COLLEGE OF NATURAL SCIENCES
DEPARTMENT OF STATISTICS

# Spatial Pattern and Determinants of Maternal Death in Ethiopia: Analysis based on 2016 EDHS data

By: Mulata Worku

A Thesis Submitted to the Department of Statistics, School of Graduate Studies, College of Natural Science, Jimma University as a Partial Fulfillment for the Requirements for the Degree of Masters of Science in Biostatistics

February 2020
Jimma, Ethiopia

# Spatial Pattern and Determinants of Maternal Death in Ethiopia: Analysis based on 2016 EDHS data

By: Mulata Worku

Advisor: Akalu Banbeta (PhdScholar)

Co-advisor: Fikadu Zawudie (MSc.)

A Thesis Submitted to the Department of Statistics, School of Graduate Studies, College of Natural Science, Jimma University as a Partial Fulfillment for the Requirements for the Degree of Masters of Science in Biostatistics

February 2020

Jimma, Ethiopia

**Approval Sheet**

This is to certify that the thesis titled " **Spatial Pattern and Determinants of Maternal death in Ethiopia: Analysis based on 2016 EDHS data**" submitted in partial fulfillment of the requirement for the degree of Master of Science in Biostatistics to the college of natural science Jimma University, and is record of original research carried out by **Mulata Worku Nagari**, ID.No: RM1072/2010, under my supervision and no part of the thesis has been submitted for another degree or diploma. The assistance and the help received during the course of this investigation have been duly acknowledged. Therefore, I recommended that would be accepted as fulfilling the thesis requirement.

Akalu Banbeta (PhD Scholar) ---------------------------------------------- ---------------------------

Main advisor signature Date

Fikadu Zawudie (MSc) ---------------------------------------------------- --------------------------

Co-advisor Signature Date

As the members of the board of examiners of MSc. thesis open defense examination of **Mulata Worku Nagari**, we certify that we have read and evaluated the thesis and examined the candidate. We recommend that the thesis has been accepted as it fulfills the requirements for the degree of Master of Science in Biostatistics.

| | | |
|---|---|---|
| ------------------------------------------------- | ---------------------------- | - -------------------------- |
| Name of chairman | Signature | Date |
| ------------------------------------------------- | -------------------------- | --------------------- |
| Name of Main Advisor | Signature | Date |
| ------------------------------------------------- | ------------------------- | ----------------------- |
| Name of Co-advisor | Signature | Date |
| ------------------------------------------------ | ------------------------- | ------------------------ |
| Name of internal Examiner | Signature | Date |
| ----------------------------------------------- | --------------------------------- | ------------------------ |
| Name of External Examiner | Signature | Date |

**Statement of author**

The research work described in this thesis was carried out in the School of Graduate Studies, College of Natural Science, Jimma University as a Partial Fulfillment for the Requirements for the Degree of Masters of Science in Biostatistics.

I, Mulata Worku Nagari, declare that this thesis is my own, unaided work. It has not been submitted in any form for any degree or diploma to any other University. Where use has been made of the work of others, it is duly acknowledged.

Mulata Worku

Date: _____________

Signature: _________

February 2020

Jimma, Ethiopia

## ACKNOWLEDGMENTS

Before everything, I would like to give thanks to the Almighty God, who makes everything possible and passing me through many challenge and obstacle by his power.

I would like to express my gratitude to my thesis advisor, Mr. Akalu Banbeta (Ph.D. scholar), for his unflinching and generous supervision as well as easy accessibility at all times. His constructive comments and advice have helped me in widening my research abilities. I am also grateful to my co-advisor, Mr. Fikadu Zawude (MSc) for his encouragement, supervision and constructive comments that develop my research abilities and in improving the quality of the thesis.

I would like to express my thanks to Mettu University for offering me the opportunity and financial support during my study at Jimma University. Further more special dedication goes to Addis Ababa Central statistical Agency for their support any information (Data) which was very crucial for my study.

Finally, I really would like to express my deep gratitude and appreciation to my Mother Dureti Gobena and my Father Worku Nagari who suffer so much for me. I am deeply indebted to them for their patience, concern and love. I cannot praise them enough for their support. To them all I have dedicated this dissertation.

## Table of Contents pages

## LIST OF TABLES

## LIST OF FIGURES

## LIST OF ABBREVIATIONS

| | |
|---|---|
| AIDS | Acute Immune Deficiency Syndrome |
| ANC | Antenatal Care |
| CSA | Central statistical agency |
| CI | Confidence interval |
| DHS | Demographic Health Survey |
| EA | Enumeration Area |
| EDHS | Ethiopian Demographic Health Survey |
| IHDSS | Ifakara Health and demographic surveillance site (IHDSS) |
| MCMC | Markov Chain Monte Carlo |
| MDG | Millennium Development Goal |
| MMR | Maternal Mortality Ratio |
| MRF | Markov random field |
| MoHSS | Ministry of Health and Social Services |
| OR | Odds ratio |
| SDG | Sustainable Development Goal |
| UN | United Nation |
| UNICEF | United Nations International Children's Emergency Fund |
| WHO | World Health Organization |

## *ABSTRACT*

***Introduction:*** *Maternal death is one of the basic problems of women on their health during their reproductive ages (15-49). Worldwide, 99% of deaths of women in their reproductive ages are due to childbirth and pregnancy complications. In a world Ethiopia was one of the countries affected by this problem.*

***Objectives:*** *This study has been aimed to analyze the spatial pattern of maternal death in Ethiopia and identify the factors that affect the maternal death.*

***Methods****: This study was conducted in Ethiopia and the data was basically secondary which is obtained from 2016 Ethiopian Demographic and Health survey (EDHS). The Bayesian Geo-additive regression model is used to identify the major risk factors and spatial effects (spatial pattern) on maternal death in Ethiopia. The DIC model selection criterion is used for model selection.*

***Results****: A total of 10,009 women aged 15 to 49 were included in the study of which 1.43% died due to childbirth or pregnancy related complications. Based on the DIC, Bayesian Geo-additive regression model suits data over the two usual generalized linear regression models fitted in this study (Bayesian generalized linear model and semi parametric regression model).Based on the results obtained using Bayesian Geo-additive regression model, place of delivery, number of antenatal care visit, marital status, wealth index and continuous covariate (age of mother and number of birth order) significantly determines maternal death. Based on the evidences of spatial variation in a model, higher risk of maternal death is found in Afar, Somali, Benishangul gumuz and Gambela regions.*

***Conclusions*****:** *The results of this study suggested that there are complex social, demographic and geographic processes operating in maternal mortality. This result can be more clearly understood using the appropriate statistical models. There have been geographical differences in patterns of maternal death.*

**Key words:** Maternal death, Bayesian Geo-additive regression model, spatial pattern

# CHAPTER ONE

# 1. INTRODUCTION

## 1.1 Background of the problem

Maternal death(mortality) is the death of a woman while pregnant or within 42 days of termination of pregnancy, irrespective of the duration and site of the pregnancy, from any cause related to or aggravated by the pregnancy or its poor management but not from accidental or incidental causes (WHO, ICD 10).

Document of World Health Organization (WHO) from Millennium Development Goals (MDG) to Sustainable Development Goals (SDG) indicate, an estimation of 303, 000 maternal deaths occurred worldwide from pregnancy and its complications which is equivalent to 830 mothers dying every day (more than 1 life lost every 2 minutes) in 2015. Nearly, all of these deaths are preventable in nature if appropriate interventions are taken (WHO, 2015).

Almost 99% of maternal deaths were reported from the developing regions that showed the largest discrepancy between developed and developing countries (WHO, 1990 to 2015).Among the developing countries, sub-Saharan Africa alone accounts for approximately 66% followed by Southern Asia, 22% of maternal deaths. Ethiopia is also categorized under the countries with high maternal mortality (WHO, 1990 to 2015).

Over the years, many studies has been conducted and assess geographical variation (spatial pattern) and risk factor on maternal mortality between the region of countries, large disparities within region of countries, between people with high and low incomes, and between rural and urban populations.

Alfred Kwesi (2012) made a study on spatial patterns and trends of maternal mortality over a five year period and their associated risk factors in Ifakara Health and demographic surveillance site (IHDSS) in Tanzania and showed that declining trend of maternal mortality in southern rural Tanzania, there are marked geographical differences in maternal mortality, with variations across a relatively small geographical area.

Samson G. (2018) conduct study on variation of maternal death at national and sub-national levels of Ethiopia by using Generalized Estimating Equation (GEE) and showed that declining trend of maternal mortality and variation of maternal death in Ethiopia

Many researchers used different statistical method like non-spatial model to identify the determining factors of maternal death which cannot able to consider the violation of independency assumption of observation between clusters of our study area. In short it means that those studies have no power to handle the spatial variation when our data may geographically correlate. In this study the researchers used death related to pregnancy (maternal death) as response variable to identify spatial pattern and determinants of maternal death by including spatial effect by using geo additive regression model in Bayesian approach.

**Why Bayesian approach?** In recent years Bayesian methods have been applied extensively for modeling of both point and areal type spatial data because they allow flexible modeling and inference and provide computational advantages via the implementation of Markov chain Monte Carlo (MCMC) methods (Gelfand and Smith, 1990).

With this study, the reason why the Bayesian approach is preferred over the usual frequentist technique is that, Bayesian models are given preference because the technique is more robust and precise than the traditional (classical) statistics since it is usually criticized based on the priors and information from the likelihood. Thus, this collective information has been strengthening the better determination of the parameter (Ojo OB, 2017).

A Geo-additive regression model is one family of Structural additive regression model, which have capability to analyze the spatial distribution and risk factors of the study variable while accounting for possible non-linear covariate effects (Kammann and Wand (2003)).

**Over view of the study**: This paper is organized as follows: In the first chapter of this study we describe brief Introduction about the Study, Statement of the problem, Objectives of the study, and Significance of the study. In the second chapter we review literature, in third chapter data and methods of data analysis, in fourth chapter results and discussion and in fifth chapter we conclude and recommend to the study.

### 1.2 Statement of the problem

Maternal death remains unacceptably high and has become major challenge in most developing countries including Ethiopia (Maternal death of 412/100,000 live births) where maternal mortality and morbidity levels are among the highest in the world (WHO, 2016). Many scholars believe and recommend the need to conduct in-depth studies on the various aspects of maternal mortality causes and factors with identification of hot spot area in different demographic, economic and socio-cultural settings. So far, there are no such detailed studies conducted in Ethiopia, to explore all aspects of maternal mortality, particularly the effects of socio-economic factors and identification of locations those highly experience maternal mortality burden which is useful tool for monitoring maternal mortality.

Different studies have been conducted to identify important risk factors of maternal death in Ethiopia by using different statistical models. Jarso S.*et.al* (2019) used multivariable logistic regression to determine the risk factors of maternal mortality. Another study used multilevel logistic regression to identify the risk factor of maternal mortality (Weyesa, 2015). But in all previous studies maternal death risk factors were examined using non-spatial statistics analysis and parametric methods of estimation.But, the factors affecting Maternal death might have both spatial variability and nonlinear relationships with maternal death. These effects were not done previously. Thus, with this study the authors fill all those gaps and able to address spatial pattern and Determinants of the maternal death by using Geo-additive regression model based on Bayesian setting by considering the non-linear effect of metrical covariate and spatial effect.

Hence, the studies were addressed the following basic research questions:

- Which variables are significant affect on maternal death from the study variables?
- Which regions are most severed (highly associated) with risk of maternal death?
- How do we apply the appropriate statistical model that includes the effect of spatial dependence and metrical covariates?

### 1.3 objectives of the study

**General objectives**: To analyze the spatial pattern of maternal death in Ethiopia and to identify variables that affects the maternal death.

**Specific objectives:**

- To identify socio-economic and demographic factors associated with maternal death.
- To identify the spatial variation of maternal death between Regions of Ethiopia.
- To explore appropriate statistical model which can capture the risk factors of maternal death by accommodating, both spatial effect and non-linear effect of metrical covariates.

### 1.4. Significance of the study

The purpose of this study was identifying the major contributing socio-economic and demographic risk factors that can determine maternal mortality and provides an empirical maternal mortality risk map that can be used for intervention by the government or other concerned body. Understanding the different factors that can determine maternal mortality and identifying areas that are likely to have higher maternal mortality risks provides basic information to policy makers and researchers for further studies and helps to take measures in those areas that require special attention.

# CHAPTER TWO

# LITRATURE REVIEW

## 2.1 Maternal death in a World

Each year 536,000 maternal deaths occur worldwide from complications arising from pregnancy, and most of these deaths occur in sub-Saharan Africa (WHO 2005). Developing countries accounted for 99% (533,000) of the deaths. Slightly more than half of the maternal deaths (270,000) occurred in the sub-Saharan Africa region alone, followed by South Asia (188,000). Thus, sub-Saharan Africa and South Asia accounted for 86% of global maternal deaths. Additionally the joint report by WHO, UNICEF, World Bank, and others stated, "Maternal Mortality Ratio (MMR) has fallen globally by 45% from 523,000 (380/100,000) in 1990 to 289,000 (210/100,000) in 2013. In 2013, only 1% of maternal deaths occurred in developed regions, whereas the rest (99%) occurred in developing countries (WHO, 2015).

## 2.2. Geographical Variation of Maternal death

Many authors in various disciplines discussed geographical distribution of health event like mortality, morbidity, disease ,etc as a key Element in epidemiologic research, depending on importance given to the description of health events such as subject(patient), place and time. Researchers have been focusing on the relationship between demographic factors and health that extremely determine geographical distribution of health out come. The description of spatial patterns of maternal mortality can be defined as geographical epidemiology. Health event (disease, mortality, morbidity, etc) mapping has a long history and method of descriptive analysis was first used as an attempt to identify health event patterns or describe rates or intensity of spread.

Michelle M. Schmitz (2018) conduct study on Understanding Local Geospatial Variation in Maternal Mortality Outcomes in Uganda's Western Region and found that Moderate evidence of global autocorrelation in the raw rates, indicating a spatial dependency in maternal mortality (Moran I=0.1572,p-value=0.041) and explore spatial pattern of maternal mortality and found higher maternal mortality incidence in the eastern areas of Kyenjojo, Kamwenge and Kibaale, and lower-than-average rates in southern Kabarole district.

Kolo (2017) conduct institutional based study on spatial and temporal variation of maternal death in Borno state of Nigeria and concluded that maternal mortality is widely spread in Borno state, with Northern Borno having the highest MMR (1373/100,000), even though it had the lowest number of women who delivered at health facilities. This is followed by central Borno, which is also very high and as well account for the highest total number of deaths in the state. Southern Borno recorded the lowest MMR (894/100,000), which is significant compare to other regions in the country. Despite the security challenges facing the state, the estimated average MMR in the state(1149/100,000) almost double the country's estimated MMR of 576/100,000.

Tlou B.*et al.* (2017) made study on Space-time patterns in maternal and mother mortality in a rural South African population with high HIV prevalence (2000–2014). The result show a significant spatial cluster of mother deaths in childhood ($p = 0.022$) in a peri-urban community near the national road.

### 2.3. Determinants of Maternal death

Many authors study that, the impact of demographic, socio-economic and environmental characteristic variables on maternal mortality. Socio-economic factors are usually considered as distant predictors of maternal mortality (McCarthy and Maine, 1992). That supposes that their influence on maternal mortality pass through others determinants such as intermediate or close determinants. The following variables are some of them:

**Age of mother:** Pregnancy is a leading cause of death for young women aged 15 to 19 worldwide; with complications of childbirth and unsafe abortion being the major factors (Guerrareyes, 2013). One studies conducted in Tanzania case of Ifakara Health and demographic surveillance site (IHDSS) suggest as women aged 20 to 29 years were 43% less likely to experience maternal mortality compared to those younger than 20 years (HR: 0.57, 95% CI: 0.16 - 1.99). Women aged 30 to 39 years were 3% more likely to experience maternal death compared with those younger than 20 years (HR:1.03, 95% CI: 0.30 - 3.49). On the other hand, being older (40-49 years) was protective: such women were 80% less likely to die from maternal causes of death compared to those younger than 20 years (HR:0.20, 95% CI: 0.04 - 1.00)( Alfred Kwesi, 2012).

**Birth order:** Birth order which translated into the number of children already in the family has a significant effect on death of mother due to related pregnancy. Sharmistha S.*et.al* (2004) conduct the research which was entitled as factors influencing maternal health care in Nepal: the role of socioeconomic interaction and found that birth order was negatively associated with health care of mother which is indirectly affect maternal mortality also. This study showed that mother with higher birth order less liable to receive maternal health care like prenatal care and post natal care than mother with lower birth order which mean that women with higher birth order faces death due to pregnancy related because of they did get less maternal health care.

**Women Marital status**: Women's marital status appears in existent research as a statistical significant risk factor of maternal mortality. Using a data of West Germany from 1980 to 1997, Razum&Jahn (2000) proved that non married women are more exposed to maternal mortality than married women. Even if their explanation was base on the availability and accessibility to good health services, the situation is more complicated in the case of Africa. Indeed, in Africa in general, several single women in reproductive ages are under family and social pressure regarding marriage issues. Due to the strength of cultural values in most of developing countries, marriage still remains the only acceptable frame of getting intercourses and pregnancy. Therefore, unmarried women could be more exposed to maternal mortality through unsafe abortion Because young women getting pregnant out of the societal legitimate frame are generally banned from their household and then maternal mortality. In other hands, married women are more likely to get pregnant, particularly during the first year of their marriage and therefore more exposed to complications and maternal mortality.

In addition according to International, P. (2013) women who had ever been married had a protective effect of 62% compared to women who had never been married.

**Mother Place of residence:** Unlike urban communities, rural communities are at high risk of having home births, which is similar to findings in other studies; the nature of urban and rural areas explains this discrepancy. Urban areas are accessible to health facilities, with a higher proportion of informed and educated people, and better infrastructure (Bicego, 2002). When a woman experiences a complication during pregnancy, she needs immediate medical care. However, families living in these remote communities have a long journey to these medical centers and cannot bring these mothers to the clinics in time, this show as the majority of maternal mortality occurs in rural communities in developing countries (Dahiru, 2015). Women living in rural areas experience higher maternal mortality than women living in urban (Liu, 2011).

**Educational Level of Mother's:** Primary education or none were significant predictors of increased risk of maternal death as compare with secondary or more than (Tlou B.et al (2017)). According to study Jat T. *et al*. (2011) the risk of maternal death is 2.7 times higher among women with no education, and two times higher among women with one to six years of education than for women with more than 12 years of education. The same idea was raised from the study made in Nigeria which showed that secondary educated mothers reduces the risk of maternal mortality by 95% compared to none educated mothers and primary educated mothers reduces the risk of maternal mortality by 53% compared to none educated.

**Exposure to mass media:** Watching television, listening to radio and reading newspaper are considered to measure exposure to media in the analysis. Exposure to information on television and radio and in the print media can increase knowledge and awareness of new ideas, social changes, and opportunities and can affect an individual's perceptions and behavior, including those about health. Nearly three in four (74%) women and 62% of men have no access to radio, television, or newspapers on a weekly basis (EDHS 2016).

**Contraceptive use**: Ahmed *et al*. (2012) found that contraceptive use is efficient for the primary prevention of maternal mortality in developing countries by about 44%.The use of contraceptives reduces unwanted pregnancies, lower rates of abortion, decreases the rate of baby dumping and reduces the risk of premature deaths (MoHSS, 2009).Family planning contributes to reducing maternal mortality by reducing the number of births and, thus, the number of times a woman is exposed to the risk of mortality.

**Antenatal Care:** Antenatal care (ANC) from a skilled provider is important to monitor pregnancy and reduce morbidity and mortality risks for the mother and child during pregnancy, delivery, and the postnatal period (within 42 days after delivery). Antenatal care (ANC) utilization is the potential risk factor of maternal mortality (Samson G.(2018)).

Jarso S.*et al* (2019) made study on Determinants of Maternal Death in a Pastoralist Area of Borena Zone and showed that Mothers who were not attending ANC were 5 times more at risk for death than those who attend (OR 5.3, 95% CI 2.3–12.1).

**Delivery Place:** Maternal and neonatal mortality can be reduced by increasing institutional deliveries in both urban and rural areas. In a world among 132,352,900 births, it is estimated that 34% of mothers deliver with no skilled attendant; this means there are 45 million births occurring at home without skilled health personnel each year. Skilled attendants assist in more than 99% of births in developed countries compared with 62% in developing countries. Mothers who gave birth at home/on transit were twice to die compared to health facility delivery (OR 2.6, 95% CI 2.4–6) that were contributing factors of maternal deaths( Jarso S.*et al* (2019)).

**Wealth index**: As studies developed on risk factor that determine maternal mortality in the rural Tanzania status of women by linking poverty and maternal deaths has indicated that with increasing poverty, the proportion of women dying of non-maternal causes generally increased, and the proportion dying of maternal causes increased consistently. This is because social status of women in developing countries limits their access to basic education or economic resources, which in turn affects their ability to make decisions related to their health. In case of this study 32 to 34% of maternal deaths occurred among women from the poorest quintile of the population (International P. 2013).

**Water sanitation:** The lack of access to clean water and basic sanitation may contribute to increased maternal and neo-natal mortality at three different points. First, it may affect the health of the woman and the fetus during the pregnancy. When a pregnant woman drinks polluted water, she is exposed to a host of bacterial, viral, and parasitic infections (Benova L *et al*. (2014)).In many instances, women contract various diarrheal diseases including dysentery, cholera, or typhoid. These diseases may directly kill a woman or weaken her immune system, which leads to complications during birth (Save the Children, 2014). Certain diseases, like Hepatitis E, are more commonly transmitted when a community lacks access to basic sanitation facilities. Such waterborne diseases tend to have more severe consequences for pregnant women than for the broader population (Mamaye, 2015). Second, pregnant women who habitually travel long distances to collect water often experience weight loss and issues during birth (Dankelman I.*et al*. (1988)). Third, a lack of clean water and sanitation prevents essential hygiene practices including hand washing by birth attendants, delivering the infant on a clean surface, sterilizing equipment for cord cutting, and providing clean blankets which can lead to sepsis, tetanus and, ultimately, death (Blencowe H, *et al*, 2011)

### 2.4. Overview of Bayesian Spatial Modeling

Of late, there has been a rising interest in the development and application of spatial statistical methods for analysis of geographically correlated data. This can be attributed to the increasing availability of geo-referenced data in many fields of study, for example public health and ecology.

Geo-statistics is concerned with the problem of producing a map of a quantity of interest over a particular geographical region based on, usually noisy, measurement taken at a set of locations in the region. The aim of such a map is to describe and analyze the geographical pattern of the phenomenon of interest. Geo-statistical methodologies are born and apply in areas such as environmental studies and epidemiology, where the spatial information is traditionally recorded and available. However, in the last years the diffusion of spatially detailed statistical data is considerably increased and these kind of procedures possibly with appropriate modification can be used as well in other fields of application, for example to study demographic and socioeconomic characteristics of a population living in a certain region.

Basically, to obtain a surface estimate we can exploit the exact knowledge of the spatial coordinates (latitude and longitude) of the studied phenomenon by using bivariate smoothing techniques, such as kernel estimate or kriging (Cressie, 1993; Ruppert *et al.,* 2003). However, usually the spatial information alone does not properly explain the pattern of the response variable and we need to introduce some covariates in a more complex model.

Geo-additive models, introduced by Kammann and Wand (2003), answer this problem as they analyze the spatial distribution of the study variable while accounting for possible non-linear covariate effects.

Although there is no study done on maternal death by using geo-additive regression model, many studies were conducted by using geo-additive regression model on health event area. Adebayo S,*et al.*(2009) conducted study by using Geo-additive Regression Modeling of Levels and Trends of Fertility in Nigeria and observed that spatial variation of fertility in district of Nigeria. This study showed that Akwa Ibom, Anambra, Bayelsa, and Rivers states are significantly associated with high fertility; while Kwara and Zamfara states are significantly associated with low fertility after adjusting for possible determinants of fertility.

Kandala N. (2006) conducted a study to model Bayesian Geo-additive regression on childhood morbidity in Malawi with Bayex software to assess the spatial variation in the prevalence of cough among children under five. The study observed that high risk of cough for children mainly in the central districts of Malawi. In particular, children from the Ntchisi district have the highest probability of having cough compared to children from the rest of the country.

Kazembe LN. (2010) conducted by using geo-additive logistic regression model on Neonatal Mortality using the BayesX software in Malawi and identified number of districts, particularly in Lilongwe, Kasungu and Mchinji in the central region, Mwanza and Chikwawa in the southern region, and Karonga, Rumphi and Chitipa in the northern region of the country.

Dawit Getnet *et al.* (2015) undertook Geo- additive regression models with spatial correlation to estimate under-five mortality risk factors in Ethiopia and spatial pattern of under-five mortality. This study shows that, in general, Tigray, Afar, Somali and Benshangul-Gumuz regions had the highest risk followed by Amhara region and lower risk in Oromia region.

# CHAPTER THREE

# DATA AND METHODOLOGY

## 3.1 Description of the Study Area

This study was conducted in Ethiopia. Ethiopia is located in the horn of Africa and bordered by Eritrea to the North, Djibouti and Somalia to the East, Sudan and South Sudan to the West, and Kenya to the South. Administratively, Ethiopia is organized into nine state of regions: Tigray, Affar, Amhara, Oromiya, Somali, Benishangul-Gumuz, Southern Nations Nationalities and Peoples (SNNP), Gambela, and Harari; and two administrative cities: Addis Ababa and Dire Dawa (CSA Ethiopia and ICF International ,2012)

## 3.2 Source of Data

The data used for this study was 2016 Ethiopia Demographic and Health Survey (2016 EDHS). The data was implemented by the Central Statistical Agency (CSA) at the request of the Ministry of Health (MoH) and Data collection took place from January 18, 2016, to June 27, 2016 with national representative of 18,008 households were selected based on a nationally representative sample that provides estimates at the national and regional levels and for urban and rural areas. The Woman's Questionnaire from 15683 was used to collect information from all women age 15-49 from the selected households. Since the data of EDHS has many missed cases across some variables, after clearing those missing values, a total of the sample of 10,009 women between the ages of 15-49 years in Ethiopia was included in this study.

Administratively, Ethiopia is divided into nine geographical regions and two administrative cities. Each region in Ethiopia is divided into zones, and zones, into administrative units called woredas. Each woreda is further subdivided into the lowest administrative unit, called kebele.

The 2016 EDHS samples were selected using a stratified, two-stage cluster design and Eas were the sampling units for the first stage. The sample for the 2016 EDHS was designed to provide estimates of key indicators for the country as a whole, for urban and rural areas separately, and for each of the nine regions and the two administrative cities. The 2016 EDHS sample was stratified and selected in two stages. Each region was stratified into urban and rural areas, yielding 21 sampling strata. Samples of Eas were selected independently in each stratum in two stages. Implicit stratification and proportional allocation were achieved at each of the lower administrative levels by sorting the sampling frame within each sampling stratum before sample selection, according to administrative units in different levels, and by using a probability proportional to size selection at the first stage of sampling. For the first stage, the 2016 EDHS sample included 645 Eas (202 Eas in urban areas and 443 Eas in rural areas), were selected with probability proportional to the EA size (based on the 2007 PHC) and with independent selection in each sampling stratum. In the second stage of selection, a fixed number of 28 households per cluster were selected with an equal probability systematic selection from the newly created household listing. All women age 15-49 and all men age 15-59 who were either permanent residents of the selected households or visitors who stayed in the household the night before the survey were eligible to be interviewed (CSA, 2016).

### 3.3 Study variables

**Dependant variable**

The dependant variable in this study was maternal death at a reproductive age and it is dichotomous (binary) variable, which is coded as zero if death has not occurred with related pregnancy and 1 if death is occurred due to pregnancy.

$$Y_{ij} = \begin{cases} 1, \; if \; death \; is \; occured \; with \; related \; pregenancy \\ 0, \; if \; \; no \; death \; is \; occured \, with \; related \; pregenancy \end{cases}$$

**Independent variables:** In this study three class of covariate were considered.

1) **Metrical (continuous ) covariate:**
- ✓ Age of mother
- ✓ Number of Birth order of mother

2) **Spatial covariate**
- ✓ Regions where mother reside

3) **Categorical covariate:** Many categorical covariates were used as predictors of maternal death. Those are place of delivery, Number of Antenatal care visit, Place of residence, mothers education level , Marital status, Wealth index, Contraceptive use , Exposure to mass media, Marital status of women and source of drinkable water. Different authors use different variable according to the area of their study, therefore these study was also developed by taking different variables regarding on different study (Alfred Kwesi ,2012), (Ahmed, 2012), (Navaneetham K. &., 2002), (Bayati, 2016), (Liu, 2011).

Table 3.1.Categorical Covariates with their coding

| No. | Variables | Categories |
|---|---|---|
| 1 | place of delivery | 0= Home |
| | | 1= Health facilities |
| 2 | Number of Antenatal care visit | 0= No antenatal visit |
| | | 1= 1-3 Visit |
| | | 4 and above |
| 3 | Place of residence | 0= Urban |
| | | 1= Rural |
| 4 | mothers educational level | 0= Not educated |
| | | 1= Primary |
| | | 2= Secondary |
| | | 3=higher |
| 5 | Marital status | 0= Not married |
| | | 1=Married |
| | | 2= Separated/living with the partner |
| | | 3= widowed/divorced |
| 6 | Wealth index | 0= Poor |
| | | 1= Middle |
| | | 2= Rich |
| 7 | Exposure to mass media | 0=not at all |
| | | 1=less than one week and above |
| 8 | Source of drinkable water | 0= piped water |
| | | 1= Tube well water |
| | | 2= surface water and other |

### 3.4 Method of Data Analysis

#### 3.4.1. Spatial data Analysis

Spatial data analysis is defined as an analysis that involves the accurate description of data relating to a process operating in space, the exploration of patterns and relationship in such data, and the search for explanation of such patterns and relationships (Bailey and Gatrell (1995, p.21)).

Over the last twenty years, spatial data analysis has become a relevant instrument in most areas of observational sciences, from epidemiology to environmental to social sciences, since the focus on geographical locations and on possible spatial patterns and relationships can help our understanding of the studied phenomena. Therefore in this study exploratory spatial data analysis like Moran's I and Geary's C were used to detect the presence of spatial dependency in our data and Bayesian Geo-additive regression model which account spatial effect and non- linear effect of metrical covariate in the model was used. To analyze spatial pattern data relevant software such as GeoDa, ArcGIS and R 3.61 version program were used for practical modeling and analysis.

Obviously, not all data that can be located in space need to be subject to this kind of analysis. Spatial data analysis is involved when data are spatially located and explicit consideration is given to the possible importance of their spatial distribution in the analysis or in the interpretation of results. Bailey and Gatrell (1995) define different classes of spatial data analysis and for each one they outline specific objectives of analysis:

**Point pattern:** which is composed by a set of point events, or point pattern and we want to investigate whether the proximity of the events, that is their spatial configuration, represents a significant pattern. Sometimes these points have some attributes associated with them distinguishing one kind of event from another, but it is the spatial arrangement of the events themselves that is of interest.

**Point referenced** (geo-statistics): this is one class of spatial data that comprise again a set of point locations, but the pattern of these locations is not itself the subject of analysis. This time, the locations are simply the sampled points at which a continuous variable is measured and the aim of the analysis is to understand the process generating these values and to use the information to model the variable of interest. This kind of data is common in the environmental sciences and we refer to it as spatially continuous data, while the analysis techniques are usually known collectively as geo-statistics.

**Area data:** that is the data that have been aggregated to a set of areal units, such as districts, municipalities, census enumeration districts, and so on. One or more variables are measured over this set of zones and the analysis object is to understand the spatial arrangement of these values, to detect patterns and to examine relationships among the set of variables.

### Spatial Association

Spatial association is a basic concept in understanding and analyzing a spatial phenomenon. It enables to assess statistically the degree of spatial dependence in the data. Finding the degree of spatial association (correlation) among data representing related location is fundamental to the statistical analysis of dependence and heterogeneity in spatial patterns. Pattern is characteristic of the spatial arrangement of objects given by their spacing in relation to each other.

Spatial dependence indicates that near places are more likely to be related than distant ones and usually most geographical patterns of interest involve groupings of similar values in clusters.

Tobler (1970) has referred to "The first law of geography: everything is related to everything else but near things are more related than distant things". These ideas explain why spatial dependence has to be an issue in investigating the spatial pattern of maternal death in Ethiopia.

The presence of spatial dependency in the data can be detected by using different technique of statistical measurement. Among them the most common one is measures of spatial autocorrelation. Spatial autocorrelation uses a measure known as spatial autocorrelation coefficient to measure and test the presence of spatial dependency in the data.

Spatial autocorrelation tests whether or not the observed value of a variable at one locality is independent of values of that variable at neighboring localities. A positive spatial autocorrelation refers to a map pattern where geographic features of similar value tend to cluster on a map, whereas a negative spatial autocorrelation indicates a map pattern in which geographic units of similar values scatter throughout the map. When no statistically significant spatial autocorrelation exists, the pattern of spatial distribution is considered to be random.

Spatial autocorrelation can be measured globally or locally. But the most common used are the global measure of spatial auto correlation such as Moran's I and Geary C.

**Moran's I**

Moran's I Global measures summarize spatial association with respect to the whole region. Spatial autocorrelation index measures spatial association in the data considering simultaneously both location and attribute information. The Moran I on the other hand, described the overall spatial dependence of variable of interest over the study area whose values ranged from -1 to +1. A positive spatial autocorrelation $(+1)$ referred to a map pattern where geographical features of similar value tended to cluster together, a negative spatial autocorrelation showed a map pattern in which geographical units of similar values were scattered throughout the map and a statistically insignificant spatial autocorrelation depicted a random distribution

The general formula for computing Moran's I is (Moran, 1950):

$$I = \frac{N}{\sum_i \sum_j w_{ij}} \frac{\sum_i \sum_j w_{ij}\,(y_i - \bar{y})(y_j - \bar{y})}{\sum_i (y_i - \bar{y})^2}$$

Where $N$ is the number of spatial units indexed by $i$ and $j$; $y$ is the variable of interest; $\bar{y}$ is the mean of $y$; and $w_{ij}$ is an element of a matrix of spatial weights.

The observed value of $I$ can be compared to its distribution under the null hypothesis of no spatial autocorrelation or no clustering i.e. when the values of $yi$ are independent of the values $yj(i \neq j)$ at neighboring locations. This is equivalent to say that under the reference null distribution, data are randomly distributed over locations. Therefore, inference can be based on the standardized version of I, namely:

$$Z(I) = \frac{I - E(I)}{\sqrt{var(I)}}$$

The expected value of Moran's I under the null hypothesis of no spatial autocorrelation is

$$E(I) = \frac{-1}{N - 1}$$

Its variance equals $var(I) = E(I^2) - \left(E(I)\right)^2 = \frac{N^2(N-1)S_1 - N(N-1)S_2 - 2S_0^2}{(N+1)(N-1)S_0^2}$

Where $S_0 = \sum_{i \neq j}^{n} wij$ , $S_1 = \frac{1}{2}\sum_i \sum_j \left(w_{ij} + w_{ji}\right)^2$ , $S_2 = \sum_i \left(\sum_j w_{ij} - \sum_j w_{ji}\right)^2$

In this study, the global Moran's I test statistic was used to test the null hypothesis:

$Ho$: No significant clustering of maternal death (no spatial auto correlation) in the entire study region. The mean found by Moran's I coefficient analysis is used to identify the presence of spatial autocorrelation by comparing Moran's I calculated with the tabulated value.

**Geary's C**

The Geary C statistic is useful in identifying local patterns of health event distribution (distribution of maternal death in case of this study). Geary's C interactions are not the cross product of the deviations from the mean, but the deviations in intensities of each observation location with one another.

Geary's C can be computed as follows:

$$C = \frac{(N-1)\sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij} \left(y_i - \mathrm{y_j}\right)^2}{2\left(\sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij}\right)\sum_{i=1}^{N}(y_i - \bar{y})^2}$$

Usually the values of C range between 0 and 2. Values of C between 1 and 2 indicate presence of negative spatial autocorrelation while values between 0 and 1 indicate presence of positive spatial autocorrelation and a value of 1 showing a random pattern.

Moran's I gives a more global indicator, whereas the Geary's coefficient is more sensitive to differences in small neighborhoods.

Testing the significance is done by using the standardized version of C, namely:

$$Z(C) = \frac{C - E(C)}{\sqrt{var(C)}}$$

With $E(C) = 1$ , $var(C) = \frac{((2S1 + S2)(n-1) - 4SO^2)}{2(n+1)SO}$

#### 3.4.1.1 Bayesian Geo-additive Regression Model for Binary Spatial data

A common way to build regression models extending the classical linear model for Gaussian responses to more general situations such as binary responses are generalized linear models, originally introduced by Nelder and Wedderburn (1972). While being flexible in terms of the supported response distributions, generalized linear models obey rather strong assumptions considering the linearity of the influence of covariates and the independence of the observations. However in most practical situation, we are facing at least one of the following problems in which generalized linear model is not appropriate to fit data:

- ❖ For the continuous covariates in the data set, the assumption of a strictly linear effect on the predictor may not be appropriate.
- ❖ Observations may be spatially correlated.
- ❖ Heterogeneity among individuals or units may not be sufficiently described by covariates because of the hierarchical structure of the data.

To overcome these difficulties a model with predictors that contain spatial effect and non-linear effect of metrical covariate called Geo-additive regression model was proposed (Kammann E&Wand M. (2003)).Geo-additive model is one class of structural additive regression (STAR) model which based on the framework of Bayesian generalized linear models (GLMs, see e.g., McCullagh and Nelder 1989 and Fahrmeir and Tutz 2001).

#### 3.4.1.2 Model building

Consider regression models, where observations $(y_{i,}x_i, w_i)$;); i= 1…n on response $y_i$, a vector $x_{i=}$=$(x_1 \ldots\ldots\ldots x_p\ )\ '$ of metrical covariates (continuous covariate), spatial covariates (regions) in case of this study and a vector $w_i = (w_{1\ ,} \ldots\ldots w_r)\ '$ of further covariate, in which categorical variables are often given.

Where, the response Yi is a binary indicator of the response variable, survival status of mother (death, live), for the $i^{th}$ mother in this study.

$i.e. Yi{\sim}Binomial\ (ni,\ \pi i)$

The generalized additive modeling framework (Hastie T.*et al* (1990)) assumes that, given $x_i$ and $w_i$, the distribution of the response yi belongs to an exponential family, with mean $\mu_i = (y_i|x_i, w_i)$ linked to an additive semi-parametric predictor $\mu_i = h(\eta_i)$.

Traditionally the effect of the covariates on the response $y_{i,}$ is modeled by a linear predictor:

$$\eta_i = x'_i\,\beta + w'_i\gamma \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 3.1$$

Where h is a known response function.

The predictor on equation 3.1 above, which is the predictor of generalized linear model, can also assess the socioeconomic, demographic and geographic effects that are highly related with maternal death in Ethiopia. But In this Predictor, spatial variability was not included and as well as non-linearity of metrical covariate was not assumed. Therefore in order to account spatial dependency with non linearity assumption of metrical covariate, equation (3.1) was extended to a geo-additive model by accommodating the spatial variability and non-linear effect of metrical covariate as follows.

$$\eta_i = f_1(x_{i1}) + \cdots + f_p\left(x_{ip}\right) + f_{spat}(s_i) + w'_i\gamma \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 3.2$$

Where, $f_1$,. ...., $f_p$ are nonlinear smooth effects of the metrical (continuous) covariates and $f_{spat}$ is an additional spatially correlated effect of the location $s_i$an observation pertains to.

In a further step, we may split up the spatial effect $f_{spat}$into a spatially correlated (structured) and an uncorrelated (unstructured) effect.

$$f_{spat}(s_i) = f_{str}(s_i) + f_{u\,nstr}(s_i) \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 3.3$$

Here, spatial variability is usually contained many unobserved influences exist only locally (Adebayo S.et al. (2004)). Therefore, equation (*3.1*) was extended as:

$$\eta_i^{\,geo} = f_1(x_{i1}) + \cdots + f_p\left(x_{ip}\right) + f_{str}(s_i) + f_{u\,nstr}(s_i) + w'_i\boldsymbol{\gamma} \ldots\ldots\ldots\ldots\ldots\ldots\ldots 3.4$$

Where:

- $\eta_i$ is a linear predictor of survival status of mother on socioeconomic, demographic and geographic variables in this study.
- $\gamma$ is unknown parameters(fixed effect parameter) corresponding categorical predictor.
- $f_{str}(s_i)$ is spatially correlated (structured) effect, $f_{u\,nstr}(s_i)$spatially uncorrelated (unstructured) effect

As a side effect, we are able to assess to some extent the amount of spatial dependency in the data by observing which one of the two effects is larger. If the unstructured effect exceeds the structured effect, the spatial dependency is smaller and vice versa (Fahrmeir L,&Lang S.(2001)). It should be noted that all functions are centered about zero for identification purpose, thus fixed effects parameters automatically include an intercept term $\gamma_o$ such models are common in spatial epidemiology (Besag J *et al.* 1991). Therefore the predictor above equation 3.4 is the predictor of geo-additive regression model.

Therefore in this study the following two models model 1 which represent fixed effect of all covariate and model 2 which represent fixed effect of categorical covariate plus non-linear effect of continuous covariate without including spatial effect were fitted together with model3 (geo-additive regression model) for response variable $Y_i$(Survival status of mother) conditionally on $S_i$ (region), with a logit link function given with predictors increasing in complexity in order to show the appropriateness of geo-additive model over the rest of the models.

On other hand showing the appropriateness of geo-additive model means it is the same with showing the importance of including spatial covariate and metrical covariate in non-linear form.

$\boldsymbol{Model}$ **1**: $\eta_i = x'_i \beta + w'_i\gamma$............................ (Generalized linear model)

$\boldsymbol{Model}$ **2**: $\eta_i = f_1(x_{i1}) + \cdots + f_p(x_{ip}) + w'_i\boldsymbol{\gamma}$........ (Semi parametric regression model)

$\boldsymbol{Model}$ **3**: $\eta_i = f_1(x_{i1}) + \cdots + f_p(x_{ip}) + f_{spa}(s_i) + w'_i\boldsymbol{\gamma}$...... (Geo-additive regression model)

**The aim of above each model is:**

**Model 1:** In order to investigate the appropriateness of the nonlinear effect of continuous covariate and the inclusion of the geographic locale in Model 3, Model 1 was estimated by assuming a linear effect for continuous covariate and ignoring the regional-specific (spatial) effects.

**Model 2**: It assumes non linear effect of metrical covariate and include categorical variable without including region specific effect (spatial effect).

**Model 3**: This model account unobserved heterogeneity that might exist in the data, all of which cannot be captured by covariates, with the introduction of the spatial effects (region) and the assumption of nonlinear effect of metrical covariates.

#### 3.4.1.3 Prior assumptions

In a Bayesian approach unknown functions $f_1$ ,.........$f_p$ $f_{str}$ , $f_{u\,nstr}$ $a$nd the fixed effects parameters $\boldsymbol{\gamma}$ as well as the variance parameter $\sigma^2$ are considered random variables and have to be supplemented with appropriate prior assumptions.

**Priors for Fixed Effects**

In the absence of any prior knowledge, we assume independent diffuse priors $\gamma_j \propto constant$ ,j=1…r for the parameters of fixed effects. Diffuse priors are the suitable alternative for fixed effects parameters if there is no any prior knowledge (Echavarria L (2004), Fahrmeir L, Lang S (2001))

**Priors for Metrical (Continuous) Effects**

Several alternatives are available for the priors of the unknown (smooth) functions $f_1 \ldots . f_p$. Among them the two main approaches for Bayesian semi-parametric modeling are basis functions approaches with adaptive knot selection and approaches based on smoothness priors. This study, focus on the latter one.

Several alternatives have been proposed for specifying a smoothness prior for the effect of a metrical covariate. Among them: random walk priors [Fahrmeir L, Lang S, Lang S, Brezger A], Bayesian smoothing Splines and Bayesian P-splines (Lang S, Brezger A) were the common one. However our focus in this study was second order random walk prior which is one type of random walk priors.

**Second order random walk prior:** Let us consider the case of a metrical covariate x with equally-spaced observations $x_i, \mathrm{i} = 1 \ldots\ldots \mathrm{m}, \mathrm{m} \leq \mathrm{n}$ .Then $x_{(1)} < \cdots < \cdots x_{(m)}$ defines the ordered sequence of distinct covariate values. Here m denotes the number of different observations for x in the data set. A common approach in dynamic or state space models is to estimate one parameter f(t) for each distinct x(t).

, i.e. Define f(t) = f($x_{(t)}$) and let $\mathrm{f} = (\mathrm{f}(1), .., \mathrm{f}(\mathrm{t}), .., \mathrm{f}(\mathrm{m}))^T$ denote the vector of function evaluation. Then the second order random walk is given by:

$$f(t) = 2f(t-1) - f(t-2) + u(t) \ldots\ldots\ldots 3.5$$

Where: $u(t) \sim N(0, \tau^2)$ is a Gaussian error.

The diffuse priors can be given as $f(1) \propto constant$ and $f(2) \propto constant$ for first values respectively.

**Priors for Spatial Effects**

When we come to spatial effect, Consider first that the spatial index s ∈ {1,..,S} represents a location or site in connected geographical regions. It is assumed that neighboring sites that share boundaries are more homogenous than any other arbitrary sites. Therefore, for a valid prior definition a set of neighbors must be defined for each site s. Hence sites s and t are neighbors if they share a common boundary. Depending on the application, the spatial effect may be further split into a spatially correlated (structured) and an uncorrelated (unstructured) effect .i.e. $f_{spa} = f_{str} + f_{u\ nstr}$ A rationale is that a spatial effect is usually a surrogate of many unobserved influential factors, some of them may obey a strong spatial structure while others may exist only locally. Therefore in this study for the spatially correlated effect $f_{str}\ (si)$, we choose Markov random field priors common in spatial statistics (Besag J. *et al*. (1991)).

The spatial smoothness prior of function evaluations $f_{str}\ (s)$is given as:

$$f_{str}\ (s) \big/ f_{str}\ (t), t \neq s, \tau^2_{str} \sim N\left(\sum_{t' \in \delta s} \frac{f_{str}\ (t)}{N_s}, \frac{\tau^2_{str}}{N_s}\right) \ldots\ldots\ldots\ldots 3.6$$

Where: Ns are the number of adjacent (neighboring) regions; $t \in \delta s$ represent region $t$ is a neighbor of region s.

For the spatially uncorrelated (unstructured) effect,$f_{u\ nstr}$ as the common assumptions of Gaussian, are assumed to be i.i.d (Kandala N, Lang S, Klasen S, Fahrmeir L (2001) , Kneib T, Fahrmeir L (2007))

i.e. $$f_{u\ nstr}(s) \big/ \tau^2_{u\ nstr} \sim N\ (0, {\tau^2}_{u\ nstr}) \ldots\ldots\ldots\ldots 3.7$$

For a fully Bayesian analysis, variance or smoothness parameters ${\tau^2}_j$, j=1…..p, structured, unstructured, are also considered as unknown and estimated simultaneously with corresponding unknown functions $f_j$.

For the variance $\tau_j^2$ inverse Gamma hyper-prior $IG\ (a_j,\ b_j)$ are allocated to $\tau_j^2$. with known hyper parameters $a_j\ and\ b_j$.Standard choices for the hyper parameters $a_j\ and\ b_j$.are $a = 1\ and\ b = 0.05\ or\ a = b = 0.001$ (Kandala N. 2006p). For our case we used a=b=0.001 for $\tau_j^2$ as a standard option. In some data situations (e.g., for small sample sizes), the estimated nonlinear functions $f_j$ may depend considerably on the particular choice of hyper parameters. It is therefore good practice to estimate all models under consideration using a (small) number of different choices for $a_j$ and $b_j$ to assess the dependence of results on minor changes in the model assumptions.

#### 3.4.1.5 Posterior inference

The type of inferential concept used for estimation of regression parameters in this study was full Bayesian inference via Markov chain Monte Carlo (MCMC).Bayesian inference is based on posterior distributions and is carried out using MCMC simulation techniques so that samples are drawn from full conditionals of single parameters or block parameters given the rest. Let $\alpha$ denote the vector of all unknown parameters in the model (i.e. $\alpha = (f_j, f_{spa})$ , $\gamma$ and $\tau$ represent the vector of all variance components. Then, under usual conditional independence assumptions, for the binomial logit model:

$$P(\alpha/y) \propto \prod_{i=1}^{n} L_i(y_i, \eta_i) \prod_{j}^{p}\{p(\beta_j/\tau_j^2)p(\tau_j^2)\}\, p(f_{str}/\tau_j^2)p(f_{u\ nstr}/\tau_{u\ nstr}^2) \prod_{j=1}^{r} p(\gamma_j)\, p(\sigma^2)$$

Where, $\beta_j$, j=1……p, are the vectors of regression coefficients corresponding to the functions $f_j$. For updating the parameters in an MCMC sampler, we used Metropolis–Hastings algorithm based on iteratively weighted least squares (IWLS) proposals introduced by Gamerman (1997).

**Metropolis-Hastings algorithm**

Metropolis–Hastings algorithm is a Markov chain Monte Carlo (MCMC) method for obtaining a sequence of random samples from a probability distribution. The Metropolis–Hastings algorithm works by generating a sequence of sample values. In such a way that, as more and more sample values are produced, the distribution of values more closely approximates the desired distribution $p(x)$. In this thesis the posterior doesn't look like any distribution we know (no Conjugacy) and some (or all) of the full conditionals do not look like any distributions we know (no Gibbs sampling for those whose full conditionals we don't know). Thus why we were interested to use Metropolis–Hastings algorithm The Metropolis-Hastings Algorithm follows the following steps:

Initialize $\theta^0$

Start $b = 0$

Set $B$ number of iterations.

While $b < B$ do

Set $\theta = \theta^b$ select a component $i$

Propose new variable for component $i$ from proposal distribution $\left(\frac{\theta_i}{\theta^b}\right)$ ) Set $\theta^{b+1}$ with the probability

$\alpha = \min\left(1, \frac{\pi(\theta)q\left(\frac{\theta_i}{\theta^b}\right)}{\pi(\theta^b)q\left(\frac{\theta_i}{\theta^b}\right)}\right)$ Otherwise set $\theta^{b+1} = \theta^b$ Set b = b + 1, end while once convergence is reached, all simulation values are from the target posterior distribution and a sufficient number would be drawn so that all areas of the posterior will be also explored.

### 3.4.2 Model Comparison

The classical approach to model comparison involves a trade-off between how well the model fits the data and the level of complexity. Spiegelhalter *et al.* (2002) devised a selection criterion which was based on Bayesian measures of model complexity and how good a fit the model is for the data. The measure of complexity which we adopted in this work is suggested by Spiegelhalter *et al*. (2002). A complexity measure pD is suggested by using an information theoretic argument to get more effective number of parameters in a model, as the difference between the posterior mean of the deviance and the deviance at the posterior estimates of the parameters of interest.

pD is assumed to be an approximate trace of the product of Fisher's information and the posterior covariance matrix. It could be obtained through a Markov Chain Mont Carlo analysis. In the case of normal models, pD corresponds to the trace of 'hat' matrix projection observations on to fitted values.In an exponential family model, $\overline{D}$ which calls for a posterior mean deviance, can be taken as a measure of fit. Assume that f(y) is a fully specified standardizing term, then:

$$pD = \overline{D}\,(\theta) - D(\theta) \ldots\ldots\ldots\ldots 3.9$$

Where D(θ) =−2logp(y |θ)+2logf(y), is a Bayesian deviance. Deviance Information Criteria (DIC), which could be used for model comparison, is computed by adding the fit $\overline{D}$ to a complexity pD.

DIC is defined as a "plug in" estimate of fit plus twice the effective number of parameter, as follows:

$$\text{DIC} = \overline{D}\,(\theta)\text{+2pD} = D + \text{pD} \ldots\ldots\ldots\ldots\ldots 3.10$$

Where, the posterior mean of the deviance $\overline{D}\,(\theta)$ is penalized by the effective number of model parameters pD. See Spiegelhalter *et al.* (2002) for more details.

Therefore in order to select the best model among several fitted model in this study, Deviance information criterion (DIC) was used. In a Bayesian frame work the Deviance Information Criterion is a widely used statistic for comparing models. The deviance information criterion (DIC) is a hierarchical modeling generalization of the AIC (Akaike information criterion) and BIC (Bayesian information criterion, also known as the Schwarz criterion).

It is particularly useful in Bayesian model selection problems where the posterior distributions of the models have been obtained by Markov chain Monte Carlo (MCMC) simulation. Like AIC and BIC it is an asymptotic approximation as the sample size becomes large. It is only valid when the posterior distribution is approximately multivariate normal. The advantage of DIC over other criteria, for Bayesian model selection, is that the DIC is easily calculated from the samples generated by a Markov chain Monte Carlo simulation. Assessing goodness of fit involves investigating how close the values are predicted by the model with that of observed values.

### 3.4.3 Model Diagnostic

Model diagnosis was performed based on MCMC post-estimation diagnosis, to examine the convergence of MCMC. Among several ways of test of convergence, the most popular and straight forward convergence assessment methods have been used for this study .The following two methods were more likely considered for this study.

**1. Trace plots**: Iteration numbers on $x$-axis and parameter value on $y$-axis are commonly used to assess convergence. If the plot looks like a horizontal band, with no long upward or down ward trends, then we have evidence that the chain has converged. The posterior distribution is obtained by sampling toward the end of this longer iteration sequence when the posterior distribution is stationary, as determined by an examination of trace plots of the iteration history of selected model quantities.

**2. Autocorrelation plot**: High correlation between the parameters of a chain tends to give slow convergence, whereas high autocorrelation within a single parameter chain leads to slow mixing and possibly individual non convergence to the limiting distribution because the chain tend to explore less space in finite time. In analyzing Markov chain autocorrelation, it is helpful to identify lags in the series in order to calculate the longer run trends in correlation, and in particular whether they decrease with increasing lags.

# CHAPTER FOUR

# 4. RESULTS AND DISCUSSIONS

## 4.1. Descriptive statistics

Table 4.1 below shows the characteristics of Survival status of mother by different covariates for the sample of 10,009 women aged 15-49 in Ethiopia because of 2016 EDHS data has many missed or not illegible cases across some variables. The dependant variable considered in this study was the Survival status of mother (died if death is related to pregnancy or alive otherwise).The result revealed that the percentage of the death of the mother with related to pregnancy was varied from one region to the other region in Ethiopia. The highest percentage of maternal mortality was observed in Afar (3.4%) followed by Somalia (2.8%) while the lowest percentage of maternal death was recorded in Addis Ababa (0.4%) and followed by Dire Dawa (0.7%).

Concerning to place of delivery the percentage of maternal death those who deliver at home was 1.7% and 1% those who deliver at health facilities. In the same manner the percentage of maternal death who are follow no antenatal visits,1-2 visit and 3 and above was 1.8%, 1%, 1.8% respectively.

The percentage of maternal death also varies according to wealth index and educational level of mother as a higher percentage of the death of mother related to pregnancy was observed in poor wealth index (1.9%), while the lowest percentage of the death of mother related to pregnancy was observed in medium wealth index (1.1%) and regard to educational level of mother maternal death rates are 1.7%, 1.1%, 1.3%, 0.9% for mothers of no education, primary education, secondary and higher, respectively. This implies that maternal death was highest for mothers of no education and lower for mother of higher level of education.

Again based up on covariates exposure to mass media, source of drinkable water and marital status of mother, the percentage of maternal death is high for a mother who follow mass media not at all (1.5%) while low for Less than one week and above (1.4%), high for mother who use source of drinkable water Surface water and other (1.5%), while low for piped water and tube well water (1.4%).

Table 4.1: Description of the Socio-economic, Demographic and Environmental factors of maternal death in the Regional state of Ethiopia.

| Covariate | No.of mother (N) | Survival status of mother | |
|---|---|---|---|
| | | Death N (%) | Live N (%) |
| **Region** | | | |
| Addis Ababa | 890 | 4(0.4%) | 886(99.6%) |
| Afar | 701 | 24(3.4%) | 677(96.6%) |
| Amhara | 1156 | 14(1.2%) | 1142(98.8%) |
| Benishangul Gumuz | 795 | 14(1.8%) | 781(98.2%) |
| Dire Dawa | 696 | 5(0.7%) | 691(99.3%) |
| Gambela Peoples | 745 | 11(1.5%) | 734(98.5%) |
| Harari Peoples | 617 | 7(1.1%) | 610(98.9%) |
| Oromia | 1087 | 16(1.5%) | 1071(98.5%) |
| Somali | 976 | 27(2.8%) | 949(97.2%) |
| SNNP | 1283 | 12(0.9%) | 1271(99.1%) |
| Tigray | 1063 | 10(0.9%) | 1053(99.1%) |
| **Place of delivery** | | | |
| Home | 6169 | 105(1.7%) | 6064(96.3%) |
| Health facility | 3840 | 39(1%) | 3801(99%) |
| **No. of antenatal care visit** | | | |
| No antenatal visit | 2825 | 50(1.8%) | 2775(98.2%) |
| 1-3 | 4093 | 39(1%) | 4054(99%) |
| 4 and above | 3091 | 55(1.8%) | 3036(98.2%) |
| **Source of drinkable water** | | | |
| piped water | 3767 | 54(1.4%) | 3713(98.6%) |
| Tube well water | 3587 | 49(1.4%) | 3538(98.6%) |
| Surface water and other | 2655 | 41(1.5%) | 2614 (98.5%) |
| **Exposure to mass media** | | | |
| not at all | 5819 | 85(1.5%) | 5734(98.5%) |
| Less than one week and above | 4190 | 59(1.4%) | 4131(98.6%) |

| | | | |
|---|---|---|---|
| **Educational level of mother** | | | |
| No education | 5512 | 96(1.7%) | 5416(98.3%) |
| Primary | 2937 | 32(1.1%) | 2905(98.9%) |
| Secondary | 544 | 7(1.3%) | 537(98.7%) |
| Higher | 1016 | 9(0.9%) | 1007(99.1%) |
| **Wealth index** | | | |
| Poor | 3089 | 58(1.9%) | 3031(98.1%) |
| Medium | 4229 | 46(1.1%) | 4183(98.9%) |
| Rich | 2691 | 40(1.5%) | 2651(98.5%) |
| **Marital status** | | | |
| Not married | 967 | 6(0.6%) | 961(99.4%) |
| Married | 927 | 9(1.0%) | 918(99.0%) |
| Living with partner/separated | 310 | 7(2.3% \|) | 303(97.7%) |
| Widowed/Divorced | 7805 | 122(1.6%) | 7683(98.4%) |
| **Contraceptive use** | | | |
| Not use | 7316 | 107(1.5%) | 7209(98.5%) |
| Use | 2693 | 37(1.4%) | 2656(98.6%) |
| **Residence place** | | | |
| Urban | 2865 | 41(1.2%) | 2824( 98.8% ) |
| Rural | 7144 | 103(1.4%) | 7041(98.6%) |

### 4.2 Result of Chi-square test

In this study chi-square test statistic were used to measures the strength of association between two variables (response variable and independent variables) which are categorical in nature and test whether there is variation on maternal death between regions of Ethiopia. From summary presented on table 4.7 (see appendix A), we have seen that, a test of association carried out using the Pearson chi-square test at 5% level of significance for each predictor. The result indicates that mother's survival status related to pregnancy was strongly associated with place of delivery; region, number of antenatal visit, wealth index, marital status and educational level of mother at 5% significance level.

On other hand from the chi-square test statistic between survival status of mother related to pregnancy and the region where mother reside ( $x^2$ = 116.29, df=10, p=0.000,α= 0.05), we have enough evidence to reject the null hypothesis and conclude that there is heterogeneity of maternal death between regions of Ethiopia.

### 4.3. Results of Spatial Dependency test

Moran's and Geary's indices of spatial autocorrelation were used to measure the degree of correlation of maternal death among neighboring regions in Ethiopia. The detail of them was written briefly in methodology part.

From the results of Moran's and Geary's index in the Table 4.2 below we can conclude that the presence of spatial dependency (I = 0.260 P-value=0.03462 for Moran's, I=0.487, P-vale=0.04813 for Gery'c) between neighboring Region of maternal death in Ethiopia.

Table 4.2 .Results of Global Moran's I and Geary's C statistics

| Measures | Observed | p-value |
|---|---|---|
| Moran's *I* | 0.260 | 0.03462 |
| Geary's *C* | 0.487 | 0.04813 |

### 4.4 Model Based Data Analysis

The objective of this study was to identify the effect of large set of bio-demographic and socio-economic variable, including covariates such as number of antenatal care visit, place of delivery, residence place, educational level of mother, wealth index , source of drinking water, exposure to mass media, number of birth order of mother, and age of mother on Maternal death. Starting with very simple models(Bayesian generalized linear model), we increase complexity to show what can be gained by more sophisticated approaches, and then we end up with the analysis using models that included the significant effect of categorical covariates as well as the nonlinear effects of metrical covariate and the spatial effects which is called Geo-additive regression model.

### Bayesian Generalized linear Model (M1)

The summary presented in Table 4.3 below shows that M1 (Bayesian Generalized linear model) which represent fixed effect of categorical covariate plus metrical covariate without assuming non-linear effect of those continuous covariates such as age of mother , number of birth order of mother and without including spatial effect. From the results of model 1, we have seen that, most of the covariates such as, delivery place (health facility), Number of Antenatal care visit (1-3, 4 and above), wealth index (rich), marital status of mother, age of mother, and number of birth order of mother have a significant effect on maternal death, hence their correspondence 95% confidence interval does not include zero. But many of the covariates such as, source of drinkable water, exposure to mass media, residence place, educational level of mother and contraceptive usage have non-significant effect on the maternal death.

On other hand the coefficient of each significant covariate were interpreted by using exponentiation value as follows: The coefficient value of the delivery place categories of health facility was OR= exp (-0.4499) = 0.63 with 95% CI (0.43, 0.90). This is interpreted as the odds of maternal death for mothers who deliver at health facility was 0.63 times less likely than the odds of mothers who deliver at home.

For number of antenatal care visit covariate, the coefficient value for 1-3 categories was OR= exp (-1.2839) = 0.12 with 95% CI(0.048, 0.30),which mean that, the odds of maternal death for mothers who visit 1-3 number of antenatal care was 0.12 times less likely than the odds of mothers who have no antenatal care visit and also the odds of maternal death those who visit 4 and above number of antenatal care(OR =exp(-2.1607)=0.28, 95%CI: 0.13,0.62) was 0.28 times less likely than the odds of mothers who have no antenatal care visit.

In case of coefficients value for rich category of wealth index (OR = 0.25, 95%CI: 0.11, 0.55), which mean that, the odds of maternal death for rich women was 0.25 times likely than that of poor women. Similarly, the odds of maternal death for separated/no longer living together women was 4.25 times (OR=4.25, 95%CI: 1.40, 13.00) and Divorced/widowed women was 2.5 times (OR= 2.50, 95%CI: 1.2, 5.8) more likely than the odds of not married women respectively.

Table 4.3: Posterior mean estimate of model1 (Bayesian Generalized linear model).

| **Variable** | **Category** | **mean** | **Std error** | **2.5% quantile** | **97.5% quantile** |
|---|---|---|---|---|---|
| Intercept | | -5.7612 | 0.7589 | -7.1236 | -4.1337 |
| Delivery place | Home(ref) | | | | |
| | Health facility | -0.4499 | 0.1921 | -0.8433 | -0.0875 |
| | No antenatal care visit(ref) | | | | |
| No.Antenatal care visit | 1-3 | -2.1607 | 0.4526 | -3.0315 | -1.2041 |
| | 4 and above | -1.2839 | 0.3902 | -2.0612 | -0.4821 |
| Source of drinkable water | piped water(ref) | | | | |
| | Tube well water | -0.0546 | 0.2528 | -0.5034 | 0.4595 |
| | Surface_water_and_other | -0.2002 | 0.2354 | -0.6410 | 0.2657 |
| Exposure to mass media | not at all (ref) | | | | |
| | Less than one_week_and_above | 0.2161 | 0.2002 | -0.1785 | 0.6175 |
| Educational level of mother | No education(ref) | | | | |
| | Primary | -0.1711 | 0.2242 | -0.6029 | 0.2917 |
| | Secondary | -0.0828 | 0.4442 | -1.0304 | 0.7022 |
| | Higher | -0.3572 | 0.3871 | -1.0999 | 0.3917 |
| Wealth index | Poor(ref) | | | | |
| | Medium | 0.2834 | 0.4608 | -0.6468 | 1.1257 |
| | Rich | -1.4300 | 0.4057 | -2.1684 | --0.5899 |
| Marital status | Not married(ref) | | | | |
| | Married | 0.6749 | 0.5081 | -0.3645 | 1.6598 |
| | Living with partner/separated | 1.4460 | 0.5760 | 0.3035 | 2.5651 |
| | Widowed/Divorced | 0.9179 | 0.4061 | 0.1397 | 1.7629 |
| Contraceptive use | Not use (ref) | | | | |
| | Use | 0.0363 | 0.2093 | -0.3781 | 0.4297 |
| Residence place | Urban(ref) | | | | |
| | Rural | -0.4335 | 0.2626 | -0.9436 | 0.1018 |
| Maternal Age | | 0.0715 | 0.0194 | 0.0319 | 0.1083 |
| Number of Birth order | | 0.1649 | 0.0327 | 0.0983 | 0.2277 |

### Bayesian Semi-parametric Regression Model

Table 4.4 below shows that the results of M2 (Bayesian semi-parametric regression model) which represent fixed effect of categorical covariate plus non-linear effect of continuous covariate (age of mother and number of birth order of mother) but not including spatial effect. The result revealed that the same result with model 1, except the posterior estimated value of age of mother and number of birth order of mother become bigger than in model 1 which can show that the variation of estimated value of those continuous covariate when their effect was assumed linearly and non-linearly in model 1 and model 2 respectively.

Table 4.4 Posterior mean estimate of model 2 (Bayesian semi-parametric regression model)

| **Variable** | **Category** | **Mean** | **Std error** | **2.5% quantile** | **97.5% quantile** |
|---|---|---|---|---|---|
| Intercept | | -2.0165 | 0.8522 | -3.8570 | -0.5573 |
| Delivery place | Home(ref) | | | | |
| | Health facility | -0.4317 | 0.1994 | -0.8146 | -0.0540 |
| No. of Antenatal care visit | No antenatal care visit(ref) | | | | |
| | 1-3 | -2.3784 | 0.4761 | -3.2997 | -1.3824 |
| | 4 and above | -1.2763 | 0.4718 | -2.1181 | -0.3258 |
| Source of drinkable water | piped water(ref) | | | | |
| | Tube well water | 0.0235 | 0.2592 | -0.4481 | 0.5306 |
| | Surface_water_and_other | -0.1893 | 0.2525 | -0.6523 | 0.2995 |
| Exposure to mass media | not at all (ref) | | | | |
| | Less than one_week_and_above | 0.2456 | 0.2130 | -0.1790 | 0.6634 |
| Educational level of mother | No education(ref) | | | | |
| | Primary | -0.0694 | 0.2423 | -0.5315 | 0.4247 |
| | Secondary | -0.6708 | 0.4980 | -1.7046 | 0.2384 |
| | Higher | -0.8027 | 0.4302 | -1.6861 | -0.0142 |
| Wealth index | Poor(ref) | | | | |
| | Medium | 0.4702 | 0.5189 | -0.5921 | 1.4834 |
| | Rich | -1.3667 | 0.4960 | -2.2850 | -0.3369 |
| Marital status | Not married(ref) | | | | |
| | Married | 1.0340 | 0.5638 | -0.0021 | 2.1884 |

| | | | | | |
|---|---|---|---|---|---|
| | Living with partner/separated | 1.4419 | 0.6172 | 0.2698 | 2.6290 |
| | Widowed/Divorced | 0.9614 | 0.4397 | 0.2322 | 1.9240 |
| Contraceptive use | Not use (ref) | | | | |
| | Use | 0.2042 | 0.2088 | -0.2091 | 0.6110 |
| Residence place | Urban(ref) | | | | |
| | Rural | -0.1150 | 0.2735 | -0.6343 | 0.4134 |
| | **Smooth terms variances** | | | | |
| Maternal age | Continuous | 2.0457 | 1.3509 | 0.6179 | 5.7285 |
| Number of Birth order | Continuous | 0.1649 | 0.0327 | 0.0983 | 0.2277 |

**Bayesian Geo-additive Regression Model**

Table 4.5 below shows that the result of M3 (Bayesian geo-additive regression model) which represent fixed effect of categorical covariate plus non-linear effect of metrical covariate plus spatial effect (both structured and un-structured spatial effect). The result shows that most of the covariates have a significant effect on the survival status of mother.

The fixed effect results of this model indicate that place of delivery, Number of antenatal care visit, wealth index of women, and marital status of mother were found to have a significant effect on maternal death and metrical covariate age of mother and number of birth order of mother were significantly affect survival status of mother; since their corresponding credible interval has not included zero. In contrast of this source of drinkable water, exposure to mass media, contraceptive usage, educational level of mother and residence place were insignificant effect on maternal death; since their corresponding credible interval included zero. For the convenience and get the easy understanding of the interpretation, the researcher has interpreted the exponentiation value of the coefficients of the significant variables as follows:

The coefficient value of the delivery place categories of health facility was OR= exp (-0.4447) = 0.64 with 95% CI (0.42, 0.94). This is interpreted as mothers who deliver at health facility had 46% lower odds to die than those who deliver at home. Regarding to number of antenatal visit the coefficient value for 1-3 categories was OR= exp (-2.3707) =0.093 with 95% CI (0.034,0.281),which means that mothers who visit 1-3 number of antenatal care had 9.3% lower odds to die than those who has no antenatal care visit. Similarly, those who visit 4 and above number of antenatal care had 24% lower odds to die than those who have no antenatal care visit (OR = 0.24, 95%CI: 0.098, 0.614).

Concerning to coefficients value for Rich category of wealth index (OR = 0.25, 95%CI: 0.102, 0.614), which mean that, the odds of pregnancy-related death for rich women were 25% less likely than that of poor women.

Another finding of this study showed that marital status of mother has a significant contribution to maternal death. The odds of maternal death for Divorced/widowed women was 2.86 times (OR=2.86, 95%CI: 1.26, 7.93) and separated/no longer living together women was 5.52 times (OR= 5.52, 95%CI: 1.57, 18.82) more likely than the odds of not married women respectively.

Regarding to spatial effect both structural and un-structural spatial effect were significantly affect maternal death in this study, hence their corresponding 95% credible intervals ((0.023, 3.7446), (0.01, 1.5022)) for structured and un-structured spatial effect respectively which does not included zero. From the posterior mean estimate the structured spatially correlated effect (0.9207) clearly exceeds the unstructured effect (0.2886) which indicates that the spatial dependency is bigger in our data. More, the spatial and metrical covariate effect on maternal death were briefly explained by it is visualization using map effect and graph on section 4.4.1 below.

Table 4.5: Posterior mean estimate of M3 (Bayesian Geo-additive Regression model)

| **Variable** | **Category** | **Mean** | **Std error** | **2.5% quantile** | **97.5% quantile** |
|---|---|---|---|---|---|
| Intercept | | -2.1640 | 0.8597 | -4.0041 | -0.6274 |
| Delivery place | Home(ref) | | | | |
| | Health facility | -0.4447 | 0.2078 | -0.8650 | -0.0537 |
| No.Antenatal care visit | No antenatal care visit(ref) | | | | |
| | 1-3 | -2.3707 | 0.5528 | -3.3762 | -1.2664 |
| | 4 and above | -1.3991 | 0.4631 | -2.3161 | -0.4873 |
| Source of drinkable water | piped water(ref) | | | | |
| | Tube well water | -0.1926 | 0.2556 | -0.6726 | 0.3141 |
| | Surface_water_and_other | -0.2922 | 0.2370 | -0.7462 | 0.1800 |
| Exposure to mass media | not at all (ref) | | | | |
| | Less than one_week_and_above | 0.2556 | 0.2230 | -0.1890 | 0.6734 |

| | | | | | |
|---|---|---|---|---|---|
| Educational level of | No education(ref) | | | | |
| mother | Primary | 0.0167 | 0.2404 | -0.4215 | 0.4819 |
| | Secondary | --0.3230 | 0.4670 | -1.2623 | 0.5292 |
| | Higher | -0.5915 | 0.4373 | -1.4996 | 0.2530 |
| Wealth index | Poor(ref) | | | | |
| | Medium | 0.5547 | 0.5080 | 0.5136 | 1.4594 |
| | Rich | -1.4006 | 0.4728 | -2.2815 | -0.4798 |
| Marital status | Not married(ref) | | | | |
| | Married | 1.3089 | 0.5878 | 0.2421 | 2.4845 |
| | Living with partner/separated | 1.7086 | 0.6316 | 0.4531 | 2.9354 |
| | Widowed/Divorced | 1.0521 | 0.4836 | 0.2291 | 2.0712 |
| Contraceptive use | Not use (ref) | | | | |
| | Use | 0.4058 | 0.2202 | -0.0298 | 0.8362 |
| Residence place | Urban(ref) | | | | |
| | Rural | -0.2728 | 0.2758 | -0.7988 | 0.2620 |
| | **Smooth terms variances** | | | | |
| Maternal age | Continuous | 2.7172 | 1.5719 | 0.7250 | 6.5865 |
| Number of Birth order | Continuous | 0.0257 | 0.0498 | 0.015 | 0.1391 |
| sx(Region):mrf | | 0.9207 | 1.0730 | 0.023 | 3.7446 |
| sx(Region):re | | 0.2886 | 0.4090 | 0.01 | 1.5022 |

### 4.4.1 Visualization of the effect of Metrical Covariate and Spatial effect

Figures 4.1, below display the posterior mean of nonlinear effect of age of mother on maternal death for models M3 (Geo-additive regression model).From the pattern of this graph we have seen that the effect of mother's age is comparably higher in age interval (<20), then somewhat decrease at age interval (20-29) and again increase at age interval (30-39). In short, from this graph we understand that younger mother (>20) and mother with age interval (30-39) were at higher risk of maternal death, as compared to mothers with age interval (20-29).

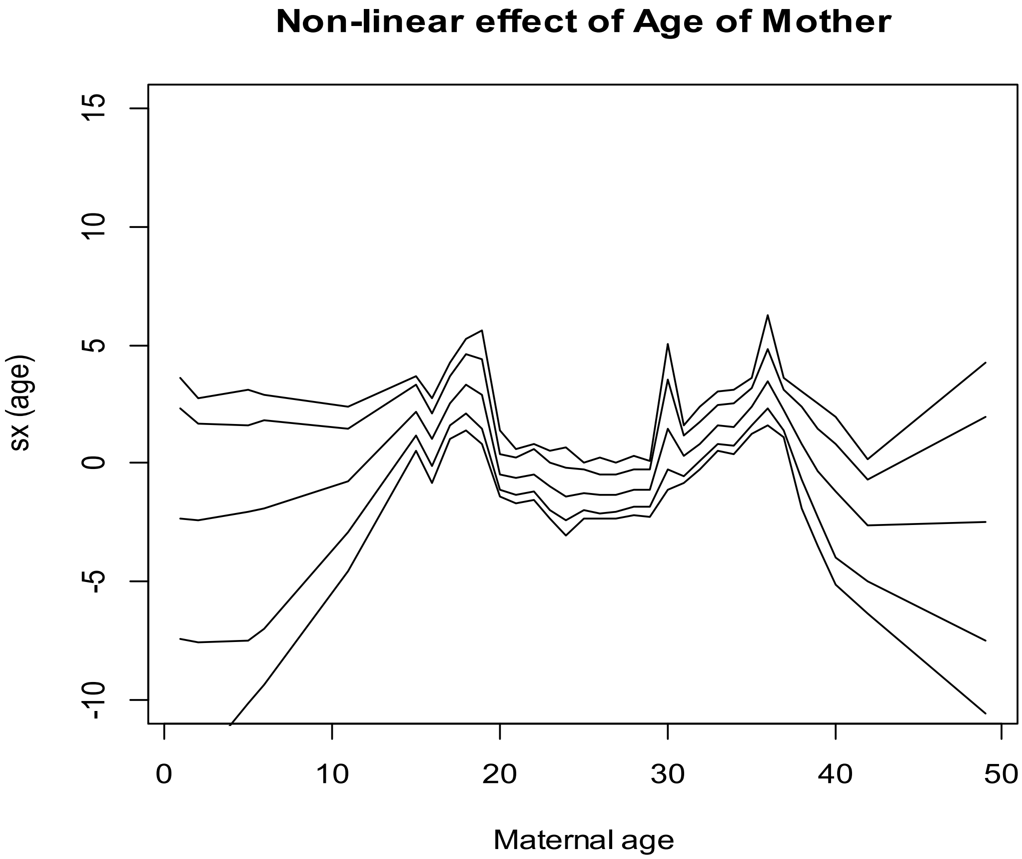



Figure 4.1: Non-linear effects of maternal age

Figures 4.2 below show the nonlinear effects of number of birth order of mother on maternal death for models M4 (Geo-additive regression model). The figure reveals that as number of birth order of mother increase, mothers were faced with high risk of pregnancy related death.

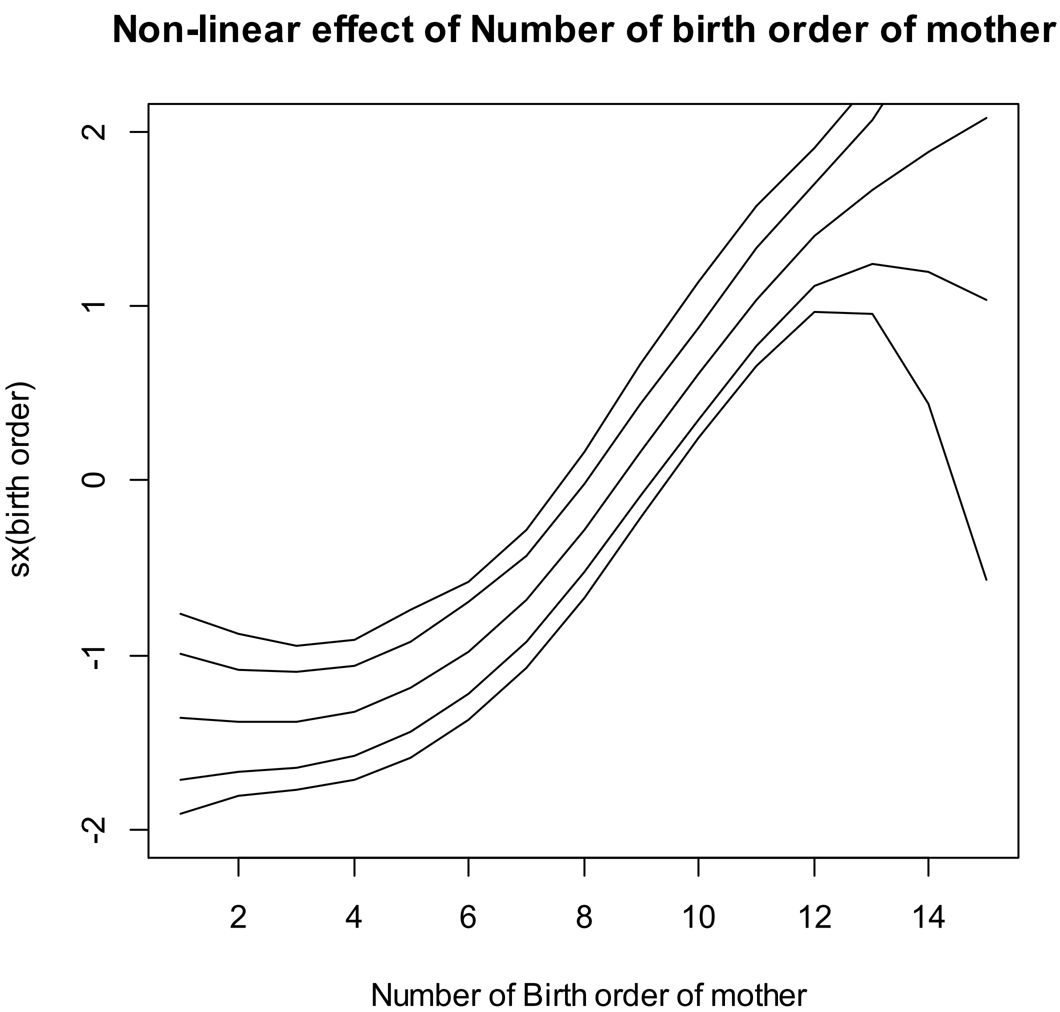



Figure 4.2: Non-linear effects of Number of Birth order of Mothers.

With regard to spatial effects, figures 4.3 and 4.4 below depict the estimated posterior mean of the structured spatial effects and unstructured spatial effect by using map effect on the maternal death in Ethiopia respectively. Therefore from the pattern of both map effect below, it is clearly seen that there was spatial (Geographical) variation between regions of Ethiopia. Blue refers to a positive spatial effect signifying lower mortality, while red refers to a negative effect signifying higher mortality. As the figure indicates, the highest maternal death was presented in Affar, Somali and followed by Benishangul gumuz and Gambela.

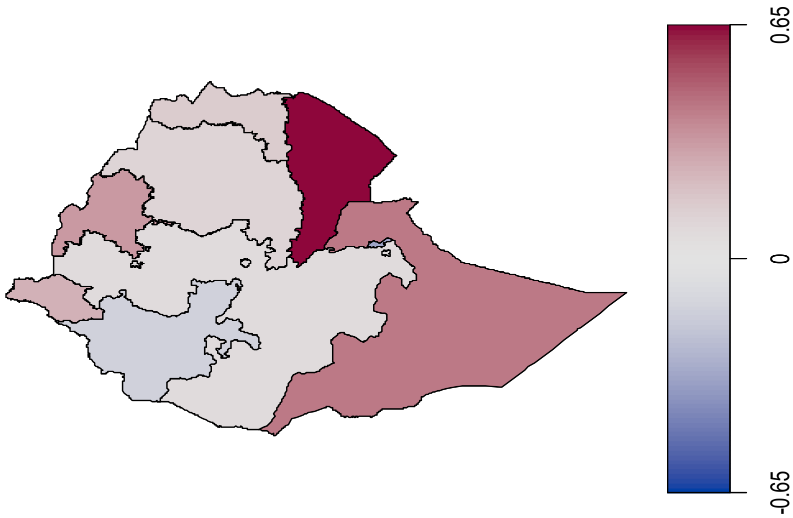



Figure 4.3: Maps of Ethiopia for maternal death showing structured spatial effects in model M3.

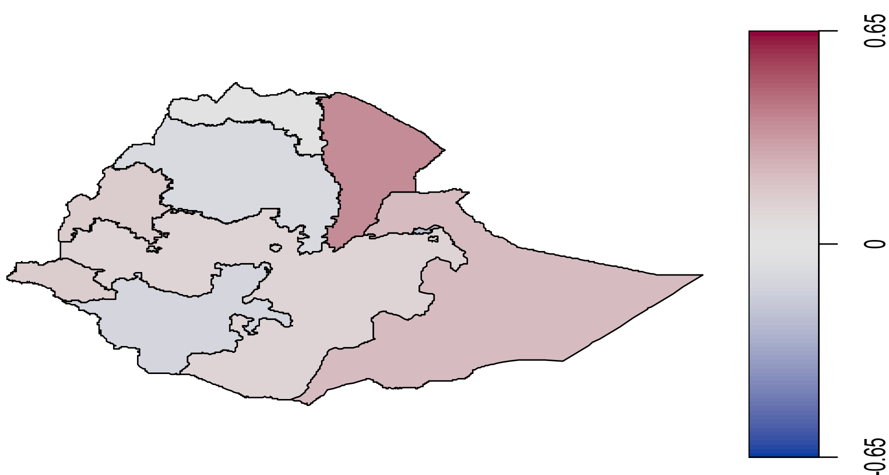



Figure 4.4: Maps of Ethiopia for maternal death showing unstructured spatial effects in M3

### 4.5 Model Comparison Results

From the result of Table 4.6 below, the DIC value for model 1(Bayesian Generalized linear model), model2 (Bayesian semi parametric regression model) and model 3(Bayesian Geo-additive regression model) are 1454.870, 1301.140, 1260.838 respectively. From this result we have seen that the DIC value for Bayesian Geo-additive regression model is small as compared to the rest of model 1 and model 2. This indicates that as model complexity increase from simple model to more sophisticated model, the DIC value decrease. Therefore M3(Bayesian Geo-additive regression model ) which represent fixed effect of categorical covariate plus non-linear effect of metrical covariate plus spatial effect is good fit the data in terms of DIC, as compared to Model 1 and Model 2.

On other hand from model comparison result we conclude that the importance of including spatial effect and non-linear effect of metrical covariate. For instance when we compare model1 which represent fixed effect of categorical covariate and metrical covariate without assuming its non-linear effect with model2 which represent fixed effect of categorical variable plus metrical covariate with non-linear effect assumption model2 is better than model1 because of it has small DIC value. This indicate that metrical covariate (age of mother and number of birth order of mother) affects survival status of mother non-linearly rather than linearly. Again in model3 when we add spatial effect on model2, it become better than model2, hence it is DIC value is smaller as compared to the rest models. This result also indicate that the importance of including spatial effect with non-linear effect of metrical covariate in model to identify risk factor of maternal death in Ethiopia.

Table 4.6: Summary of the Deviance Information Criterion (DIC) for models M1 to M3

| Model | Deviance | pD | DIC |
|---|---|---|---|
| M1 | 1419.240 | 17.815 | 1454.870 |
| M2 | 1228.863 | 36.138249 | 1301.140 |
| M3(Geo-additive model) | 1170.949 | 44.944653 | 1260.838 |

### 4.6 Assessment of Model Convergence

There are several methods to check for convergence. Among them trace plot and autocorrelation plot for monitoring convergence have been presented below.

**Trace plots**: This is the graph which could be plotted the iterations versus the generated values. In this graph convergence can be attained if all values are within a zone without strong periodicities (up and down periods). Therefore, the trace plots in the figure 4.1 below are all straight line which did not show up and down periods. This is an indication that all posterior estimates were converged. Not all trace plots are presented here; the rest plots can be found in appendices (see appendix B).

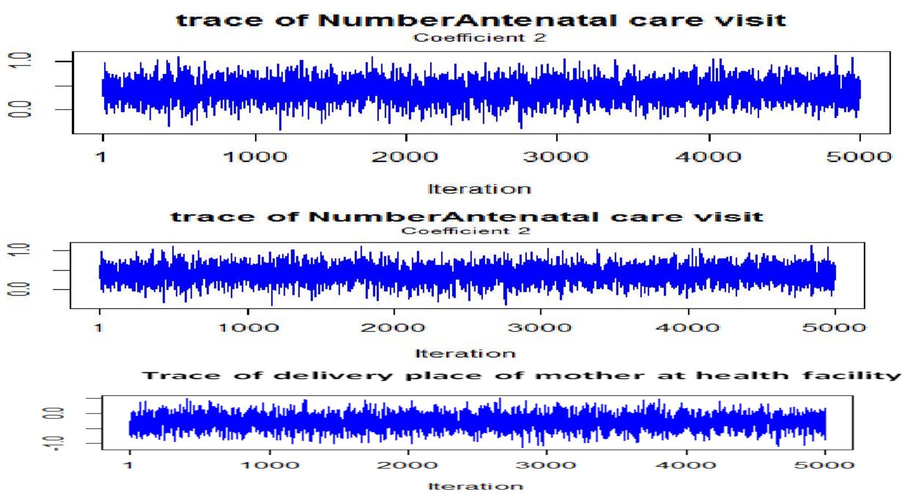



Figure 4.5: Trace plots for convergence of coefficients for Number of antenatal care visit and delivery place of mother.

**4. Autocorrelation plot**: Autocorrelation plot produces lag-autocorrelations for the monitored parameters within each chain. In Markov chain autocorrelation analysis, it is necessary to identify lags in the series in order to calculate the long-run trends in correlation, and in particular whether they decrease with increasing lags. In Figure 4.6 below, Autocorrelations for all lags were closed to zero. So the figure has an evidence of convergence. Not all Autocorrelation plots are presented here; the rest plots can be found in appendices (see appendix C).

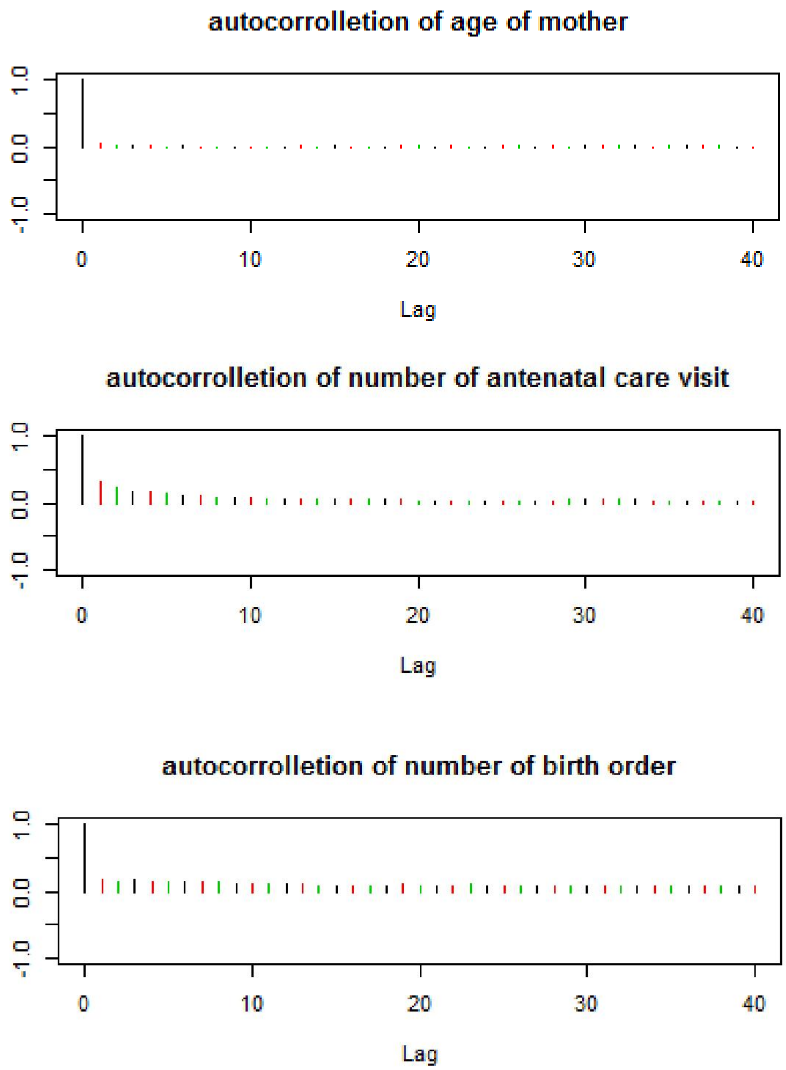



Figure 4.6: Convergence of autocorrelation plots for coefficients age of mother, Number of antenatal care visit, and Number of birth order of mother.

### 4.7 Discussions

The descriptive results of this study indicated that the percentage of maternal death was higher in eastern part of Ethiopia (Afar (3.4%)), Somali (2.8%)) and lower in two administrative city of Ethiopia, Addis Ababa (0.4%) and Dire Dawa (0.7%). Similarly, the percentage of maternal death for urban women was less than that of rural, place of delivery in health facilities was less than that of home; wealth index rich was less than wealth index poor and the percentage of maternal death for mother who visit Antenatal care 4 and above less than for mother who were no Antenatal care visit.

Before passing to model based data analysis the presence of spatial dependency were tested by using exploratory spatial data analysis like Moran I(I = 0.260 P-value=0.03462) and Geary'c (I=0.487, P-vale=0.04813). From this result we have enough evidence as maternal death between neighboring regions of Ethiopia was spatially correlated. This idea was supported by previous study which conducted in Uganda western region by (Michelle M. Schmitz (2018)).

Regarding to model based data analysis part, three models (M1, M2, and M3) which represent Bayesian Generalized linear model, Bayesian semi parametric regression model and Bayesian Geo-additive regression model were used in the analyses. At the first step of the model fit the Bayesian Generalized linear model without assuming non-linear effect of metrical covariate and spatial effect has been fitted. From the result of the model adequacy Bayesian Geo-additive regression model is the best fitted model and it was selected for discussion purpose in this study.

After investigating influential factors for the pregnancy related death of mother, the fixed-effects of Bayesian Geo-additive Regression model showed the importance of delivery place of mother, Number of antenatal care visits during pregnancy, wealth index rich category, marital status of mother, age of mother and number of birth order of mother.

Concerning to delivery place of mother, the findings revealed that mothers who were delivered at health facility were in lower risk of pregnancy related death than mothers who delivered at home. On the other hand it means that the odds of maternal death in health facilities was (OR= exp (-0.4447) = 0.64) times less likely than the odds of maternal at home. This is also consistent with the previous study by (Navaneetham K. &., 2002).

We also found that Pregnancy related death was lower among mothers who had 1-3 and 4& above Number of antenatal care visit than among mothers who haven't antenatal care visit. This can be also interpreted as (OR= exp (-2.1607 ) =0.093) for mothers who had 1-3, exp(-1.3991 )= 0.24 for mothers who had 4& above) ,which mean that the odds of pregnancy related death of mother who had 1-3 was 0.093 times less likely than the odds of mothers who haven't antenatal care visit and the odds of pregnancy related death of mother who had 4 & above number of antenatal care was 0.24 times less likely than the odds of mothers who haven't antenatal care visit. More over this finding was almost supported by previous study by (Jarso S.*et al*, 2019)

With regard to wealth index, our result revealed that mothers from rich wealth index were faced lower risk of maternal death than the mother from poor wealth index (low income) which is also consistent with the previous study by (Bayati, 2016).

Marital status of mother was another significant factor for pregnancy-related death. Under this, we observed that women with no longer living together/separated and widowed/divorced women was positively affected pregnancy-related death which mean that women who were no longer living together/separated and widowed/divorced faces higher risk of pregnancy related death as compared with not married women's. This result was also supported by previous study (Illah, 2010).

This study also found an interesting result which show that the non-linear effect of continuous variable such as age of mother, number of birth order of mother and spatial effects on pregnancy related death in Ethiopia.

The effects of age of mother indicate a continuous worsening of pregnancy related death during the age of mother less than 20 years old and between age interval (30-39) which indicate that mother younger than 20 years old and age interval (30-39) were experienced high risk of pregnancy related death (maternal death) in relative to mother age interval (20-29).This result has the same idea with previous study done by (Alfred Kwesi ,2012).

Number of Birth order of mother has also a slight effect on the pregnancy related death, as mother with higher birth order (>5), was associated with high risk of maternal death than mother with lower birth order. This result was also consistent with previous study (Michel G.*et.al*, 1997).

Spatial effect is one of the very important factors included in this study. Bayesian geo-additive regression model result revealed that, the estimates of the presumed spatial correlated regional level random effects in fact showed strong evidence of spatial dependence in Ethiopia on maternal death. As the results showed the eastern part of countries such as Afar and Somali were associated with a higher risk of maternal death. In addition, some regions like Benishangul gumuz and Gambela were also experienced with high risk maternal death.

# CHAPTER FIVE

# 5. CONCLUSIONS AND RECOMMENDATIONS

## 5.1 Conclusions

The study primarily aimed to see spatial pattern and the effect of socioeconomic, bio-demographic and health factors on maternal death using the 2016 EDHS data. This is achieved through the use of three different fully Bayesian based on Markov chain Monte Carlo techniques in increasing complexity from simple to more sophisticated model namely Bayesian Generalized linear model, Bayesian semi parametric regression model and Bayesian Geo-additive regression model.

Our findings support the notion that maternal mortality is a basic problem of public health issue in the country with spatial variation across different regions. The results of this study suggested that there are complex social, demographic and geographic processes operating in maternal mortality. This result can be more clearly understood using the appropriate statistical models.

Spatial pattern of maternal death in Ethiopia was explored and the hot spot regions were identified. This pattern suggests that the highest risk of maternal death was found in the Afar, Somalia, Gambela and Benishangul Gumuz regions.

## 5.2 Recommendations

Based on the findings of this study, the researcher recommended the following points for Government and other concerned bodies:

1. Even though maternal death was becoming decrease and decrease throughout the year, still there is a risk of maternal death in some region of Ethiopia. Thus government and other concerned sectors should have to focus on controlling maternal death with special focus to regions that have a high severity of this problem.
2. The data of this study was basically secondary which have the problem of missing data and the expected potential variables were also not availed ,hence it is great headache for researcher for further study. Thus the concerned bodies that collect such type of data should have to focus on data collection process in order to handle all the potential variables and collecting full information for each individual.
3. The government and other concerned bodies should have to take attention to control the significant factors that associated to maternal death by sharing knowledge on necessity of delivering at health facility rather than at home, usage of antenatal care follow up and

concerned body should have to do on the way of up grading the economic status of the mother.

4. The posterior estimate of this study is totally determined with the methods of MCMC which actually not fast as compared with INLA Method. Thus, the researchers should strongly recommend so that to fit the same model by the both method and compare two of them.
5. Different study showed that time is also one of the basic factors which affect maternal death. Thus, the researchers should strongly recommend so that to do the same study by using space-time modeling by including time effect in model.

**Limitation of the study**

The data used for this study is obtained from the Ethiopian demographic health survey (EDHS) of 2016), that have not full information, due to some women were not eligible under some covariates and the non-exhaustiveness of the factors which might be relevant to maternal mortality, but the variables are not included in the EDHS. The other limitation of this study is that, it is not possible to produce risk maps at the lowest administrative units as the EDHS data is not sampled by clustering the region/country by the lowest administrative units, called *Wereda*.

## Reference

Adebayo S, Fahrmeir L. Klasen S.(2004).Analyzing infant mortality with geoadditive categorical regression models: a case study for Niger ia. *Economics and Human Biology* 2: 229-244.

Ahmed, S. Li, Q., Liu, L & Tsui, A.O.(2012). Maternal deaths averted by contraceptive use: an analysis of 172 countries. *Lancet:* 380(9837):111-25 .

Alfred Kwesi (2012). Spatial patterns and trends of maternal mortality over a Five year period and their associated risk factors in ifakara Health and demographic surveillance site (IHDSS).

Bailey, T. C. and Gatrell, A. C. (1995). Intereactive Spatial Data Analysis, Longman, Harlow.

Bayati, M. V. ( 2016). Determinants of maternal mortality in Eastern Mediterranean region: A panel data analysis. *Medical journal of the Islamic Republic of Iran*, 30, 360.

Benova, L.,Cumming,O.,&Campbell O. M. R. (2014).Systematic review and meta-analysis: Association between water and sanitation environment and maternal mortality. *Tropical Medicine and International Health, electronically published ahead of printing* .

Besag J, York Y, Mollie A.(1991).Bayesian image restoration with two applications in spatial statistics (with discussion) *.Annals of the Institute of Statistical Mathematics;43:1–59* .

Bicego G, Boerma JT, Ronsmans.(2002).The effect of AIDS on maternal mortality in Malawi and Zimbabwe. *AIDS , 16(7):1078-1081* .

Bloncowe H.,Cousens S .(2011). Clean birth and postnatal care practices to reduce neonatal deaths from sepsis and tetanus a systematic review and Delphi estimation of mortality effect .

Colston, J. M., & Burgert, C. R. (2010). Using geospatial analysis to inform decision making in targeting health facility-based programs.

Cressie, N. (1993). Statistics for spatial data, Wiley, New York .

CSA [Ethiopia] and ICF International.(2012).Ethiopia Demographic and health Survey 2011Addis Ababa, Ethiopia and Calverton Maryland, USA. *:Centeral Statistic Agency and ICF International, 1– 430*.

Dahiru, T. &.(2015). Determinants of antenatal care, institutional delivery and postnatal care services utilization in Nigeria. *Pan African Medical Journal*, 21(1).

Dankelman, I., & Davidson, J.(1988).Women and environment in the third world. *Alliance for the future London: Earth scan Publications Limited* .

Dawit Getnet Ayele, Temesgen Zewotir, Henry Mwambi. (2016).Structured additive regression models with spatial correlation to estimate under-five mortality risk factors in Ethiopia.

Department of Public Information – UN Web Services Section. Goal 3: Ensure healthy lives and promote wellbeing for all at all ages. Sustainable Development Goals: 17 Goals to Transform Our World. Retrieved from http://www.un.org/sustainabledevelopment/health/.

Echavarria L. (2004).Semiparametric Bayesian count data models. LudwigMaximiliams-University of Munich.

Ethiopian demographic and health survey. (2011).

Ethiopian demographic and health survey. (2016).

Fahrmeir L, Lang S.( *2001).* Bayesian inference for generalized additive mixed models based on Markov random field priors. *Applied Statistics(JRSS C) 2;50:201–220* .

Fahrmeir, L. & Tutz, G. (2001).Multivariate Statistical Modeling Based on Generalized Linear Models, Springer, New York.

Gamerman, D. (1997). Efficient Sampling from the posterior distribution in generalized linear models. *Statist. Comput. 7, 57-68.*

Goodchild M.. (1987). A spatial analytical perspective on geographical information systems. *International Journal of GIS 6(5), 407-423* .

Hastie T, Tibshirani R.Generalized Additive Models. Chapman & Hall: London, 1990.

Illah, E. O. (2010.). Assessment of risk factors associated with maternal mortality in rural Tanzania (Doctoral dissertation).

International, P. (2013). Maternal Health in Rufiji HDSS Causes and Risk Factors for Maternal.

Jarso S,Yusuf H,Achamyelesh G. (2019). ;Determinants of Maternal Death in a Pastoralist Area of Borena Zone, Oromia Region, Ethiopia: Unmatched Case-Control Study .

Jat T, Sebastian M, Ng N. Factors affecting the use of maternal health services in Madhya Pradesh state of India: a multilevel analysis. *International Journal for Equity in Health 2011; 10:59. doi: 10.1186/ 14759276-10-59 PMID: 22142036Mortality in Rural Tanzania - Case of Rufiji Health and Demographic Surveillance Site ( HDSS )', 17(September), pp. 119–130* .

Kammann EE, Wand MP.(2003).Geoadditive models. *Journal of the Royal Statistical Society;52:1–18.*

Kneib T. (2006) Mixed model based inference in structured additive regression. PhD thesis. Munich: University of Munich. Institute of Statistics.

Kandala N .(2006). Bayesian geo-additive modelling of childhoodmorbidity in Malawi .

Kandala N Lang S, Klasen S, Fahrmeir L.(2001).Semiparametric analysis of the socio-demographic and spatial determinants of undernutrition in two African countries. *Collaborative Research Center, University of Munich, 386 Kneib* .

Kazembe LN, Mpeketula PMG .(2010). Quantifying Spatial Disparities in Neonatal Mortality Using a Structured Additive Regression Model. PloS ONE 5(6): e11180. Doi:10.1371/*journal.pone*.0011180

Kneib T, Fahrmeir L .(2007). A mixed model approach for geoadditive hazard regression. *Scandinavian Journal of Statistics 34: 207-228* .

Kolo.(2017). Spatial and Temporal Variations of Maternal Deaths in Borno State of Nigeria. *An Institutional Based Study ; IJTDH, 27(1): 1-14, 2017; Article no.IJTDH.36007* .

Lang S, Brezger A. Bayesian P-splines.*Journal of Computational and Graphical Statistics2004;13:183–217* .

Liu, Z. (2011). RH knowledge and service utilization among unmarried rural-to-urban migrants in three major cities, China. *BMC public health*, 11(1), 74.

Mamaye. (2016). Facts and figures on the link between water, sanitation and hygiene (WASH) and maternal and newborn health (MNH). *London:Evidence for Action* .

McCullagh P, Nelder JA .(1989).Generalized Linear Models,2nd edition. Chapman & Hall, London

McCarthy, J. and Maine, D. (1992). A framework for analyzing the determinants of maternal mortality. *Studies in Family Planning, 23(1):23 { 33* .

Michel G., Khadidiatou M., Mohamed D., Paul C.(1997) .Risk Factors for Maternal Mortality: A Case-Control Study in Dakar Hospitals (Senegal). *African Journal of Reproductive Health / La Revue Africaine de la Santé Reproductive*, Vol. 1, No. 1

Michelle M. Schmitz(2018). Understanding Local Geospatial Variation in Maternal Mortality Outcomes: An Analysis in Uganda's Western Region

Ministry of Health and Social Services (MoHSS). (2009). National maternal and Child Conference Report.

Navaneetham, K. &. (2002). Utilization of maternal health care services in Southern India. *Social science & medicine,* 55(10), 1849-1869.

Nelder, J. A. & Wedderburn R. W. M. (1972). Generalized linear models. *Journal of the Royal Statistical Society Series A 135, 370–384.*

Nove A, Matthews Z, Neal S, Camacho AV.( 2014*)*.Maternal mortality in adolescents compared with women of other ages: evidence from 144 countries. *Lancet Glob Health ;* published online Feb 6. http://dx. Doi.org/10.1016/S2214-109X(13)70179-7 .

Odland, M. (1987). Spatial Autocorrelation. Sage: Beverly, CA.

Ojo OB, L. S. (2017). Bayesian generalized linear mixed modeling of Tuberculosis using informative priors. PloS ONE . journal. Pone.0172580, 12(3): e0172580. Doi:10.1371.

Omo-Aghoja, E. A. (2008). Maternal Mortality at the Central Hospital, Benin City Nigeria African. *Journal of Reproductive Health,* 12 (3).

Prince, O. (2010). Statistical methods of disease mapping. *Journal of Royal Statistical* .

Razum, O. and Jahn, A. (2000). Marital status and maternal mortality. *Lancet, 355(9221):2169 { 2169* .

Royston E, Armstrong S .(1989). Preventing maternal deaths. Geneva. *World Health Organization*

Samson B, Adebayo and Ezra Gayawan.(2009). Stepwise Geoadditive Regression Modelling of Levels and Trends of Fertility in Nigeria: Guiding Tools towards Attaining MDGs

Schabenberger O, Gotway CA. *(2005)* Statistical Metods for Spatial Data Analysis. *Chapman and Hall/CRC; Goovaerts P, Jacquez GM, Greiling D:Exploring scale-dependentcorrelations between cancer mortality rates using factorial kriging andpopulation-weighted* .

Stephan, F. (1934). Sampling errors and the interpretation of social data ordered in time and space. *Journal of the American Statistical Association 29,165-166* .

Spiegelhalter D., Best N., Carlin B., and Van der Line A. (2002). Bayesian Measures of Models Complexity. *Journal of the Royal Stat. Soc. B(64):1-34* .

Sharmistha Self and Richard Grabowski. 2004. Factors influencing maternal health care in Nepal: the role of socioeconomic interaction. *Asia-Pacific Sustainable Development Journal. Vol. 25, No. 2*

Tlou,B.,Sartorious B.,Tanser F.(2017). Space-time patterns in maternal and mother mortality in a rural South African population with high HIV prevalence (2000–2014) ratio at national and sub-national levels and its application for describing sub-national variations of maternal death in Ethiopia.

Tobler, R. (1970). A computer movie simulating urban growth in the Detroit region. *Economic geography supplements 46,234-240.*

Tsegay,Y, Gebrehiwot T, Goicolea I, Edin K, Lemma H and Sebastian MS . (2013). Determinants of antenatal and delivery care utilization in Tigray region, Ethiopia: a crosssectional study. *International journal for equity in health 12: 30* .

Weyesa, J. B. (2015). Prevalence and risk factors associated with maternal mortality in MizanAman Hospital, Bench Maji, Southwest Ethiopia. . J Women's Health Care, 4(274), 2167-0420.

WHO,International Classification of Diseases (ICD) 10 .(2010). Geneva, Switzerland.

WHO.(2011).Guidelines on preventing early pregnancy and poor reproductive outcomes among adolescents in developing countries. Geneva: World Health Organization.

WHO, &. U. (2015). Trends in maternal mortality: 1990-2015: estimates from. WHO, UNICEF, UNFPA, World Bank Group and the United Nations Population Division. WHO, U. U. (2012). Trends in Maternal Mortality: 1990 to 2010. *WHO, UNICEF,UNFPA and World bank estimates. , 1–72* .

WHO,(2016). Trends in Maternal Mortality: 1990 to 2015, Estimates by WHO, UNICEF,UNFPA, The World Bank and the United Nations Population Division.

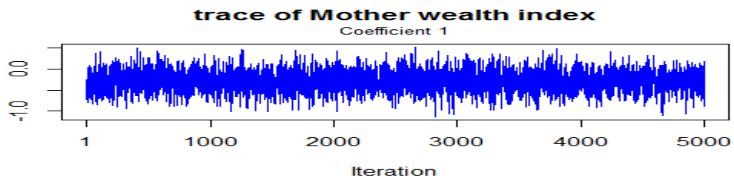
# Appendices

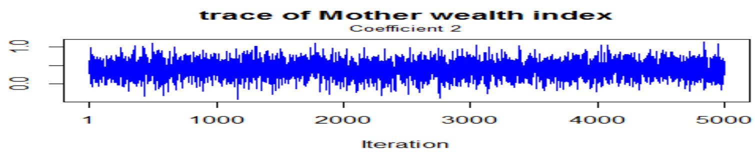
## Appendices A

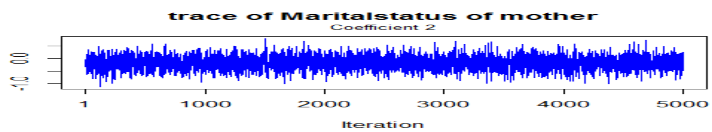
Table 4.7:Bivariate Association between maternal mortality and its determinants.

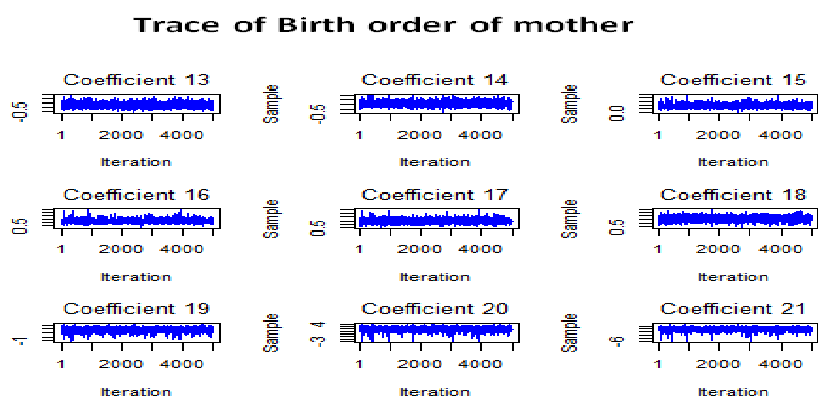
| Independent variables | Survival status of mother | |
|---|---|---|
| | Pearson's Chi-squared | p-value |
| Region | 45.87537 | 0.000 |
| Residence | 0.001652532 | 0.967 |
| Marital status | 8.315827 | 0.0399 |
| Delivery place | 7.387923 | 0.006 |
| ANC visits | 11.52873 | 0.003 |
| Source of drinkable water | 0.3428717 | 0.842 |
| Mother education level | 8.371329 | 0.038 |
| Wealth index | 7.913881 | 0.019 |
| Exposure to mass media | 0.0176922 | 0.894 |
| Contraceptive usage | 0.05547198 | 0.813 |

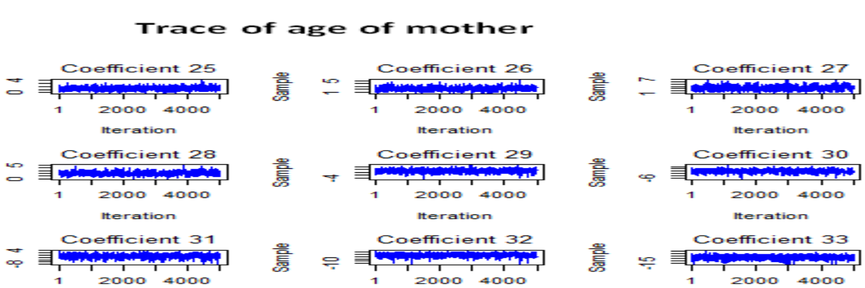
Appendices B: Trace plot of coefficients for convergence



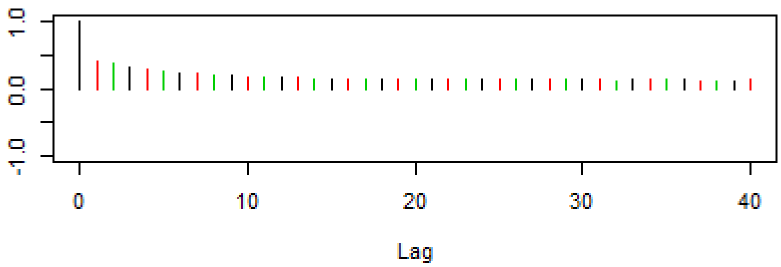



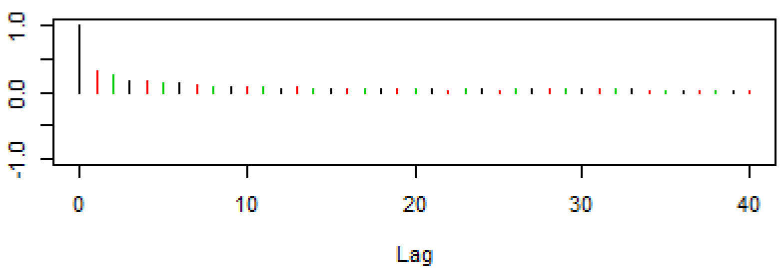



Appendices C**:** Autocorrelation plot of coefficients for convergence

**autocorrolletion of delivery place of mother**




**autocorrolletion of wealth index of mother**




**autocorrolletion of Marital status of mother**