(dic = TRUE,
        cpo = TRUE), E = log(nagem), control.predictor(compute = TRUE))
summary(resultUHB)
pdresultUHB <- resultUHB$dic$p.eff  #25.03
exp(resultUHB$summary.fixed)
## save the file to csv
write.csv(data.frame(resultUHB$summary.fixed), "results2_14498.08.csv")
```

C.1.5 Spatial Temporal Model -INLA

```
############### MOdel
############### 3#############################################################
############### spatial model structured and unstrustured with
############### the temporal component included
formulaUH <- cumulitive_count ~ EVI_VALUE + rain_mm +
    gender + severe_disease + total_admission + admdays +
    nweight + f(Adj_ID, model = "bym", graph = klf.adj,
    scale.model = TRUE, hyper = list(prec.unstruct = list(prior = "loggamma",
        param = c(0.0111, 0.001)), prec.spatial = list(prior = "loggamma",
        param = c(0.0011, 0.001)))) + f(count_adm,
    model = "ar1")
resultUH <- inla(formulaUH, family = "nbinomial", data = admData2,
    control.compute = list(dic = TRUE, cpo = TRUE),
    E = log(nagem_int), control.predictor(compute = TRUE))
summary(resultUH)
pdresultUH <- resultUH$dic$p.eff  #35.50
exp(resultUH$summary.fixed)
write.csv(data.frame(resultUH$summary.fixed), "nm_results2_13640.2.csv")
#### The computation of the posterior mean for the
```

```
#### random effects <f0><U+009D><U+009D><U+0083> is
#### performed in two steps as we have more than one
#### parameter: we extract the marginal posterior
#### distribution for each element of the random
#### effect
csi <- resultUH$marginals.random$Adj_ID[1:40]

## then apply the exponential transformation and
## calculate the posterior mean for each of them
## using the lapply function.
zeta <- lapply(csi, function(x) inla.emarginal(exp,
    x))
## define the cut offs for your risk ratio
zeta.cutoff <- c(0.83, 0.9, 0.95, 0.999, 1, 1.01, 1.05,
    1.1, 1.2)

# Transform zeta in categorical variable
cat.zeta <- cut(unlist(zeta), breaks = zeta.cutoff,
    include.lowest = TRUE)

# Create a dataframe with all the information
# needed for the map
maps.cat.zeta <- data.frame(unique(admData2$Adj_ID),
    cat.zeta = cat.zeta)

# Add the categorized zeta to the kilifi spatial
# polygon
data.kilifi <- attr(kilifi_sub, "data")
attr(kilifi_sub, "data") <- merge(data.kilifi, maps.cat.zeta,
    by.x = "Adj_ID", by.y = "unique.admData2.Adj_ID.")
```

```
## mapping the risk ratio spplot(obj=kilifi_sub,
## zcol= 'cat.zeta',
## col.regions=gray(seq(0.9,0.1,length=4)), asp=1)
png(filename = paste0("figure4A", "img.png"), width = 19.45,
    height = 22.4, units = "cm", res = 300)

spplot(obj = kilifi_sub, zcol = "cat.zeta", col.regions = diverge_hsv(8),
    scales = list(draw = TRUE), asp = 1)
dev.off()

### mapping the convergence plots
plot(resultUH, plot.fixed.effects = TRUE, constant = FALSE,
    plot.lincomb = TRUE, plot.random.effects = TRUE,
    plot.hyperparameters = TRUE, plot.predictor = TRUE,
    plot.q = TRUE, plot.cpo = TRUE, single = TRUE)
plot(resultUH, plot.fixed.effects = TRUE, constant = FALSE,
    plot.cpo = F, single = F)

save.image("stModel.RDA")
```

C.2 Mortality Analysis

C.2.1 Stata Multi-level model

```
    xtmelogit noutcome whz06 i.nsex  i.severe_disease EVI_VALUE rain_mm  ///
    total_admission admdays timeTR age_yr || sublocs: ,    nolog
```

C.2.2 Bayesian Multilevel Random Effects Model (INLA)

```
formulaMorta0 <- noutcome ~ whz06 + nsex + severe_disease + EVI_VALUE +
    rain_mm + total_admission + admdays + timeR + age_yr + f(Adj_ID,
    model = "iid", prior = "normal", param = c(0, 0.001), initial = 1)
```

```
# f(count_adm, model = 'ar1', replicate = Adj_ID3)
resultMorta0 <- inla(formulaMorta0, family = "binomial", data = admData,
    control.compute = list(dic = TRUE, cpo = TRUE), control.predictor(compute = TRUE))
summary(resultMorta0)
```

C.2.3 Bayesian Spatial Random Effects Model (INLA)

```
formulaMorta <- noutcome ~ whz06 + nsex + severe_disease +
    EVI_VALUE + rain_mm + total_admission + admdays +
    timeR + age_yr + f(Adj_ID, model = "bym", graph = klf.adj,
    scale.model = TRUE, hyper = list(prec.unstruct = list(prior = "loggamma",
        param = c(1, 0.001)), prec.spatial = list(prior = "loggamma",
        param = c(1, 0.001))))


# f(count_adm, model = 'ar1', replicate = Adj_ID3)
resultMorta <- inla(formulaMorta, family = "binomial",
    data = admData, control.compute = list(dic = TRUE,
        cpo = TRUE), control.predictor(compute = TRUE))
summary(resultMorta)
```

Appendix 7 - UNICEF Conceptual Framework

The UNICEF conceptual framework, which the nutrition community has been using for programming for the past 25 years, identifies three levels of causes of undernutrition.

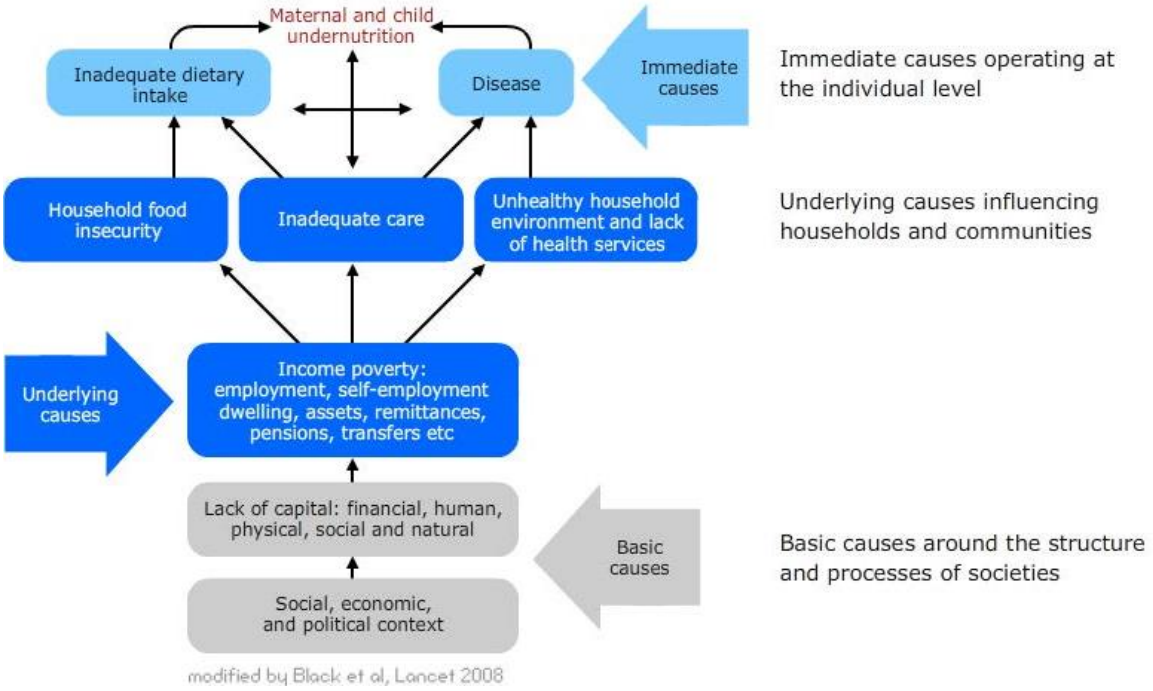


Figure 16.12: UNICEF Conceptual Framework for malnutrition