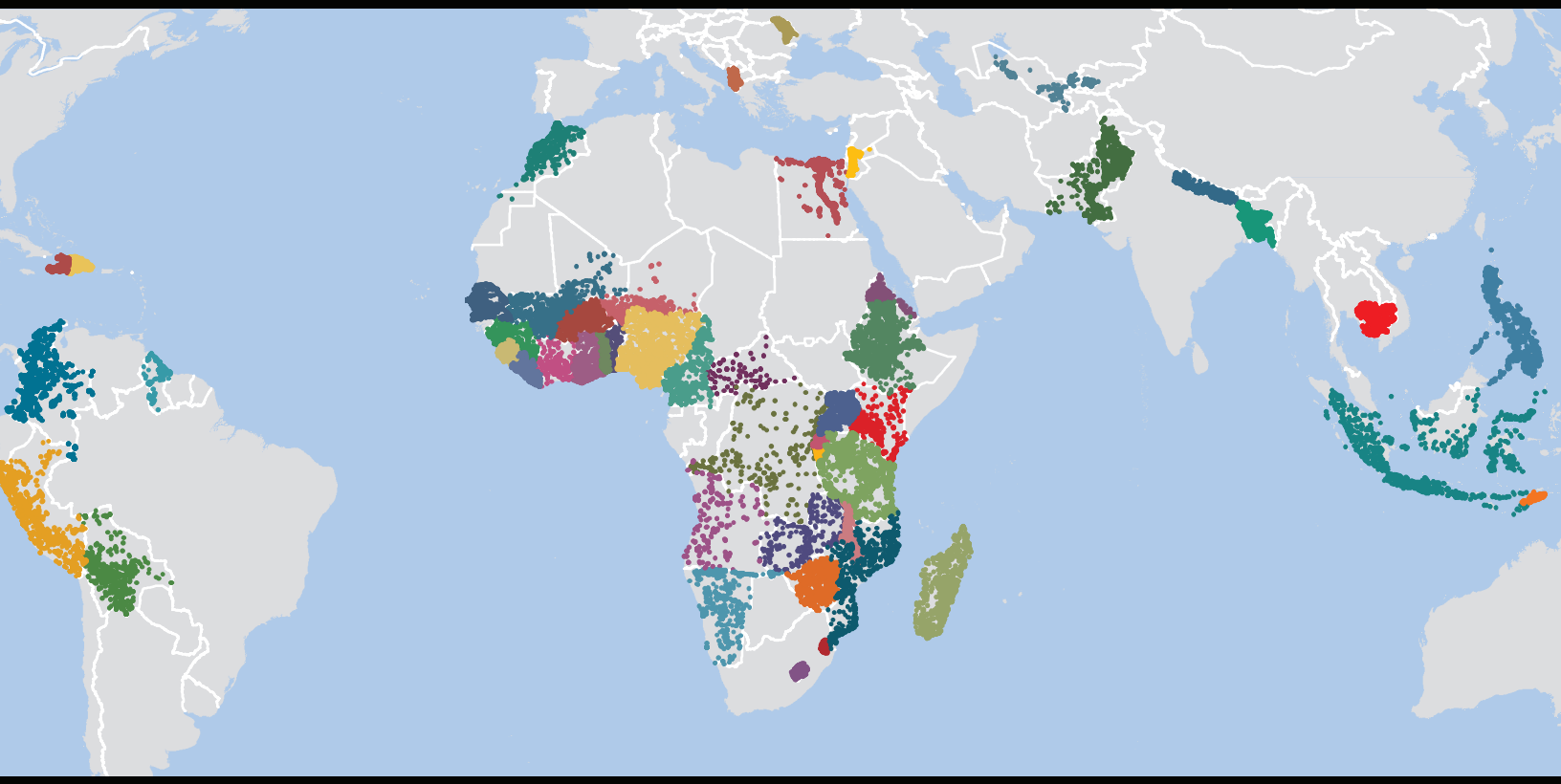
USAID
FROM THE AMERICAN PEOPLE

GEOSPATIAL MODELING OF CHILD MORTALITY ACROSS 27 COUNTRIES IN SUB-SAHARAN AFRICA

DHS SPATIAL ANALYSIS REPORTS 13



AUGUST 2016
This publication was produced for review by the United States Agency for International Development (USAID). The report was prepared by Carla Pezzulo, Tomas Bird, Edson C. Utazi, Alessandro Sorichetta, Andrew J. Tatem, Jennifer Yourkavitch, and Clara R. Burgert-Brucker.

DHS Spatial Analysis Reports No. 13

Geospatial Modeling of Child Mortality across 27 Countries in Sub-Saharan Africa

Carla Pezzulo[1,2]
Tomas Bird[1,2]
Edson C. Utazi[3,1]
Alessandro Sorichetta[1,2]
Andrew J. Tatem[1,2]
Jennifer Yourkavitch[4]
Clara R. Burgert-Brucker[4]

ICF International
Rockville, Maryland, USA

August 2016

[1] WorldPop, Department of Geography and Environment, University of Southampton, Southampton, UK
[2] Flowminder Foundation, Stockholm, Sweden
[3] Southampton Statistical Sciences Research Institute, University of Southampton, Southampton, UK
[4] ICF International, Rockville, MD

Corresponding author: Clara R. Burgert-Brucker, International Health and Development, ICF International, 530 Gaither Road, Suite 500, Rockville, MD 20850, USA; phone: +1-301-572-0446; fax: +1-301-407-6501; email: clara.burgert@icfi.com

Acknowledgment: The authors would like to acknowledge Tom Pullum and Trevor Croft for their assistance and guidance in developing this report, and Shea Rutstein, Danzhen You, and Jon Pedersen for their review of this report.

Editor: Diane Stoy
Document Production: Natalie La Roche

This study was carried out with support provided by the United States Agency for International Development (USAID) through The DHS Program (#AID-OAA-C-13-00095). The views expressed are those of the authors and do not necessarily reflect the views of USAID or the United States Government.

The DHS Program assists countries worldwide in the collection and use of data to monitor and evaluate population, health, and nutrition programs. For additional information about The DHS Program contact: DHS Program, ICF International, 530 Gaither Road, Suite 500, Rockville, MD 20850, USA. Phone: 301-407-6500; fax: 301-407-6501; email: reports@dhsprogram.com; Internet: www.dhsprogram.com.

Recommended citation:

Carla Pezzulo, Tomas Bird, Edson C. Utazi, Alessandro Sorichetta, Andrew J. Tatem, Jennifer Yourkavitch, and Clara R. Burgert-Brucker. 2016. *Geospatial Modeling of Child Mortality across 27 Countries in Sub-Saharan Africa.* DHS Spatial Analysis Reports No. 13. Rockville, Maryland, USA: ICF International.

Contents

Tables ..... **v**
Figures ..... **v**
Abbreviations ..... **vii**
Preface ..... **ix**
Abstract ..... **xi**
Executive Summary ..... **xiii**
1. Introduction **1**
1.1. Background 1
1.2. Rationale for DHS Spatial Analysis Report 13 7
2. Data **9**
2.1. Study Area and Data Description 9
2.2. Data Preparation 15
3. Analysis Methods **27**
3.1. Overview of Analysis Plan 27
3.2. Covariate and Model Selection 27
3.3. Spatial Analysis Using a Conditional Autoregressive (CAR) Model 29
4. Results **31**
4.1. Results from Covariate and Model Selection 31
4.2. Results from Spatial Analysis using a Conditional Autoregressive (CAR) Model 34
5. Discussion **43**
6. Policy Implications **47**
7. Summary and Conclusions **49**
References **51**
Appendix A **61**
A.1 United Nations Geographical Region Composition for the 27 Sub-Saharan African Countries included in Our Analysis 61
A.2. Geospatial Datasets 62
Appendix B – Detailed Methods and Results **87**
B.1. Conditional Autoregressive (CAR) Model 87

Tables

Table 1. DHS indicators tested in analysis and their relevance to child mortality 4
Table 2. Spatially relevant factors tested in the analysis and their relevance to child mortality 6
Table 3. Countries, number of DHS sub-national areas in each country, and year of the DHS surveys used in the framework of this study 11
Table 4. Full range of geospatial datasets and derived covariates tested for modeling child mortality in the study area 24
Table 5. Summary of single covariate models. Coefficients are presented as both unstandardized (i.e., native units) or standardized..... 31
Table 6. Top interaction models 34
Table 7. Posterior estimates of the parameters of the spatial model..... 35
Table A.1. List of countries in analysis by United Nations Regions..... 61
Table B.1. Top 10 models resulting from the model selection process..... 88
Table B.2. Posterior means of the parameters of the spatial model without region-level interactions 88

Figures

Figure 1. Conceptual model describing prevention and treatment interventions for, as well as distal factors affecting, the main causes of death for children under age 5 in sub-Saharan Africa, as identified by Liu et al. (2012)..... 2
Figure 2. Focus countries and year of the DHS survey used for this study 9
Figure 3a. Percentage of literate women (age 15-49) in the 225 DHS sub-national areas 16
Figure 3b. Percentage of births with a birth interval less than 36 months between birth and conception in the 255 DHS sub-national areas 16
Figure 3c. Percentage of children born in last 5 years with access to a health facility in the 225 DHS sub-national areas (access defined as delivery in a health facility) 17
Figure 3d. Percentage of stunting children age (12-59 months) in the 255 DHS sub-national areas..... 17
Figure 3e. Percentage of live children age 12-23 months who received the measles vaccination at any time prior to the survey in the 255 DHS sub-national areas..... 18
Figure 3f. Percentage of population practicing open defecation in the 225 DHS sub-national areas 18
Figure 4a. Population density in each DHS sub-national area..... 20
Figure 4b. Satellite-based night-time lights' intensity spatially averaged within each DHS sub-national area..... 20
Figure 4c. Number of dominant ethnic groups, indicating ethnic variability, in each DHS sub-national area..... 21
Figure 4d. Spatially averaged economic activity estimated using the grid-based equivalent of the purchasing parity gross domestic product, within each DHS sub-national area..... 21
Figure 4e. Annual mean temperature spatially averaged within each DHS sub-national area 22
Figure 4f. Population-weighted mean *Plasmodium falciparum* prevalence in each DHS sub-national area 22
Figure 5. United Nations geoscheme regions used in the analysis 23
Figure 6. Flow diagram of analysis methods adopted in the study..... 27
Figure 7a. Plots of selected DHS variables against child mortality rates. The blue lines are linear regression lines while the red lines are nonparametric *loess* fits to the data. 32
Figure 7b. Plots of selected geospatial covariates against child mortality rates. The blue lines are linear regression lines while the red lines are nonparametric *loess* fits to the data. 33

Figure 8a. Observed mortality rates and 95% confidence interval width for the 255 DHS sub-national areas in sub-Saharan Africa 38
Figure 8b. Predicted mortality rates for the 255 DHS sub-national areas in sub-Saharan Africa 38
Figure 8c. Predicted standard deviations of the child mortality rates for the 255 DHS sub-national areas in sub-Saharan Africa 39
Figure 8d. Observed minus predicted child mortality rates for the 255 DHS sub-national areas in sub-Saharan Africa 39
Figure 9. A plot of observed vs. predicted child mortality rates 40
Figure 10. Classification of child mortality regions 41
Figure A.1. Population distribution 62
Figure A.2a. MODIS-based spatial distribution of built-up areas 63
Figure A.2b. GHSL-based spatial distribution of built-up areas 64
Figure A.3. Spatial pattern of VIIRS-based night-time lights' intensity 65
Figure A.4a. Travel-time to major centers with more than 50,000 inhabitants 66
Figure A.4b. Travel-time to major centers with more than 100,000 inhabitants 67
Figure A.4c. Travel-time to major centers with more than 500,000 inhabitants 67
Figure A.5a. Spatial distribution of dominant ethnic groups 68
Figure A.5b. Legend of Figure A.5a 69
Figure A.6. Location of conflict events and associated fatalities between 2010 and 2013 70
Figure A.7. Spatial distribution of the gross cell product representing the regional equivalent of gross domestic product 71
Figure A.8a. Sheep density 72
Figure A.8b. Cattle density 73
Figure A.8c. Pig density 73
Figure A.8d. Goat density 74
Figure A.9a. 30-year average annual temperature 75
Figure A.9b. 30-year average annual rainfall 75
Figure A.10. 14-year average MODIS-based daytime Land Surface Temperature 76
Figure A.11. 50 year-based Aridity Index 77
Figure A.12. Spatial distribution of the growing season length 78
Figure A.13. Relative frequencies of drought occurrences and their spatial distribution 79
Figure A.14. 14-year average MODIS-based Enhanced Vegetation Index 80
Figure A.15a. Topographic elevation 81
Figure A.15b. Topographic slope 82
Figure A.15c. Topographic roughness calculated using a neighborhood window of 3x3 grid cells 82
Figure A.15d. Topographic roughness calculated using neighborhood window of 5x5 grid cells 83
Figure A.15e. Topographic roughness calculated using neighborhood window of 11x11 grid cells 83
Figure A.16. MERIS-based landcover classification 84
Figure A.17. 15-year average *Plasmodium falciparum* prevalence distribution 85

Abbreviations

4q1	Child mortality rate (calculated as the number of deaths among children between the age of 1 and 4 in the five-year period preceding the survey, which is expressed per 1,000 children surviving to age 12 months)
ACLED	Armed Conflict Location & Event Data Project
AIC	Akaike information criterion
CAFR	Central Africa Region
CAR	Conditional autoregressive model
CGIAR	Consortium of International Agricultural Research Centers
CIESIN	Center for International Earth Science Information Network
CPR	Contraceptive Prevalence Rate
DHS	Demographic and Health Surveys
DPT	Diphtheria, pertussis, and tetanus
DRC	Democratic Republic of the Congo
EAFR	Eastern Africa Region
EBF	Exclusive breastfeeding
ESA	European Space Agency
EVI	Enhanced Vegetation Index
FAO	Food and Agriculture Organization
GIS	Geographic Information System
GLW2	Gridded Livestock of the World, version 2.0
GREG	Geo-referencing of ethnic groups
LST	Land Surface Temperature
MCMC	Markov chain Monte Carlo methods
MDG	Millennium Development Goals
MODIS	Moderate Resolution Imaging Spectroradiometer
NPP	National Polar-orbiting Partnership
NOAA	National Oceanic and Atmospheric Administration
SAR12	DHS Spatial Analysis Reports 12
SDG	Sustainable Development Goal
UN	United Nations
UNDP	United Nations Development Programme
UNFPA	United Nations Population Fund
UNICEF	United Nations Children's Fund
USAID	United States Agency for International Development
WAFR	Western Africa Region
WHO	World Health Organization

Preface

The Demographic and Health Surveys (DHS) Program is one of the principal sources of international data on fertility, family planning, maternal and child health, nutrition, mortality, environmental health, HIV/AIDS, malaria, and provision of health services.

The DHS Spatial Analysis Reports supplement the other series of DHS reports to meet the increasing interest in a spatial perspective on demographic and health data. The principal objectives of all DHS report series are to provide information for policy formulation at the international level and to examine individual country results in an international context.

The topics in the DHS Spatial Analysis Reports are selected by The DHS Program in consultation with the U.S. Agency for International Development. A range of methodologies are used, including geostatistical and multivariate statistical techniques.

It is hoped that the DHS Spatial Analysis Reports series will be useful to researchers, policymakers, and survey specialists, particularly those engaged in work in low- and middle-income countries, and will be used to enhance the quality and analysis of survey data.

Sunita Kishor
Director, The DHS Program

Abstract

Preventable mortality of children has been targeted as one of the UN's Sustainable Development Goals for the 2015-30 period. Global decreases in child mortality (4q1) have been seen, although sub-Saharan Africa remains an area of concern, with child mortality rates remaining high relative to global averages or even increasing in some cases. Furthermore, the spatial distribution of child mortality in sub-Saharan Africa is highly heterogeneous. Thus, research that identifies primary risk factors and protective measures in the geographic context of sub-Saharan Africa is needed. In this study, household survey data collected by The Demographic and Health Surveys (DHS) Program aggregated at DHS sub-national area scale are used to evaluate the spatial distribution of child mortality (age 1 to 4) across 27 sub-Saharan Africa countries in relation to a number of demographic and health indicators collected in the DHS surveys. In addition, this report controls for spatial variation in potential environmental drivers of child mortality by modeling it against a suite of geospatial datasets. These datasets vary across the study area in an autoregressive spatial model that accounts for the spatial autocorrelation present in the data.

This study shows that socio-demographic factors such as birth interval, stunting, access to health facilities and literacy, along with geospatial factors such as prevalence of *Plasmodium falciparum* malaria, variety of ethnic groups, mean temperature, and intensity of lights at night can explain up to 60% of the variance in child mortality across 255 DHS sub-national areas in the 27 countries. Additionally, three regions - Western, Central, and Eastern Africa - have markedly different mortality rates. By identifying the relative importance of policy-relevant socio-demographic and environmental factors, this study highlights priorities for research and programs targeting child mortality over the next decade.

Executive Summary

In 2015, the United Nations established the Sustainable Development Goals (SDGs), which followed the eight Millennium Development Goals (MDGs). The aim of the third SDG is ending preventable deaths of newborns and children under age 5, with all countries focused on reducing neonatal mortality at least as low as 12 per 1,000 live births and under-five mortality at least as low as 25 per 1,000 live births (United Nations General Assembly 2015). Sub-Saharan Africa carries about half of the burden of the world's under-five deaths (United Nations 2015).

The third SDG recognizes the importance of addressing different aspects of child and maternal health in order to decrease mortality rates. This will involve reducing risk factors and increasing protections (United Nations General Assembly 2015). Part of the complexity in understanding child mortality across such a large region as sub-Saharan Africa lies in the heterogeneity across the region in mortality rates and associated risk factors (Balk et al. 2004). In addition to variability in socio-economic drivers, sub-national inequalities in health (Arku et al. 2016; Colson et al. 2015; Roberts et al. 2015; Tatem et al. 2012; Wollum et al. 2015), geographic differences in environmental conditions, and spatially varying factors can also affect under-five mortality (Balk et al. 2003; Balk et al. 2004).

This report builds on the work reported in The Demographic and Health Surveys (DHS) Spatial Analysis Report 12 (SAR 12, Burgert-Brucker et al. 2015) that explained spatial patterns in child mortality rates (4q1 - defined as the number of deaths among children between age 1 and 4 in the five-year period preceding the survey per 1,000 children surviving to age 12 months[1]). In SAR 12, descriptive geospatial analyses identified spatial clustering in nine maternal and child health indicators within and among 27 countries at the DHS sub-national area level in sub-Saharan Africa. This report adds geospatial factors to the usual demographic and socio-economic determinants of child mortality in order to (i) identify the geospatial, demographic, and socio-economic covariates (or combinations of them) that might best explain the observed patterns of child mortality at the DHS sub-national area level, (ii) explain patterns in geographic regions, and (iii) account for spatial clustering of the data identified in SAR 12 by using spatially explicit models.

Plausible factors that may affect or be related to child mortality have been assembled and summarized at the DHS sub-national area level. Polygons that represent DHS sub-national areas were obtained from the DHS Spatial Data Repository (ICF International 2008-2015b). A suite of determinants of child mortality were calculated within each of these areas based on the DHS data including socio-economic (housing and literacy), health (vaccination coverage), and demographic (birth interval) factors. Additional geographic information was extracted from a set of geospatial data including indicators such as night-time light intensity, travel time to major population centers, and malaria prevalence. These geospatial covariates were then aggregated within each DHS sub-national area. Covariate and model selection processes were first performed to explore which individual (and combinations of) DHS and geographic covariates could explain the variation in 4q1 in a linear model setting. An autoregressive spatial model was then applied to account for the spatial autocorrelation present in the data and to test the associations among demographic, socio-economic, geospatial factors, and child mortality.

[1] Child mortality rates (age 1 to 4) includes deaths reported at age 12-59 months and excludes infant deaths (deaths occurring between age 0 and 11 months). The DHS Program uses a synthetic cohort life table approach to directly estimate the under-five mortality rate. Details may be found in the Guide to DHS Statistics, pp.90-95 (http://www.dhsprogram.com/publications/publication-DHSG1-DHS-Questionnaires-and-Manuals.cfm) and further described in Pullum et al. 2013.

The results show how a large degree (60%) of the regional variation in child mortality can be explained by the socio-demographic and environmental indicators in the analysis, the interactions among them, and the patterns in child mortality. Socio-demographic factors such as birth interval, stunting, access to health facilities, and literacy, as well as geospatial factors such as the prevalence of *Plasmodium falciparum* malaria, ethnic group variety, mean annual temperature, night-time lights, and economic activity explain a considerable portion of the variance in child mortality across 255 DHS sub-national areas in the 27 countries.

Interestingly, the role of interaction terms suggests that risk factors can be attenuated by protective factors and vice versa. For example, increased literacy reduces the effect of a too-short birth interval on child mortality. Conversely, an increase in literacy is associated with an increase in the effect of access to a health facility on child mortality. The results indicate that literacy rate is a strong predictor of women's reproductive and health behavior, influences access to health facilities, and has an overarching role as a determinant of child mortality. This suggests the need for integrated, multi-dimensional approaches to planning and implementing interventions that are designed to reduce child mortality (age 1 to 4) in sub-Saharan Africa.

After confirming patterns of positive spatial autocorrelation in child mortality rates across the 255 DHS sub-national areas, the results also show differences in child mortality rates across the Western, Central, and Eastern Africa regions, with the highest child mortality rates in Central and Western Africa and the lowest in Eastern Africa.

1. Introduction

1.1. Background

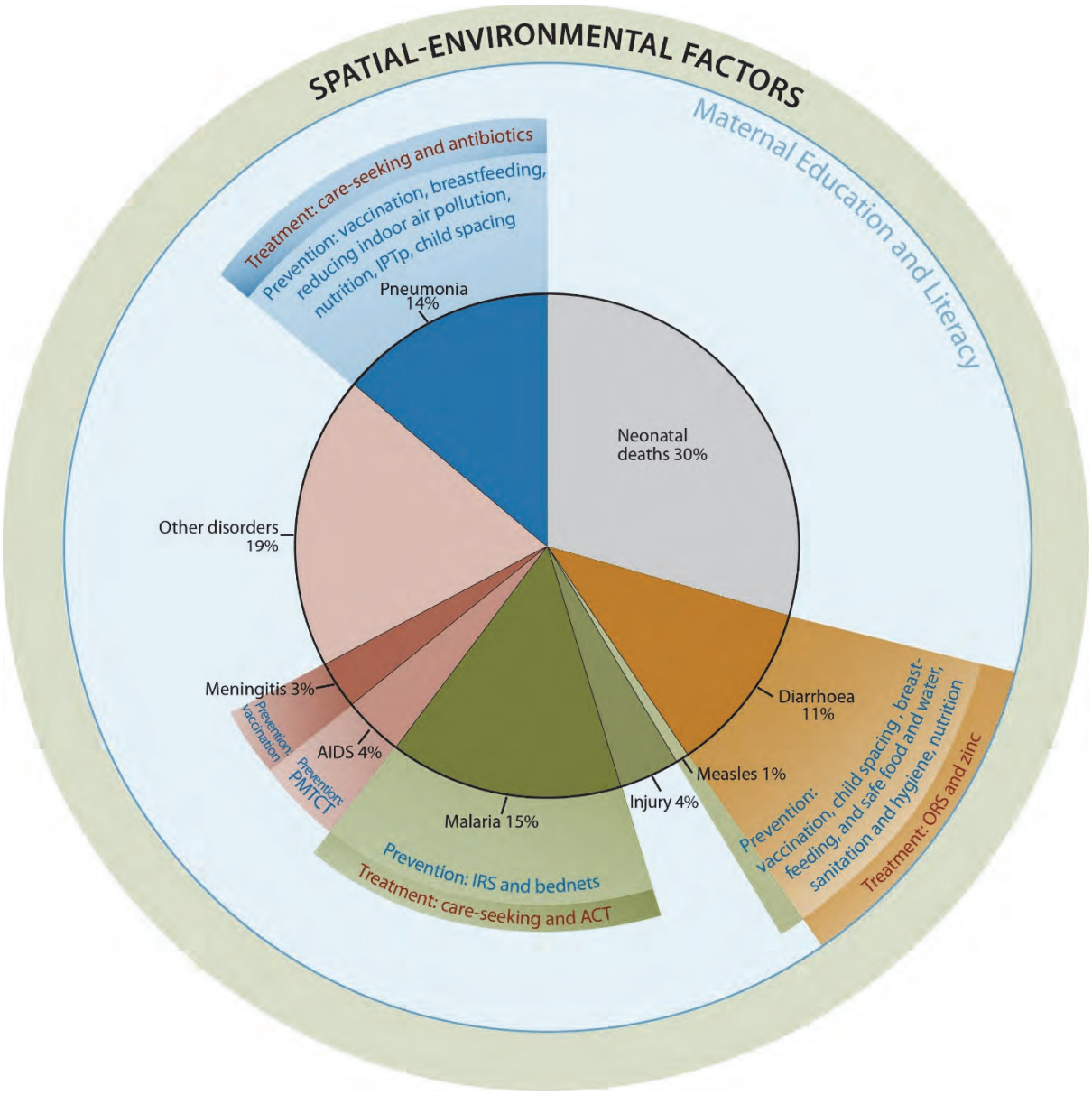
1.1.1. Policy relevance and socio-demographic determinants of child mortality

In 2015, the United Nations established the Sustainable Development Goals (SDGs), following the eight Millennium Development Goals (MDGs). The third SDG is to end preventable deaths of newborns and children under age 5, with all countries aiming to reduce neonatal mortality at least as low as 12 per 1,000 live births and under-five mortality at least as low as 25 per 1,000 live births (United Nations General Assembly 2015). In Africa, progress has been made in many areas to improve survival rates of children (UNICEF et al. 2015; WHO 2016a) through the reduction of diseases and improved access to health facilities and quality health care for mothers (WHO 2016a). In sub-Saharan Africa, under-five mortality rates have fallen from 179 deaths per 1,000 live births in 1990 to 86 deaths per 1,000 live births in 2015 (United Nations 2015). Despite this substantial improvement, sub-Saharan Africa carries about half of the burden of the world's under-five deaths - 3 million in 2015.

The SDG 3 recognizes the importance of addressing different aspects of child and maternal health to further decrease mortality rates. This would involve reducing risk factors such as the prevalence of tropical diseases to increasing protections such as access to high quality maternal health care (United Nations General Assembly 2015). About half of under-five child deaths are due to diseases that are preventable or treatable through simple, affordable interventions (WHO 2016a). A causal pie by Liu et al. (2012) identified the causes of child mortality in sub-Saharan Africa. To conceptualize the relationships among potential covariates related to child mortality for the current study, protective factors (behaviors or conditions that can reduce the risk of mortality through prevention or treatment) were overlaid onto the causal pie and identified distal factors related to child mortality, including mothers' education and the spatial environment. Figure 1 presents a conceptual model that illustrates the conditions driving child mortality in sub-Saharan Africa.

Prevalent causes of death in children are preventable diseases such as pneumonia, diarrhea, and malaria (Figure 1) (UNICEF 2015; WHO 2015b). Risk factors that increase mortality from these diseases include nutrition-related factors such as stunting (Bhutta et al. 2008; Black et al. 2008; Bryce et al. 2005). Lack of adequate sanitation and hygiene has been identified as one of the major causes of diarrheal deaths (Anand and Roy 2016; Fink, Günther, and Hill 2011; WHO 2009), which represent 11% of child deaths in sub-Saharan Africa (Liu et al. 2012, and Figure 1).

Figure 1. Conceptual model describing prevention and treatment interventions for, as well as distal factors affecting, the main causes of death for children under age 5 in sub-Saharan Africa, as identified by Liu et al. (2012).



Conversely, strengthening the health system and infrastructure to provide high quality services (Lavy et al. 1996; WHO 2016a) and improving access to health facilities (Karra, Fink, and Canning 2016) and interventions (Eisele et al. 2012) will help reduce deaths. Maternal education also plays a crucial contextual role as a protective factor against the mortality of children born in rural areas and poor households, or for a mother with no access to basic education (WHO 2016a). Caldwell (1979) suggested that mothers with higher levels of education have improved health knowledge and greater control over health choices for their children, which enhance child survival. Moreover, a mother's education is a protective factor that enables the prevention and management of childhood diseases (UNFPA 2007). Table 1 lists the demographic and socio-economic factors relevant to child mortality that were tested in this analysis.

1.1.2. Spatial heterogeneity in child mortality and driving factors

Part of the complexity in understanding child mortality across a large region such as sub-Saharan Africa lies in the heterogeneity across the area in terms of mortality rates and factors that contribute to those rates (Balk et al. 2004). Overall, countries in Western and Central Africa have the highest burden of under-five deaths. Perhaps unsurprisingly, as compared to Eastern and Southern Africa, these regions also have the slowest rates of improvement in many of the key drivers of mortality such as stunting and the lowest rates of insecticide-treated bed net use (UNICEF 2015). Inequalities in health have also been detected at sub-national levels in a wide range of literature (Arku et al. 2016; Colson et al. 2015; Roberts et al. 2015; Tatem et al. 2012; Wollum et al. 2015) that stresses the need for analysis of health indicators in small geographic units to aid in the identification of areas of greatest need and a more targeted, efficient allocation of resources (Arku et al. 2016; Burgert-Brucker et al. 2015; DHS Spatial Interpolation Working Group 2014; Rosero-Bixby 2004).

Geographic differences in environmental conditions can also add to the variability present in socio-economic drivers. Several studies have explored environmental effects on under-five mortality and have highlighted how geography affects health inequalities (Balk et al. 2003; Balk et al. 2004). Balk et al. (2003) analyzed the role of environmental factors that are explicitly spatial, together with usual proximate and socio-economic determinants, in predicting infant and child mortality in West Africa. Some spatial factors explored in their study (proximity to urban areas and population density) explain a good deal of the country-specific variation in mortality. Among other findings, Balk et al. detected a correlation between arid and semi-arid areas and higher risk of under-five deaths.

Spatially varying factors can have an indirect effect on under-five mortality. As described in Balk et al. (2003), population density was highly correlated with transmission of diseases (Root 1997) while overcrowding and slum conditions have negative effects on both infant and child survival (Defo 1994; Gupta and Baghel 1999; Woods 2003). However, areas with high population density are usually urban where access to improved water and sanitation is higher; this reduces exposure to vector borne and other diseases (Hay et al. 2005; Qi et al. 2012; Rutstein 2000; Tatem et al. 2013) and makes sharing of information and resources easier. Children residing in urban areas have, on average, higher access to water and sanitation and better nutritional status than those in rural areas; therefore, these children have better chances of survival (Balk et al. 2003, 2005).

Urbanization has been historically linked to development. Urban populations, excluding those living in slums (Vlahov et al. 2007), generally have better access to healthcare and better quality care services, including those for mothers and children. This facilitates access to vaccination and other treatments that could prevent risk of child mortality (Dye 2008; Rutherford, Mulholland, and Hill 2010). Urban areas are generally more economically prosperous than rural areas, and this results in improved living conditions such as better housing that reduce the risk of contracting infectious diseases (Hay et al. 2005), and a higher likelihood of affording healthcare and treatment and reducing the risk of child deaths (Rutherford, Mulholland, and Hill 2010). Better access to schools, higher levels of education, and increased parental employment rates are also strongly associated with reduced child mortality (Gakidou et al. 2010).

Climatic and environmental factors such as precipitation, vegetation density, and length of growing season are strongly associated with the prevalence of vector-borne diseases such as malaria because these factors are related to favorable habitats for disease vectors (Tottrup et al. 2009). Such factors can be related to overall health through increased malnutrition (Curtis and Hossein 1998; de Sherbinin 2011; Jankowska et al. 2012). Vegetation density is associated with intensity of poverty in West Africa (Sedda et al. 2015). Spatial variations in rainfall and temperature are associated with increasing of cases of diarrhea (Balk et al. 2003; Bandyopadhyay, Kanji, and Wang 2012), the spread of infectious diseases in Mali (Findley et al.

2002), and child mortality in Burkina Faso, especially in the context of specific agro-climatic regions (Dos Santos and Henry 2008).

Temperature and other climatic factors have been widely used in malaria prediction modeling efforts (Gething et al. 2011; Weiss et al. 2014) to predict seasonality of malaria incidence (Cairns et al. 2012) and dengue risk (Brady et al. 2014). Environmental data have also been used for mapping the distribution of infectious diseases and malaria (Hay et al. 2006), as discussed in Kraemer et al. (2016) and Clements et al. (2013). Moreover, climate and environment have an impact on agricultural production, which in turn affects household livelihoods (Balk et al. 2003; Mabhaudhi, Chibarabada, and Modi 2016). Table 2 further describes some of the most relevant spatial environmental factors that influence child mortality that were tested in this analysis.

Table 1. DHS indicators tested in analysis and their relevance to child mortality

Indicator	Significance	Status in sub-Saharan Africa
Child Mortality rate[2] (4q1)	Outcome indicator	Under-five mortality rate sub-Saharan Africa was 86 per 1,000 live births in 2015. The region carries half of the burden of the world's under-five deaths—3 million in 2015. These rates are expected to rise in the future (United Nations 2015).
Female Literacy	Education of women helps determine child survival via increased health knowledge, greater control over health choices for children, and influencing prevention and proactive management of childhood diseases (Caldwell 1979; Hobcraft 1993; UNFPA 2007).	Sub-Saharan Africa still has the lowest literacy rates and widest gender gap in literacy rates in the world (UNESCO UIS 2015).
Contraceptive prevalence, (CPR), any modern method	Family planning can prevent closely spaced and ill-timed pregnancies and births, which contribute to some of the world's highest infant mortality rates. Chances of survival and improvement of health status of children can be prevented by the use of modern contraception (Cleland et al. 2012; Rutstein 2011; WHO 2015a).	Contraceptive use among women aged 15 to 49 doubled between 1990 and 2015, from 13% to 28% (United Nations 2015).
Birth interval	Birth intervals shorter than 36 months have been shown to increase the risk of mortality, under-nutrition, and pregnancy morbidity (Gribble, Murray, and Menotti 2009; Rutstein 2008, 2011). Family planning brings large potential health and survival benefits for children, mainly as a result of wider intervals between births (Cleland et al. 2012; Cleland et al. 2006).	Trends in birth intervals for 46 countries showed that mothers would prefer a median birth interval of 41.5 months - over 9 months longer than their actual median interval of 32.5 months in the five years preceding the surveys. (Rutstein 2011)

Continued

[2] Child mortality rates (age 1 to 4) includes deaths reported as age 12– 59 months. Therefore, infant deaths (before age 1) are not included in the metric used in this report. The DHS Program uses a synthetic cohort life table approach to directly estimate the under-five mortality rate. Details may be found in the Guide to DHS Statistics, pp.90-95 (http://www.dhsprogram.com/publications/publication-DHSG1-DHS-Questionnaires-and-Manuals.cfm) and further described in Pullum et al. 2013.

Table 1—*Continued*

Indicator	Significance	Status in Sub-Saharan Africa
Access to a health facility	In the context of child mortality (excluding infant mortality), delivery in a health facility is used as proximate indication of access to health facilities and affordability (Karra, Fink, and Canning 2016), service availability and presence of health infrastructures (WHO 2016a) and perceived quality of healthcare (Thaddeus and Maine 1994). This broader definition of access to health facilities also captures the idea of care seeking behavior.	About a quarter of births take place without the assistance of a skilled birth attendant. In 2015 alone, this translated into more than 35 million unattended births in South Asia and sub-Saharan Africa (UNICEF 2016).
Exclusive breastfeeding (EBF)	Exclusive breastfeeding reduces risks of ARI/pneumonia, diarrheal disease and malnutrition (Black et al. 2008; Jones et al. 2003; WHO Collaborative Study Team 2000), which are leading causes of child mortality (Liu et al. 2012). Exclusive breastfeeding is also associated with increased spacing between pregnancies.	Prevalence of EBF in West and Central Africa is estimated at 28% in 2010, and 47% in Eastern and Southern Africa (Cai, Wardlaw, and Brown 2012).
Stunting prevalence in children	Undernutrition is the underlying cause of 3.1 million child deaths in 2011 (Black et al. 2011).	Stunting prevalence in 2014 was 32% in Africa (UNICEF, WHO, and The World Bank 2015).
Measles vaccination coverage[3]	Measles is still a leading cause of death for children under age five and can lead to serious complications, including severe diarrhea and pneumonia (Burgert-Brucker et al. 2015; WHO 2015b). Immunization coverage of at least 90% is required for herd immunity (Luna, Monga, and Morgan 2014).	"In 2014, about 85% of the world's children received one dose of measles vaccine by their first birthday through routine health services – up from 73% in 2000" (WHO 2015b).
Diphtheria-tetanus-pertussis (DTP3) immunization coverage	Diphtheria, pertussis, and tetanus (DPT) are vaccine-preventable diseases that cause substantial global disease burden among children under age five (Burgert-Brucker et al. 2015).	"As of 2014, DPT3 immunization coverage among one-year-olds across Africa was 75%, the lowest among all WHO regions" (WHO 2014).
Coverage of improved water sources	The majority of diarrhea deaths are attributed to unimproved water supply or unimproved sanitation, making it among the most important cause of under-five deaths (Anand and Roy 2016; Fink, Günther, and Hill 2011; WHO 2009).	In sub-Saharan Africa, the coverage of improved water sources is about 68% with widespread disparities existing between countries and across regions (WHO 2016b).
Sanitation practices	Open defecation practice facilitates the transmission of diarrheal diseases – one of the leading causes of mortality in children under-five in sub-Saharan Africa (Galan, Kim, and Graham 2013).	In sub-Saharan Africa, it is estimated that 215 million people continue to engage in open defecation (United Nations 2015).

[3] Two doses of measles vaccination are recommended.

Table 2. Spatially relevant factors tested in the analysis and their relevance to child mortality

Spatial environmental factors relevant to child mortality	Significance
Population density	As described in Balk et al. (2003), population density is highly correlated with the transmission of diseases (Root 1997) where, for example, overcrowding has negative effects on both infant and child mortality (Defo 1994). Slum areas, where maternal services are usually inadequate, have higher mortality rates (Gupta and Baghel 1999).
Urbanicity and night-time lights	Urban areas have lower risk of contracting infectious diseases (Hay et al. 2005) and are usually associated with lower mortality rates than rural areas. By contrast, very sparsely populated areas may have inadequate infrastructure to support prenatal and delivery services (Balk et al. 2003). Balk et al. (2005) found strong relationships among urbanicity, population density, and child malnutrition. Moreover, urban residence is overall associated with development, with urban areas being more economically prosperous than rural areas, and providing better access to health care services (Rutherford, Mulholland, and Hill 2010), vaccination coverage, and schools (Gakidou et al. 2010). Intensity of lights at night is often considered as a proxy of population agglomerations and urbanicity, and can therefore capture urban-rural differences as well as different levels of wealth and access (Xie et al. 2015).
Travel time to major population centers	Entwisle et al. (1997) found significant effects of travel time on contraceptive choice and accessibility to towns. Populations living in remote and isolated areas are typically poorer (Tatem et al. 2014) and have lower access to population centers, resources, and vaccination coverage (Metcalf et al. 2015). Pozzi, Robinson, and Nelson (2010) found that distance and travel time to population centers in the Horn of Africa are highly correlated with wealth indices: “Welfare decreases rapidly as access to population centers gets worse” (Pozzi, Robinson, and Nelson 2010).
Ethnicity	A study conducted in 11 sub-Saharan African countries showed large disparities in early child survival chances among ethnic groups in a wide range of African countries (Brockerhoff and Hewett 2000). Other studies (Platas 2010) have found associations between ethnic group variety, provision of health services, and health outcomes; these show that greater ethnic diversity is associated with poorer health outcomes, including higher infant and child mortality and lower public health expenditure.
Conflicts	Conflict areas are usually deprived and can lack basic access to clean water, sanitation, and adequate food. Moreover, some association has been found between conflict zones and malaria prevalence (Sedda, Qi, and Tatem 2015).
Economic activity	The global distribution of per-capita gross domestic product is correlated with malaria and poverty; malaria-endemic countries also have lower rates of economic growth (Gallup and Sachs 2001; Sachs and Malaney 2002). The development of health systems and poverty reduction are major drivers for the reduction of health inequalities (Marmot 2005; Wagstaff 2002); on the other hand, the distribution of mortality aids in understanding inequalities (Sen 1998).
Livestock density	Ownership of livestock is usually associated with rural residency, where households depend on agriculture for part of their livelihood (FAO 2011). Livestock is recognized as an indispensable asset of the poor that contributes to their nutrition and health (Randolph et al. 2007). Ownership of livestock is associated with lower prevalence of child stunting in three countries in Eastern Africa (Mosites et al. 2015). In a more recent work conducted in Western Kenya, Mosites et al. (2016) did not find associations between ownership and child growth, although they saw some correlation between presence of livestock diseases and lower growth rates in some groups.
Temperature	Bandyopadhyay, Kanji, and Wang (2012) found that raised temperatures raise diarrheal prevalence in dry seasons in sub-Saharan Africa (in Balk et al. 2003).

Continued

Table 2—*Continued*

Spatial environmental factors relevant to child mortality	Significance
Rainfall	Findley et al. (2002) detected effects of rainfall on the incidence of infectious diseases (in Balk et al. 2004); Adams (1994) and colleagues (Sauerborn, Adams, and Hien 1996) linked climate anomalies in rainfall-dependent agricultural communities with health and nutrition of household members (in Balk et al. 2004). Balk et al. (2003) observed that variation in average daily rainfall affected the survival of children age 1-4, and interpreted this as a link between agricultural production and children dietary needs. Together with raised temperatures, rainfall shortage raised diarrheal prevalence in dry seasons (Bandyopadhyay, Kanji, and Wang 2012). Dos Santos and Henry (2008) also found a relationship between rainfall and child mortality in Burkina Faso, especially in specific agro-climatic regions.
Aridity / Growing season length	Curtis and Hossein (1998) found that aridity and length of growing season affect child malnutrition. Balk et al. (2003) also found that children living in areas with the shortest growing seasons, classified as arid and semiarid, had a higher risk of death.
Drought	de Sherbinin (2011) found that, among other geographic and socio-economic variables, drought prevalence was significantly correlated with malnutrition in sub-Saharan Africa.
Vegetation density	Vegetation indices such as the Enhanced Vegetation Index (EVI) are considered proxies of vector habitat; disease transmission and physical environment are strongly correlated with disease environments in determining the intensity of exposure to infectious diseases or providing favorable habitats for disease vectors (Tottrup et al. 2009). Tottrup et al. (2009) also found that districts with dense vegetation, high rainfall, and low elevation experienced the lowest reductions in child mortality. Finally, vegetation was associated with intensity of poverty in West Africa (Sedda et al. 2015).
Topography	Topographic spatial data (elevation, slope, and roughness) are usually used as proxies for natural barriers to geographical access to city centers, markets, and health care (Ombok et al. 2010), and for different exposures to infectious diseases, given the spatial variability of malaria prevalence (Gemperli et al. 2004). Webb (1998) found an association between topographic conditions and child stunting, while a similar association between elevation and underweight rates was reported by Balk et al. (2005). However, de Sherbinin (2011) found no association between elevation and malnutrition.
Land cover	Land cover serves as proxy for vector habitats and ecological factors (Balk et al. 2003; Mutuku et al. 2009)
Malaria prevalence	Mosley and Chen (1984) described a child survival conceptual framework that highlighted the relationship between environmental contamination and disease transmission (exposure to malaria). Balk et al. (2003) found correlations between high transmission of malaria and child mortality. Snow et al. (1999) used climatic data and probability models to estimate populations exposed to the risk of infection of malaria in Africa. Gemperli et al. (2004) found a spatial pattern of infant mortality related to the foci of malaria transmission in Mali.

1.2. Rationale for DHS Spatial Analysis Report 13

This report builds upon the work reported in Demographic and Health Surveys (DHS) Spatial Analysis Report 12 (SAR 12) (Burgert-Brucker et al. 2015), in which descriptive geospatial analyses were applied to nine maternal and child health indicators within and among countries, using DHS data for 27 countries in sub-Saharan Africa. Burgert-Brucker et al. (2015) were able to (i) explore sub-national geographic variability of nine maternal and child health indicators across the 27 countries; (ii) identify areas of high or

low coverage through comparison with neighboring areas by using sub-national area level autocorrelation for each indicator as well as spatial clusters; (iii) establish the significance of geography in analyzing differences in the coverage of health indicators. This work set the stage for the current analysis of differences in coverage among contiguous political sub-national areas.

While SAR 12 identified spatial clustering for some of the DHS indicators across 27 sub-Saharan DHS sub-national areas, this report aims to add spatial environmental factors to the usual demographic and socio-economic determinants of child mortality. The report also aims to further identify the spatial or demographic and socio-economic covariates (or combinations of covariates) that might best explain observed patterns of child mortality at DHS sub-national area level, particularly those that may have implications for future interventions and policy relevance. Finally, this report will also account for spatial clustering of the data identified in SAR 12 using spatially explicit models. More specifically, this report aims to understand the extent to which patterns of child mortality can be explained by combining the effects of geospatial covariates with the usual metrics of socio-economic and demographic determinants of child mortality by:

1. Accounting for geospatial factors that might be associated with patterns in child mortality on their own;

2. Understanding how socio-economic, environmental, and geographic variables interact to produce observed patterns in child mortality;

3. Defining patterns in child mortality that are likely due to factors that transcend country boundaries;

4. Explaining patterns in terms of geographic regions.

These findings will be used to make policy recommendations that address the combinations of factors identified within the study.

In this report, child mortality rates are calculated as the number of deaths among children between age 1 and 4 in the five-year period preceding the survey per 1,000 children surviving to age 12 months[4]. This includes deaths reported at age 12–59 months. Therefore, neonatal deaths and other deaths before age 1 are not included in this metric. In general, the literature shows that factors associated with child survivorship after the first year of life are different from those associated with infant survivorship (Balk et al. 2003). Since this analysis only considers children between ages 1 and 4, only factors associated with the probability of dying between the age 1 and 4 are considered. Among those factors, as also discussed in (Balk et al. 2003), spatial environmental factors are expected to act more strongly on child mortality than infant mortality.

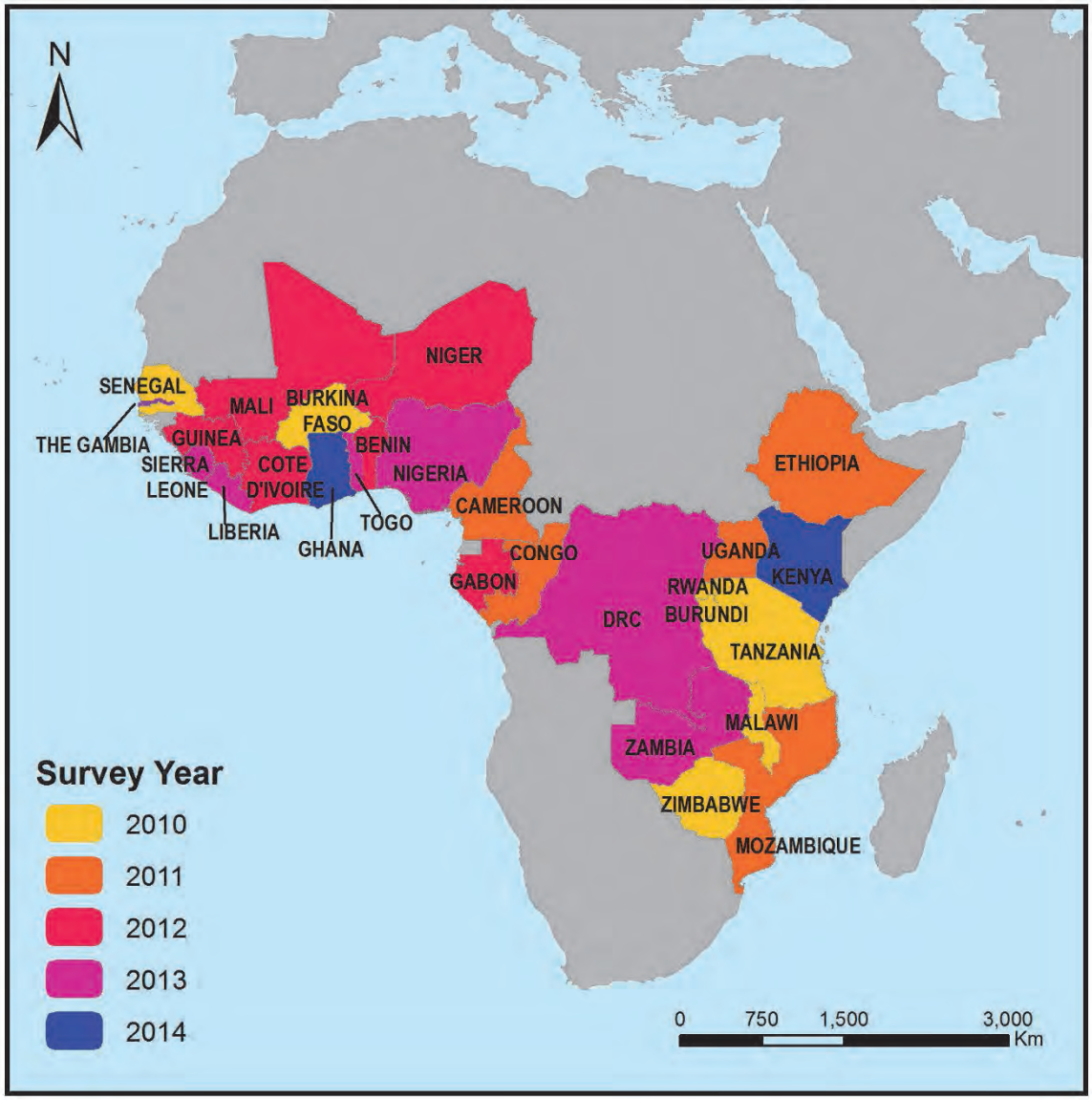
[4] Child mortality rates (age 1 to 4) include deaths reported as age 12– 59 months. Therefore, infant deaths (before age 1) are not included in the metric used in this report. The DHS Program uses a synthetic cohort life table approach to directly estimate the under-five mortality rate. Details may be found in the Guide to DHS Statistics, pp.90-95 (http://www.dhsprogram.com/publications/publication-DHSG1-DHS-Questionnaires-and-Manuals.cfm) and further described in Pullum et al. 2013.

2. Data

2.1. Study Area and Data Description

This report focuses on 27 countries located in sub-Saharan Africa (Figure 2): Benin, Burkina Faso, Burundi, Cameroon, Republic of Congo, Côte d'Ivoire, Democratic Republic of Congo (DRC), Ethiopia, Gabon, The Gambia, Ghana, Guinea, Kenya, Liberia, Malawi, Mali, Mozambique, Niger, Nigeria, Rwanda, Senegal, Sierra Leone, Tanzania, Togo, Uganda, Zambia, and Zimbabwe.

Figure 2. Focus countries and year of the DHS survey used for this study



These countries were selected as the focus countries for the SAR 12 report because they had recent DHS survey data (from after 2010) and were contiguous in space. This report uses the same countries but includes updated data when available (Ghana and Kenya; Figure 2). In addition, these 27 countries have high policy relevance both in the framework of the new Agenda for Sustainable Development Goals (United Nations

General Assembly 2015) and the Strategic Plan 2016-2020 of the Partnership for Maternal, Newborn & Child Health (PMNCH 2015). Investigating further aspects of child mortality in these 27 countries would contribute to tracking progress of countries in the achievement of child mortality related goals.

Polygons representing DHS sub-national areas were obtained from the DHS Spatial Data Repository (ICF International 2008-2015b). Each country contains between 3 and 26 sub-national areas, with a mean number of areas of 9.4 and median of 10 for a total of 255 sub-national areas across the 27 countries (Table 3; Burgert-Brucker et al. 2015). The DHS sub-national areas correspond to administrative level-one areas (provinces) or a combination of such areas within each survey country.

Determinants of child mortality also include environmental and physical (climate, land cover, and accessibility to cities), socio-economic (housing and poverty), health (vaccination coverage) and demographic (birth interval) factors. Thus, this analysis is based on DHS survey data and additional information extracted from a variety of spatial data sources. In this report, a number of plausible factors that may affect or be related to child mortality have been assembled and summarized at the DHS sub-national area level.

Data from the most recent DHS surveys conducted in these 27 countries between 2010 and 2014 were used to compute the key child mortality rate indicator (Pullum et al. 2013; Rutstein and Rojas 2006) and a set of other indicators known in the literature to be associated with child mortality (ICF International 2008-2015a, Table 1): contraceptive prevalence (any modern method), access to a health facility (delivery in a health facility is used as a proxy for access to health facilities), female literacy, exclusive breastfeeding (EBF), measles vaccination, DPT-3 immunization coverage, coverage of improved water sources, sanitation practices, stunting prevalence in children, and birth interval. All DHS surveys used in this report are representative at national and the DHS sub-national area level. Sampling weights were applied to calculate appropriate estimates for each indicator at the DHS sub-national area level with Stata 14 software.

As with the DHS indicator estimates, geospatial datasets that represent factors associated with or having an impact on child mortality rates in sub-Saharan Africa (Table 2) were used to extract and summarize covariates at the DHS sub-national area level. A full list of the geospatial datasets assembled in this study's framework and the derived geospatial covariates tested in the analysis appear in Table 4.

Table 3. Countries, number of DHS sub-national areas in each country, and year of the DHS surveys used in the framework of this study

Country	DHS sub-national areas	DHS Survey Year
Benin	12	2012
Burkina Faso	13	2010
Burundi	5	2010
Cameroon	12	2011
Congo	12	2012
Côte d'Ivoire	11	2012
The Democratic Republic of the Congo	11	2014
Ethiopia	11	2011
Gabon	10	2012
The Gambia	8	2013
Ghana	10	2014
Guinea	8	2012
Kenya	8	2014
Liberia	5	2013
Malawi	3	2010
Mali	6	2013
Mozambique	11	2011
Niger	8	2012
Nigeria	6	2013
Rwanda	5	2010
Senegal	14	2011
Sierra Leone	4	2013
United Republic of Tanzania	26	2010
Togo	6	2013-14
Uganda	10	2011
Zambia	10	2013
Zimbabwe	10	2011

2.1.1. DHS indicators

Definitions of the DHS indicators used in the analysis are described briefly below. Section 2.2.1 describes how these indicators were summarized at the DHS sub-national area level for use as covariates in the analysis. All covariates were estimated using the proper DHS sample weights for each survey.

Child mortality rate (4q1)

Number of deaths among children between age 1 and 4 in the five-year period preceding the survey per 1,000 children surviving to age 12 months[5]. This rate includes deaths reported at age 12–59 months and excludes infant deaths (age 0 to 11 months). In this analysis, rates were calculated by DHS sub-national areas and were used as the response variable (outcome).

[5] Child mortality rates (age 1 to 4) includes deaths reported as age 12–59 months. Therefore, infant deaths (before age 1) are not included in the metric used in this report. The DHS Program uses a synthetic cohort life table approach to directly estimate the under-five mortality rate. Details may be found in the Guide to DHS Statistics, pp.90-95 (http://www.dhsprogram.com/publications/publication-DHSG1-DHS-Questionnaires-and-Manuals.cfm) and further described in Pullum et al. 2013.

Female literacy

Percentage of women age 15-49 who attended primary schooling and can read a whole or part of a sample sentence. A respondent who attended secondary education or higher is coded as literate, as well as respondents who could read a whole sentence.

Contraceptive prevalence (CPR), any modern method

Percentage of all women age 15-49 who report current use of a modern contraceptive method.[6]

Preceding Birth interval

Percentage of births with a birth interval less than 36 months between the preceding birth and conception of the index child.

Access to a health facility

Percentage of children born in the last 5 years who were delivered in a health facility. In the context of child mortality (age 1 to 4), this variable is used as a proxy for (and in this work referred interchangeably with) access to a health facility.

Exclusive breastfeeding (EBF)

Percentage of last-born infants under age 6 months who are living with the mother, breastfeeding, and have not had any water, liquids, or solids in the day or night before the interview.

Stunting prevalence in children

Stunting is defined as those children whose height is two standard deviations (SD) or more below the median height for children of the same sex and age in an internationally standard index, defined in the WHO Child Growth Standards that were adopted in 2006.

Measles vaccination coverage

Percentage of live children age 12-23 months who received the measles vaccination at any time before the survey.

Diphtheria-tetanus-pertussis (DTP3) immunization coverage

Percentage of live children age 12-23 months who received three doses of DPT vaccine at any time before the survey.

Coverage of improved water sources

Percentage of de jure population whose main source of drinking water is a household connection (piped), public standpipe, borehole, protected dug well or spring, rainwater collection, or bottled water.

[6] Modern contraceptive methods include female sterilization, male sterilization, the contraceptive pill (oral contraceptives), intrauterine contraceptive device (IUD), injectables (Depo-Provera), implants (Norplant), female condom, male condom, diaphragm, contraceptive foam and contraceptive jelly, lactational amenorrhea method (LAM), country-specific modern methods and other modern contraceptive methods mentioned by the respondent (including cervical cap, contraceptive sponge, and others). Abortion, menstrual regulation, and withdrawal are NOT considered modern contraceptive methods.

Sanitation practices

Percentage of de jure population that use fields, bushes, forests, open bodies of water, or other open spaces rather than using the toilet to defecate (United Nations 2016).

2.1.2. Geospatial datasets

Geospatial datasets that represent factors potentially correlated with child mortality in the study area (Table 2) were obtained from multiple existing sources or produced ad hoc for this study (topographic slope and roughness). The datasets are different in format (grid or polygons), type (categorical or continuous), accuracy, survey scale, spatial resolution, and temporal and spatial coverage.

All geospatial datasets in this study are briefly described below and summarized in Appendix A and Table 4, which also provides source links for obtaining additional information about the methodology and data used to produce them. (See Section 2.2.2 below for a description of how the geospatial covariates in the analysis were extracted from these datasets and summarized at the DHS sub-national area level.)

Population density

The WorldPop (www.worldpop.org) 2010 UN adjusted People Per Pixel raster datasets (Linard et al. 2012; Stevens et al. 2015), which are publicly available for all countries in this study and have a resolution of approximately 100 m at the equator, were used to represent the population distribution within the study area (Figure A.1 in Appendix A).

Urbanicity

The publicly available 2000/2001 "MODIS 500m Global Urban Extent" raster dataset (Schneider, Friedl, and Potere 2009, 2010), with a resolution of approximately 500 m near the equator, was used to represent the extent of built-up and settled areas within the study area (Figure A.2a in Appendix A).

The European Commission Joint Research Centre "Global Human Settlement Layer" raster dataset (courtesy of M. Pesaresi), with a resolution of 38 m (Figure A.2b in Appendix A), provided an alternative representation of the extent of built-up areas within the study area.

Night-time lights

The publicly available global NOAA Suomi National Polar-orbiting Partnership (NPP) composite night-time light raster dataset (Elvidge et al. 2013) for April 2012, with a spatial resolution of approximately 500 m near the equator, was used to measure the night light intensity in the study area (Figure A.3 in Appendix A).

Travel time to major population centers

The publicly available global European Union Joint Research Centre "travel time to major cities" raster datasets (Uchida and Nelson 2008), representing the time required to travel between two grid cells with a spatial resolution of approximately 1km near the equator, were used to estimate the accessibility, via all transport methods, to settlements with more than 50,000, 100,000, and 500,000 inhabitants located within the study area (Figure A.4a, A.4b, A.4c in Appendix A).

Ethnicity

The 2010 global "Geo-Referencing of Ethnic Groups (GREG)" vector dataset (Weidmann, Rød, and Cederman 2010), was used to map the spatial distribution of the various ethnic groups within the study area (Figure A.5 in Appendix A).

Conflicts

The publicly available "Armed Conflict Location & Event Data Project (ACLED)" database (www.acleddata.com), with disaggregated conflict and protest data that included event locations and associated fatalities, was used to map the level of political instability and violence within the study area (Figure A.6 in Appendix A).

Economic activity

The global "Geographically-based Economic (G-Econ)" raster dataset (Nordhaus and Chen 2009), representing the regional equivalent of the gross domestic product at a spatial resolution of approximately 5 km near the equator, was used to depict the level of economic activity in 2005 within the study area (Figure A.7 in Appendix A).

Livestock density

The publicly available FAO "Gridded Livestock of the World, v2.0 (GLW2)" (Robinson et al. 2014), with a spatial resolution of approximately 1 km near the equator, were used to represent the cattle, goat, pig, and sheep density within the study area (Figure A.8a, A.8b, A.8c, A8d in Appendix A).

Temperature and rainfall

The publicly available WorldClim - Global Climate Data "Annual Mean Temperature (BIO1)" and "Annual Total Precipitation (BIO12)" raster datasets (Hijmans et al. 2005), both with a spatial resolution of approximately 1km at the equator, were used to represent the current conditions in climate conditions within the study area (Figure A.9a and A.9b in Appendix A).

MODIS-based daytime Land Surface Temperature (LST) monthly composite raster datasets (MOD11C3), with a spatial resolution of approximately 5.6 km near the equator, were also used to represent the current annual average temperature in the study area (Figure A.10 in Appendix A).

Aridity

The publicly available CGIAR Global Aridity Index raster dataset (Zomer et al. 2008), with a spatial resolution of approximately 1km near the equator, was used to represent the current aridity conditions in the study area (Figure A.11 in Appendix A).

Growing season length

The publicly available FAO "Global length of growing periods" raster dataset (dataset lgp_pl71available at http://www.fao.org/geonetwork/srv/en/main.home), with a resolution of approximately 50 km near the equator, was used to characterize the agricultural setting within the study area (Figure A.12 in Appendix A).

Drought

The CIESIN Global Drought Hazard Frequency and Distribution, v1 raster dataset (Dilley et al. 2005), with a spatial resolution of about 5 km near the equator, was used to estimate the relative frequencies of drought occurrences within the study area (Figure A.13 in Appendix A).

Vegetation density

MODIS-based Enhanced Vegetation Index (EVI) monthly composite raster datasets (MOD13C2), with a spatial resolution of approximately 5.6 km near the equator, were used to represent the vegetation conditions in the study area (Figure A.14 in Appendix A).

Topography

The publicly available Global 30 Arc-Second Elevation raster dataset (GTOPO30; http://webgis.wr.usgs.gov/globalgis/gtopo30/gtopo30.htm), with a spatial resolution of approximately 1 km near the equator, was used to represent the topographic characteristics of the study area (Figure A.15a, A.15b, A.15c, A.15d, A.15e in Appendix A).

Land cover

The publicly available ESA 2009 GlobCover raster dataset (Bontemps et al. 2011), with a resolution of approximately 300 m near the equator, was used to represent the land cover within the study area (Figure A.16 in Appendix A).

Malaria prevalence

The publicly available Malaria Atlas Project "*Pf*PR2-10 in Africa 2000-2015" raster datasets, with a resolution of approximately 5 km near the equator, were used to depict the *Plasmodium falciparum* infection prevalence in children age 2-10 within the study area (Figure A.17 in Appendix A).

2.2. Data Preparation

All DHS indicators and geospatial datasets that represented the factors in this study were summarized at the DHS sub-national area level with the methodology described below.

2.2.1. DHS covariates aggregation

The units of analysis of this work are DHS sub-national areas, which correspond to administrative level-one areas (provinces) or a combination of such areas within each survey country. The data are statistically representative estimates of each indicator of interest for the population within each DHS sub-national areas at the time of each survey. Sampling weights were applied to calculate appropriate estimates for each indicator at the sub-national area level.

In addition to the child mortality rate indicator, which represents a rate expressed as number of deaths per 1,000 children surviving to age 12 months by DHS sub-national area, all other DHS indicators described above were calculated and used as percentages by DHS sub-national area.

Figures 3(a-f) below shows the spatial distribution at the DHS sub-national area level for the six most relevant DHS indicators selected for use in the final statistical model (birth interval, sanitation, female literacy, access to a health facility (delivery in a health facility used as a proxy for access to health facilities), measles vaccination, and stunting). In Figures 3(a-f), the scale ranges vary for each indicator. A common color scheme is used for all indicators in these figures, with relatively better conditions shown in green and worse conditions in red. For example, reddish colors are associated with higher values of stunting prevalence in children and with lower values of female literacy.

Figure 3a. Percentage of literate women (age 15-49) in the 225 DHS sub-national areas

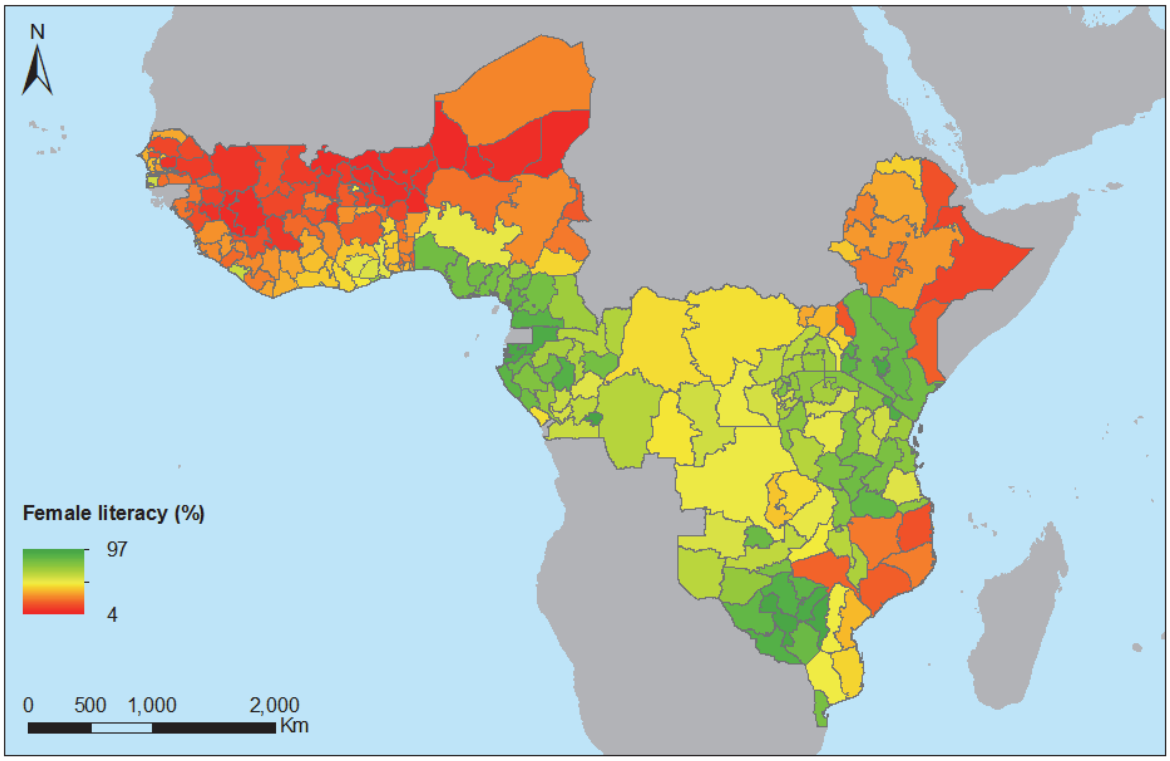


Figure 3b. Percentage of births with a birth interval less than 36 months between birth and conception in the 255 DHS sub-national areas

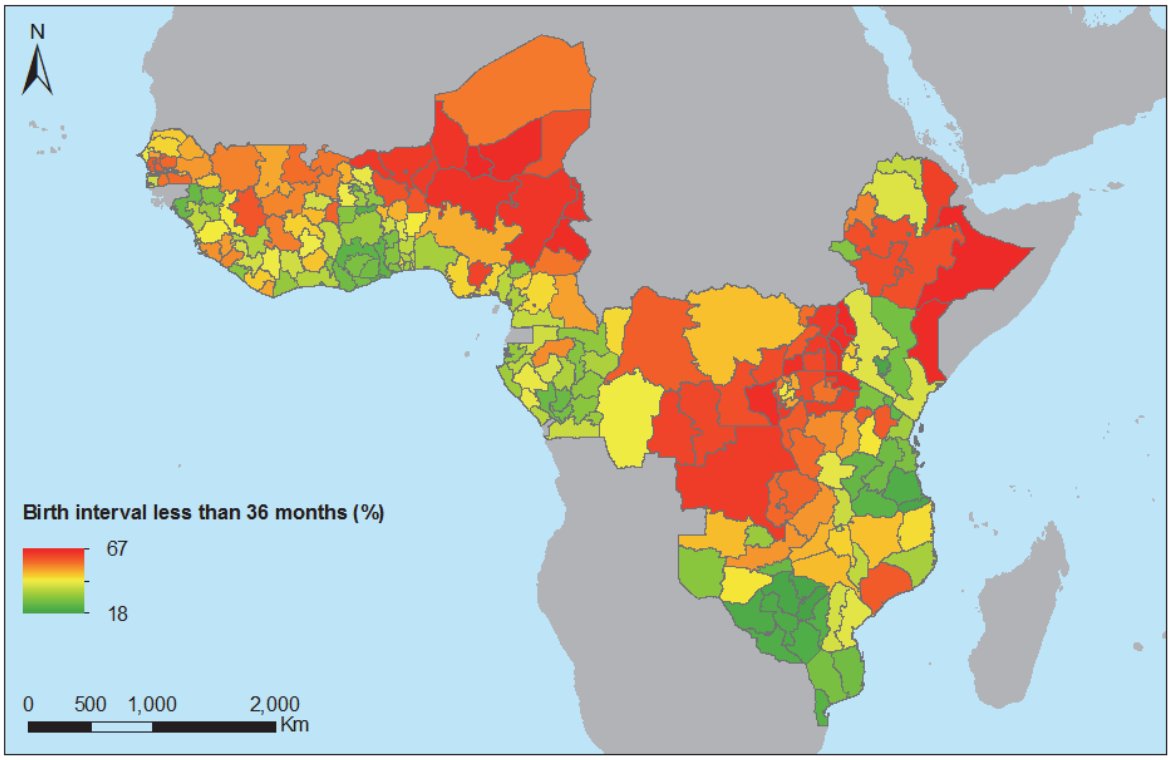


Figure 3c. Percentage of children born in last 5 years with access to a health facility in the 225 DHS sub-national areas (access defined as delivery in a health facility)

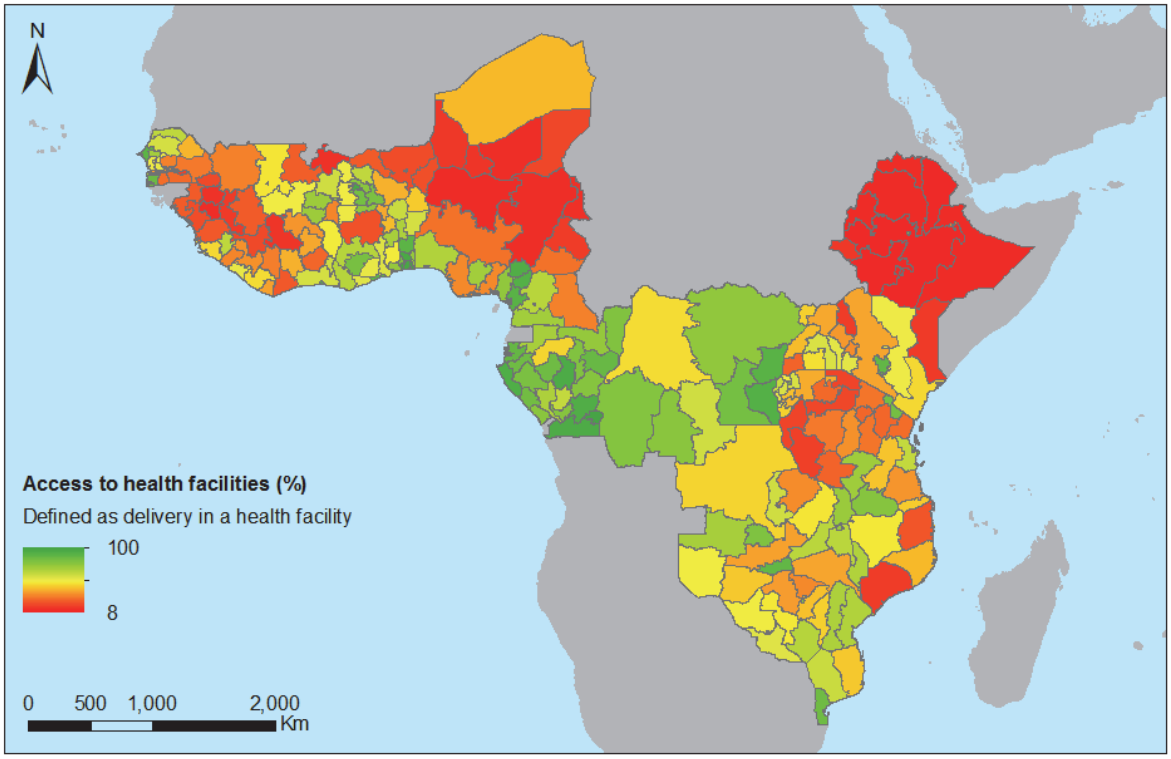


Figure 3d. Percentage of stunting children age (12-59 months) in the 255 DHS sub-national areas

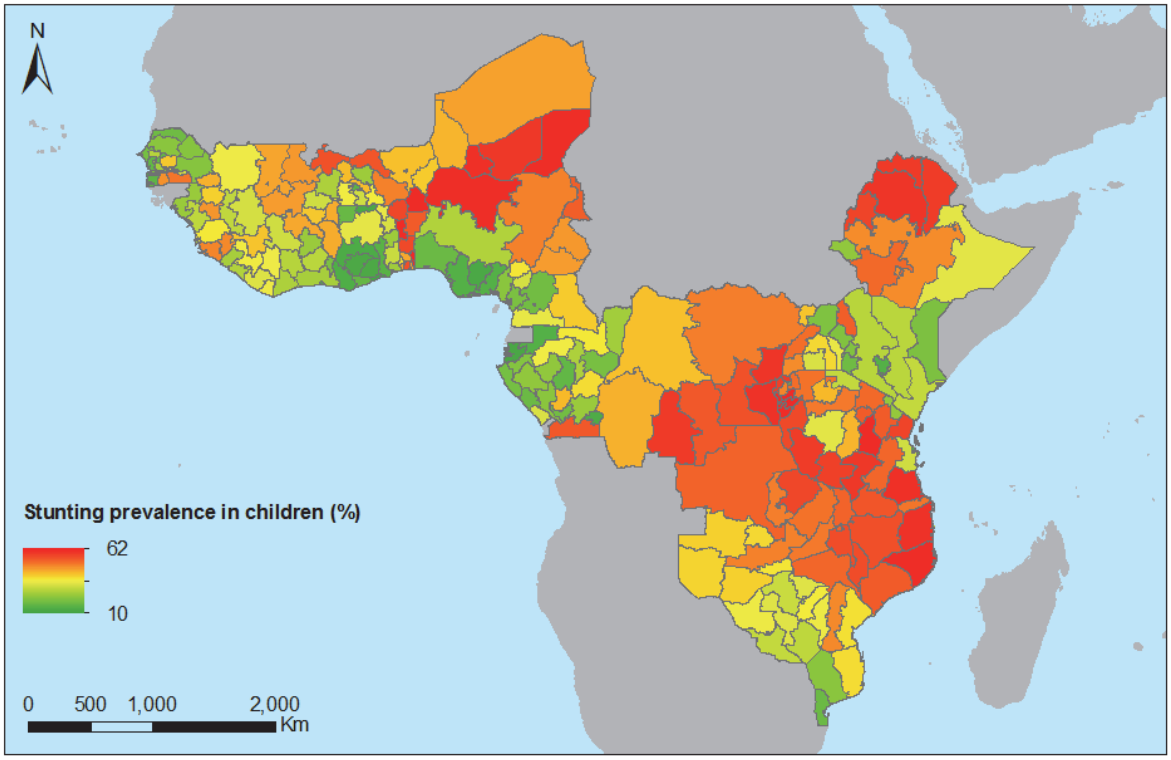


Figure 3e. Percentage of live children age 12-23 months who received the measles vaccination at any time prior to the survey in the 255 DHS sub-national areas

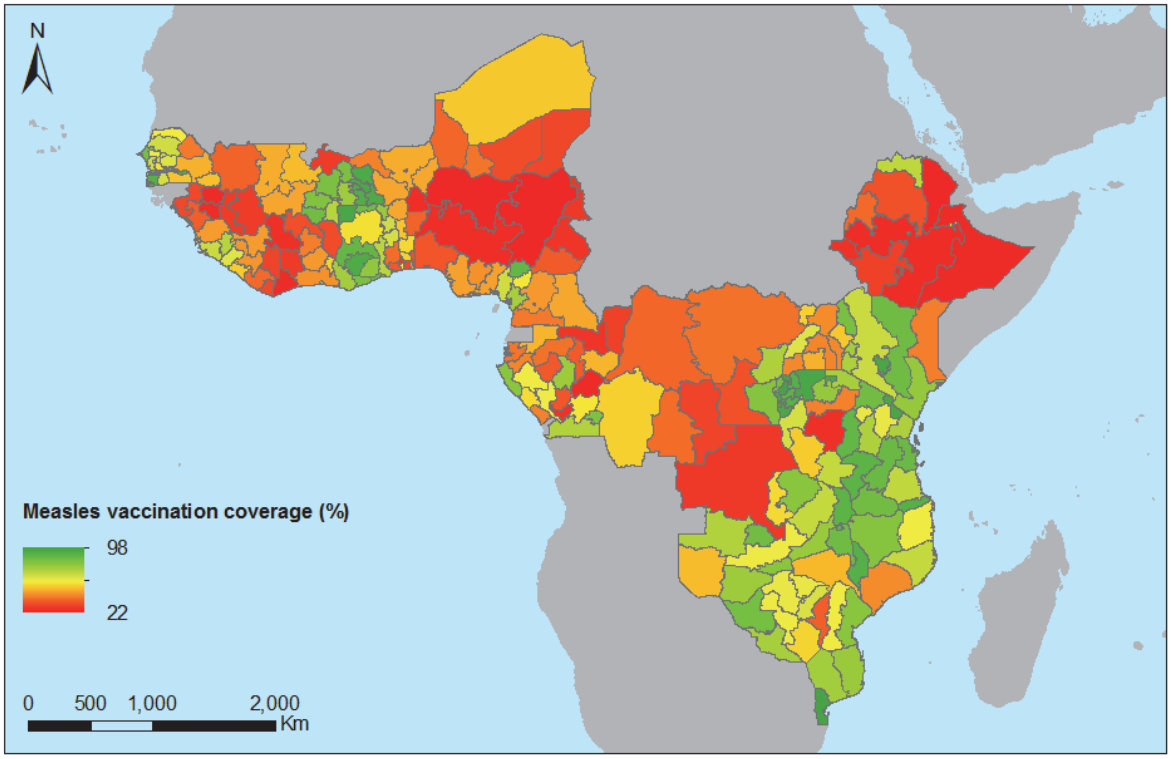
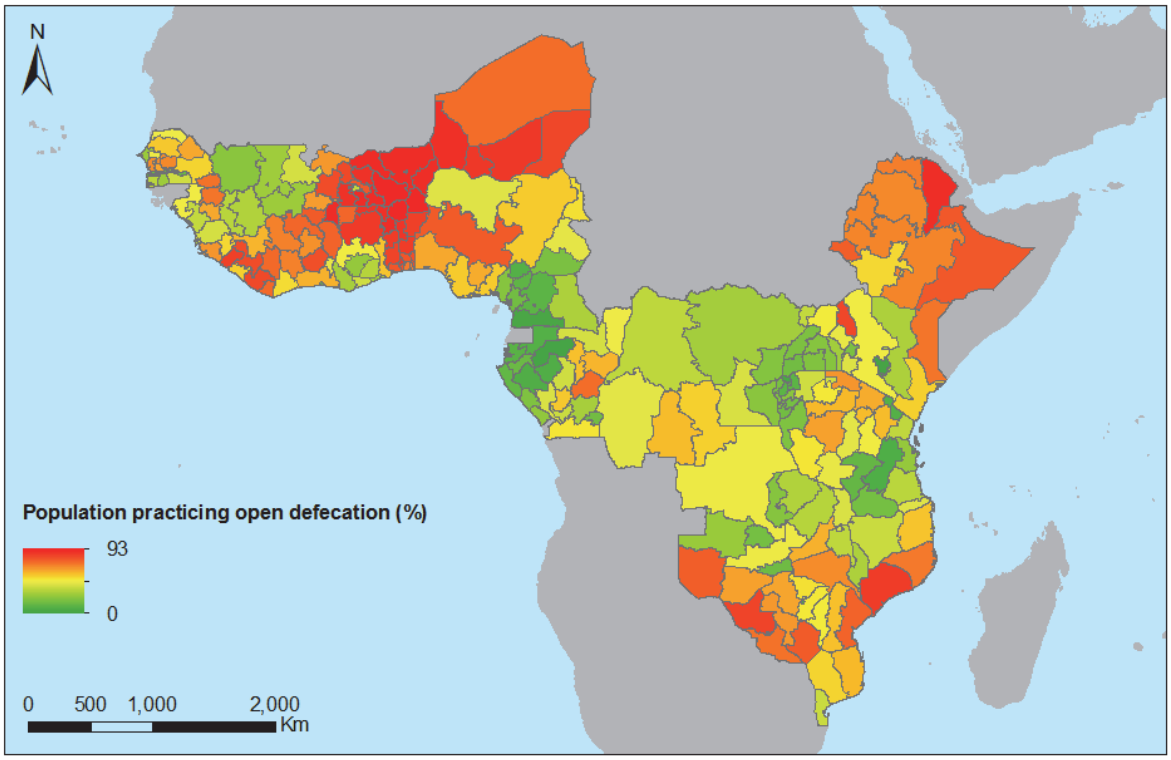


Figure 3f. Percentage of population practicing open defecation in the 225 DHS sub-national areas



2.2.2. Geospatial covariates preparation

In order to extract covariates from the geospatial datasets described in subsection 2.1.2 and to summarize these covariates at the DHS sub-national area level (at the same spatial detail of the DHS covariates), the DHS sub-national areas (ICF International, 2008-2015b), were used to calculate GIS-based zonal statistics and append the results to the corresponding aggregate DHS covariates.

The WorldPop 2010 UN adjusted People Per Pixel datasets were mosaicked and used to calculate the population density of each DHS polygon by dividing the total population in each polygon by the surface area of the corresponding polygon.

The mosaicked WorldPop dataset was also combined with the temporally averaged "*Pf*PR2-10 in Africa 2000-2015" datasets to calculate the population-weighted mean malaria prevalence value for each DHS polygon. This involved multiplying the two raster datasets and dividing the total number of infected people within each polygon (calculated with the output raster generated by the multiplication) by the total population in the same polygon (calculated with the mosaicked WorldPop dataset).

The "MODIS 500m Global Urban Extent" and "Global Human Settlement Layer" were used to calculate two different estimates of the percentage of urban area in each DHS polygon. These estimates were subsequently combined with the mosaicked WorldPop to produce two estimates of the percentage of urban population living within each polygon.

The georeferenced event locations and the associated fatalities were extracted from the ACLED database and were used to calculate the total number of events and fatalities in each DHS polygon. The total number of events and fatalities was then divided by the area and total population of the corresponding polygon, respectively.

The "Annual Total Precipitation (BIO12)" was used to calculate the annual rainfall rate, in each DHS polygon, defined as the average volume of water in the form of rain (mm) that falls per unit of area (~km^2) and per unit of time (year). This involved calculating the total amount of annual rain falling in each DHS polygon and dividing this total by the area of each polygon.

Geospatial covariate values to be used in the statistical analysis were summarized for each DHS polygon in two different ways according to the geospatial dataset type from which they were extracted (Table 4). For continuous covariates (night-time lights, travel time to major population centers, economic activity, livestock density, temperature, aridity, drought, EVI, topographic elevation, slope, and roughness), the mean (referred as spatial average in Figure 4b, 4d, 4e, and in Table 4) and median values were calculated and extracted. In particular, for each DHS polygon, (i) the mean was calculated by averaging the values of all grid cells located within it, while (ii) the median was calculated by identifying the value dividing into two equal parts the ordered values of the grid cell located within the polygon. For categorical covariates (landcover and a 30 arc seconds rasterized version of the GREG vector dataset), the modal value (the value that appears most often) and the variety of the values (the number of unique values) were considered.

In order to calculate and extract the mean and median values from the "global length of growing periods" dataset, the dataset was reclassified into a continuous dataset by replacing, for each grid cell, the day interval (representing the growing season length) with its corresponding mean value. Figure 4(a-f) below shows the spatial distribution at the DHS sub-national area level for the most relevant geospatial covariates used in the analysis. Finally, because policy actions can often be implemented across large regions and patterns in 4q1 appear to vary within distinct areas across sub-Saharan Africa (UNICEF 2015), three broad regions (East, Central, and West) were defined, as described by the United Nations geoscheme (Figure 5 and Table A.1 in Appendix A).

Figure 4a. Population density in each DHS sub-national area

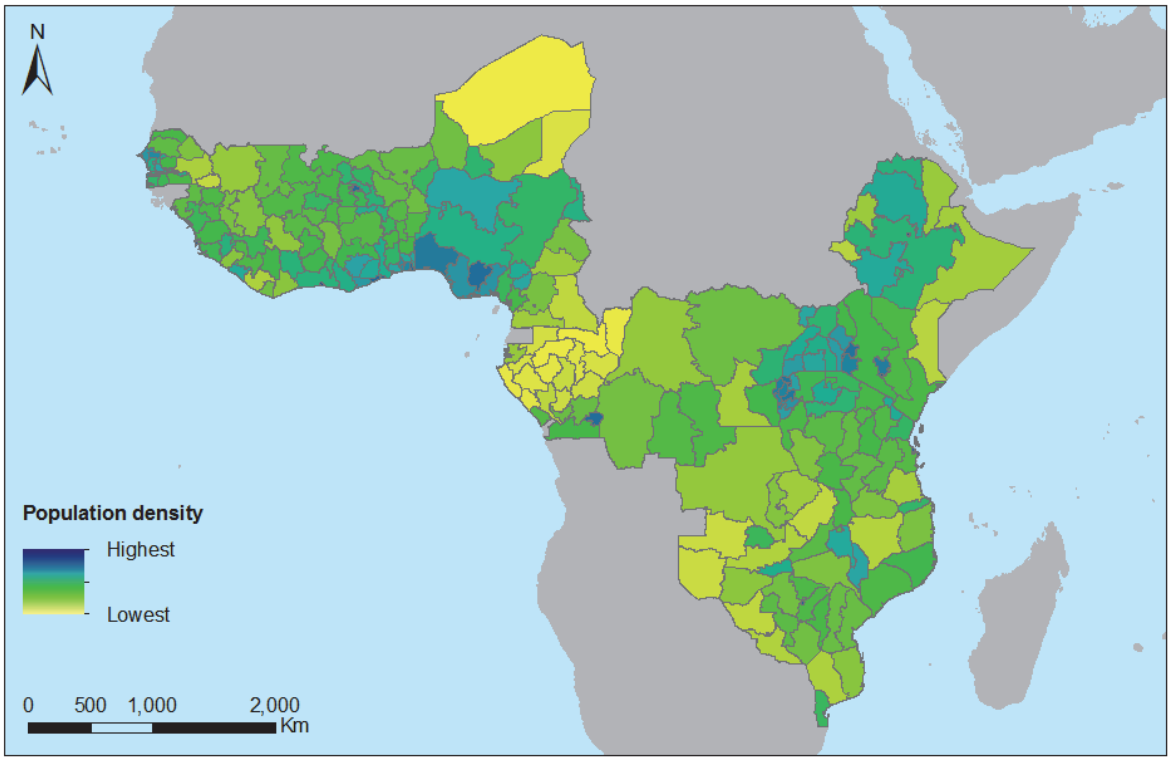


Figure 4b. Satellite-based night-time lights' intensity spatially averaged within each DHS sub-national area

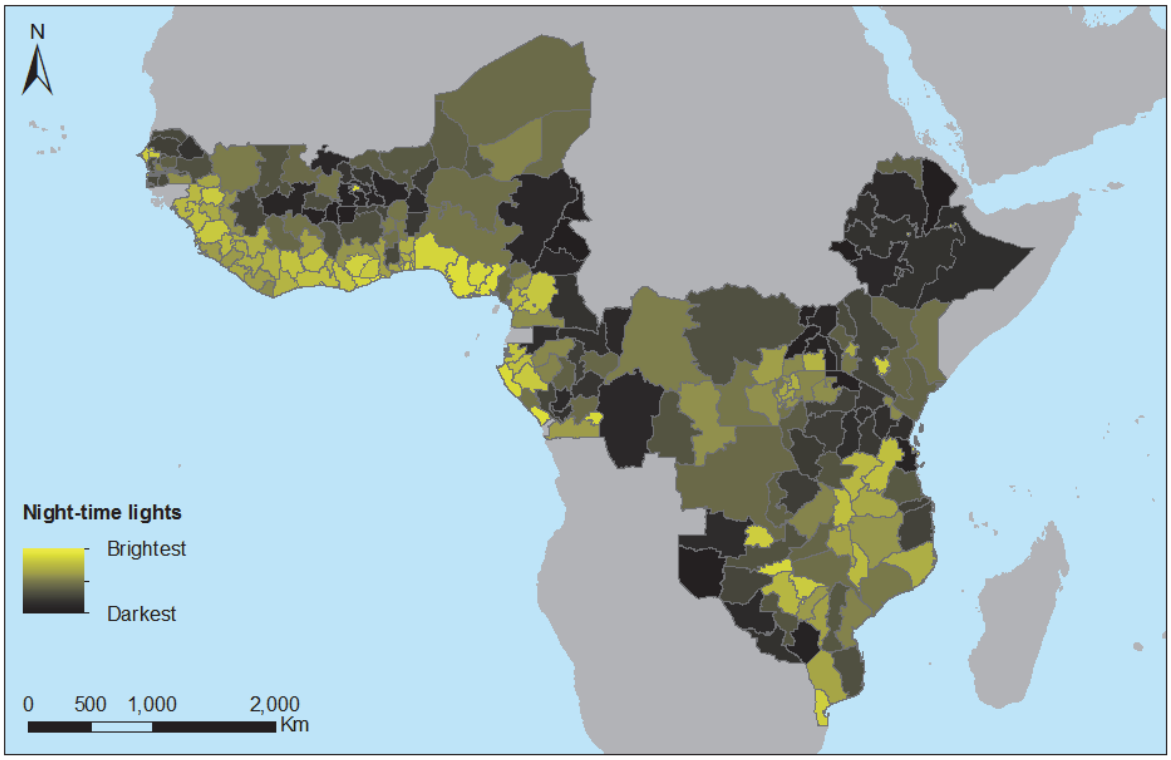


Figure 4c. Number of dominant ethnic groups, indicating ethnic variability, in each DHS sub-national area

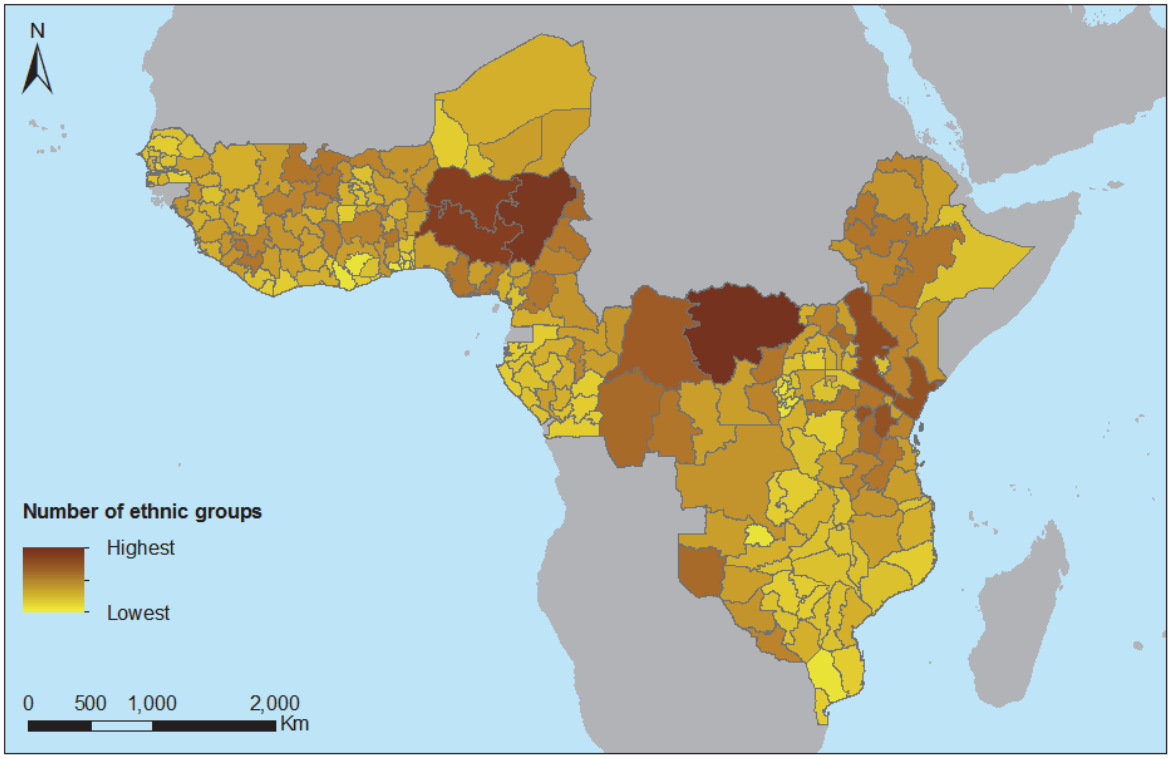


Figure 4d. Spatially averaged economic activity estimated using the grid-based equivalent of the purchasing parity gross domestic product, within each DHS sub-national area

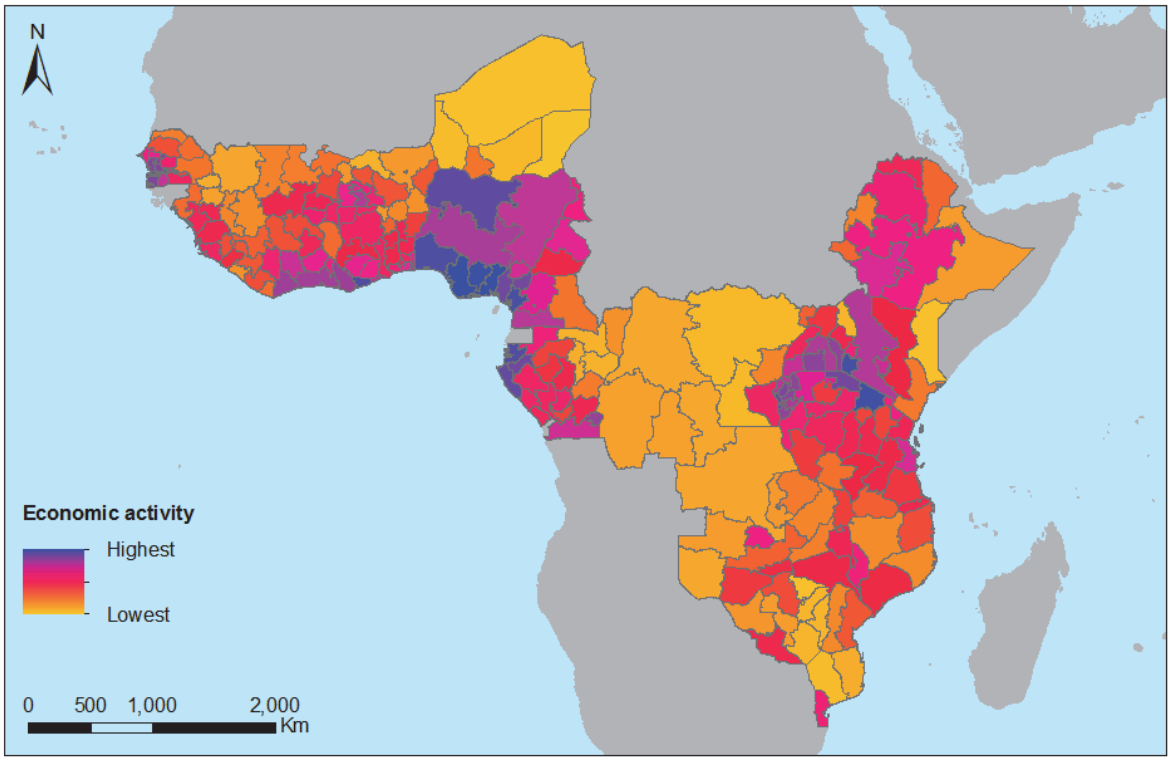


Figure 4e. Annual mean temperature spatially averaged within each DHS sub-national area

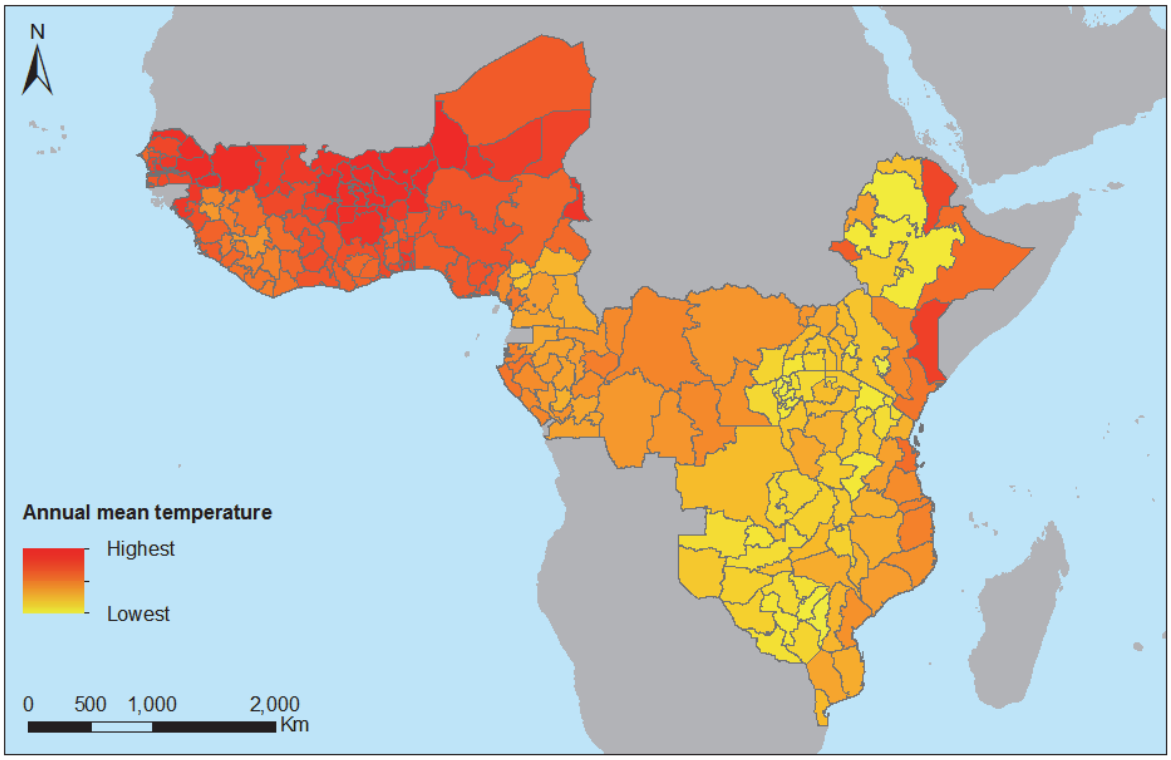


Figure 4f. Population-weighted mean *Plasmodium falciparum* prevalence in each DHS sub-national area

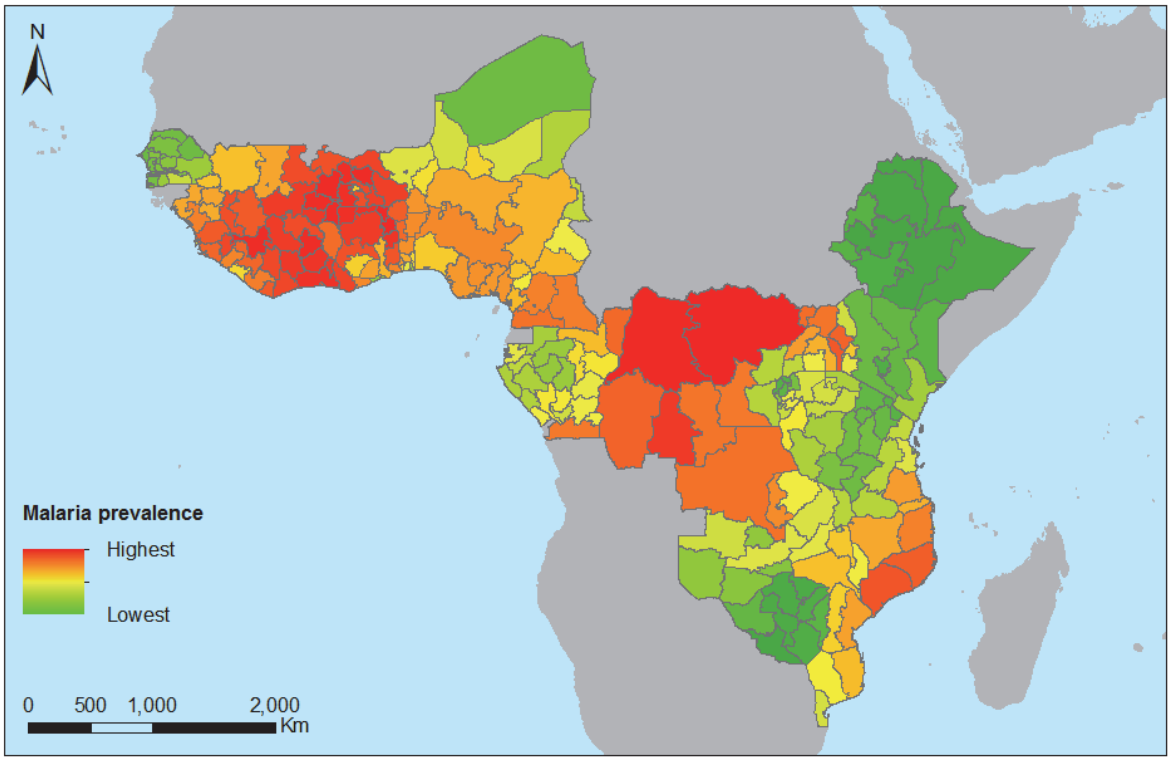


Figure 5. United Nations geoscheme regions used in the analysis

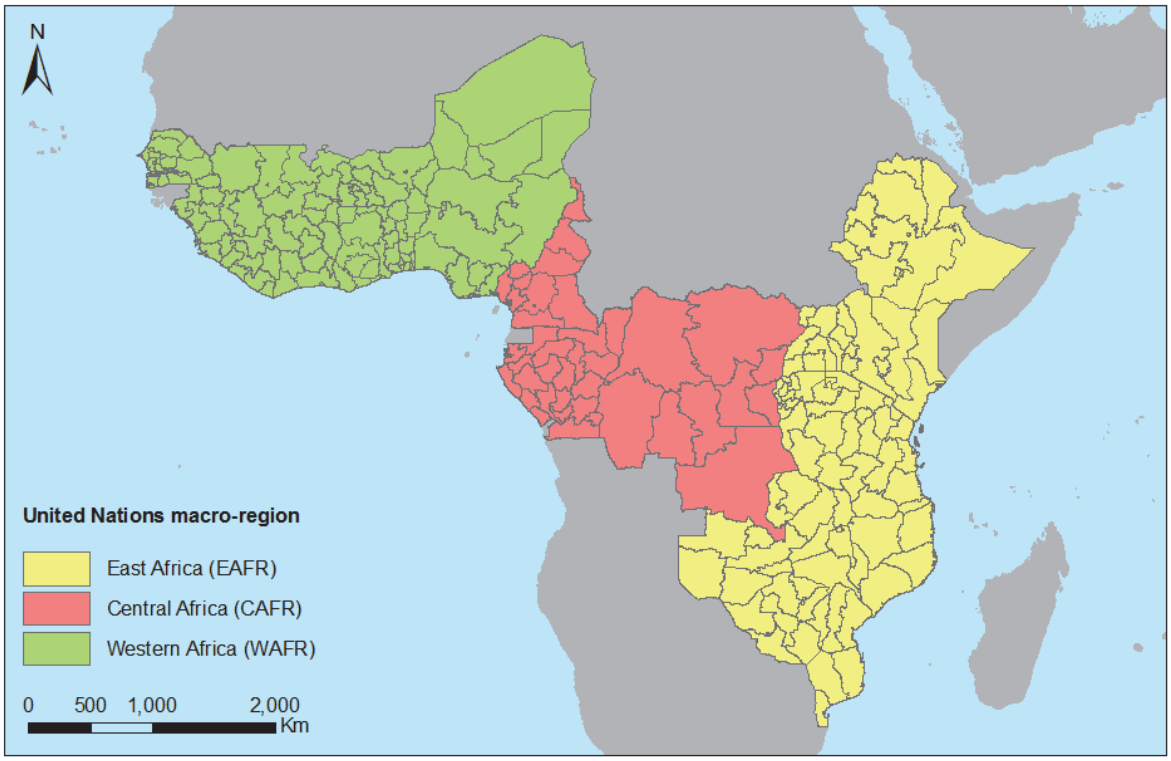


Table 4. Full range of geospatial datasets and derived covariates tested for modeling child mortality in the study area

Geospatial dataset	Temporal coverage	Type	Format	Native reso-lution	Source	Derived geospatial covariate	Covariate temporal coverage	Covariate type	Covariate Spatial detail
Population distribution (WorldPop People Per Pixel)	2010	Continuous	Raster	3 arc seconds	WorldPop Project http://www.worldpop.org.uk/data/data_sources/	Population density	2010	Continuous	DHS sub-national area
Urbanicity (MODIS 500m Global Urban Extent)	2000/ 2001	Categorical (binary)	Raster	15 arc seconds	SAGE https://nelson.wisc.edu/sage/data-and-models/schneider.php	Percentage of urban area	2000/ 2001	Continuous	DHS sub-national area
						Percentage of WorldPop urban population	2000/ 2001-2010	Continuous	DHS sub-national area
Urbanicity (Global Human Settlement Layer)	2014	Categorical (binary)	Raster	38m	EC-JRC (personal communication) http://ghslsys.jrc.ec.europa.eu/	Proportion of urban area	2014	Continuous	DHS sub-national area
						Proportion of WorldPop urban population	2010-2014	Continuous	DHS sub-national area
Night-time lights (Suomi NPP VIIRS)	2012	Continuous	Raster	15 arc seconds	NOAA http://ngdc.noaa.gov/eog/viirs/download_monthly.html	Spatially averaged night-lights' intensity	2012	Continuous	DHS sub-national area
Travel time to major population centers with more than 50/100/500k inhabitants	2000	Continuous	Raster	30 arc seconds	EC-JRC http://forobs.jrc.ec.europa.eu/products/gam/	Spatially averaged travel time to populated places	2000	Continuous	DHS sub-national area
Ethnicity (Geo-referencing of Ethnic Groups - GREG)	2010	Categorical	Vector	Polygon locations	ETH International Conflict Research http://www.icr.ethz.ch/data/other/greg	Modal ethnic group	2010	Categorical	DHS sub-national area
						Ethnic group variety	2010	Continuous	DHS sub-national area
Conflicts (Armed Conflict Location & Event Data - ACLED)	1997-2016	Continuous/ categorical	Table	Point locations	ACLED Project http://www.acleddata.com/data/acled-versions-1-5-data-1997-2014/	Conflict density	2010-2013	Continuous	DHS sub-national area
						Fatalities due to conflicts (as a proportion of the total population)	2010-2013	Continuous	DHS sub-national area

Continued

Table 4—*Continued*

Geospatial dataset	Temporal coverage	Type	Format	Native reso-lution	Source	Derived geospatial covariate	Covariate temporal coverage	Covariate type	Covariate Spatial detail
Economic activity (Geographically-based Economic Data - G-Econ)	2005	Continuous	Raster	2.5 arc minutes	Yale University http://gecon.yale.edu/	Spatially averaged gross cell product	2005	Continuous	DHS sub-national area
Livestock density (Gridded Livestock of the World, v2.0)	2006	Continuous	Raster	30 arc seconds	Livestock Geo-Wiki http://livestock.geo-wiki.org/	Sheep/Pig/ Cattle/Goat density	2006	Continuous	DHS sub-national area
Temperature (WorldClim BIO_1)	Current (~1960-1990)	Continuous	Raster	30 arc seconds	WorldClim http://www.worldclim.org/current	Spatially averaged annual mean temperature	Current (~1960-1990)	Continuous	DHS sub-national area
Temperature (MODIS Land Surface Temperature - MOD11C3)	2001-2015	Continuous	Raster	30 arc seconds	NASA https://lpdaac.usgs.gov/dataset_discovery/modis/modis_products_table/mod11a1	Spatially averaged LST	2001-2015	Continuous	DHS sub-national area
Rainfall (WorldClim BIO_{12})	Current (~1960-1990)	Continuous	Raster	30 arc seconds	WorldClim http://www.worldclim.org/current	Spatially averaged annual rainfall rate	Current (~1960-1990)	Continuous	DHS sub-national area
Aridity	1950-2000	Continuous	Raster	30 arc seconds	CGIAR-CSI http://www.cgiar-csi.org/data/global-aridity-and-pet-database	Spatially averaged Aridity Index	1950-2000	Continuous	DHS sub-national area
Growing season length	2000	Categorical	Raster	0.5 degree	FAO http://www.fao.org/geonetwork/srv/en/main.home	Spatially averaged growing season length	2000	Continuous	DHS sub-national area
Drought (Global Drought Hazard Frequency and Distribution, v1)	1980-2000	Continuous	Raster	2.5 arc minutes	CIESIN http://sedac.ciesin.columbia.edu/data/set/ndh-drought-hazard-frequency-distribution	Spatially averaged relative-frequency of drought occurrence	1980-2000	Continuous	DHS sub-national area
Vegetation density (MODIS Enhanced Vegetation Index - MOD13C2)	2001-2015	Continuous	Raster	30 arc seconds	NASA https://lpdaac.usgs.gov/dataset_discovery/modis/modis_products_table/mod13a3	Spatially averaged EVI	2001-2015	Continuous	DHS sub-national area

Continued

Table 4—*Continued*

Geospatial dataset	Temporal coverage	Type	Format	Native reso-lution	Source	Derived geospatial covariate	Covariate temporal coverage	Covariate type	Covariate Spatial detail
Topographic elevation (GTOPO)	2000	Continuous	Raster	30 arc seconds	USGS https://lta.cr.usgs.gov/GTOPO30	Spatially averaged topographic elevation	2000	Continuous	DHS sub-national area
Topographic slope (GTOPO-derived)	2000	Continuous	Raster	30 arc seconds	-	Spatially averaged topographic slope	2000	Continuous	DHS sub-national area
Topographic roughness (GTOPO-derived using different window sizes of 3x3, 5x5, and 11x11 grid cells)	2000	Continuous	Raster	30 arc seconds	-	Spatially averaged topographic roughness	2000	Continuous	DHS sub-national area
Land cover (MERIS-based GlobCover)	2009	Categorical	Raster	10 arc seconds	ESA http://due.esrin.esa.int/page_globcover.php	Modal landcover	2009	Categorical	DHS sub-national area
						Landcover variety	2009	Continuous	DHS sub-national area
Malaria prevalence (*PfPR2-10 in Africa 2000-2015*)	2000-2015	Continuous	Raster	2.5 arc minutes	Malaria Atlas Project http://www.map.ox.ac.uk/	Population-weighted mean *Pf* prevalence	2000-2015	Continuous	DHS sub-national area

3. Analysis Methods

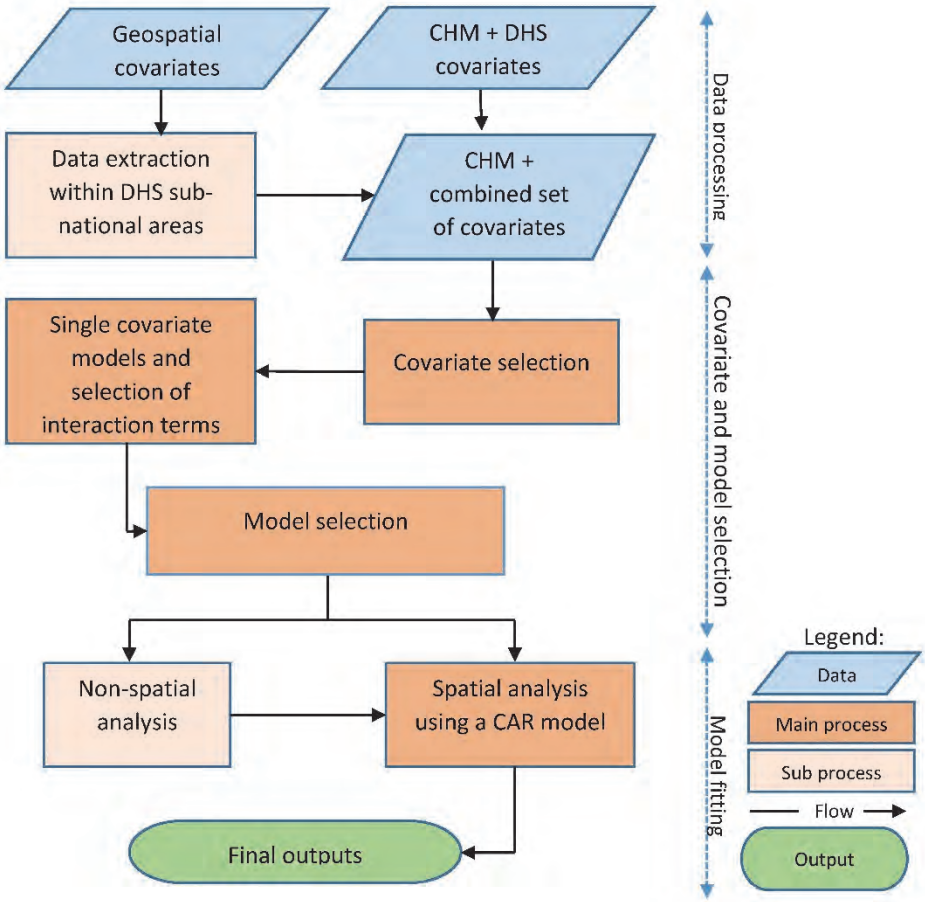
3.1. Overview of Analysis Plan

This section describes the spatial and non-spatial analysis methods used in the study. In general, the analysis involved three main steps. The first step, data processing, has been described in the previous section. The next step involved covariate and model selection in which individual (and combinations of) DHS and geographic covariates that were capable of explaining variation in 4q1 in a linear model setting were explored. This step included an exhaustive model selection process. Finally, a spatial model along with the top linear model identified in the model selection process was fitted. These can be visualized in the flow diagram in Figure 6 and are described in detail in the following sections.

3.2. Covariate and Model Selection

The exploratory analysis methods included 1) selection of the most relevant set of covariates, 2) selection of top interaction terms and 3) the model selection procedure that resulted in the best combination of covariate terms for use in the final analysis.

Figure 6. Flow diagram of analysis methods adopted in the study



3.2.1. Covariate selection

To select the best set from the 10 DHS and over 30 geospatial covariates, both statistical and policy relevance considerations were explored. First, the relationships between each covariate and 4q1 was visualized with scatter plot diagrams to ensure assumptions of linearity were met. The latter was implemented by fitting nonparametric smooth curves (specifically, locally weighted regression (*loess*) curves; Cleveland and Devlin 1988) to the points in the scatter plots, as well as simple linear regressions for comparison.

To identify potential issues arising from collinearity between covariates, all covariate pairs were compared for the correlation between them. The pairs with a correlation of greater than 0.5 (in absolute value) were identified for potential exclusion. As an additional step to address the problem of collinearity between the covariates, the variance inflation factor (VIF) of each variable was calculated. The VIF measures the increase in the variance of an estimated regression coefficient due to (multi) collinearity. For this analysis, collinearity was considered high for covariates with a VIF greater than 4 - a value that represents a two-fold increase in the standard error of a regression coefficient as a result of collinearity. When choosing between collinear variables after the VIF analysis, the decision was based on the variables with the highest R^2 statistic when compared to 4q1, as well as the policy relevance of each variable. In a few cases, collinear covariates with moderate VIFs and high policy relevance were retained.

3.2.2. Single covariate models, selection of interaction terms and best non-spatial model

To determine an optimal subset of covariates for modeling 4q1, single covariate Gaussian regression models were fitted to the screened set of covariates and the AIC (Akaike information criteria), R^2 and predictive R^2 (PR^2) statistics were calculated. The AIC is a measure of the relative quality of fit of a statistical model; the smaller the AIC, the better the fit. The predictive R^2 statistic was calculated with Monte Carlo cross-validation to determine the performance of the models for the out-of-sample prediction. The out-of-sample prediction, with a randomly selected 80% of the data, was used for model training while the remaining 20% was used for validation. This process was repeated 10 times for 10 different splits of the data and measures of model fit averaged across these iterations. All the fitted models were then ranked based on the R^2 statistic. All covariates whose R^2 statistics were less than 5% were eliminated after this ranking.

To identify potentially important interactions between covariates, models that included interactions between pairs of covariates were explored and measures of model fit recorded. To reduce the candidate set of interaction terms from the selection process and to maintain interpretability of the resulting models, particular emphasis was placed on interactions with potential policy relevance. By default, interaction models include terms for each of the covariates in the interaction, as well as a term for the interaction. To ensure that the interaction significantly improved the fit of the model, the interaction model was also compared with a model that included only the two main effects. Only interaction models that resulted in a greater than 2 improvement in AIC score were retained. All interaction models with an R^2 of less than 0.25 were excluded. In addition, we investigated potential differences in the effects of each covariate on 4q1 between East, Central, and Western Africa by testing models that included an interaction with region.

With the set of interaction terms identified (excluding selected region-level interactions), an exhaustive model selection procedure was used to test all possible combinations of multiple (first and second order) terms in a model. Inclusion of additional covariates was stopped when there was no discernible improvement to model fit with the inclusion of greater numbers of covariates. To guard against overfitting, the difference between PR^2 and R^2 scores was monitored; models in which PR^2 was lower than R^2 by greater than 5% were eliminated. Models were assessed by AIC, R^2 and PR^2 and ranked by R^2; the model with the lowest AIC score was retained.

3.3. Spatial Analysis Using a Conditional Autoregressive (CAR) Model

In a non-spatial modeling framework, the observed data are assumed to be identically and independently distributed. However, data collected over space such as the 4q1 typically exhibit spatial autocorrelation. Although a proportion of this spatial autocorrelation may be accounted for by the inclusion of spatially-varying covariates in standard linear regression models, it is often the case that spatial dependence remains in the residuals that result from fitting such models. To test for spatial autocorrelation in the residuals, Moran's I permutation test was used (Banerjee, Carlin, and Gelfand 2015, p.75). The Moran's I statistic is similar to a correlation coefficient. Positive values indicate that neighboring observations tend to be more alike whereas negative values indicate that nearby values tend to be dissimilar. A value of zero indicates a random spatial pattern. Spatial random effects were modeled with a conditional autoregressive (CAR; e.g., Leroux, Lei, and Breslow (2000) fitted in a Bayesian framework using the CARBayes software package in R; Lee 2013). A full description and implementation details of the model are provided in Appendix B.

4. Results

4.1. Results from Covariate and Model Selection

4.1.1. Covariate selection

The plots of each covariate against 4q1, together with the loess fits, showed that the relationships could be reasonably modeled as linear (see Figures 7a and 7b). Testing for highly correlated covariates restricted the list to 29 covariates for the analysis (Table 5).

4.1.2. Single covariate models, interaction terms and best non-spatial model

Table 5 shows the rankings of the 29 retained covariates in terms of R^2. The unstandardized and standardized (centered and scaled by the standard deviation of the covariate) coefficients of the covariates are also reported. Note that standardized coefficients are only included to allow comparison between the relative magnitudes of the effects of the covariates on 4q1. This comparison is untenable using the unstandardized coefficients because units/scales on which the covariates were measured were different. For example, the unstandardized coefficient of malaria prevalence, measured in proportions, is estimated at 55.02. This coefficient will, however, change to 0.55 when this covariate is expressed in percentages to make it comparable with the DHS covariates. The R^2 statistics show that some of the covariates offer little or no explanatory power for modeling 4q1. Using the criterion of an R^2 value of at least 5%, the first 12 covariates were selected as the best subset to be tested for inclusion in the final model. This subset includes 6 DHS and 6 geospatial covariates. Figures 7a and 7b show the relationships between these covariates and 4q1.

Table 5. Summary of single covariate models. Coefficients are presented as both unstandardized (i.e. native units) or standardized.

Rank	Covariate	Unstandardized coefficient	Standardized coefficient	AIC	R^2	PR^2
1	Literacy	-0.42	-10.74	2235	0.24	0.24
2	Malaria prevalence	55.02	10.44	2240	0.23	0.23
3	Birth interval	1.04	9.46	2252	0.18	0.20
4	Access to a health facility	-0.36	-8.09	2268	0.14	0.13
5	Sanitation	0.31	7.99	2269	0.13	0.11
6	Measles vaccination	-0.56	-7.93	2269	0.13	0.13
7	Stunting	0.62	6.82	2279	0.10	0.08
8	Ethnic group variety	2.33	6.53	2281	0.09	0.10
9	Mean temperature	0.22	6.27	2283	0.08	0.08
10	Population density	0.00	-5.33	2289	0.06	0.06
11	Economic activity	-8.19	-5.14	2290	0.05	0.05
***12**	Night-time lights	-2.20	-4.75	2292	0.05	0.05
13	DPT3 immunization	-0.21	-4.50	2294	0.04	0.04
14	Breastfeeding	-0.20	-4.27	2295	0.04	0.04
15	Conflict density	-8.64	-3.15	2299	0.02	0.02
16	Modal ethnic group	-0.01	-2.89	2300	0.02	0.01
17	Landcover variety	-0.90	-2.79	2300	0.02	0.01
18	Cattle density	-0.05	-2.53	2301	0.01	0.01
19	Water sources	-0.10	-2.47	2301	0.01	0.01
20	Growing season length	0.02	2.39	2302	0.01	0.02
21	Roughness	-3.97×10^{6}	-2.35	2302	0.01	0.02
22	Pig density	-0.16	-1.92	2303	0.01	0.01
23	Travel time	0.00	-1.19	2304	0.00	0.00
24	Rainfall	0.00	0.96	2304	0.00	0.00
25	Vegetation	-7.41	-0.75	2304	0.00	0.00
26	Modal landcover	-0.01	-0.55	2304	0.00	0.00
27	Urban population	0.00	-0.48	2304	0.00	0.00
28	Conflict fatalities	-189.91	-0.13	2305	0.00	0.00
29	Drought	-0.04	-0.11	2305	0.00	0.00

*The top 12 covariates have R^2 values greater than 5%.

Figure 7a. Plots of selected DHS variables against child mortality rates. The blue lines are linear regression lines while the red lines are nonparametric *loess* fits to the data.

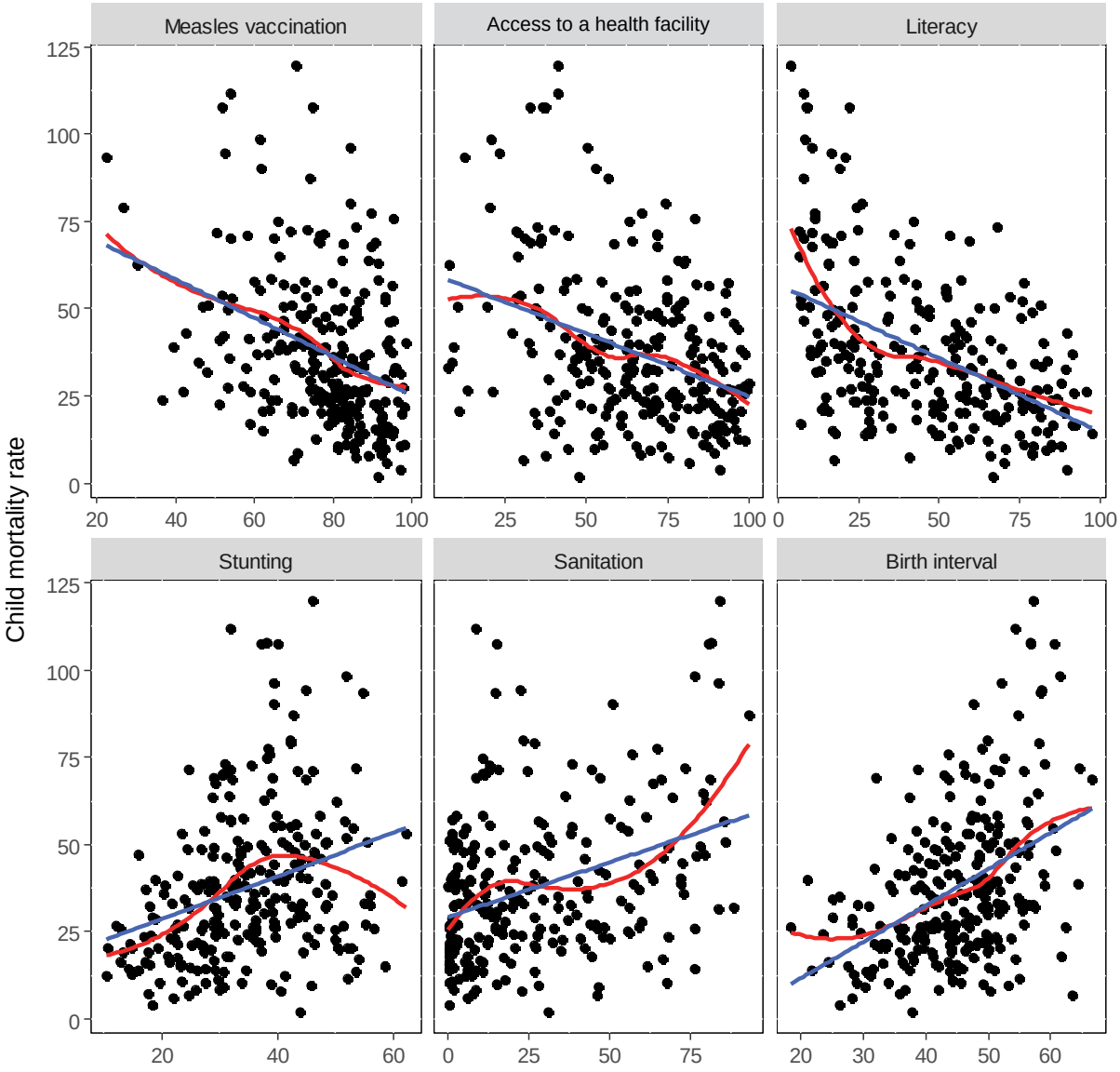
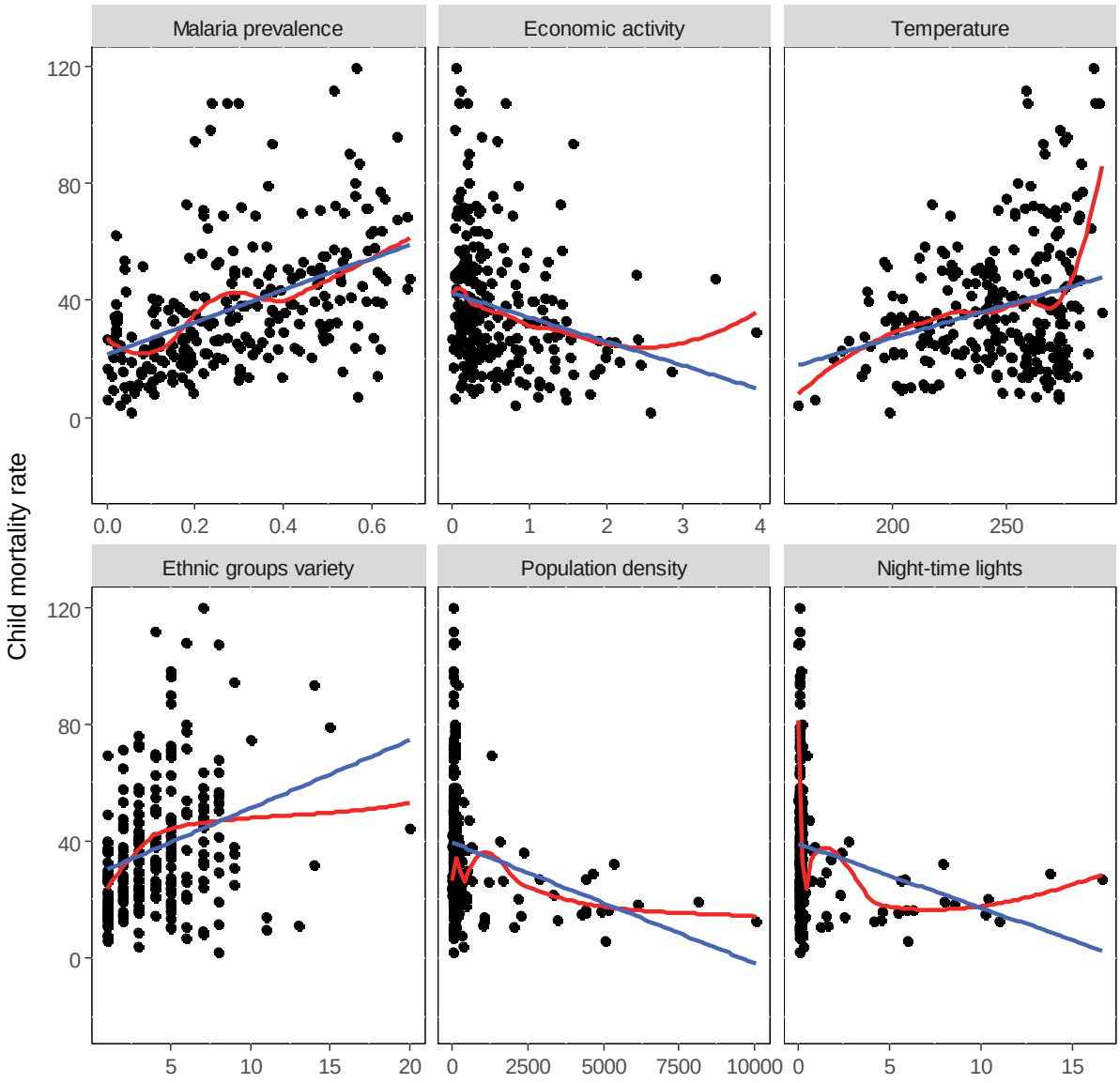


Figure 7b. Plots of selected geospatial covariates against child mortality rates. The blue lines are linear regression lines while the red lines are nonparametric *loess* fits to the data.



Relationships between the DHS variables and child mortality are reasonably linear (Figure 7a). Similarly, the linearity assumption for the geospatial covariates is also justifiable in most cases (Figure 7b). Population density and night-time lights exhibit substantial skewness. Nevertheless, residual diagnostics showed that this was not a problem in the analysis.

Table 6 shows the results of selection of interaction terms. Note that not all interactions were included in the model selection; only interactions that were relevant for policy and had R^2 values greater than 25% were retained. Most top interactions included either female literacy or birth interval in combination with other top DHS and geographic variables. Top interactions with region also included stunting and access to a health facility. When considering pairs of explanatory factors, the interaction models show significant improvement over the single covariate models. The best single covariate model explains only about 24%

of the variation in child mortality; the best interaction model accounts for as much as 40% of this variation. This clearly shows the importance of interactions between covariates when modeling child mortality.

Table 6. Top interaction models. A ':' indicates the interaction between covariates.

Interactions between covariates			Region-level interactions		
Model	**AIC**	**R^2**	**Model**	**AIC**	**R^2**
Malaria prevalence: Birth interval	2176.98	0.40	Birth interval: Region	2213.76	0.32
Literacy: Ethnic group variety	2204.61	0.34	Access to a health facility: Region	2215.08	0.32
Literacy: Birth interval	2205.36	0.33	Literacy: Region	2218.45	0.31
Literacy: Access to a health facility	2224.45	0.28	Stunting: Region	2240.61	0.25
Birth interval: Temperature	2228.78	0.27			
Literacy: Sanitation	2230.91	0.26			
Literacy: Night-time lights	2232.99	0.26			
Literacy: Economic activity	2233.00	0.26			
Birth interval: Ethnic group variety	2235.14	0.25			

In the non-spatial model selection phase, top models included up to 8 of the 12 covariate terms. Models with greater numbers of covariates exhibited only marginal improvements in AIC scores and increasing evidence of overfitting as shown by the large differences in R^2 and PR^2 values. Top models had adjusted R^2 values of up to 0.54, which indicated good explanatory power. These models also had PR^2 values of $\approx$0.53 that indicated no overfitting of the models to the data. A list of the top 10 models as ranked by the AIC is presented in Table B.1 (Appendix B).

4.2. Results from Spatial Analysis using a Conditional Autoregressive (CAR) Model

Testing for spatial autocorrelation in the 4q1 data using Moran's I yielded a statistic of 0.1164 ($p<0.005$); this indicated the presence of significant positive spatial autocorrelation in the residuals. Fitting of a spatial random effect with the top selected model yielded an adjusted R^2 of 0.60; this showed some improvement over the non-spatial model. Spatial models without region-level interactions were also fitted to provide confirmatory tests of significance for the marginal effects of the covariates with region-level interactions (see Table B.2 in Appendix B for the results of this model). Table 7 reports the posterior means, standard deviations and quantiles of the covariate effects ($\boldsymbol{\beta}$), the variance parameters (σ^2, τ^2), and the correlation parameter (ρ) of the spatial model.

Table 7. Posterior estimates of the parameters of the spatial model

Parameter	Mean	Standard Deviation	Median	95% Credible Interval
(Intercept)	27.0289	2.8163	27.2437	(21.1830, 32.0304)
Region(CAFR)	17.6602	4.9860	17.5071	(8.8964, 27.6040)
Region(WAFR)	7.2262	4.6206	7.0430	(-0.6524, 16.3675)
Stunting	0.4594	0.1935	0.4623	(0.0646, 0.8045)
Malaria prevalence	49.9618[a]	7.5869	49.9206	(35.8431, 64.4336)
Access to a health facility	-0.1821	0.1030	-0.1733	(-0.3990, 0.0049)
Birth interval	0.5418	0.2423	0.5480	(0.0787, 0.9899)
Literacy	0.1298	0.1073	0.1145	(-0.0670, 0.3620)
Temperature	0.0214	0.0641	0.0228	(-0.1094, 0.1596)
Night-time lights	0.0678	1.2112	-0.0215	(-2.1017, 2.5609)
Ethnic groups' variety	-0.7261	0.3939	-0.7016	(-1.5366, 0.0667)
Literacy: Birth interval	-0.0159	0.0076	-0.0154	(-0.0307, -0.0011)
Literacy: Access to a health facility	0.0063	0.0032	0.0064	(0.0002, 0.0129)
Birth interval: Malaria prevalence	1.6979	0.8352	1.7029	(-0.1293, 3.4090)
Birth interval: Temperature	-0.0069	0.0065	-0.0072	(-0.0186, 0.0046)
Literacy: Night-time lights	-0.0222	0.0337	-0.0201	(-0.0931, 0.0405)
Birth interval: Ethnic groups' variety	0.0790	0.0478	0.0740	(-0.0040, 0.1813)
Birth interval: Region(CAFR)	0.0918	0.5292	0.0535	(-0.9604, 1.1593)
Birth interval: Region(WAFR)	0.7182	0.4311	0.7471	(-0.1750, 1.5533)
Access to a health facility: Region(CAFR)	-0.3455	0.2330	-0.3549	(-0.7955, 0.1560)
Access to a health facility: Region(WAFR)	0.0584	0.1616	0.0502	(-0.2476, 0.3994)
Literacy: Region(CAFR)	-0.3164	0.2771	-0.3177	(-0.8563, 0.2214)
Literacy: Region(WAFR)	-0.2182	0.1527	-0.2186	(-0.5356, 0.0873)
Stunting: Region(CAFR)	-0.3964	0.4566	-0.4265	(-1.3164, 0.4256)
Stunting: Region(WAFR)	-0.3345	0.2585	-0.3226	(-0.8072, 0.1601)
$\boldsymbol{\sigma^2}$	144.8332	104.8733	204.9695	(0.0014, 262.5637)
$\boldsymbol{\tau^2}$	117.1275	155.2425	0.0231	(0.0005, 404.6400)
$\boldsymbol{\rho}$	0.3339	0.2407	0.2495	(0.0390, 0.8736)

[a] Note that malaria prevalence is measured in proportions. This estimate is equal to 0.4996 (i.e., 49.9618 divided by 100) when malaria prevalence is transformed to percentage.

A number of conclusions can be drawn from these results. First, the 95% credible intervals of both parameters of the spatial random effects (τ^2, ρ) do not include zero; this affirmed the existence of substantial (positive) spatial autocorrelation in child mortality across the DHS sub-national areas. General patterns in the model coefficients (the parameter estimates) can be interpreted as:

1. Child mortality rates vary significantly across the three regions in the study.

2. Stunting, malaria prevalence, access to a health facility, and too short birth interval have significant marginal effects on child mortality (conditional on the average values of other covariates in each case). The marginal effects of other covariates were not significant.

3. The interactions between female literacy, birth interval, and access to a health facility had a significant effect on child mortality (conditional on the average values of other covariates in each case). Other inter-covariate interactions were not significant.

4. None of the region-level variations in the effects of the tested covariates (birth interval, access to a health facility, literacy, and stunting) on child mortality were significant.

A more detailed interpretation of these patterns is described below. All coefficients are interpreted in the context of holding all other covariates at their mean values.

Intercept, CAFR, WAFR: The intercept (set to EAFR in this model) has an estimated value of 27.03. This value is the average predicted mortality rate per 1,000 children surviving to age 12 months for EAFR. Similarly, 27.03 + 17.67 = 44.69 and 27.03 + 7.26 = 34.3 are the respective average predicted mortality rates per 1,000 children surviving to age 12 months for CAFR and WAFR. The 95% credible intervals of these coefficients show that the average mortality rate for the reference region (EAFR) is significant. There is a significant difference between the average predicted mortality rates for EAFR and CAFR, but not between EAFR and WAFR. This implies that regional variation in mortality is significant. Finally, the average predicted mortality rate is highest in CAFR and lowest in EAFR.

Stunting: Estimates of regression coefficients for stunting, stunting:CAFR, and stunting:WAFR are 0.46, -0.40 and -0.33 respectively. Thus, a unit increase in the percentage of stunting prevalence in EAFR will lead to an increase of 0.46 in child mortality. Similarly, for CAFR and WAFR, these are 0.46 - 0.40 = 0.06 and 0.46 – 0.34 = 0.12, respectively. These region-level variations in the effect of stunting are not significant. Table B.2 (Appendix B) shows that for all the DHS sub-national areas, a unit increase in stunting will lead to a significant increase of 0.33 in child mortality.

Malaria Prevalence (*Plasmodium falciparum*): In the fitted model, there exists an interaction between malaria prevalence and birth interval. The effect of this interaction is estimated at 1.70. Thus, a unit increase in too short birth intervals will increase the effect of malaria prevalence (in percentage) on child mortality by 0.02 (≈ 0.17 divided by 100) and vice versa. This interaction is not significant. A percentage increase in the prevalence of *Plasmodium falciparum* infection will result in an additional 0.5 deaths in children age 1 – 5. The 95% credible interval shows that this marginal effect is significant.

Birth Interval: The interaction between short birth interval and malaria prevalence has been explained. Birth interval also interacts with literacy, temperature, and ethnic group variety. The estimates of these interaction effects are -0.02, -0.01, and 0.08, respectively. Hence, a unit increase in ethnic group variety will increase the effect of too short birth intervals on child mortality by 0.08, whereas increasing mean temperature and female literacy will decrease its effect by 0.01 and 0.02, respectively. However, only the interaction between birth interval and female literacy is significant in the model. These conclusions are relevant for the reverse cases when the roles of these covariates are interchanged. For example, a unit increase in too short birth intervals will also increase the effect of ethnic group variety on child mortality by 0.08.

The model also includes interactions between region and birth interval. The estimates of the interaction effects are 0.54 for EAFR, 0.63 (= 0.54 + 0.09) for CAFR and 1.26 (= 0.54 + 0.72) for WAFR. Thus, a unit increase in too short birth intervals will increase child mortality by 0.54, 0.63, and 1.26 in EAFR, CAFR and WAFR respectively. However, these regional variations in the effect of birth interval are not significant. Table B.2 (Appendix B) shows that a unit increase in too short birth intervals will significantly increase child mortality by 0.78.

Literacy: This covariate interacts with birth interval as previously explained. The covariate also interacts with access to a health facility and night-time lights. The respective estimates of the effects of these interactions are 0.01 and -0.02. Thus, a unit increase in night-time lights will reduce the effect of female literacy on child mortality by 0.02. On the other hand, an increase in access to a health facility will increase the effect of female literacy on child mortality by 0.01. The reverse cases of these relationships are also true. While other interaction effects are significant, the interaction between literacy rate and night-time lights is not significant.

The regional effects of this variable have also been estimated. These include 0.13 for EAFR, -0.19 (= 0.13 - 0.32) for CAFR and -0.09 (= 0.13 - 0.22) for WAFR. A unit increase in female literacy corresponds with a reduction of child mortality by 0.19 in CAFR and 0.09 in WAFR. In EAFR, there is a positive (but non-significant) relationship between female literacy and child mortality, which may be due to the attenuating effect of other covariates. Table B.2 (Appendix B) shows that the overall effect of female literacy is not significant.

Access to a Health Facility (delivery in a health facility used as a proxy for access to a health facility): This covariate interacts with literacy rate. The model also includes its interaction with region. The estimates of these regional effects are -0.18, -0.53 (= -0.18 - 0.35), and -0.12 (= -0.18 + 0.06) for EAFR, CAFR, and WAFR regions respectively. Thus, conditional on the average values of all other covariates, a unit increase in access to a health facility will reduce child mortality by 0.18 in EAFR, 0.53 in CAFR, and 0.12 in WAFR. The regional differences in the effect of access to a health facility are not significant. Table B.2 (Appendix B) shows that conditional on the average values of other covariates, a unit increase in access to a health facility corresponds to a reduction of child mortality by 0.16.

Temperature, Night-time Lights, and Ethnic Group Variety: Unit increases in mean temperature and night-time lights are associated with non-significant increases of 0.02 and 0.07 in child mortality, respectively. Similarly, a unit increase in ethnic groups' variety will lead to a non-significant reduction of child mortality by 0.73. There also exist non-significant interactions between temperature and birth interval, night-time lights and literacy, and ethnic group variety and birth interval as explained previously.

4.2.1. Model predictions of child mortality rates

Figures 8a displays the observed child mortality rates and the width of the 95% confidence interval for the indicator value; Figure 8b displays the predicted (or fitted) child mortality rates. These predicted rates represent smoothed values of the observed rates that have been adjusted for covariate effects as well as spatial autocorrelation. The maps of the observed and predicted rates both show a concentration of high mortality regions mostly in Central and Western Africa. The maps also show some evidence of clustering in the spatial distribution of child mortality with most neighboring DHS sub-national areas that exhibit similar mortality rates. Standard deviations around these predicted regional rates (Figure 8c) suggest that there should be more variability within some regions than in others.

Figure 8a. Observed mortality rates and 95% confidence interval width for the 255 DHS sub-national areas in sub-Saharan Africa

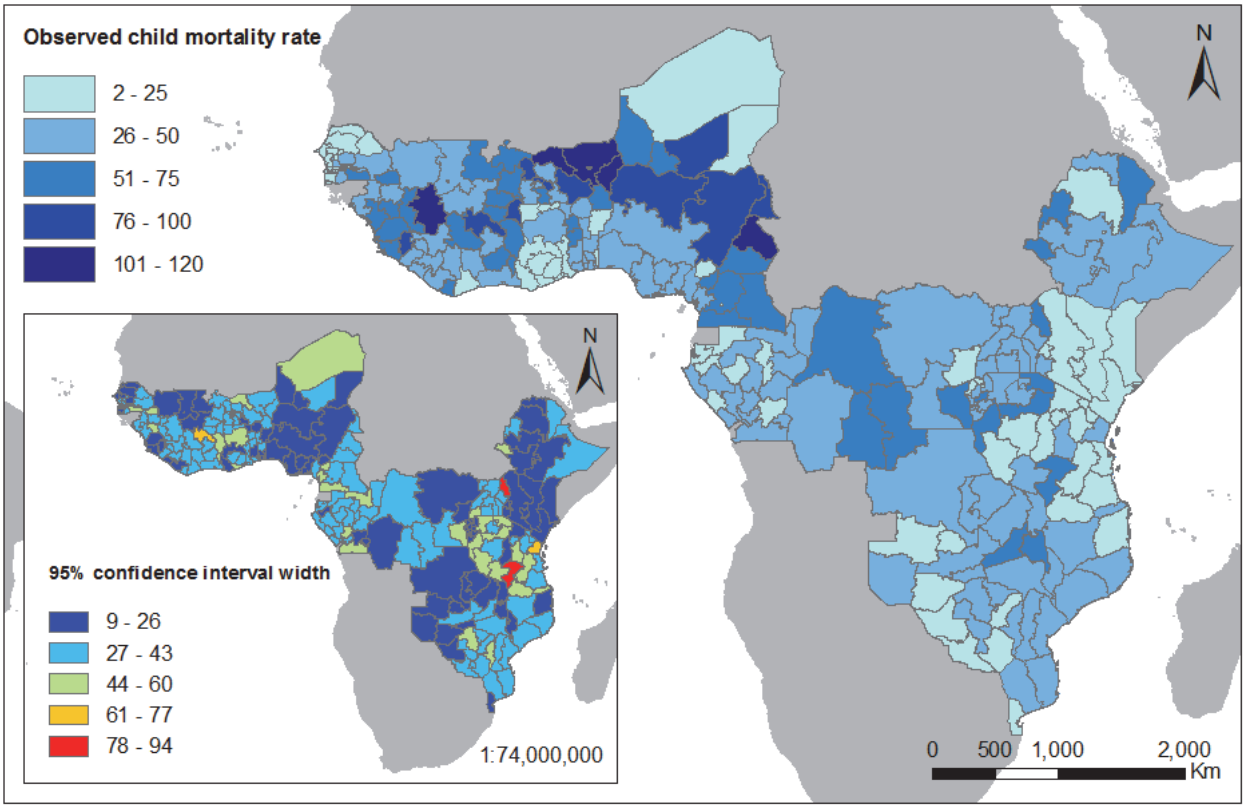


Figure 8b. Predicted mortality rates for the 255 DHS sub-national areas in sub-Saharan Africa

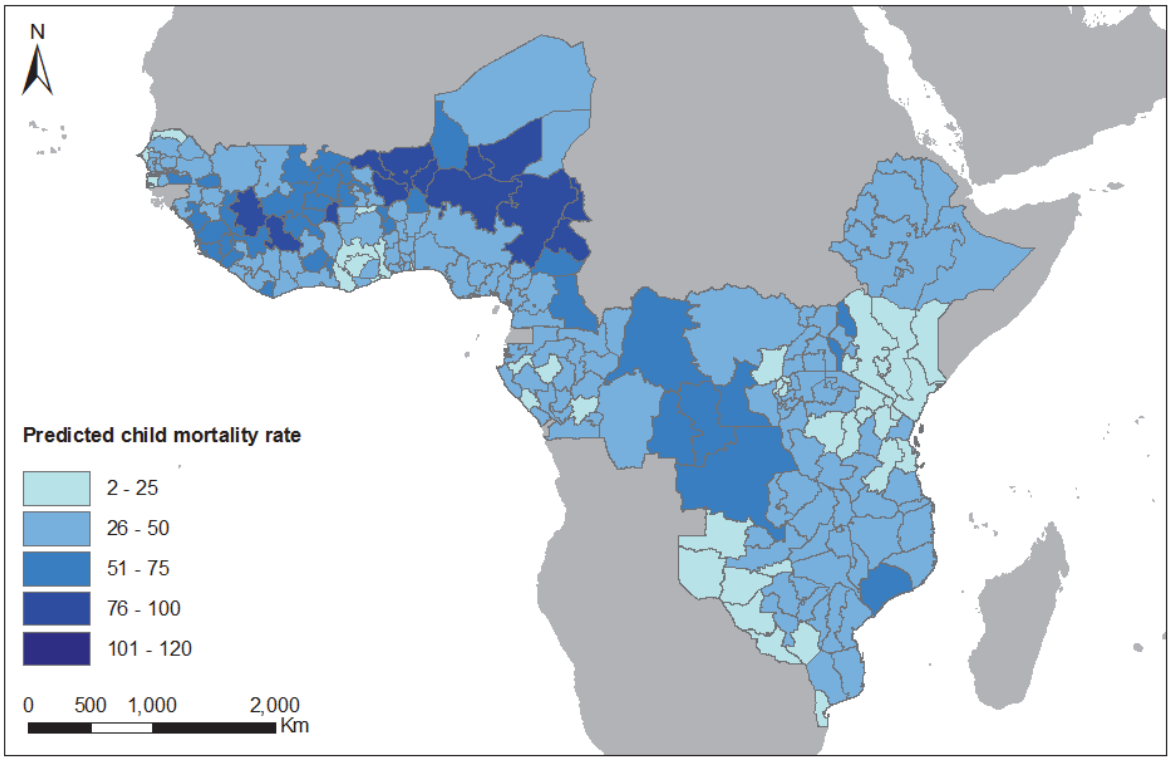


Figure 8c. Predicted standard deviations of the child mortality rates for the 255 DHS sub-national areas in sub-Saharan Africa

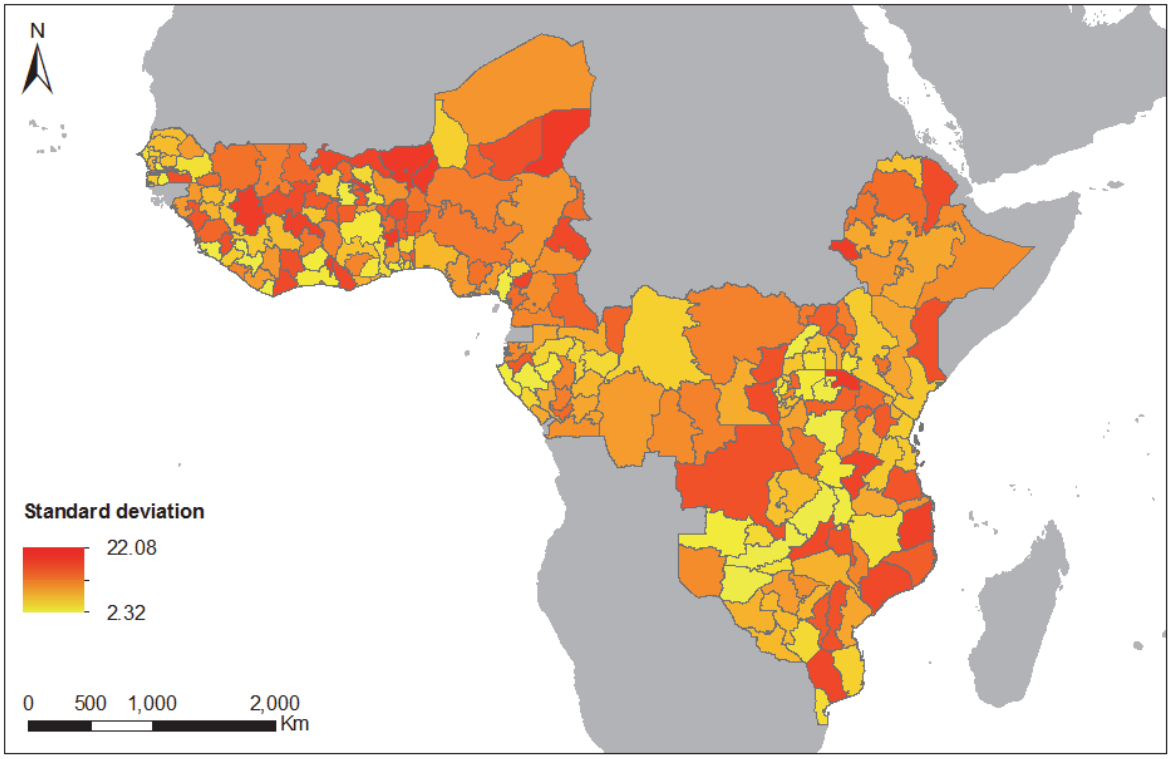
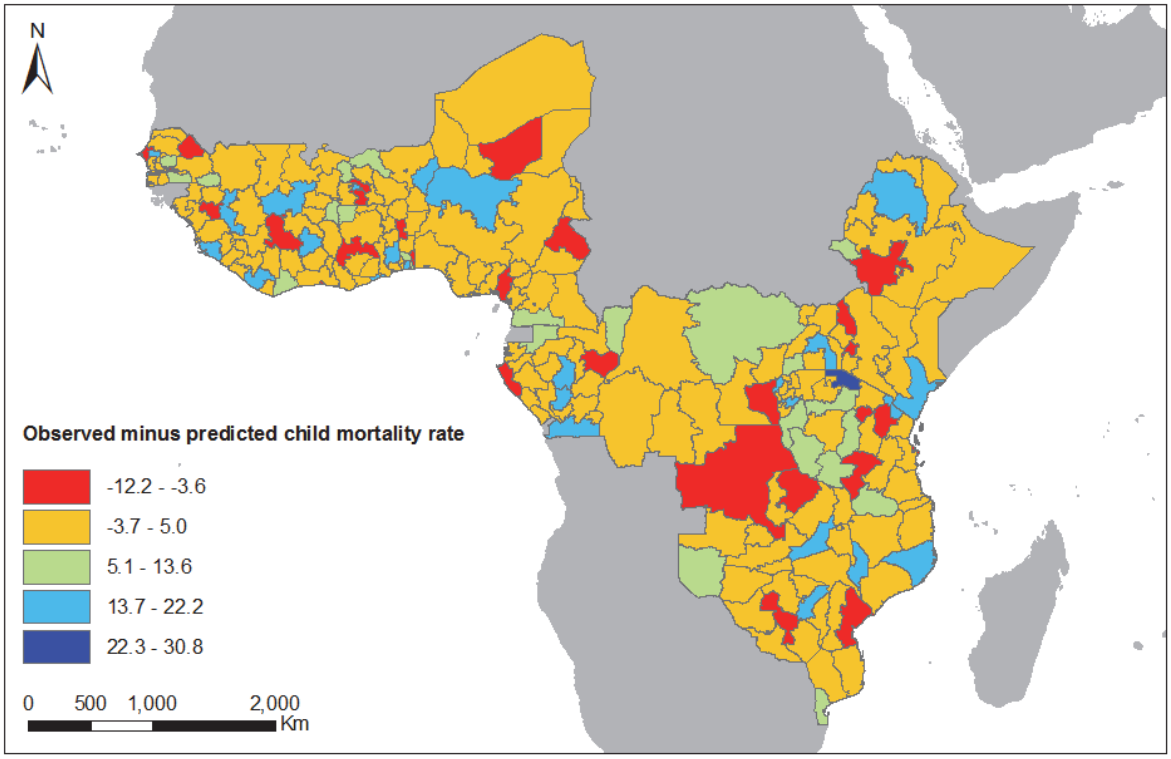


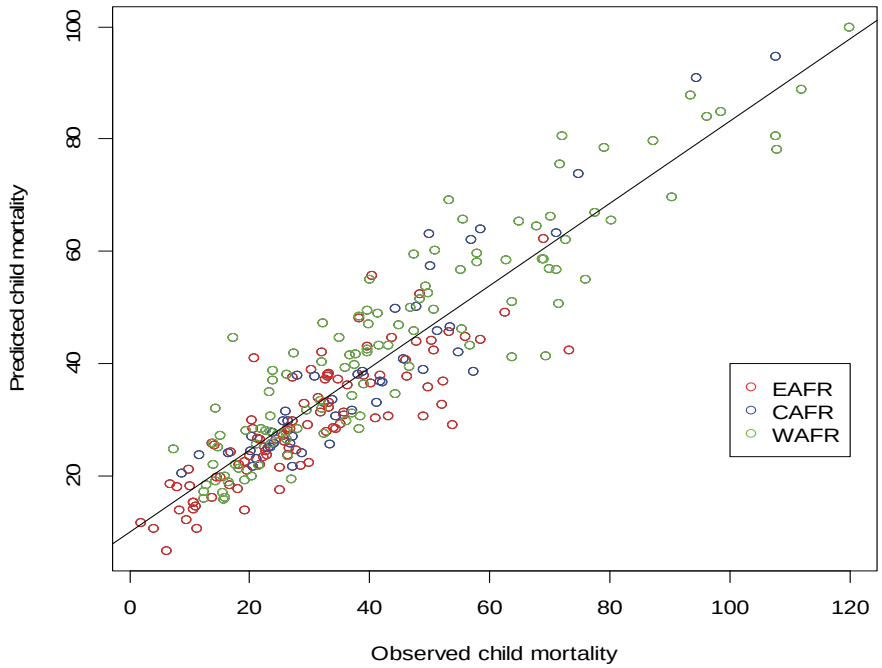
Figure 8d. Observed minus predicted child mortality rates for the 255 DHS sub-national areas in sub-Saharan Africa



The three highest predicted mortality rates occurred in a DHS sub-national area in Burkina Faso (Western Africa) and two sub-national areas in Cameroon (Central Africa). The three lowest predicted mortality rates occurred in two sub-national areas in Kenya (Eastern Africa) and a sub-national area in Ethiopia (Eastern Africa). The standard deviations of the predicted values show high uncertainty in few DHS sub-national areas but generally moderate to low uncertainty in others; this indicates a good fit to the data. In most DHS sub-national areas, differences between observed and expected 4q1 are relatively low (Figure 8d), with one exception in northern Tanzania.

Finally, Figure 9 is a scatter plot of the strong correlation between predicted mortality rates against the observed rates. This affirms the linear relationships between child mortality and the covariates used in the model. The figure reveals that high mortality rates are concentrated in Western and Central Africa as compared to Eastern Africa, where mortality values appear to be lower.

Figure 9. A plot of observed vs. predicted child mortality rates

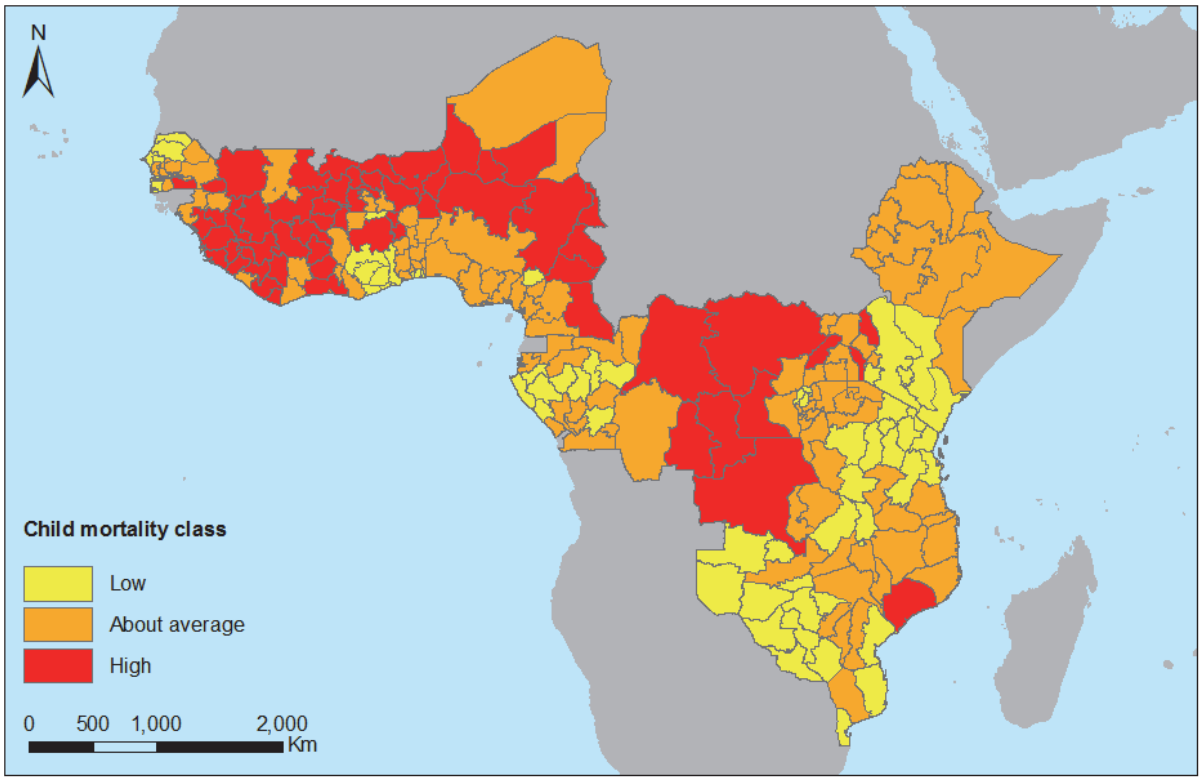


4.2.2. Classification of mortality regions

The average predicted child mortality rate for all 255 DHS sub-national areas was 37.55 deaths per 1,000 children surviving to age 12 months (37.58 deaths per 1,000 for the observed rates). However, there was a good deal of variation around this average, with some neighboring areas exhibiting strongly different differences. In this section, the DHS survey regions were classified as low, about average, and high mortality regions based on their predicted mortality rates. Low mortality regions were defined as regions whose 95% prediction intervals fell below the average predicted rate. High mortality regions were regions whose 95% prediction intervals were above the average predicted rate. Similarly, the 95% prediction intervals of the regions classified as about average included the average predicted rate.

Figure 10 maps the child mortality classes. The figure shows that high mortality regions are primarily concentrated in Western and Central Africa, while most low and about average regions are located in the Eastern region. The DHS sub-national areas are shown in yellow – significantly lower than average predicted mortality rate; red – significantly higher than average predicted mortality rate; and orange – not significantly different from average predicted mortality rate.

Figure 10. Classification of child mortality regions



5. Discussion

The DHS Spatial Analysis Report 12 (SAR12) analyzed the spatial distribution of key maternal and child health indicators for 255 DHS sub-national areas across 27 countries in sub-Saharan Africa. The report found spatial autocorrelation and cross-border relationships in the indicators. The presence of spatial autocorrelation in the indicators implies that neighboring DHS sub-national areas are more likely to share similar patterns than areas that are geographically distant. The present report builds on SAR 12 by exploring the spatial autocorrelation observed in child mortality rates to understand the determinants of its spatial distribution. Moreover, this report accounts for the effect of spatial environmental factors along with the usual demographic and socio-economic determinants on child mortality rates. This choice is supported by previous studies such as Balk et al. (2003, 2004) in which spatial environmental factors were identified as acting more strongly on child mortality than on infant mortality.

Child mortality rates in this study were defined as the number of deaths among children between the age of 1 and 4 in the five-year period preceding the survey per 1,000 children surviving to age 12 months[7]. This includes deaths reported at age 12– 59 months and excludes infant deaths (deaths occurring between age 0 and 11 months). To the authors' knowledge, this is one of the first studies to explore spatial patterns in child mortality rates across such a large region and accounting for both the effect of demographic and socio-economic determinants and spatially explicit environmental factors. Importantly, the protective and risk factors used in the models to explain variation in child mortality have direct relevance for both policy and interventions.

The results of this report show how a large degree (60%) of regional variation in child mortality can be explained by the socio-demographic and environmental indicators included in the analysis, the interactions between them, and the patterns in child mortality. While the results of this report do not necessarily help better predict child mortality rates, they do provide greater clarity about which of these factors are the most important drivers across the region.

This analysis further investigated and confirmed patterns of positive spatial autocorrelation in child mortality rates across the 255 DHS sub-national areas. On a larger scale, this spatial autocorrelation manifests as the presence of significant differences in child mortality rates across Western, Central, and Eastern Africa regions, with the highest child mortality rates in Central and Western Africa and the lowest in Eastern Africa (see Figure 8b). The results from the spatial regression model also highlight that several socio-economic and environmental factors have significant effects on child mortality in the 255 DHS sub-national areas. These effects, however, do not vary significantly across Western, Central, and Eastern regions. Therefore, the application of interventions to reduce child mortality may have similar effects on child mortality in the three regions.

Single covariate models used in this study are useful for identifying factors that can explain a meaningful portion of the variation in child mortality. These include female literacy, prevalence of malaria, birth interval, access to a health facility, sanitation practices, measles vaccination, and stunting. All these covariates had R^2 values greater than 10%. This agrees with other studies about the significance of these factors (Table 5; Balk et al. 2004). The ranking of these covariates using the R^2 statistic already makes a strong case for which protective and risk factors should take priority in addressing child mortality. Indeed,

[7] Child mortality rates (age 1 to 4) includes deaths reported as age 12– 59 months. Therefore, infant deaths (before age 1) are not included in the metric used in this report. The DHS Program uses a synthetic cohort life table approach to directly estimate the under-five mortality rate. Details may be found in the Guide to DHS Statistics, pp.90-95 (http://www.dhsprogram.com/publications/publication-DHSG1-DHS-Questionnaires-and-Manuals.cfm) and further described in Pullum et al. 2013.

many of these factors have been highlighted as priorities in the 2015 SDGs (United Nations General Assembly 2015). However, the effects of single covariates are not necessarily as useful as when other explanatory factors are considered.

The interpretation of the coefficients of the spatial model provides a perspective on the relationship between each factor and child mortality, while also accounting for all other factors present in the model. This provides a fuller picture of the geographic variation in child mortality. When interpreting marginal effects, it is important to remember that each coefficient represents the marginal increase in child mortality in response to a unit increase in the factor of interest *when all other covariates are kept at their average values*. Thus, for all 255 DHS sub-national areas, stunting showed a statistically significant and positive effect on child mortality. For each increase in a percentage point of stunting, there is an increase of 0.33 deaths per 1,000 children surviving to age 12 months. There were no regional differences observed in this effect. Malaria prevalence is also a significant factor correlated with child mortality in all the DHS sub-national areas analyzed. A unit increase in the percentage of the population with *Plasmodium falciparum* parasites in their blood (measured in the 2-10 age group) corresponds with an increase of 0.5 deaths per 1,000 children surviving to age 12 months. Similarly, a unit increase in the percentage of births with a birth interval defined as too short is related to a 0.78 increase in child mortality. The marginal effect of access to a health facility is also statistically significant; a unit increase in the percentage of women who accessed a health facility is associated with a reduction of 0.16 deaths per 1,000 children surviving to age 12 months.

Interestingly, interactions between pairs of covariates explained a large amount of the variation in child mortality. This perspective has not been well reported. The presence of interaction terms in the model indicates that risk factors can be attenuated by protective factors and vice versa. For example, although the marginal effect of literacy rates on child mortality is not significant, its combination (interactions) with other covariates such as a short birth interval and access to a health facility is statistically significant. Increased literacy reduces the effect of a too short birth interval on child mortality. Conversely, an increase in literacy is associated with an increase in the effect of access to a health facility on child mortality. Therefore, interventions to increase literacy rates will reduce the risk effect of a too short birth interval on child mortality. On the contrary, improving literacy increases the protective effect of accessing a health facility on child mortality.

In contrast with other work that assessed the demographic and socio-economic determinants of child mortality across the 255 DHS sub-national areas in sub-Saharan Africa, this study controlled for spatial environmental factors. Among the covariates tested in this study, mean temperature, night-time lights and ethnic group variety showed a significant correlation with child mortality rates in the single covariates model. While none of these terms or their interactions were significant in the final model, their inclusion helped with the overall predictive power. These factors have strong geographic variation, and can be viewed as covariates that help explain the spatial variability in child mortality.

There are several limitations to the present study that should be taken into account when interpreting the total results. First, although the sample size (N=255) was sufficient to highlight patterns across sub-Saharan Africa, regional differences in the effects of the covariates were not significant in the final model. This may be a consequence of the predictive power of the model given the data. That is, larger sample sizes or more focused studies might help clarify some factors at the regional level; this could further highlight priorities for interventions. One simple solution to this problem would be exploring the use of smaller aggregation areas for the analysis. The DHS data in this analysis included tens of thousands survey clusters aggregated at the DHS sub-national area level. This increases the support and therefore reliability of each data point, although it reduces the total number of available data points. In addition, aggregation at DHS sub-national level is likely to obscure some of the finer-scale variation in each factor. Thus, the relationship between covariates and child mortality might be clearer if it were analyzed with a finer geographical scale.

Similarly, aggregating geospatial covariates within the DHS sub-national area level may obscure the effects of these factors by not representing the local conditions that are experienced at the cluster level. For example, a region that is mostly arid and uninhabited but has a small area of high vegetation cover and a densely populated area would be represented as arid in this analysis, whereas the conditions experienced in the inhabited areas would be different. Moreover, another limitation is that some of the geospatial datasets used in this study were available only for specific time points up to 15 years prior to the temporal coverage of the DHS surveys. Indeed, although it is relatively safe to assume that time-invariant datasets (topography) did not change over time, temporally explicit datasets (economic activity) may have significantly changed over time; consequently, their use may have produced results that could be improved if more recent datasets were available.

Furthermore, it is important to note that all geospatial datasets in this study have a degree of uncertainty that will be propagated into the analysis. Each DHS variable has sampling error associated with each estimate. Additionally, some surveys included in this analysis may be have less reliable child mortality estimates as assessed by UN IGME model (www.childmortality.org). An alternative approach for analysis would involve working at the DHS survey primary sampling unit scale, for which GIS coordinates (latitude and longitude) are usually available (but not for all the 27 countries included in this work). Child mortality aggregated at finer scales would then be used as a binary outcome (detected or not detected) and modeled in a logistic regression. This approach would have the disadvantage of not being easily relatable to the numbers of deaths per 1,000, but it would present a view of overall risk with potentially greater resolution than the aggregated approaches. Further investigations are needed to understand the degree of error and uncertainty that arises when calculating a rare event such as child mortality at fine spatial scale. The modeling approach used in this study allowed us to use the maximum amount of available data despite the limited size of the DHS sub-national areas. Further analyses might investigate alternative approaches to overcome these limitations.

Despite this work showing that a great amount of variation in child mortality (60%) can be explained by a relatively small set of factors, other factors that were not included in the model may contribute in capturing some of the unexplained variance. For example, a direct measure of access to a health facility that includes the number and type of medical staff for a given area, which indicates the degree of supply of health services, was not available for the 27 countries; a proximate variable was used instead to account for access to health facilities. Moreover, none of the variables tested in this study provide a direct measure of HIV prevalence. The set of geospatial covariates tested in this work covered several different aspects that influence child mortality; however, future work could focus on testing additional geospatial covariates, such as rates of air pollution. Household wealth in this work was captured to an extent by using two household level covariates: coverage of improved water sources and sanitation practices. Future areas for investigation may include the use of a comparable wealth index that is standardized across countries and areas. Finally, the association of some variables with child mortality, such as ethnic group variety , might be a subject for future research.

It is remarkable that such strong trends across such broad regions were detected in spite of the above limitations. The emergence of these trends highlights the power of standardized surveys that are deployed across broad regions and cover such a wide range of topics. This broad perspective allows for greater understanding than that achieved by looking within a smaller region. It is important that policy-makers around the world utilize and support broad-scale data collection initiatives. Broad trends can also help to focus research on particular drivers within localized areas, and to understand finer-scale variability in the drivers and protective factors for child mortality. In addition, integrating various data sources could greatly enhance this type of analysis with survey data, data collected from study areas such as surveillance sites, civil and vital registration systems, and routine health system data.

The importance of geographic factors in predicting child mortality presents an interesting opportunity in the context of preventing child mortality. The availability of geographic data at high spatial resolutions, aided by the continuous evolution of geospatial analyses, makes it possible to use these data to predict in areas where DHS data are not available and or at higher spatial resolution in the current study. Such approaches have been successfully deployed, for example, in mapping children for the polio vaccination program in Nigeria (Alegana et al. 2015) and modeling malaria prevalence (Bhatt et al. 2015), both of which facilitate the improved targeting of interventions.

6. Policy Implications

The findings of this study have important implications for both program planning and implementation of programs designed to reduce child mortality (age 1 to 4) in sub-Saharan Africa. The results show that child mortality is best explained by a multitude of interrelated factors. This suggests that there is a role for integrated, multi-dimensional approaches to planning and implementing interventions. This is particularly true where linkages between environmental, demographic, and socio-economic factors are found. Relationships between individual and environmental factors should be further considered and addressed in order to strengthen policy and program effectiveness. Analysis of the interaction terms in this study further highlights the importance of acting not just on single factors that may affect child mortality. Instead, it is important to coordinate actions in which different factors are considered simultaneously for possibly magnifying or dampening the effects on each other.

One particular interaction worth highlighting is female literacy with birth interval. In this study, female literacy significantly influenced the magnitude of the impact of birth interval and access to health care facilities on child mortality. Therefore, a program designed to reduce child mortality by increasing birth interval and improving access to health care could magnify the effects of these interventions with the addition of an intervention that increases female literacy. Alternatively, if program funding is restricted to health interventions, knowing the level of female literacy in the program area would provide valuable information for program planning such as setting targets and creating a delivery strategy. To date, many programs and interventions are targeted at individual determinants of child mortality because of operational constraints or the areas of expertise available within NGOs or government agencies. Partnerships between agencies should therefore be considered for enhanced program impact.

This report also highlights a number of broad geographic patterns that could inform policy and programs. There were significant differences in the distribution of child mortality between Western, Central, and Eastern Africa. Generally, there are higher levels of child mortality in Central and Western Africa as compared to Eastern Africa. This pattern corresponds to elevated risk factors in Central and Western Africa such as lower availability of improved sanitation and higher malaria prevalence than in Eastern Africa. Central and Western regions also suffer from lower prevalence of protective factors such as literacy rates and access to health facilities, as compared to Eastern Africa. Although disaggregated analyses could provide more specific information, there may be reasons to consider cross-border and regional interventions along with country-level policies that reduce child mortality through specific interventions.

Finally, it should be noted that the spatial, seasonal distribution of malaria prevalence is strongly related to climate conditions. This is especially true where the disease is not effectively controlled (Gething et al. 2010; Grover-Kopec et al. 2006). Thus, considering malaria prevalence and mean annual temperature among the most relevant geospatial factors that affect child mortality suggests the potential importance of accounting for climate change in future analyses.

7. Summary and Conclusions

Sustainable Development Goal 3 will ensure healthy lives and promote well-being for all people at all ages. This goal aims to end preventable deaths of newborns and children under age 5 by 2030, with a benchmark for neonatal mortality of less than 12 per 1,000 live births and under-five mortality at least as low as 25 per 1,000 live births.

With this context in mind, this report provided insight into the drivers of child mortality in the highly heterogeneous landscape of sub-Saharan Africa, a region that carries about half of the burden of the world's under-five deaths. To do so, this report explored how the importance of policy-relevant risk and protective factors and their interactions affect child mortality in sub-Saharan Africa. This report built on the work completed in SAR 12 by combining the effects of spatial-environmental factors with socio-economic and demographic determinants of child mortality. The goal was a better understanding of the drivers of spatial heterogeneity in child mortality and identification of areas for policy interventions.

Data for 27 sub-Saharan Africa countries aggregated at DHS sub-national area scale were assembled to observe the spatial distribution of child mortality across the 27 countries and Western, Central, and Eastern Africa regions. An autoregressive spatial model was applied to account for the spatial autocorrelation present in the data and to test the association between a number of demographic, socio-economic, and spatial-environmental factors on child mortality.

The study detected strong differences in child mortality across Western, Central, and Eastern Africa regions. The study found that socio-demographic factors (birth interval, stunting, access to health facilities, and literacy), along with geospatial factors (prevalence of *Plasmodium falciparum* malaria, ethnic group variety, mean annual temperature, intensity of lights at night, and economic activity) explain a considerable portion of the variance in child mortality across 255 DHS sub-national areas in the 27 countries.

The role of interaction terms explored in this study suggests that literacy rate is a strong driver of women's reproductive and health behavior, influences access to health facilities, and has an overarching place as a determinant of child mortality. The results highlight the importance of accounting for spatial heterogeneity in addressing child mortality in multiple countries in sub-Saharan Africa, designing policy interventions that reduce inequalities, and exploring some of the observed trends at finer geographic scales.

References

Adams, A.M. 1994. "Seasonal Variations in Nutritional Risk among Children in Central Mali." *Ecology of Food and Nutrition* 33(1-2):93-106.

Alegana, V.A., P.M. Atkinson, C. Pezzulo, A. Sorichetta, D. Weiss, T. Bird, E. Erbach-Schoenberg, and A.J. Tatem. 2015. "Fine Resolution Mapping of Population Age-Structures for Health and Development Applications." *Journal of The Royal Society Interface* 12(105).

Anand, A., and N. Roy. 2016. "Transitioning Towards Sustainable Development Goals: The Role of Household Environment in Influencing Child Health in Sub-Saharan Africa and South Asia Using Recent Demographic Health Surveys." *Frontiers in Public Health* 4.

Arku, R.E., J.E. Bennett, M.C. Castro, K. Agyeman-Duah, S.E. Mintah, J.H. Ware, P. Nyarko, J.D. Spengler, S. Agyei-Mensah, and M. Ezzati. 2016. "Geographical Inequalities and Social and Environmental Risk Factors for under-Five Mortality in Ghana in 2000 and 2010: Bayesian Spatial Analysis of Census Data." *PLoS Med* 13(6):e1002038.

Balk, D., T. Pullum, A. Storeygard, F. Greenwell, and M. Neuman. 2003. *Spatial Analysis of Childhood Mortality in West Africa*. Calverton, Maryland, USA: ORC Macro and Center for International Earth Science Information Network (CIESIN), Columbia University.

Balk, D., T. Pullum, A. Storeygard, F. Greenwell, and M. Neuman. 2004. "A Spatial Analysis of Childhood Mortality in West Africa." *Population, Space and Place* 10(3):175-216.

Balk, D., A. Storeygard, M. Levy, J. Gaskell, M. Sharma, and R. Flor. 2005. "Child Hunger in the Developing World: An Analysis of Environmental and Social Correlates." *Food Policy* 30(5–6):584-611.

Bandyopadhyay, S., S. Kanji, and L. Wang. 2012. "The Impact of Rainfall and Temperature Variation on Diarrheal Prevalence in Sub-Saharan Africa." *Applied Geography* 33:63-72.

Banerjee, S., B.P. Carlin, and A.E. Gelfand. 2015. *Hierarchical Modelling and Analysis for Spatial Data (2nd Edition).* CRC Press, Boca Raton, FL.

Besag, J., J. York, and A. Mollié. 1991. "Bayesian Image Restoration, with Two Applications in Spatial Statistics." *Annals of the Institute of Statistical Mathematics* 43(1):1-20.

Bhatt, S., D.J. Weiss, E. Cameron, D. Bisanzio, B. Mappin, U. Dalrymple, K.E. Battle, C.L. Moyes, A. Henry, P.A. Eckhoff, E.A. Wenger, O. Briet, M.A. Penny, T.A. Smith, A. Bennett, J. Yukich, T.P. Eisele, J.T. Griffin, C.A. Fergus, M. Lynch, F. Lindgren, J.M. Cohen, C.L.J. Murray, D.L. Smith, S.I. Hay, R.E. Cibulskis, and P.W. Gething. 2015. "The Effect of Malaria Control on Plasmodium Falciparum in Africa between 2000 and 2015." *Nature* 526(7572):207-211.

Bhutta, Z.A., T. Ahmed, R.E. Black, S. Cousens, K. Dewey, E. Giugliani, B.A. Haider, B. Kirkwood, S.S. Morris, H.P.S. Sachdev, and M. Shekar. 2008. "What Works? Interventions for Maternal and Child Undernutrition and Survival." *The Lancet* 371(9610):417-440.

Black, R.E., L.H. Allen, Z.A. Bhutta, L.E. Caulfield, M. de Onis, M. Ezzati, C. Mathers, and J. Rivera. 2008. "Maternal and Child Undernutrition: Global and Regional Exposures and Health Consequences." *The Lancet* 371(9608):243-260.

Black, R.E., C.G. Victora, S.P. Walker, Z.A. Bhutta, P. Christian, M. de Onis, M. Ezzati, S. Grantham-McGregor, J. Katz, R. Martorell, and R. Uauy. 2011. "Maternal and Child Undernutrition and Overweight in Low-Income and Middle-Income Countries." *The Lancet* 382(9890):427-451.

Bontemps, S., P. Defourny, E. van Bogaert, V. Kalogirou, and O. Arino. 2011. *Globcover 2009: Products Description and Validation Report.* http://Due.Esrin.Esa.Int/Files/Globcover2009_Validation_Report_2.2.Pdf.

Brady, O.J., N. Golding, D.M. Pigott, M.U.G. Kraemer, J.P. Messina, R.C. Reiner Jr, T.W. Scott, D.L. Smith, P.W. Gething, and S.I. Hay. 2014. "Global Temperature Constraints on Aedes Aegypti and Ae. Albopictus Persistence and Competence for Dengue Virus Transmission." *Parasites & Vectors* 7(1):1-17.

Brockerhoff, M., and P. Hewett. 2000. "Inequality of Child Mortality among Ethnic Groups in Sub-Saharan Africa." *Bulletin of the World Health Organization* 78:30-41.

Bruk, S.I., and V.S. Apenčenko. 1964. *Atlas Narodov Mira. Glavnoe Upravlenie Geodezii I Kartografii Gosudarstvennogo Geologičeskogo Komiteta Sssr.* Institut ėtnografii im. NN Miklucho-Maklaja Akademii nauk SSSR.

Bryce, J., C. Boschi-Pinto, K. Shibuya, and R.E. Black. 2005. "Who Estimates of the Causes of Death in Children." *The Lancet* 365(9465):1147-1152.

Burgert-Brucker, C.R., J. Yourkavitch, S. Assaf, and S. Delgado. 2015. *Geographic Variation in Key Indicators of Maternal and Child Health across 27 Countries in Sub-Saharan Africa*. DHS Spatial Analysis Reports No. 12. Rockville, Maryland, USA: ICF International. Available at http://dhsprogram.com/pubs/pdf/SAR12/SAR12.pdf.

Cai, X., T. Wardlaw, and D.W. Brown. 2012. "Global Trends in Exclusive Breastfeeding." *International Breastfeeding Journal* 7(1):1-5.

Cairns, M., A. Roca-Feltrer, T. Garske, A.L. Wilson, D. Diallo, P.J. Milligan, A.C. Ghani, and B.M. Greenwood. 2012. "Estimating the Potential Public Health Impact of Seasonal Malaria Chemoprevention in African Children." *Nat Commun* 3:881.

Caldwell, J.C. 1979. "Education as a Factor in Mortality Decline an Examination of Nigerian Data." *Population Studies* 33(3):395-413.

Cleland, J., S. Bernstein, A. Ezeh, A. Faundes, A. Glasier, and J. Innis. 2006. "Family Planning: The Unfinished Agenda." *The Lancet* 368(9549):1810-1827.

Cleland, J., A. Conde-Agudelo, H. Peterson, J. Ross, and A. Tsui. 2012. "Contraception and Health." *The Lancet* 380(9837):149-156.

Clements, A.C.A., H.L. Reid, G.C. Kelly, and S.I. Hay. 2013. "Further Shrinking the Malaria Map: How Can Geospatial Science Help to Achieve Malaria Elimination?" *The Lancet Infectious Diseases* 13(8):709-718.

Cleveland, W.S., and S.J. Devlin. 1988. "Locally Weighted Regression: An Approach to Regression Analysis by Local Fitting." *Journal of the American Statistical Association* 83(403):596-610.

Colson, K.E., L. Dwyer-Lindgren, T. Achoki, N. Fullman, M. Schneider, P. Mulenga, P. Hangoma, M. Ng, F. Masiye, and E. Gakidou. 2015. "Benchmarking Health System Performance across Districts in Zambia: A Systematic Analysis of Levels and Trends in Key Maternal and Child Health Interventions from 1990 to 2010." *BMC Medicine* 13(1):1-14.

Curtis, S.L., and M. Hossein. 1998. *The Effect of Aridity Zone on Child Nutritional Status*. West Africa Spatial Analysis Prototype Exploratory Analysis. Calverton, Maryland: Macro International Inc.

de Sherbinin, A. 2011. "The Biophysical and Geographical Correlates of Child Malnutrition in Africa." *Population, Space and Place* 17(1):27-46.

Defo, B.K. 1994. "Determinants of Infant and Early Childhood Mortality in Cameroon: The Role of Socioeconomic Factors, Housing Characteristics, and Immunization Status." *Social Biology* 41(3-4):181-211.

DHS Spatial Interpolation Working Group. 2014. *Spatial Interpolation with Demographic and Health Survey Data: Key Considerations*. DHS Spatial Analysis Reports No. 9. Rockville, Maryland, USA: ICF International.

Dilley, M., R.S. Chen, U. Deichmann, A.L. Lerner-Lam, M. Arnold, J. Agwe, P. Buys, O. Kjekstad, B. Lyon, and G. Yetman. 2005. *Natural Disaster Hotspots: A Global Risk Analysis. Disaster Risk Management Series No. 5*. Washington, D.C.: The World Bank. http://go.worldbank.org/PT8XJZW3K0.

Dos Santos, S., and S. Henry. 2008. "Rainfall Variation as a Factor in Child Survival in Rural Burkina Faso: The Benefit of an Event-History Analysis." *Population, Space and Place* 14(1):1-20.

Dye, C. 2008. "Health and Urban Living." *Science* 319(5864):766-769.

Eisele, T.P., D.A. Larsen, N. Walker, R.E. Cibulskis, J.O. Yukich, C.M. Zikusooka, and R.W. Steketee. 2012. "Estimates of Child Deaths Prevented from Malaria Prevention Scale-up in Africa 2001-2010." *Malaria Journal* 11:93-93.

Elvidge, C.D., K.E. Baugh, M. Zhizhi, and F.-C. Hsu. 2013. "Why Viirs Data Are Superior to Dmsp for Mapping Nighttime Lights." *Proceedings of the Asia-Pacific Advanced Network* 35:62-69.

Entwisle, B., R.R. Rindfuss, S.J. Walsh, T.P. Evans, and S.R. Curran. 1997. "Geographic Information Systems, Spatial Network Analysis, and Contraceptive Choice." *Demography* 34(2):171-187.

FAO. 2011. *Wealth Index Mapping in the Horn of Africa, by Rogers, D.J., Wint, G.R.W., Alexander, N., Pozzi, F. & Robinson, T.P.*: Animal Production and Health Working Paper. No. 4. FAO, Rome.

Findley, S., D. Balk, M. Barlow, and N. Sogoba. 2002. *Putting Climate in the Service of Public Health in Mali*. In *Annual Meeting of the Population Association of America*. Atlanta, Georgia.

Fink, G., I. Günther, and K. Hill. 2011. "The Effect of Water and Sanitation on Child Health: Evidence from the Demographic and Health Surveys 1986–2007." *International Journal of Epidemiology* 40(5):1196-1204.

Gakidou, E., K. Cowling, R. Lozano, and C.J.L. Murray. 2010. "Increased Educational Attainment and Its Effect on Child Mortality in 175 Countries between 1970 and 2009: A Systematic Analysis." *Lancet* 376.

Galan, D.I., S.-S. Kim, and J.P. Graham. 2013. "Exploring Changes in Open Defecation Prevalence in Sub-Saharan Africa Based on National Level Indices." *BMC Public Health* 13(1):1-12.

Gallup, J.L., and J.D. Sachs. 2001. "The Economic Burden of Malaria." *Am J Trop Med Hyg* 64.

Gemperli, A., P. Vounatsou, I. Kleinschmidt, M. Bagayoko, C. Lengeler, and T. Smith. 2004. "Spatial Patterns of Infant Mortality in Mali: The Effect of Malaria Endemicity." *American Journal of Epidemiology* 159(1):64-72.

Gething, P.W., D.L. Smith, A.P. Patil, A.J. Tatem, R.W. Snow, and S.I. Hay. 2010. "Climate Change and the Global Malaria Recession." *Nature* 465(7296):342-345.

Gething, P.W., T.P. Van Boeckel, D.L. Smith, C.A. Guerra, A.P. Patil, R.W. Snow, and S.I. Hay. 2011. "Modelling the Global Constraints of Temperature on Transmission of Plasmodium Falciparum and P. Vivax." *Parasites & Vectors* 4(1):1-11.

Gribble, J.N., N.J. Murray, and E.P. Menotti. 2009. "Reconsidering Childhood Undernutrition: Can Birth Spacing Make a Difference? An Analysis of the 2002–2003 El Salvador National Family Health Survey." *Maternal & Child Nutrition* 5(1):49-63.

Grover-Kopec, E.K., M.B. Blumenthal, P. Ceccato, T. Dinku, J.A. Omumbo, and S.J. Connor. 2006. "Web-Based Climate Information Resources for Malaria Control in Africa." *Malaria Journal* 5(1):1-9.

Gupta, H.S., and A. Baghel. 1999. "Infant Mortality in the Indian Slums: Case Studies of Calcutta Metropolis and Raipur City." *International Journal of Population Geography* 5(5):353-366.

Hay, S.I., C.A. Guerra, A.J. Tatem, P.M. Atkinson, and R.W. Snow. 2005. "Tropical Infectious Diseases: Urbanization, Malaria Transmission and Disease Burden in Africa." *Nat Rev Micro* 3(1):81-90.

Hay, S.I., A.J. Tatem, A.J. Graham, S.J. Goetz, and D.J. Rogers. 2006. "Global Environmental Data for Mapping Infectious Disease Distribution." *Advances in parasitology* 62:37-77.

Hijmans, R.J., S.E. Cameron, J.L. Parra, P.G. Jones, and A. Jarvis. 2005. "Very High Resolution Interpolated Climate Surfaces for Global Land Areas." *International Journal of Climatology* 25(15):1965-1978.

Hobcraft, J. 1993. "Women's Education, Child Welfare and Child Survival: A Review of the Evidence." *Health Transition Review* 3(2):159-175.

ICF International. 2008-2015a. *Demographic and Health Surveys (Various) [Datasets]. Rockville, Maryland: Icf International [Distributor]*.

ICF International. 2008-2015b. *Spatial Data Repository, the Dhs Program (Various) [Datasets]. Rockville, Maryland: Icf International [Distributor]. Available from Spatialdata.Dhsprogram.Com.*

Jankowska, M.M., D. Lopez-Carr, C. Funk, G.J. Husak, and Z.A. Chafe. 2012. "Climate Change and Human Health: Spatial Modeling of Water Availability, Malnutrition, and Livelihoods in Mali, Africa." *Applied Geography* 33:4-15.

Jones, G., R.W. Steketee, R.E. Black, Z.A. Bhutta, and S.S. Morris. 2003. "How Many Child Deaths Can We Prevent This Year?" *The Lancet* 362(9377):65-71.

Karra, M., G. Fink, and D. Canning. 2016. "Facility Distance and Child Mortality: A Multi-Country Study of Health Facility Access, Service Utilization, and Child Health Outcomes." *International Journal of Epidemiology*.

Kraemer, M.U.G., S.I. Hay, D.M. Pigott, D.L. Smith, G.R.W. Wint, and N. Golding. 2016. "Progress and Challenges in Infectious Disease Cartography." *Trends in Parasitology* 32(1):19-29.

Lavy, V., J. Strauss, D. Thomas, and P. de Vreyer. 1996. "Quality of Health Care, Survival and Health Outcomes in Ghana." *Journal of Health Economics* 15(3):333-357.

Lee, D. 2013. "Carbayes: An R Package for Bayesian Spatial Modeling with Conditional Autoregressive Priors." *2013* 55(13):24.

Leroux, B.G., X. Lei, and N. Breslow. 2000. "Estimation of Disease Rates in Small Areas: A New Mixed Model for Spatial Dependence." In *Statistical Models in Epidemiology, the Environment, and Clinical Trials*, edited by M. Elizabeth Halloran and Donald Berry, 179-191. New York, NY: Springer New York.

Linard, C., M. Gilbert, R.W. Snow, A.M. Noor, and A.J. Tatem. 2012. "Population Distribution, Settlement Patterns and Accessibility across Africa in 2010." *PLoS ONE* 7(2):e31743.

Liu, L., H.L. Johnson, S. Cousens, J. Perin, S. Scott, J.E. Lawn, I. Rudan, H. Campbell, R. Cibulskis, M. Li, C. Mathers, and R.E. Black. 2012. "Global, Regional, and National Causes of Child Mortality:

An Updated Systematic Analysis for 2010 with Time Trends since 2000." *The Lancet* 379(9832):2151-2161.

Luna, J.W., T. Monga, and L. Morgan. 2014. *Equity Matters - Lessons from Mchip and Cshgp in Measuring and Improving Equity*. Maternal Child Health Integrated Program.

Lyon, B., and A.G. Barnston. 2005. "Enso and the Spatial Extent of Interannual Precipitation Extremes in Tropical Land Areas." *Journal of Climate* 18:5095-5109.

Mabhaudhi, T., T. Chibarabada, and A. Modi. 2016. "Water-Food-Nutrition-Health Nexus: Linking Water to Improving Food, Nutrition and Health in Sub-Saharan Africa." *International Journal of Environmental Research and Public Health* 13(1):107.

Marmot, M. 2005. "Social Determinants of Health Inequalities." *The Lancet* 365(9464):1099-1104.

Metcalf, C.J.E., A. Tatem, O.N. Bjornstad, J. Lessler, K. O'Reilly, S. Takahashi, F. Cutts, and B.T. Grenfell. 2015. "Transport Networks and Inequities in Vaccination: Remoteness Shapes Measles Vaccine Coverage and Prospects for Elimination across Africa." *Epidemiology and Infection* 143(7):1457-1466.

Middleton, N., and D. Thomas. 1997. *World Atlas of Desertification*. Vol. No. Ed. 2. Arnold, Hodder Headline, PLC.

Mosites, E., S.M. Thumbi, E. Otiang, T.F. McElwain, M. Njenga, P.M. Rabinowitz, A. Rowhani-Rahbar, M.L. Neuhouser, S. May, G.H. Palmer, and J.L. Walson. 2016. "Relations between Household Livestock Ownership, Livestock Disease, and Young Child Growth." *The Journal of Nutrition* 146(5):1118-1124.

Mosites, E.M., P.M. Rabinowitz, S.M. Thumbi, J.M. Montgomery, G.H. Palmer, S. May, A. Rowhani-Rahbar, M.L. Neuhouser, and J.L. Walson. 2015. "The Relationship between Livestock Ownership and Child Stunting in Three Countries in Eastern Africa Using National Survey Data." *PLoS ONE* 10(9):e0136686.

Mosley, W.H., and L.C. Chen. 1984. "An Analytical Framework for the Study of Child Survival in Developing Countries." *Population and Development Review* 10:25-45.

Mutuku, F., M. Bayoh, A. Hightower, J. Vulule, J. Gimnig, J. Mueke, F. Amimo, and E. Walker. 2009. "A Supervised Land Cover Classification of a Western Kenya Lowland Endemic for Human Malaria: Associations of Land Cover with Larval Anopheles Habitats." *International Journal of Health Geographics* 8(1):1-13.

Nordhaus, W., D., and X. Chen. 2009. *Geography: Graphics and Economics*. In *The B.E. Journal of Economic Analysis & Policy*.

Ombok, M., K. Adazu, F. Odhiambo, N. Bayoh, R. Kiriinya, L. Slutsker, M.J. Hamel, J. Williamson, A. Hightower, K.F. Laserson, and D.R. Feikin. 2010. "Geospatial Distribution and Determinants of Child Mortality in Rural Western Kenya 2002–2005." *Tropical Medicine & International Health* 15(4):423-433.

Pesaresi, M., G. Huadong, X. Blaes, D. Ehrlich, S. Ferri, L. Gueguen, M. Halkia, M. Kauffmann, T. Kemper, L. Lu, M.A. Marin-Herrera, G.K. Ouzounis, M. Scavazzon, P. Soille, V. Syrris, and L. Zanchetta. 2013. "A Global Human Settlement Layer from Optical Hr/Vhr Rs Data: Concept and First Results." *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, Vol. 6, No. 5, October 2013.

Platas, M.R. 2010. *Africa's Health Tragedy? Ethnic Diversity and Health Outcomes*. In *Stanford University, Department of Political Science. Prepared for delivery at the Winter 2010. Working Group on African Political Economy*, December 17-18, 2010.

PMNCH. 2015. *The Partnership for Maternal, Newborn & Child Health in Support of Every Woman Every Child. Strategic Plan 2016-2020.* Geneva, Switzerland.

Pozzi, F., T. Robinson, and A. Nelson. 2010. *Accessibility Mapping and Rural Poverty in the Horn of Africa*. IGAD LPI Working Paper No. 02-10; PPLPI Working Paper No. 47. IGAD Livestock Policy Initiative.

Pullum, T.W., B. Schoumaker, S. Becker, and S.E.K. Bradley. 2013. *An Assessment of Dhs Estimates of Fertility and under-Five Mortality*. In *International Population Conference of the International Union for the Scientific Study of Population (IUSSP), Session 132: Data quality in demographic surveys, August 28.* .

Qi, Q., C.A. Guerra, C.L. Moyes, I.A.F. Elyazar, P.W. Gething, S.I. Hay, and A.J. Tatem. 2012. "The Effects of Urbanization on Global Plasmodium Vivax Malaria Transmission." *Malaria Journal* 11(1):1-11.

Randolph, T.F., E. Schelling, D. Grace, C.F. Nicholson, J.L. Leroy, D.C. Cole, M.W. Demment, A. Omore, J. Zinsstag, and M. Ruel. 2007. "Invited Review: Role of Livestock in Human Nutrition and Health for Poverty Reduction in Developing Countries123." *Journal of Animal Science* 85(11):2788-2800.

Roberts, D.A., M. Ng, G. Ikilezi, A. Gasasira, L. Dwyer-Lindgren, N. Fullman, T. Nalugwa, M. Kamya, and E. Gakidou. 2015. "Benchmarking Health System Performance across Regions in Uganda: A Systematic Analysis of Levels and Trends in Key Maternal and Child Health Interventions, 1990–2011." *BMC Medicine* 13(1):1-16.

Robinson, T.P., G.R.W. Wint, G. Conchedda, T.P. Van Boeckel, V. Ercoli, E. Palamara, G. Cinardi, L. D'Aietti, S.I. Hay, and M. Gilbert. 2014. "Mapping the Global Distribution of Livestock." *PLoS ONE* 9(5):e96084.

Root, G. 1997. "Population Density and Spatial Differentials in Child Mortality in Zimbabwe." *Social Science & Medicine* 44(3):413-421.

Rosero-Bixby, L. 2004. "Spatial Access to Health Care in Costa Rica and Its Equity: A Gis-Based Study." *Social Science & Medicine* 58(7):1271-1284.

Rutherford, M.E., K. Mulholland, and P.C. Hill. 2010. "How Access to Health Care Relates to under-Five Mortality in Sub-Saharan Africa: Systematic Review." *Tropical Medicine & International Health* 15(5):508-519.

Rutstein, S., and G. Rojas. 2006. *Guide to Dhs Statistics*. Demographic and Health Surveys- ORC Macro. Calverton, Maryland

Rutstein, S.O. 2000. "Factors Associated with Trends in Infant and Child Mortality in Developing Countries During the 1990s." *Bulletin of the World Health Organization* 78:1256-1270.

Rutstein, S.O. 2008. *Further Evidence of the Effects of Preceding Birth Intervals on Neonatal, Infant, and under-Five-Years Mortality and Nutritional Status in Developing Countries: Evidence from the Demographic and Health Surveys*. DHS Working Papers No. 41. Calverton, Maryland, USA: Macro International. Available at http://dhsprogram.com/pubs/pdf/WP41/WP41.pdf.

Rutstein, S.O. 2011. *Trends in Birth Spacing*. DHS Comparative Reports No. 28. . Calverton, Maryland, USA: ICF Macro.

Sachs, J., and P. Malaney. 2002. "The Economic and Social Burden of Malaria." *Nature* 415(6872):680-685.

Sauerborn, R., A. Adams, and M. Hien. 1996. "Household Strategies to Cope with the Economic Costs of Illness." *Social Science & Medicine* 43(3):291-301.

Schneider, A., M.A. Friedl, and D. Potere. 2009. “A New Map of Global Urban Extent from Modis Satellite Data.” *Environmental Research Letters* 4(4):044003.

Schneider, A., M.A. Friedl, and D. Potere. 2010. “Mapping Global Urban Areas Using Modis 500-M Data: New Methods and Datasets Based on ‘Urban Ecoregions’.” *Rem Sens Env* 114.

Sedda, L., Q. Qi, and A.J. Tatem. 2015. “A Geostatistical Analysis of the Association between Armed Conflicts and Plasmodium Falciparum Malaria in Africa, 1997–2010.” *Malaria Journal* 14:500.

Sedda, L., A.J. Tatem, D.W. Morley, P.M. Atkinson, N.A. Wardrop, C. Pezzulo, A. Sorichetta, J. Kuleszo, and D.J. Rogers. 2015. “Poverty, Health and Satellite-Derived Vegetation Indices: Their Inter-Spatial Relationship in West Africa.” *International Health* 7(2):99-106.

Sen, A. 1998. “Mortality as an Indicator of Economic Success and Failure.” *The Economic Journal* 108(446):1-25.

Snow, R.W., M.H. Craig, U. Deichmann, and D. le Sueur. 1999. “A Preliminary Continental Risk Map for Malaria Mortality among African Children.” *Parasitology Today* 15(3):99-104.

Stern, H., and N. Cressie. 1999. “Inference for Extremes in Disease Mapping.” In *Disease Mapping and Risk Assessment for Public Health*, edited by A. B. Lawson, Biggeri, A., Bohning, D., Lesaffre, E., Viel, J. F., Bertollini, R., 63-84. John Wiley & Sons.

Stevens, F.R., A.E. Gaughan, C. Linard, and A.J. Tatem. 2015. “Disaggregating Census Data for Population Mapping Using Random Forests with Remotely-Sensed and Ancillary Data.” *PLoS ONE* 10(2):e0107042.

Tatem, A., S. Adamo, N. Bharti, C. Burgert, M. Castro, A. Dorelien, G. Fink, C. Linard, M. John, L. Montana, M. Montgomery, A. Nelson, A. Noor, D. Pindolia, G. Yetman, and D. Balk. 2012. “Mapping Populations at Risk: Improving Spatial Demographic Data for Infectious Disease Modeling and Metric Derivation.” *Population Health Metrics* 10(1):1-14.

Tatem, A., P. Gething, C. Pezzulo, D. Weiss, and S. Bhatt. 2014. *Development of High-Resolution Gridded Poverty Surfaces*. Development of High-Resolution Gridded Poverty Surfaces Bill and Melinda Gates Foundation Contract. Worldpop Project, University of Southampton; University of Oxford.

Tatem, A.J., P.W. Gething, D.L. Smith, and S.I. Hay. 2013. “Urbanization and the Global Malaria Recession.” *Malaria Journal* 12(1):1-11.

Thaddeus, S., and D. Maine. 1994. “Too Far to Walk: Maternal Mortality in Context.” *Social Science and Medicine* 38(8):1091- 1110.

Tottrup, C., B.P. Tersbol, W. Lindeboom, and D. Meyrowitsch. 2009. “Putting Child Mortality on a Map: Towards an Understanding of Inequity in Health.” *Tropical Medicine & International Health* 14(6):653-662.

Uchida, H., and A. Nelson. 2008. “Agglomeration Index: Towards a New Measure of Urban Concentration.” *Urbanization and development: Multidisciplinary perspectives*:41-60.

UNDP. 2014. *World Urbanization Prospects: The 2014 Revision.*: United Nations Department of Economic and Social Affairs Population Division (UNPD). CD-ROM Edition. http://esa.un.org/unpd/wup/CD-ROM/.

UNESCO UIS. 2015. *Adult and Youth Literacy*. UNESCO Institute for statistics. Available at: http://www.uis.unesco.org/literacy/Documents/fs32-2015-literacy.pdf.

UNFPA. 2007. *Potential Contributions to the Mdg Agenda from the Perspective of Icpd. Mdg 4: Reduce Child Mortality.*: United Nations Population Fund. Available at: http://www.unfpa.org.br/lacodm/arquivos/mdg4.pdf.

UNICEF. 2015. *Committing to Child Survival: A Promise Renewed*. UNICEF's Division of Data, Research, and Policy in collaboration with the Programme Division and the Secretariat for A Promise Renewed.

UNICEF. 2016. *Maternal Health - Delivery Care: Current Status and Progress*. Available at http://data.unicef.org/maternal-health/delivery-care.html.

UNICEF, WHO, and The World Bank. 2015. *Global Database on Child Growth and Malnutrition -Joint Child Malnutrition Estimates - Levels and Trends (2015 Edition).* Available at: http://www.who.int/nutgrowthdb/jme_brochure2015.pdf?ua=1.

UNICEF, WHO, World Bank, and UN-DESA Population Division. 2015. *Levels and Trends in Child Mortality 2015*. UNICEF, WHO, World Bank, UN-DESA Population Division.

United Nations. 2015. *The Millennium Development Goals Report*. United Nations.

United Nations. 2016. *Open Defecation*. Available at http://opendefecation.org/.

United Nations General Assembly. 2015. *Transforming Our World: The 2030 Agenda for Sustainable Development* A/RES/70/1. Resolution adopted by the General Assembly on 25 September 2015 - Available at: http://www.un.org/ga/search/view_doc.asp?symbol=A/RES/70/1.

United Nations Statistics Division. 2013. *United Nations Geographical Region Composition*. Available at http://unstats.un.org/unsd/methods/m49/m49regin.htm.

Vlahov, D., N. Freudenberg, F. Proietti, D. Ompad, A. Quinn, V. Nandi, and S. Galea. 2007. "Urban as a Determinant of Health." *J Urban Health* 84.

Wagstaff, A. 2002. "Poverty and Health Sector Inequalities." *Bulletin of the World Health Organization* 80:97-105.

Webb, P. 1998. *Isolating Hunger: Reaching People in Need Beyond the Mainstream*. Paper Prepared for the World Food Programme, Rome, Italy.

Weidmann, N.B., J.K. Rød, and L.-E. Cederman. 2010. "Representing Ethnic Groups in Space: A New Dataset." *Journal of Peace Research*.

Weiss, D.J., S. Bhatt, B. Mappin, T.P. Van Boeckel, D.L. Smith, S.I. Hay, and P.W. Gething. 2014. "Air Temperature Suitability for Plasmodium Falciparum Malaria Transmission in Africa 2000-2012: A High-Resolution Spatiotemporal Prediction." *Malaria Journal* 13(1):1-11.

WHO. 2009. *Global Health Risks: Mortality and Burden of Disease Attributable to Selected Major Risks*. Geneva: World Health Organization.

WHO. 2014. *Dpt3 Data by Who Region*. Global Health Observatory Data Repository. Available at: http://apps.who.int/gho/data/view.main.81200?lang=en.

WHO. 2015a. *Family Planning - Contraception Factsheet N°351* Available at http://who.int/mediacentre/factsheets/fs351/en/.

WHO. 2015b. *Measles Fact Sheet.* Available at http://www.who.int/mediacentre/factsheets/fs286/en/.

WHO. 2016a. *Children: Reducing Mortality - Fact Sheet*. Available at http://www.who.int/mediacentre/factsheets/fs178/en/.

WHO. 2016b. *Use of Improved Drinking Water Sources - Situation and Trends*. Global Health Observatory (GHO) data.

WHO Collaborative Study Team. 2000. "Effect of Breastfeeding on Infant and Child Mortality Due to Infectious Diseases in Less Developed Countries: A Pooled Analysis." *Lancet* 355:451-55.

Wollum, A., R. Burstein, N. Fullman, L. Dwyer-Lindgren, and E. Gakidou. 2015. "Benchmarking Health System Performance across States in Nigeria: A Systematic Analysis of Levels and Trends in Key Maternal and Child Health Interventions and Outcomes, 2000–2013." *BMC Medicine* 13(1):1-16.

Woods, R. 2003. "Urban-Rural Mortality Differentials: An Unresolved Debate." *Population and Development Review* 29(1):29-46.

Xie, M., N. Jean, M. Burke, D. Lobell, and S. Ermon. 2015. "Transfer Learning from Deep Features for Remote Sensing and Poverty Mapping." *arXiv preprint arXiv:1510.00098*.

Zomer, R.J., A. Trabucco, D.A. Bossio, and L.V. Verchot. 2008. "Climate Change Mitigation: A Spatial Analysis of Global Land Suitability for Clean Development Mechanism Afforestation and Reforestation." *Agriculture, Ecosystems & Environment* 126(1–2):67-80.

Appendix A

A.1 United Nations Geographical Region Composition for the 27 Sub-Saharan African Countries included in Our Analysis

Table A.1. List of Countries in analysis by United Nations Regions

WESTERN AFRICA REGION COUNTRIES (WAFR)	EASTERN AFRICA REGION COUNTRIES (EAFR)	CENTRAL AFRICA REGION COUNTRIES (CAFR)
Benin	Burundi	Cameroon
Burkina-Faso	Ethiopia	Congo
Côte d'Ivoire	Kenya	Democratic Republic of Congo (DRC)
Gambia	Malawi	Gabon
Ghana	Mozambique	
Guinea	Rwanda	
Liberia	Tanzania	
Mali	Uganda	
Niger	Zambia	
Nigeria	Zimbabwe	
Senegal		
Sierra Leone		
Togo		

Source: (United Nations Statistics Division 2013)

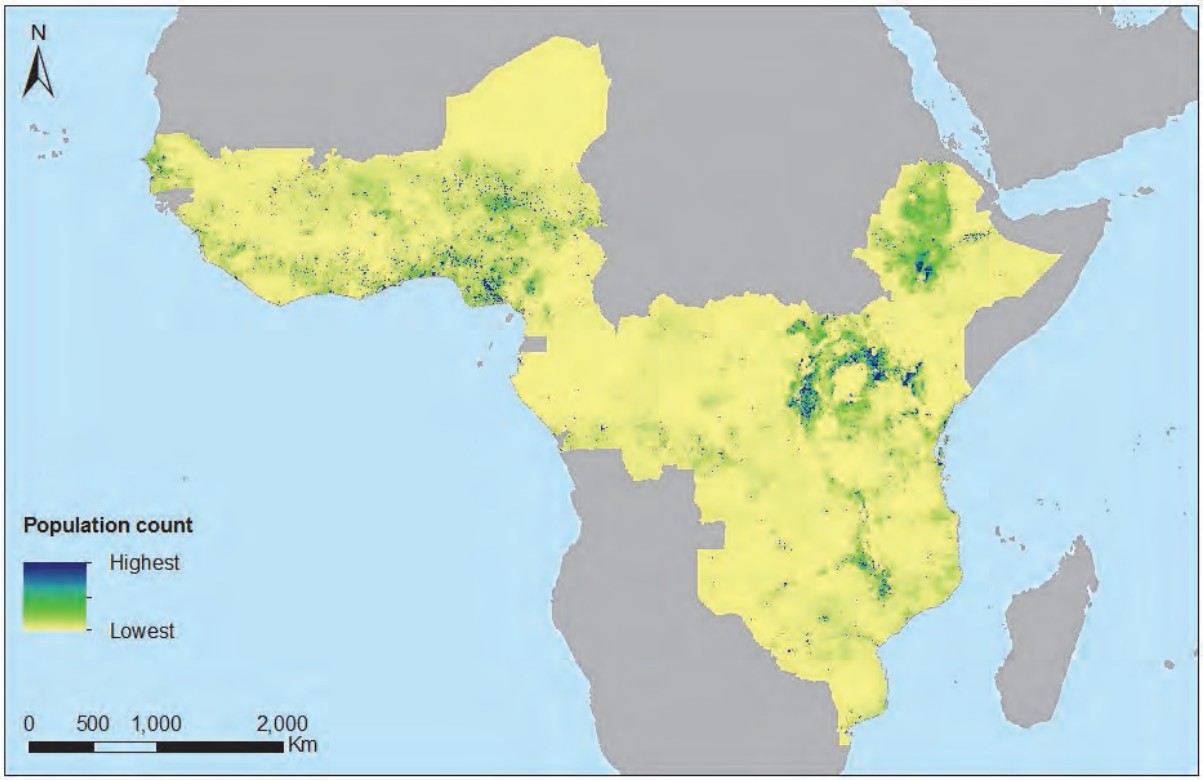
A.2. Geospatial Datasets

This section of the Appendix briefly describes how the geospatial datasets in this study were produced. The section contains the figures, referenced in Section 2.2.2, which show the spatial distribution, over the study area, of all spatial factors relevant to child mortality that were considered in the study.

Population distribution

WorldPop 2010 UN adjusted People Per Pixel raster datasets (www.worldpop.org) are produced using ancillary data to asymmetrically disaggregate administrative, unit-based population counts to a regular grid of fixed spatial resolution. National totals for 2010 were adjusted to match United Nations Population Division estimates (UNDP 2014).

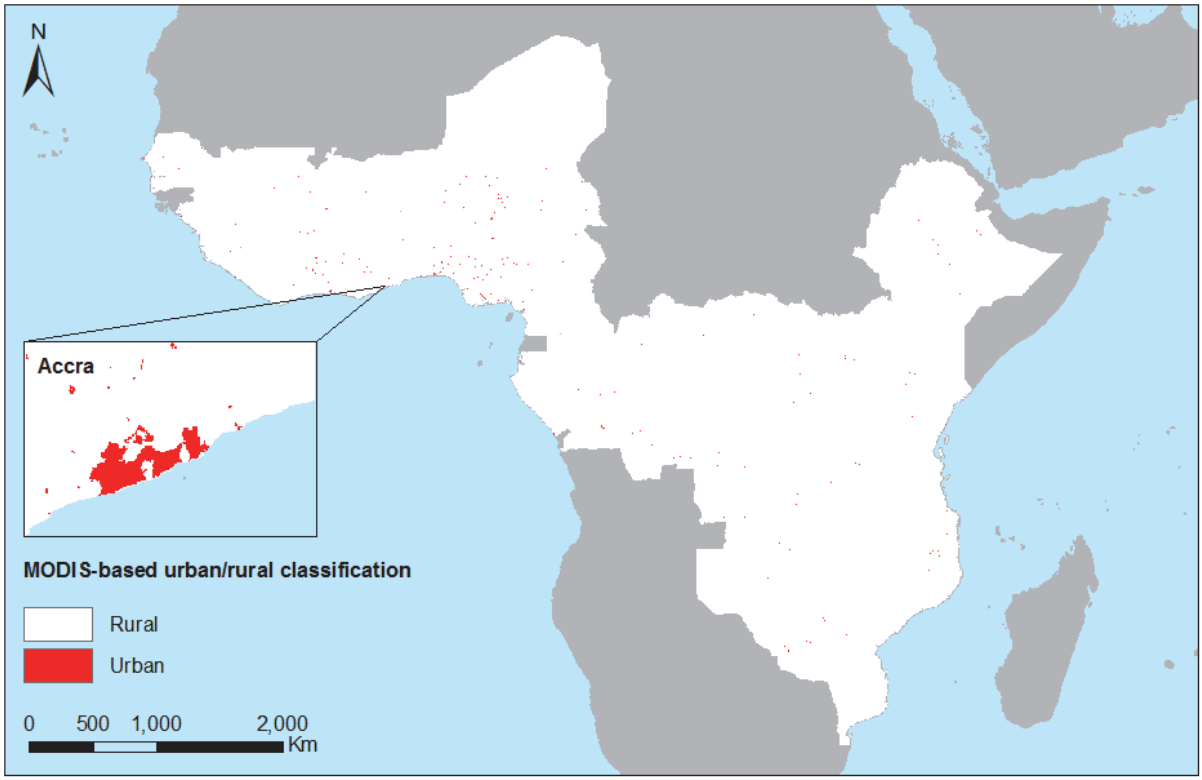
Figure A.1. Population distribution



Urbanicity (MODIS)

The "MODIS 500m Global Urban Extent" dataset (https://nelson.wisc.edu/sage/data-and-models/schneider.php) was produced with Collection 5 (C5) MODIS 500-m resolution data and a supervised decision tree classification algorithm to identify urban areas by using region-specific parameters (https://nelson.wisc.edu/sage/data-and-models/schneider-readme.php).

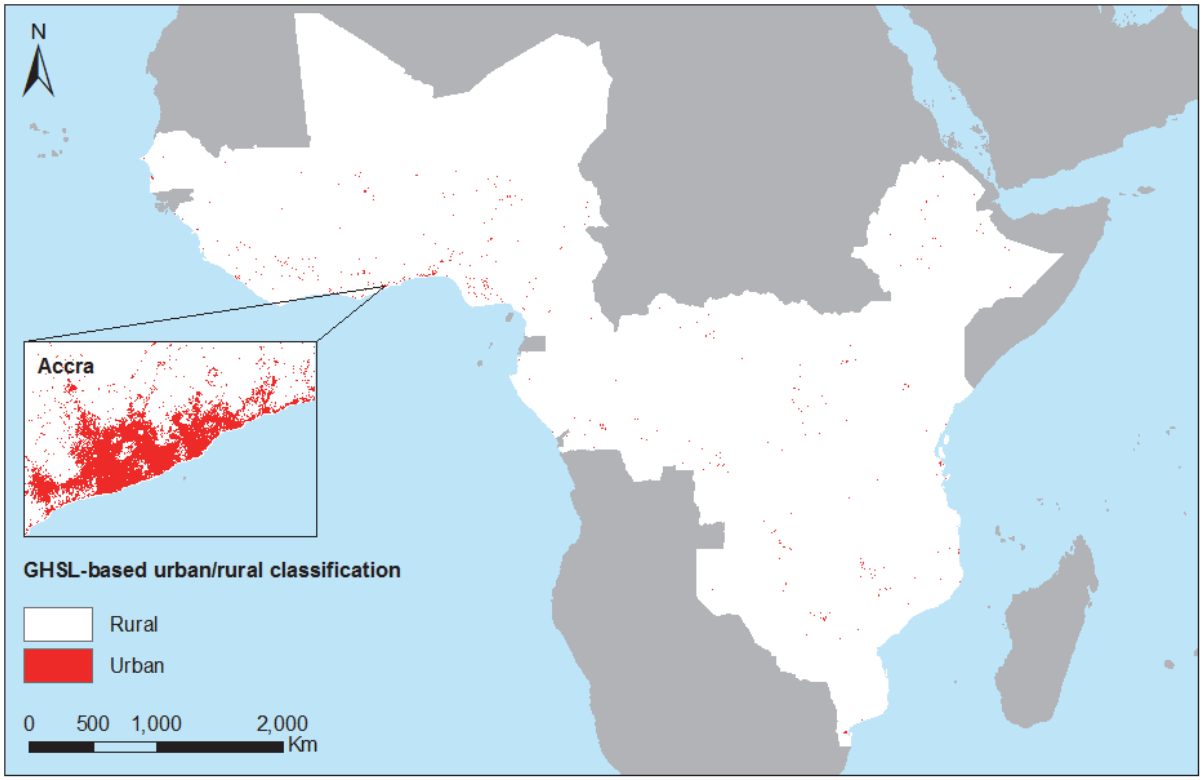
Figure A.2a. MODIS-based spatial distribution of built-up areas



Urbanicity (GHSL)

The "Global Human Settlement Layer" dataset (http://ghslsys.jrc.ec.europa.eu/) was produced by the European Union Joint Research Centre. This dataset used learning and classification techniques to process high and very high resolution images that identify areas that are densely built. Low-resolution thematic layers served as a reference (Pesaresi et al. 2013).

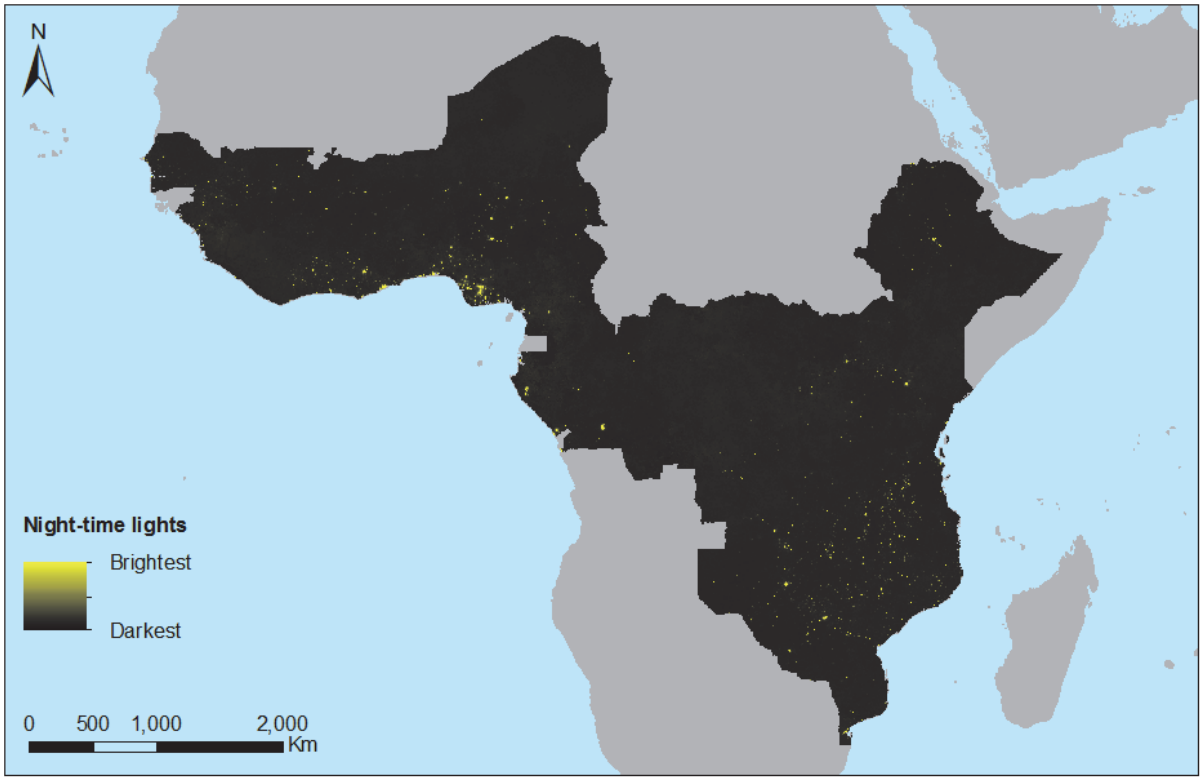
Figure A.2b. GHSL-based spatial distribution of built-up areas



Night-time lights

The NOAA Suomi National Polar-orbiting Partnership (NPP) composite night-time light dataset is produced by averaging monthly radiance data, acquired by the Visible Infrared Imaging Radiometer Suite (VIIRS) Day/Night Band (DNB). The data are filtered to exclude data affected by stray light, lightning, lunar illumination, and cloud-cover (http://ngdc.noaa.gov/eog/viirs/download_monthly.html).

Figure A.3. Spatial pattern of VIIRS-based night-time lights' intensity



Travel-times to major population centers

"Travel time to major cities" datasets (http://forobs.jrc.ec.europa.eu/products/gam/download.php) were produced by the European Union Joint Research Centre. The datasets used a cost-distance algorithm to compute the travel-time between two grid cells, expressed as a function of the cost required to travel across them (http://forobs.jrc.ec.europa.eu/products/gam/description.php).

Figure A.4a. Travel-time to major centers with more than 50,000 inhabitants

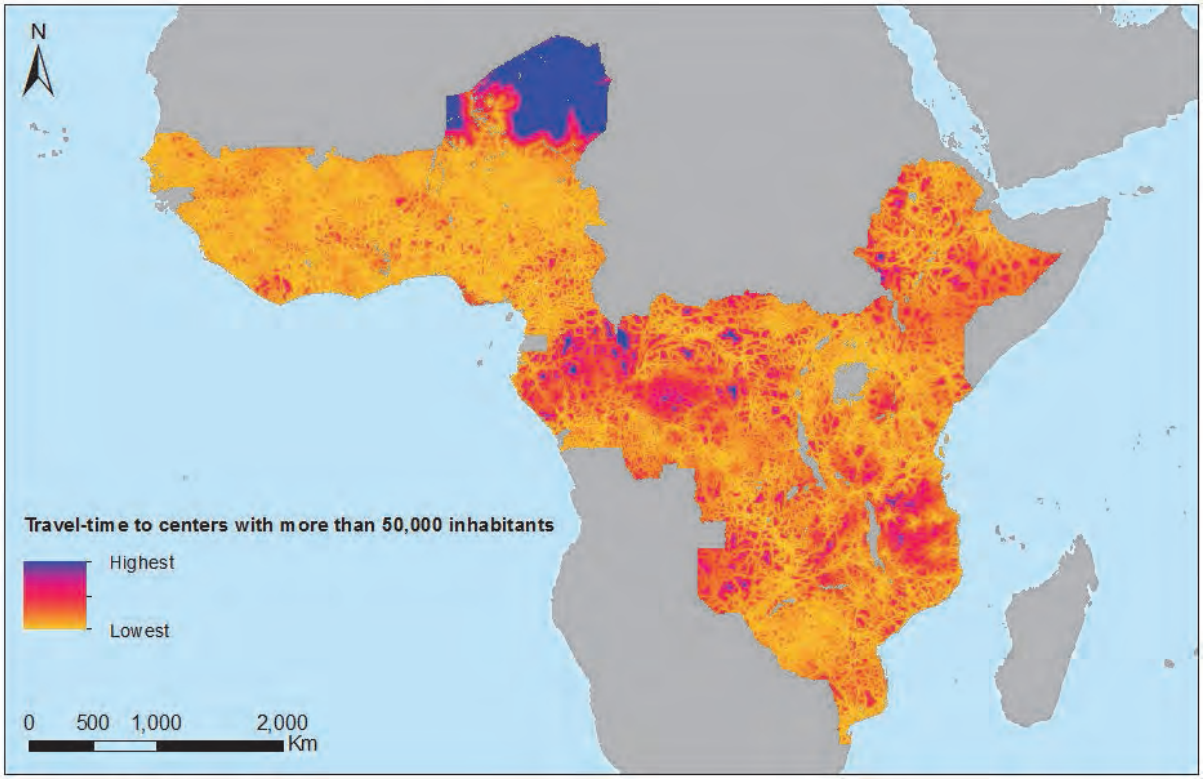


Figure A.4b. Travel-time to major centers with more than 100,000 inhabitants

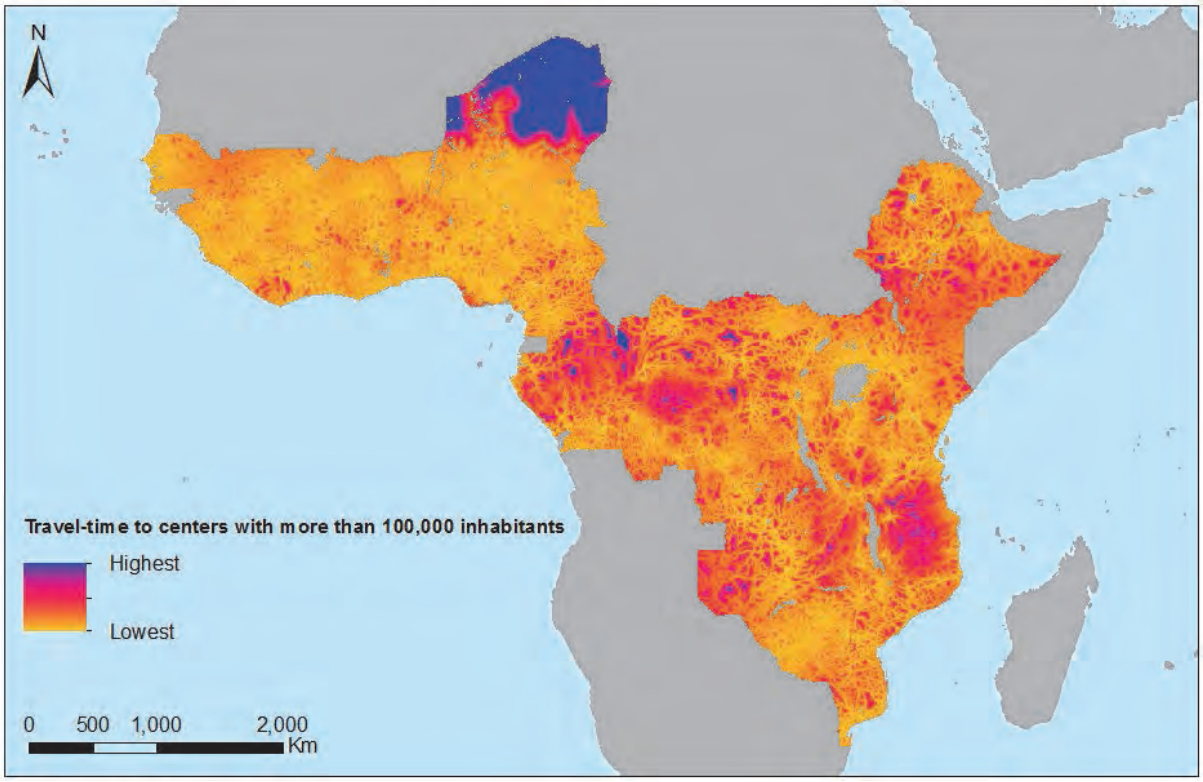
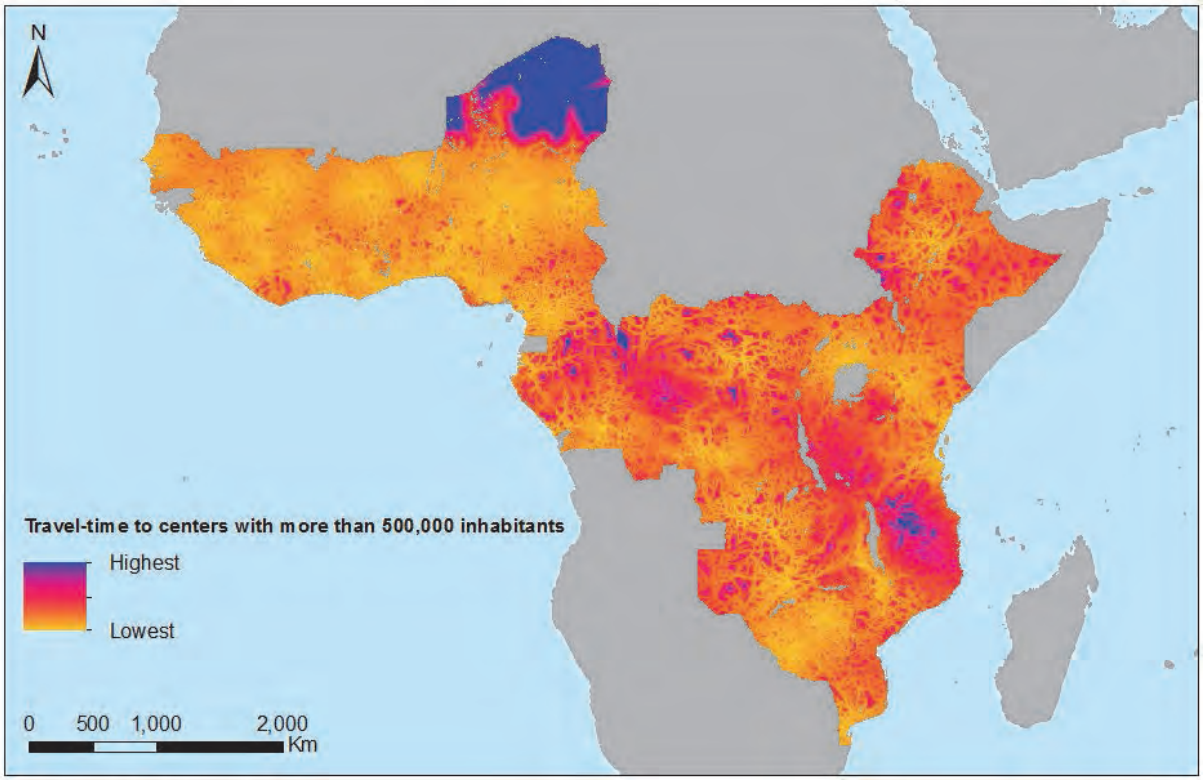


Figure A.4c. Travel-time to major centers with more than 500,000 inhabitants



Ethnicity

The "Geo-Referencing of Ethnic Groups (GREG)" dataset (http://www.icr.ethz.ch/data/other/greg) was created with spatial GIS techniques and data from the Soviet Narodov Mira atlas (Bruk and Apenčenko 1964) to represent group territories as vector polygons. For the purpose of this study, values for the Tanzanian islands of Zanzibar and Pemba (two of the DHS sub-national areas within the study area that are not available in the GREG dataset) were set to the values of the closest group polygon in the mainland.

Figure A.5a. Spatial distribution of dominant ethnic groups

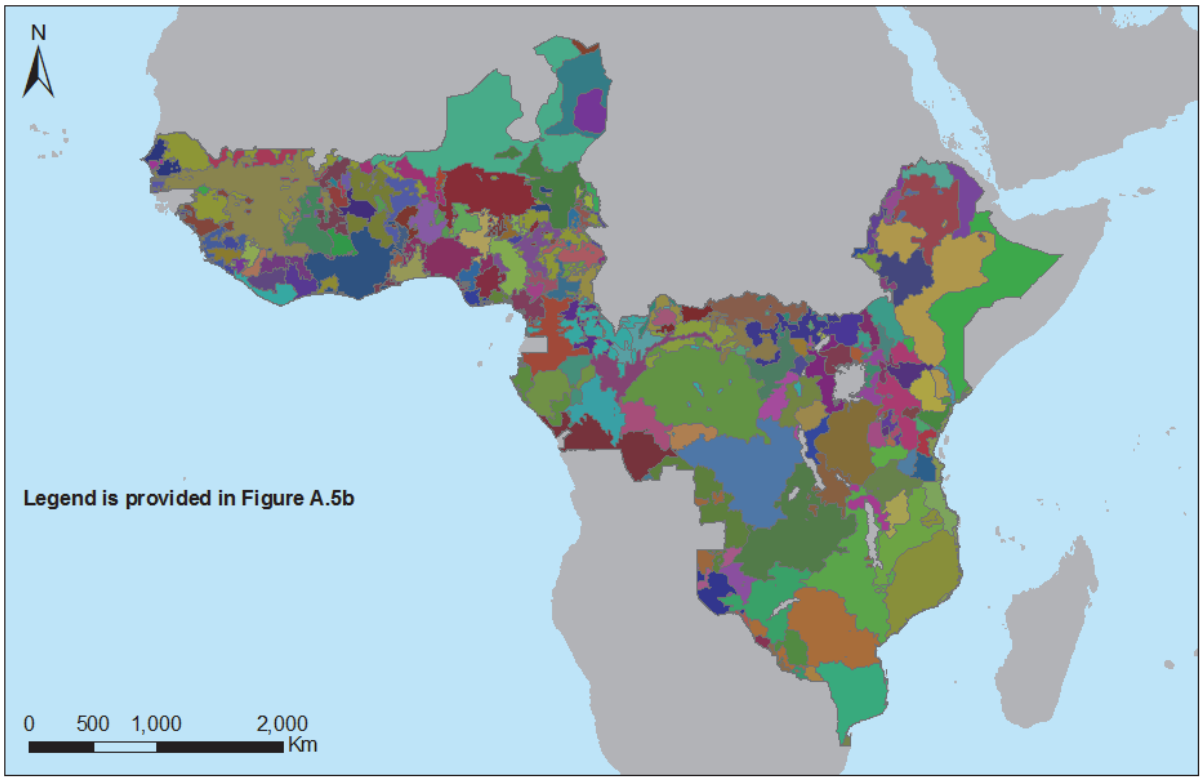


Figure A.5b. Legend of Figure A.5a

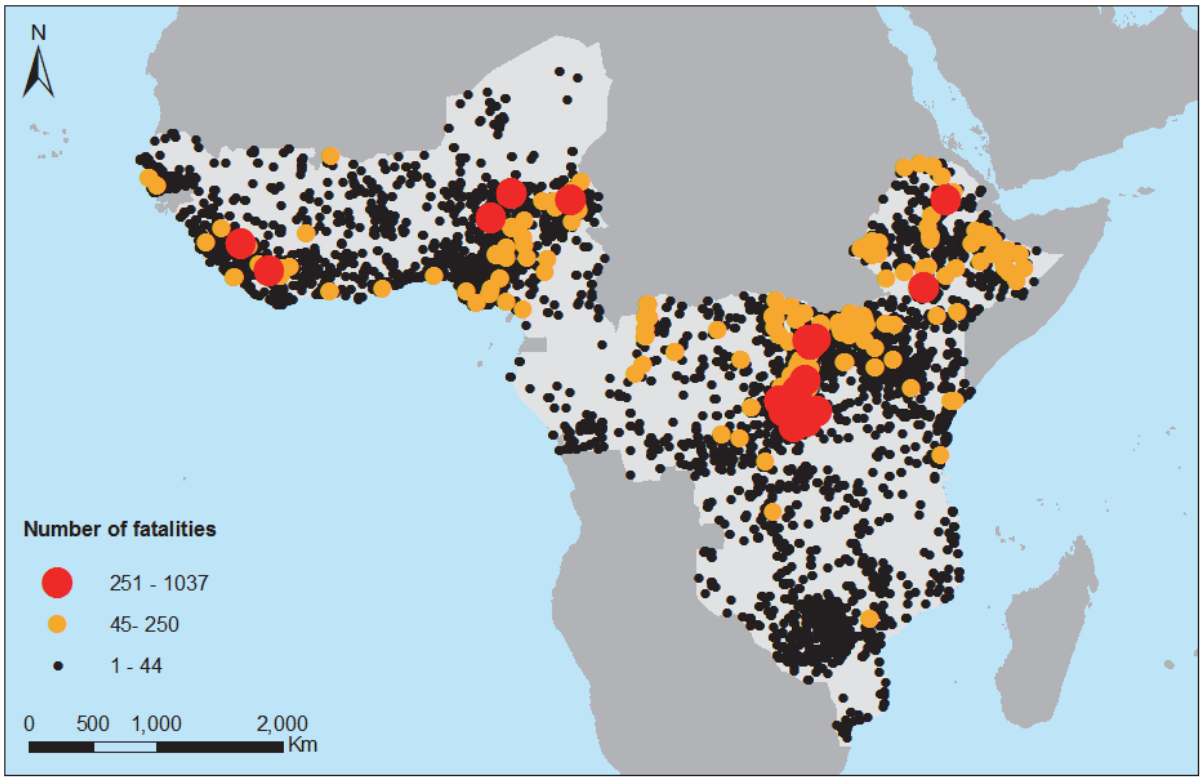
Predominant ethnic group

- Acholi
- Agau
- Akamba
- Akan
- Akikuyu
- Amhara
- Angas
- Angoni
- Anuak
- Azande
- Baboa
- Bade
- Baganda
- Bagirmi
- Baha
- Bakare
- Bakele
- Bakomo
- Bakongo
- Bakonjo
- Bakota
- Bakuba
- Balante
- Balozi
- Baluba
- Baluhya
- Balunda
- Bambundu
- Bamileke
- Banda
- Bankoya and Wambuela
- Bantu
- Banyaruanda
- Banyoro
- Barba
- Barega
- Bari
- Barundi
- Basakata
- Basoga
- Basubia
- Bateke
- Bawenda
- Bechuanas
- Bemba
- Berta
- Bete
- Bini
- Birom and Jerawa
- Bobangi and Bangala
- Bobo
- Bura, Bata and Tera
- Busa
- Bushmen
- Bute
- Chamba
- Danakil
- Diola
- Dogon
- Duala
- Ewe
- Fang
- Fulbe
- Galla
- Gbaya
- Gere
- Gola
- Grusi
- Gurma
- Hadzapi
- Hausa
- Ibibio
- Ibo
- Ijo
- Iraku
- Irangi
- Joluo
- Jukun and Idoma
- Kambari
- Kanuri
- Karamojo
- Katab
- Kisi
- Koma
- Kotoko
- Kpelle
- Kulango
- Kunama
- Lagoon tribes
- Libya Arabs
- Lobi
- Loma
- Maka
- Makonde
- Makua
- Malavi
- Mandara
- Mandingo
- Mandjak
- Mano
- Masa
- Masai
- Mashona
- Matebele
- Mba
- Mbum
- Mende
- Mongo
- Moru-Mangbetu
- Moru-Mangbetu and Sere-Mundu
- Mossi
- Mpongwe
- Mumuye
- Murle
- Nalu
- Nandi
- Ngbandi
- Ngiri
- Ngombe
- Northern Lwo
- Nuer
- Nupe
- Pedi
- Saho
- Sandawe
- Sanu
- Senufo
- Sere-Mundu
- Serer
- Shoa-Arabs
- Sidamo
- Somalis
- Somba (incl. Berba, Bilapila)
- Songai
- Soninke
- Southern Lwo
- Sudan Arabs
- Suk
- Susu
- Swahili
- Swazi
- Tatog
- Tem
- Tenda
- Teso
- Tigrai
- Tikar
- Tiv
- Togo tribes
- Tonga
- Tsonga
- Tuaregs
- Tubu
- Waboni
- Wachokwe
- Wadjagga
- Wafipa
- Wagogo
- Wahehe
- Wai
- Wakinga
- Waluchazi
- Wanyamwezi
- Wanyaturu
- Wanyika
- Wapare
- Wapokomo
- Wasagara
- Washambala
- Wateita
- Wayao
- Wayeye
- Wazaramo
- West Sahara Arabs
- Wolof
- Yoruba
- Zulus

Conflicts

For Africa, the "Armed Conflict Location & Event Data Project (ACLED)" database (http://www.acleddata.com/data/) covers the period from 1997 to the present. The database is constantly updated with data from multiple sources that include local media, national reports, humanitarian agencies, and research publications (http://www.acleddata.com/methodology/). Only data from 2010 to 2013 were used in this study.

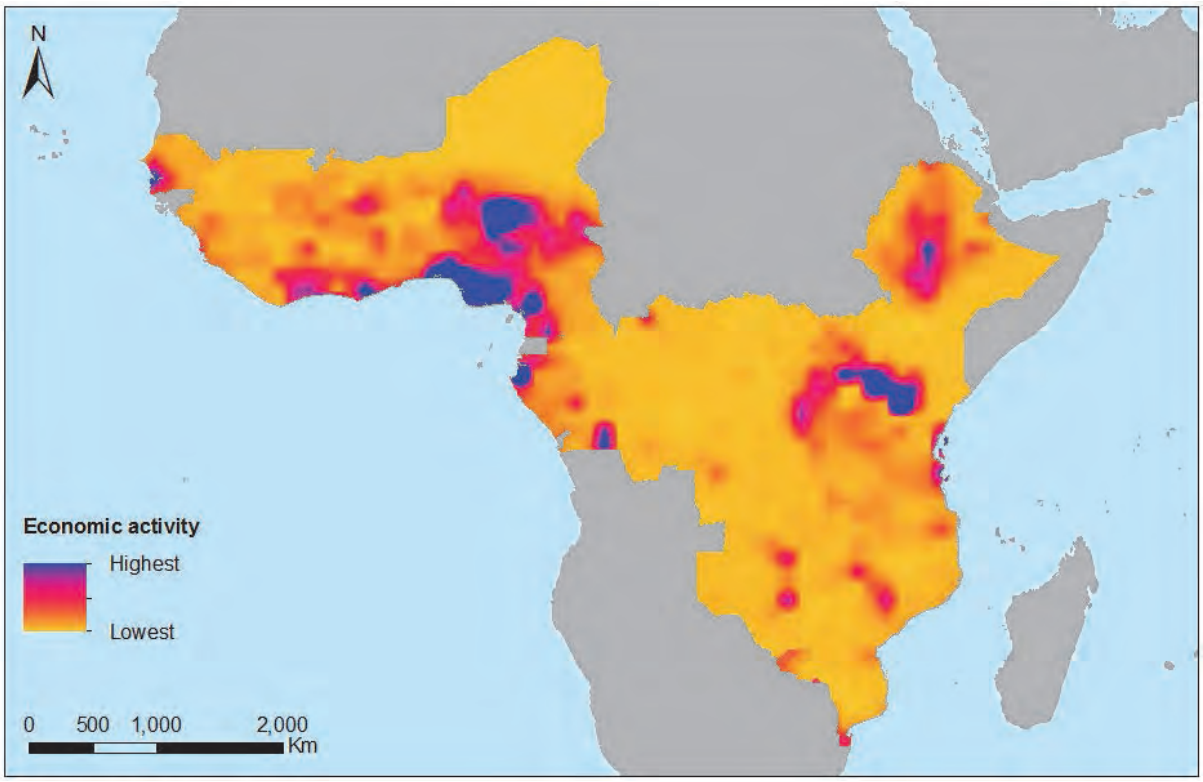
Figure A.6. Location of conflict events and associated fatalities between 2010 and 2013



Economic activity

The global "Geographically-based Economic (G-Econ)" dataset (http://gecon.yale.edu/data-and-documentation-g-econ-project) was produced with different country-based methodologies for estimating per capita gross cell product and calculating the gross cell product, with the latter representing the grid-based equivalent of the purchasing parity of gross domestic product.

Figure A.7. Spatial distribution of the gross cell product representing the regional equivalent of gross domestic product



Livestock

The FAO Gridded Livestock of the World, v2.0 (GLW2) datasets (http://livestock.geo-wiki.org/) were obtained by disaggregating sub-national livestock figures from ancillary environmental and demographic datasets.

Figure A.8a. Sheep density

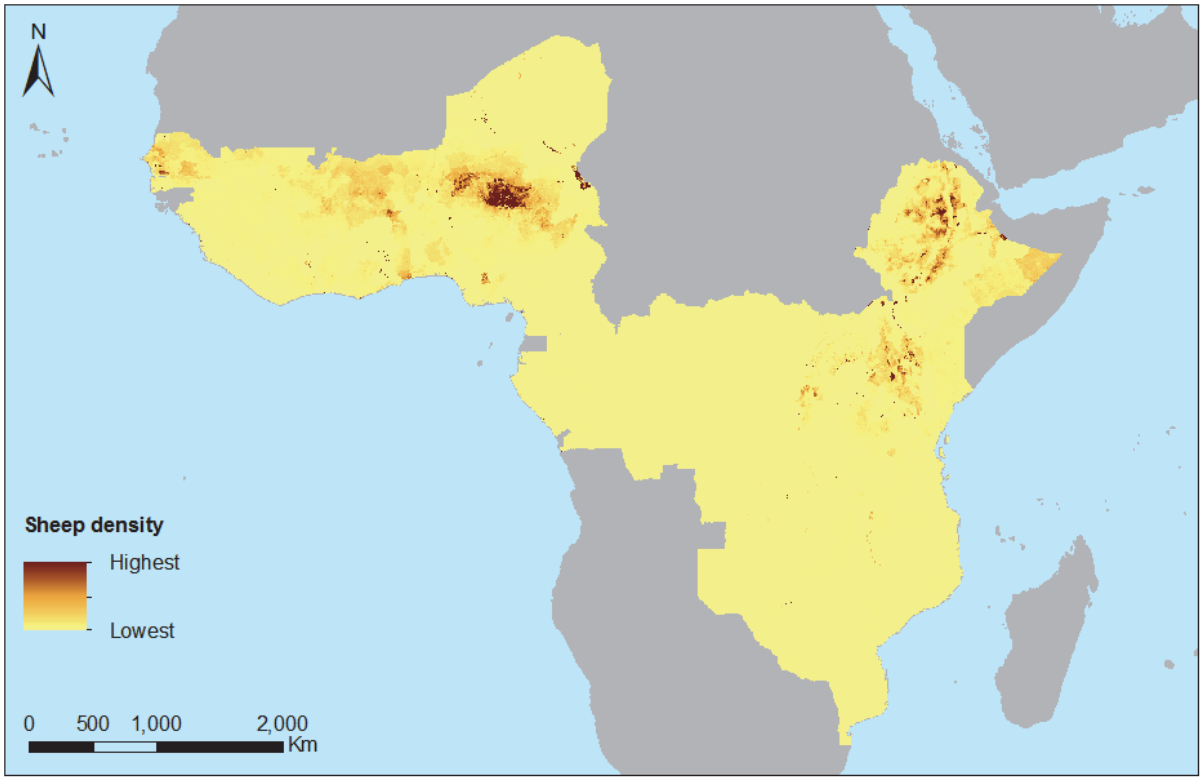


Figure A.8b. Cattle density

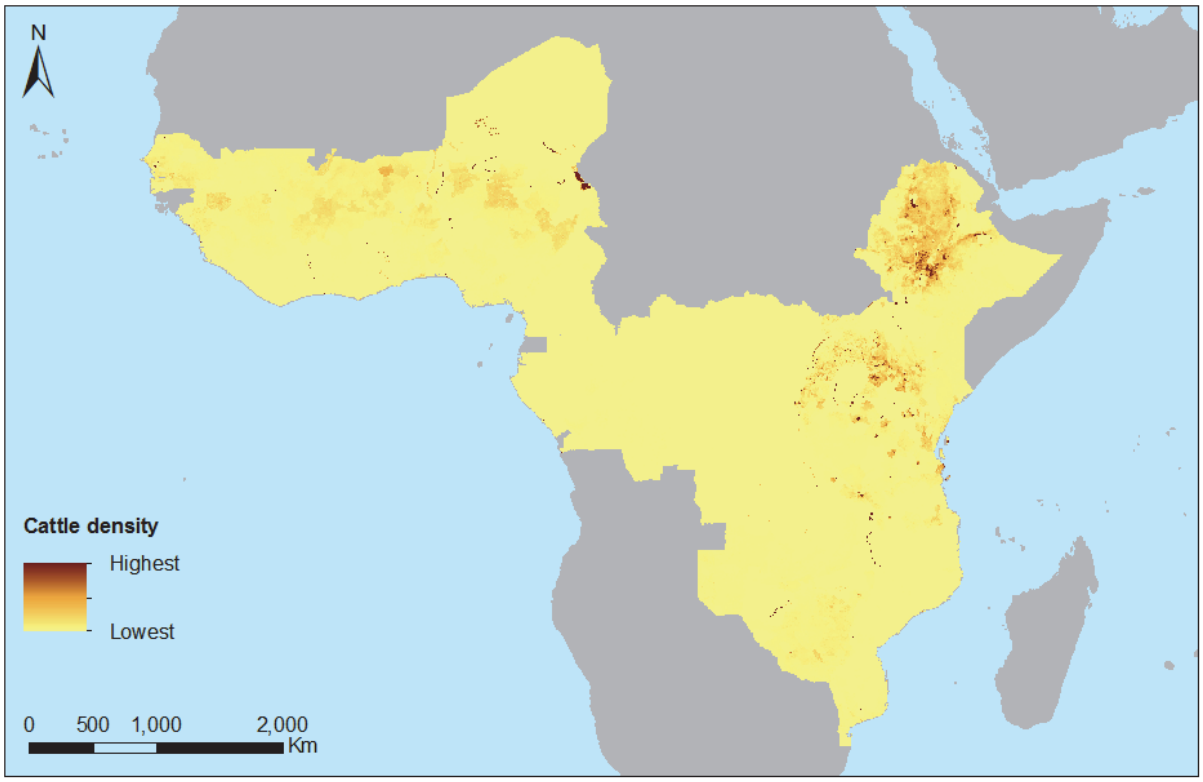


Figure A.8c. Pig density

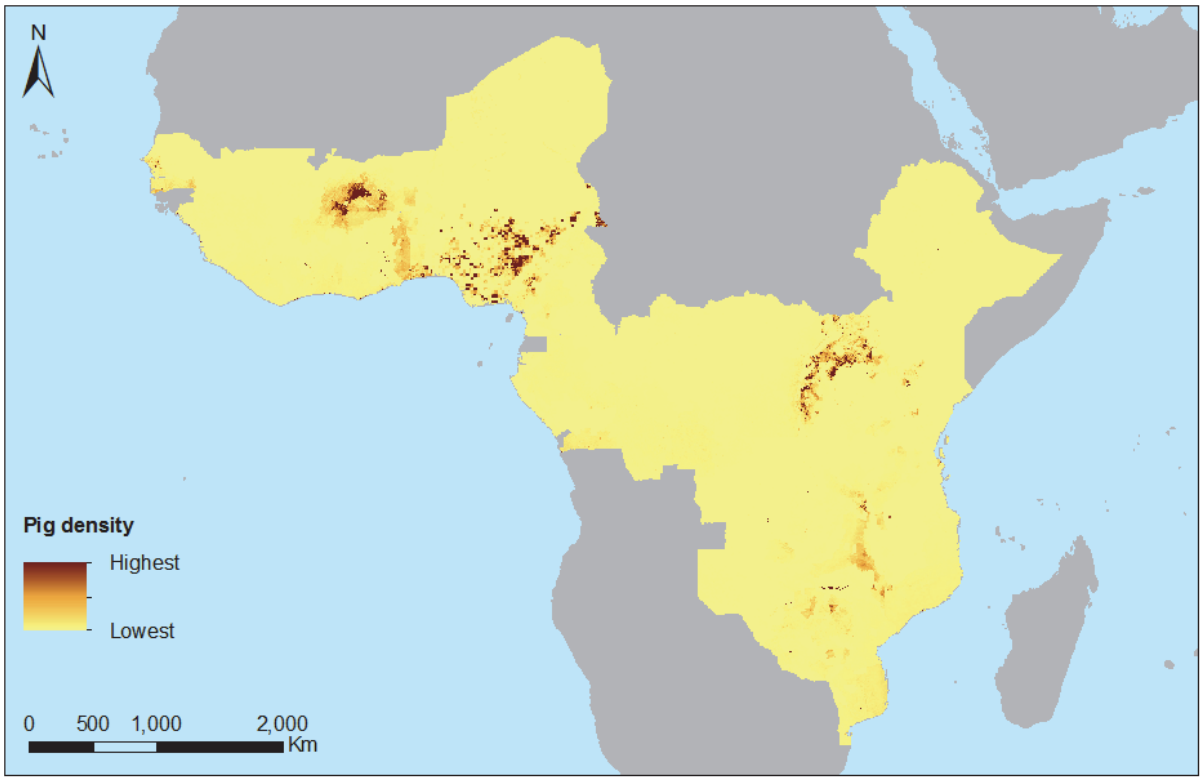
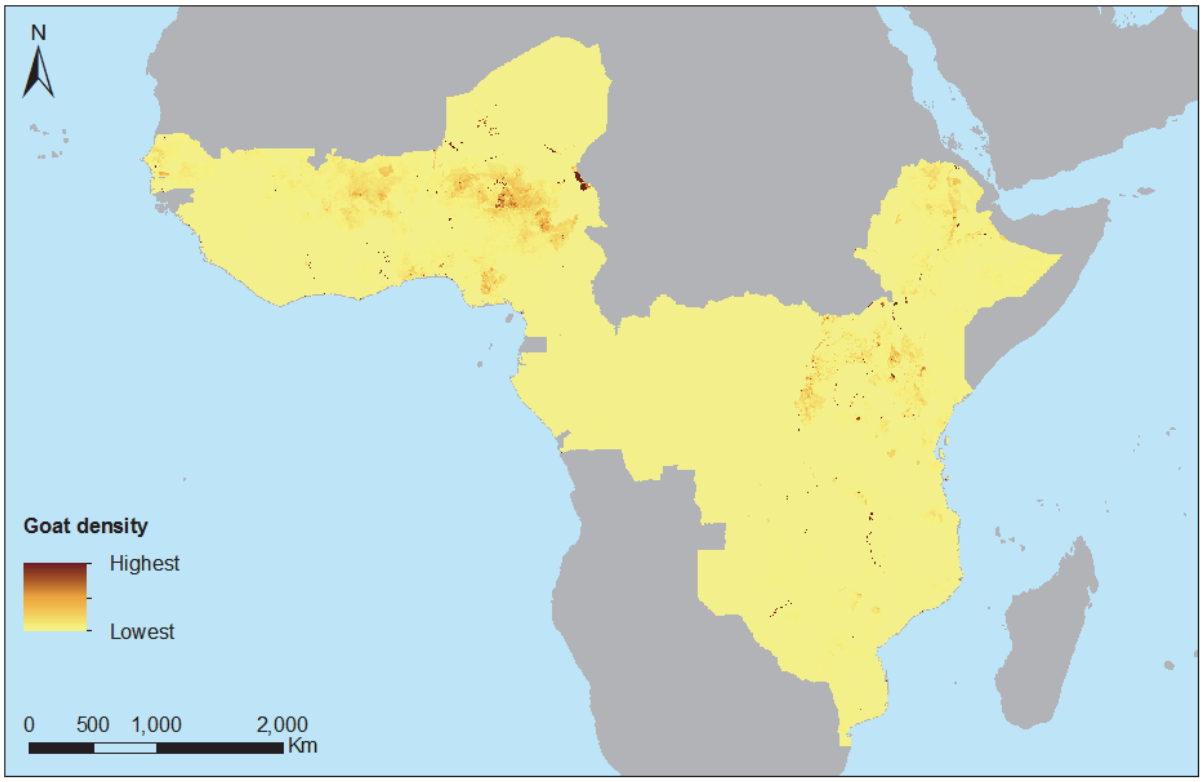


Figure A.8d. Goat density



Temperature and rainfall

WorldClim datasets (http://www.worldclim.org/current) were generated by modeling 30-years (1960-1990) of average monthly climate data from ground-based weather stations, as a function of latitude, longitude, and elevation (http://www.worldclim.org/methods1).

Figure A.9a. 30-year average annual temperature

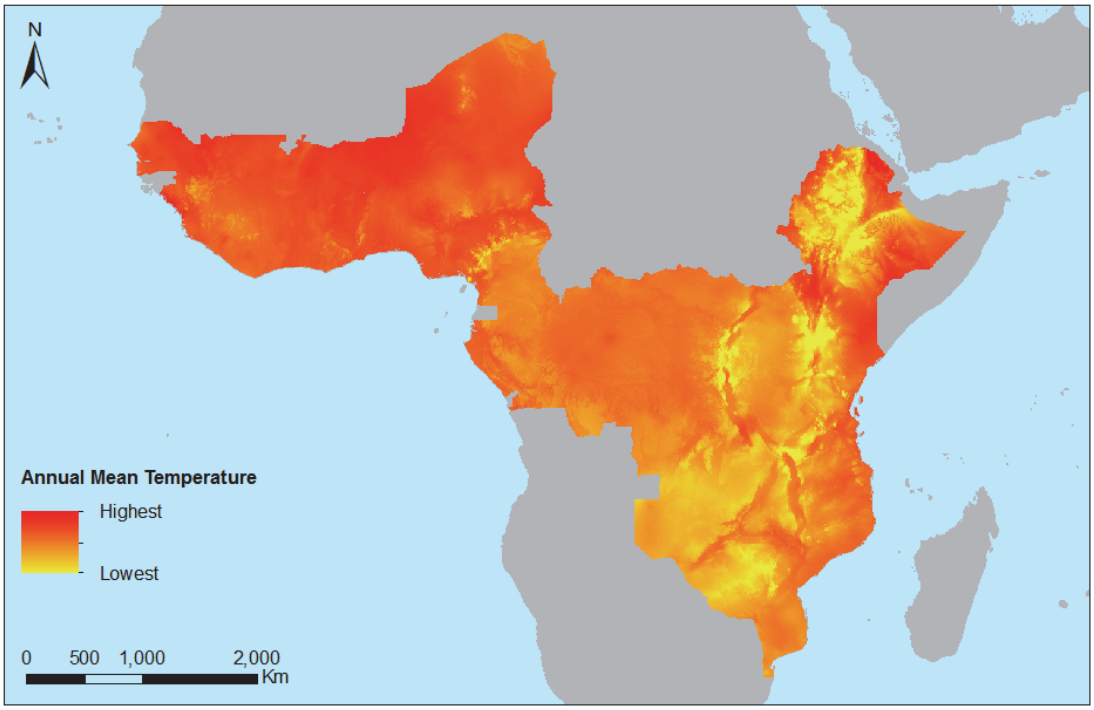
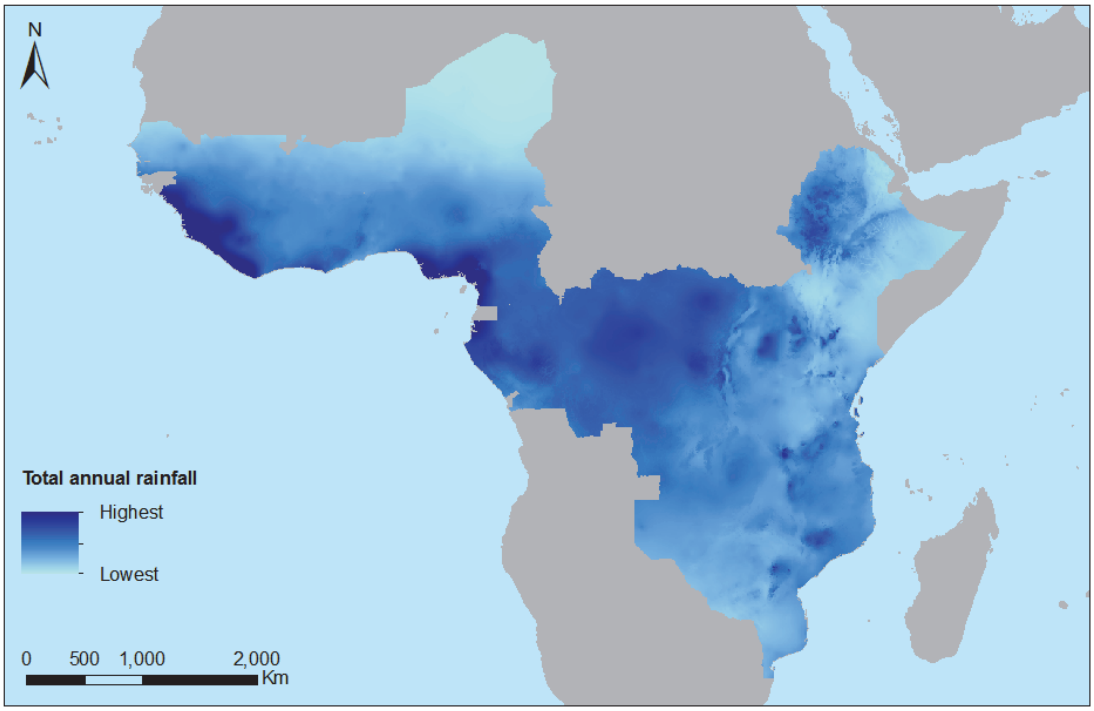


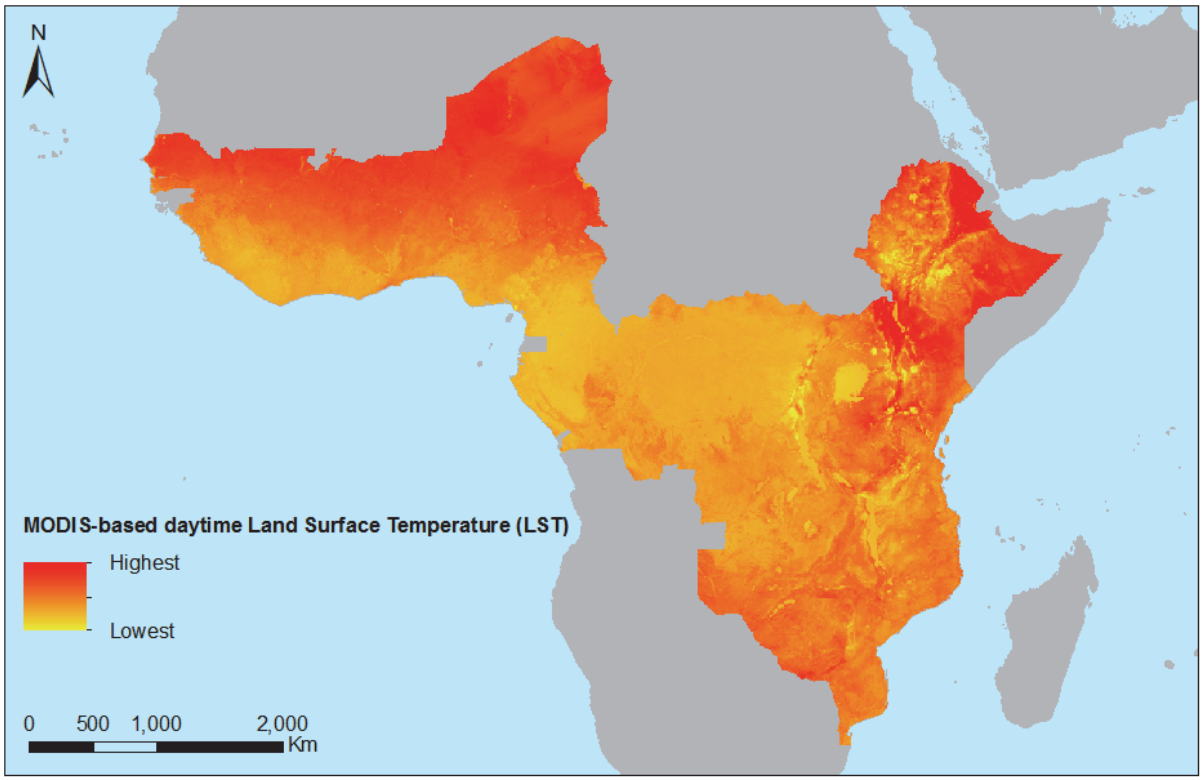
Figure A.9b. 30-year average annual rainfall



Temperature (MODIS)

The MODIS-based daytime Land Surface Temperature (LST) monthly composite datasets (https://lpdaac.usgs.gov/dataset_discovery/modis/modis_products_table/mod11c3) are derived from daily 5.6km LST datasets (MOD11C1) by averaging values of clear-sky LST values during a one-month period. In this study, all LST monthly composite datasets that covered the whole study area, from June 2001 to June 2015, were mosaicked and then averaged across time.

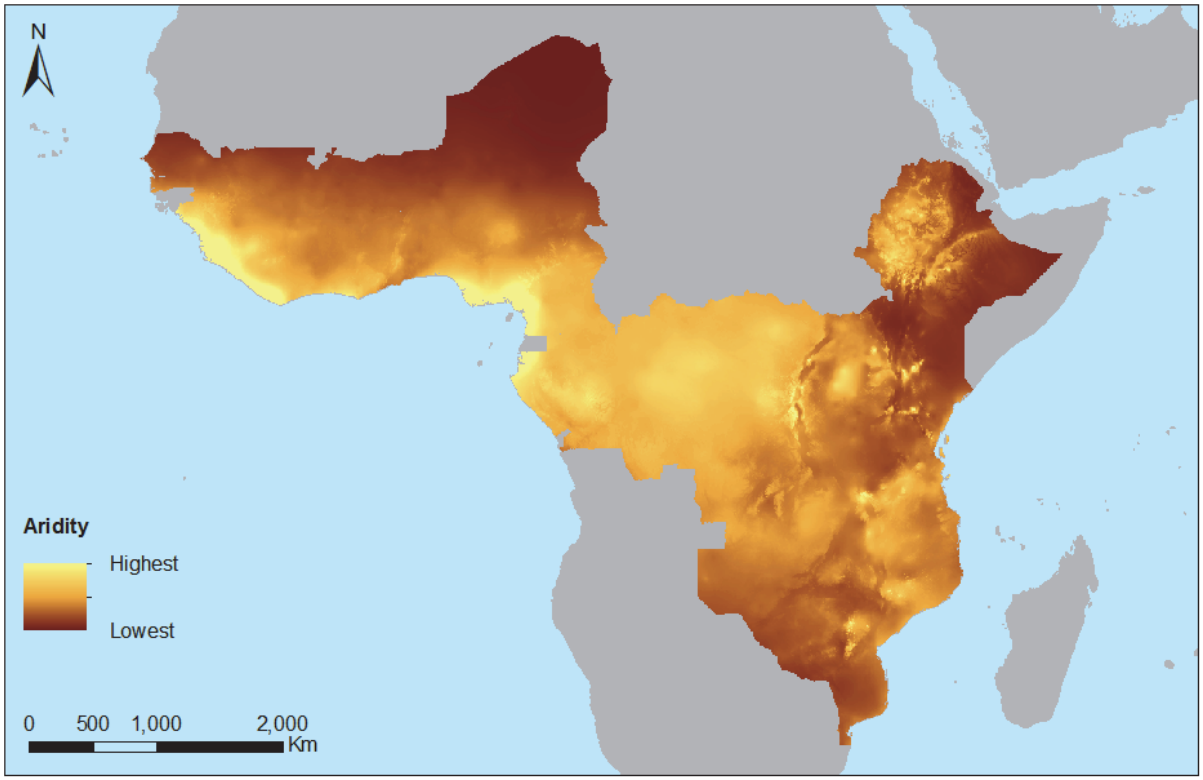
Figure A.10. 14-year average MODIS-based daytime Land Surface Temperature



Aridity

The CGIAR Global Aridity Index raster dataset (http://www.cgiar-csi.org/data/global-aridity-and-pet-database), which was produced with WorldClim Global Climate Data (Hijmans et al. 2005), is used to quantify precipitation availability over atmospheric water demand (Middleton and Thomas 1997). The index is defined as the ratio of mean annual precipitation to the mean annual potential evapo-transpiration; thus, the values increase from more arid to more humid conditions (http://www.cgiar-csi.org/wp-content/uploads/2012/11/Global-Aridity-and-Global-PET-Methodology.pdf).

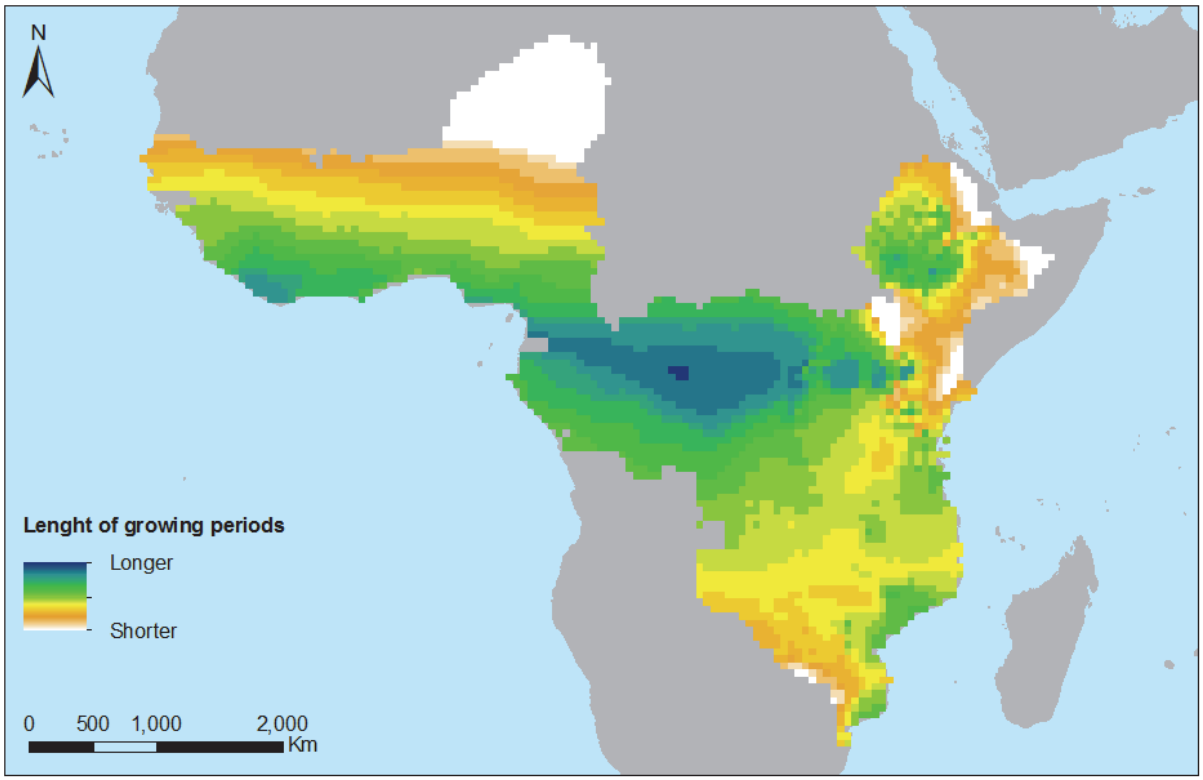
Figure A.11. 50 year-based Aridity Index



Growing season length

The FAO "Global length of growing periods" dataset (http://www.fao.org/geonetwork/srv/en/main.home) represents, for each grid cell, the number of days expressed as intervals during the year when both moisture availability and temperature are conducive to crop growth. This was produced by using the Agro-ecological Zones (AEZ) methodology for land productivity assessments (http://www.fao.org/fileadmin/user_upload/gaez/docs/GAEZ_Model_Documentation.pdf).

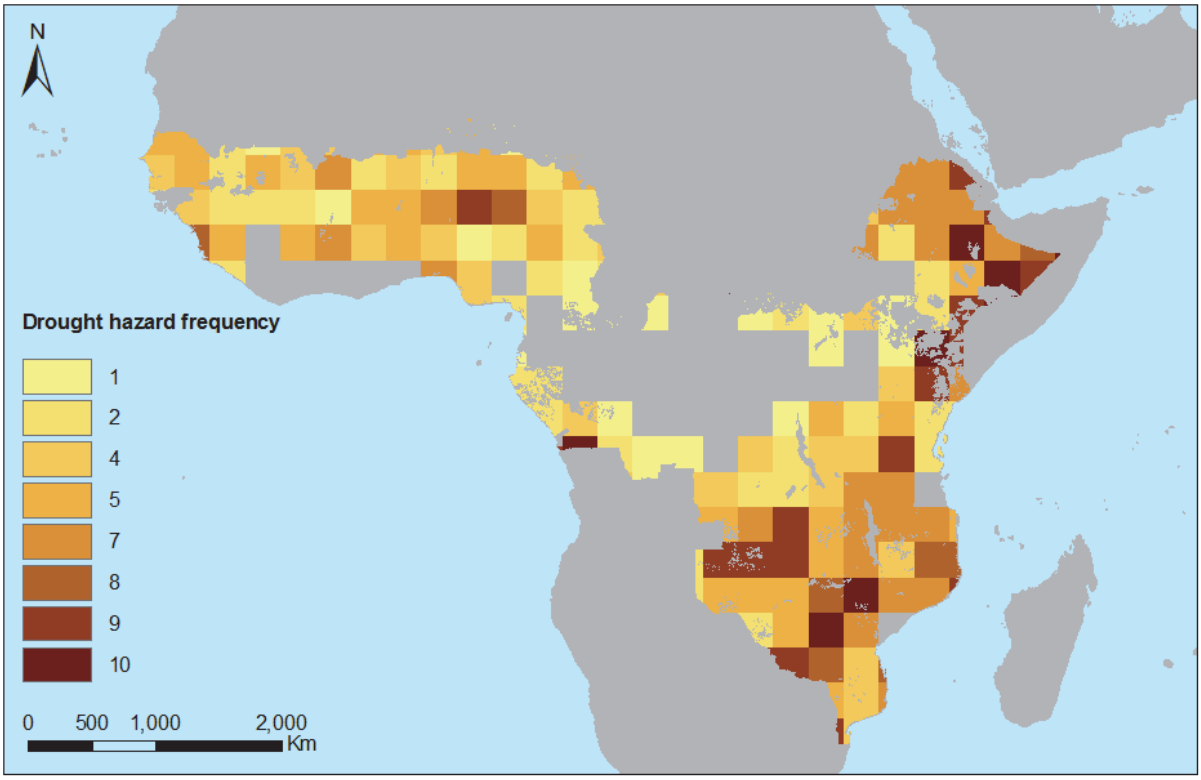
Figure A.12. Spatial distribution of the growing season length



Drought

The CIESIN Global Drought Hazard Frequency and Distribution, v1 dataset (http://sedac.ciesin.columbia.edu/data/set/ndh-drought-hazard-frequency-distribution/data-download) was created using the International Research Institute for Climate Prediction's (IRI) Weighted Anomaly of Standardized Precipitation (WASP; (Lyon and Barnston 2005) and average monthly precipitation data from 1980 through 2000 to calculate the number of drought events that occurred in each grid cell during the considered time period. According to the number of drought occurrences, grid cells are reclassified into 10 classes, with each one containing the same number of grid cells (http://sedac.ciesin.columbia.edu/data/set/ndh-drought-hazard-frequency-distribution).

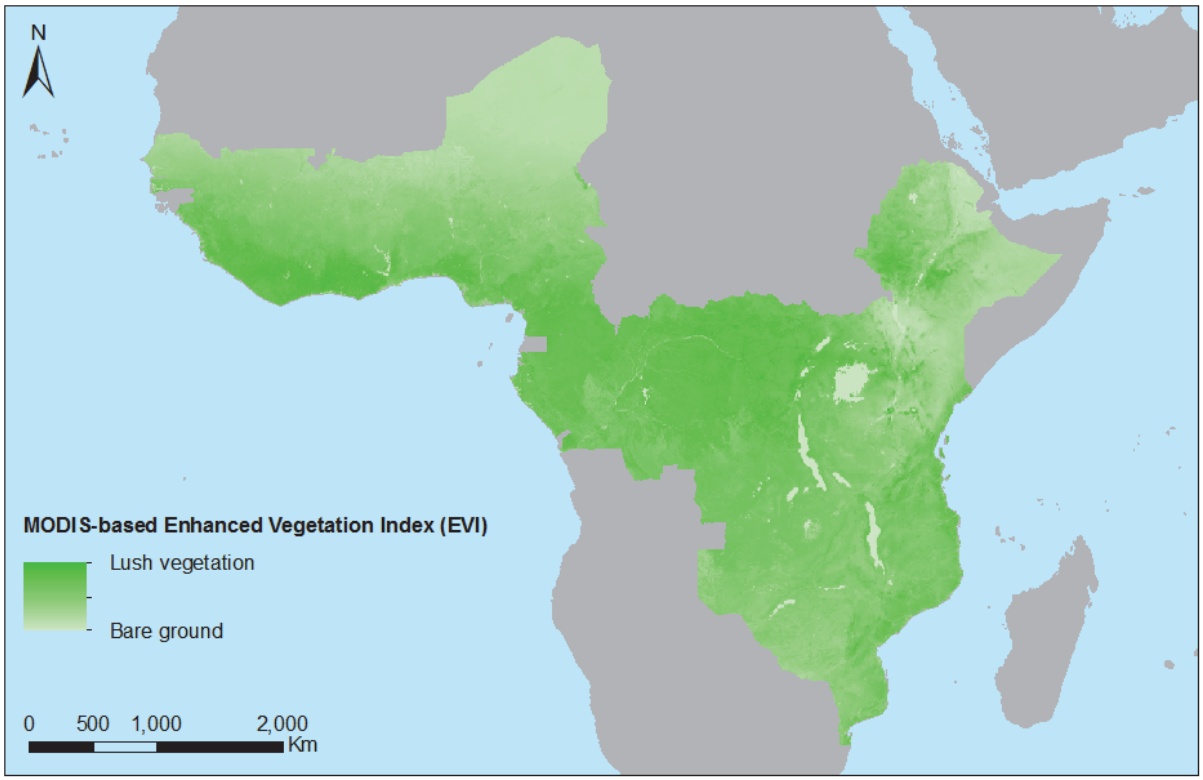
Figure A.13. Relative frequencies of drought occurrences and their spatial distribution



Vegetation

The MODIS Enhanced Vegetation Index datasets (https://lpdaac.usgs.gov/dataset_discovery/modis/modis_products_table/mod13c2), that represent a measure of photosynthetic activity ranging from 0 (no vegetation) to 1 (complete vegetation), are cloud-free spatial composites of 16-day 1 km EVI datasets (MOD13A2). In this study, all EVI monthly composite datasets that covered the whole study area, from June 2001 to June 2015, were mosaicked and then averaged across time.

Figure A.14. 14-year average MODIS-based Enhanced Vegetation Index



Topography

The GTOPO30 (https://lta.cr.usgs.gov/GTOPO30) is a digital elevation model (DEM) that was produced using multiple vector and raster sources of topographic information (https://lta.cr.usgs.gov/sites/default/files/gt30src.gif). In this study, it was used to derive a slope raster dataset and three roughness raster datasets for the study area, all of which have the same spatial resolution of the input GTOPO30 DEM. Since a raster-based roughness dataset expresses the amount of elevation difference between each grid cell and the considered surrounding area in the input DEM (grid cell values are a function of the size of the specific surrounding area selected to calculate the roughness), three different roughness datasets were produced by using three different neighborhood window sizes (3x3, 5x5, and 11x11 grid cells).

Figure A.15a. Topographic elevation

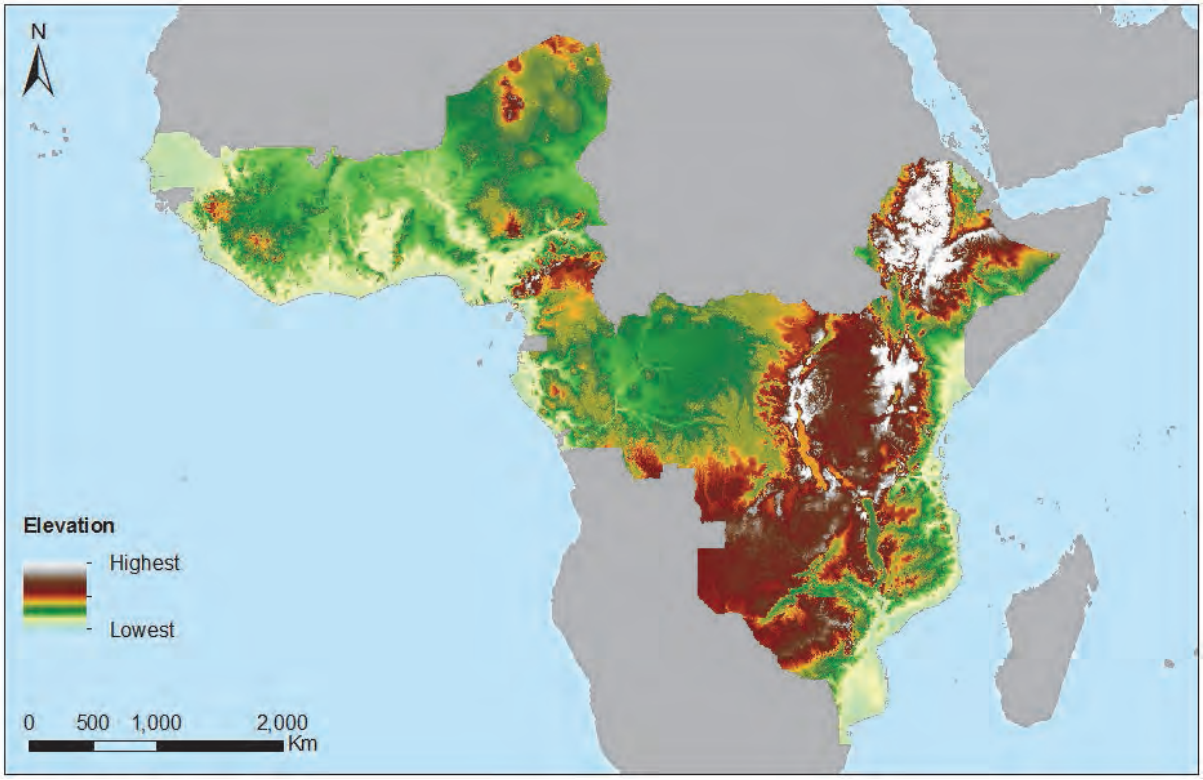


Figure A.15b. Topographic slope

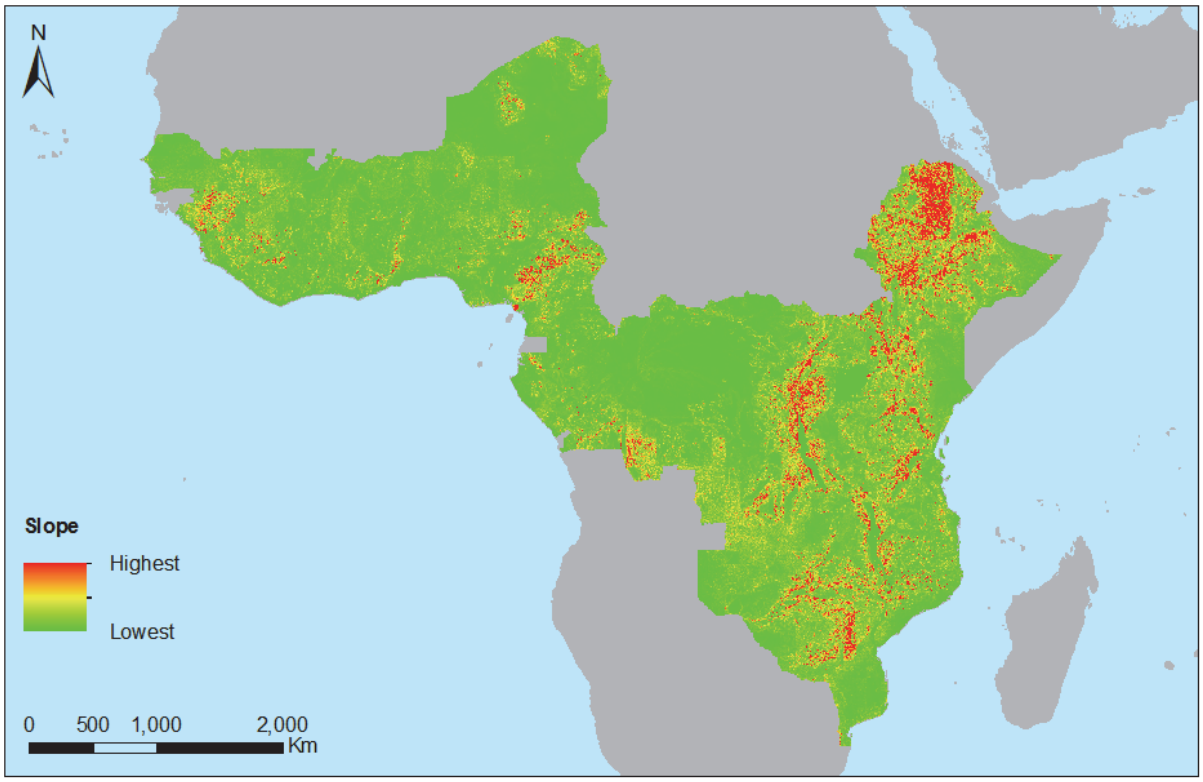


Figure A.15c. Topographic roughness calculated using a neighborhood window of 3x3 grid cells

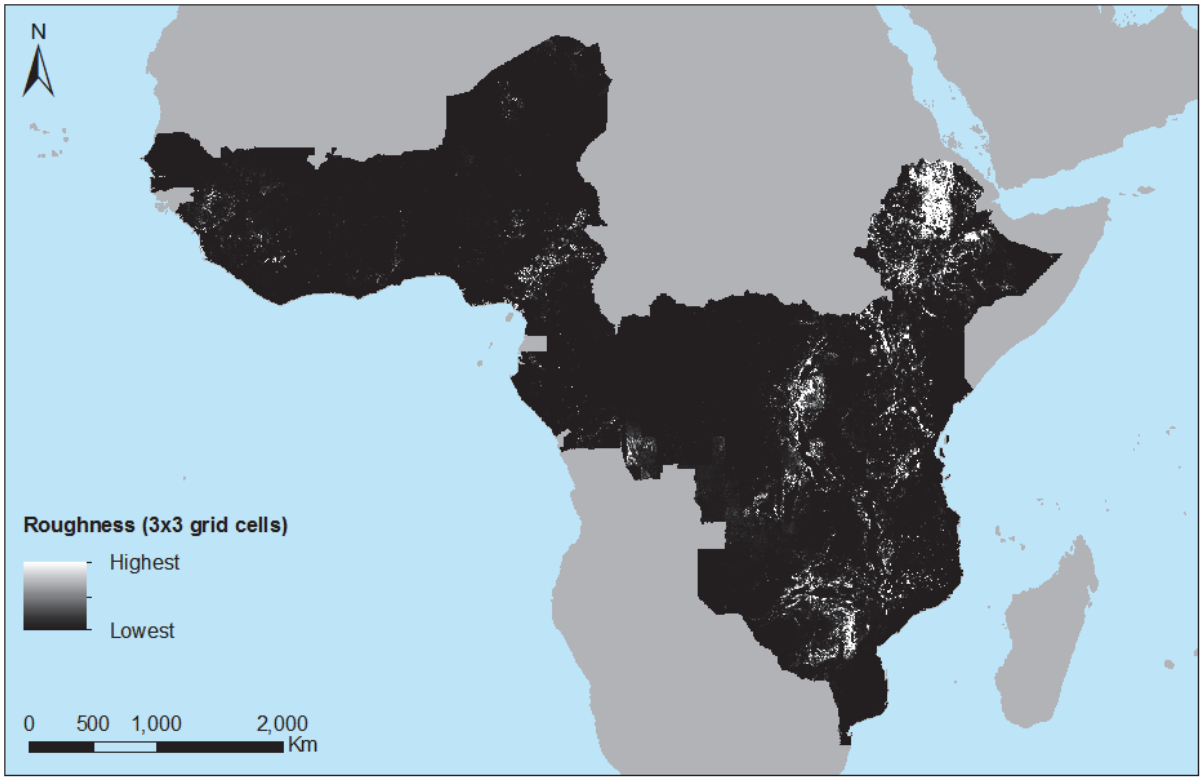


Figure A.15d. Topographic roughness calculated using neighborhood window of 5x5 grid cells

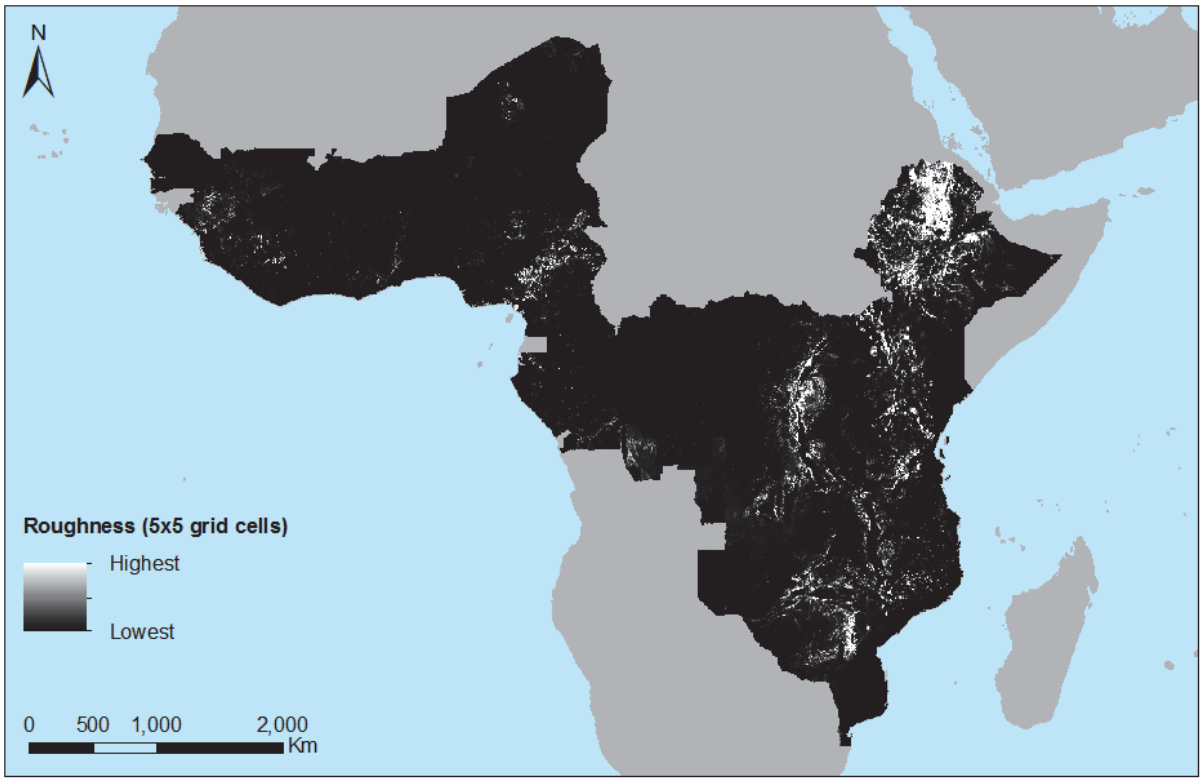
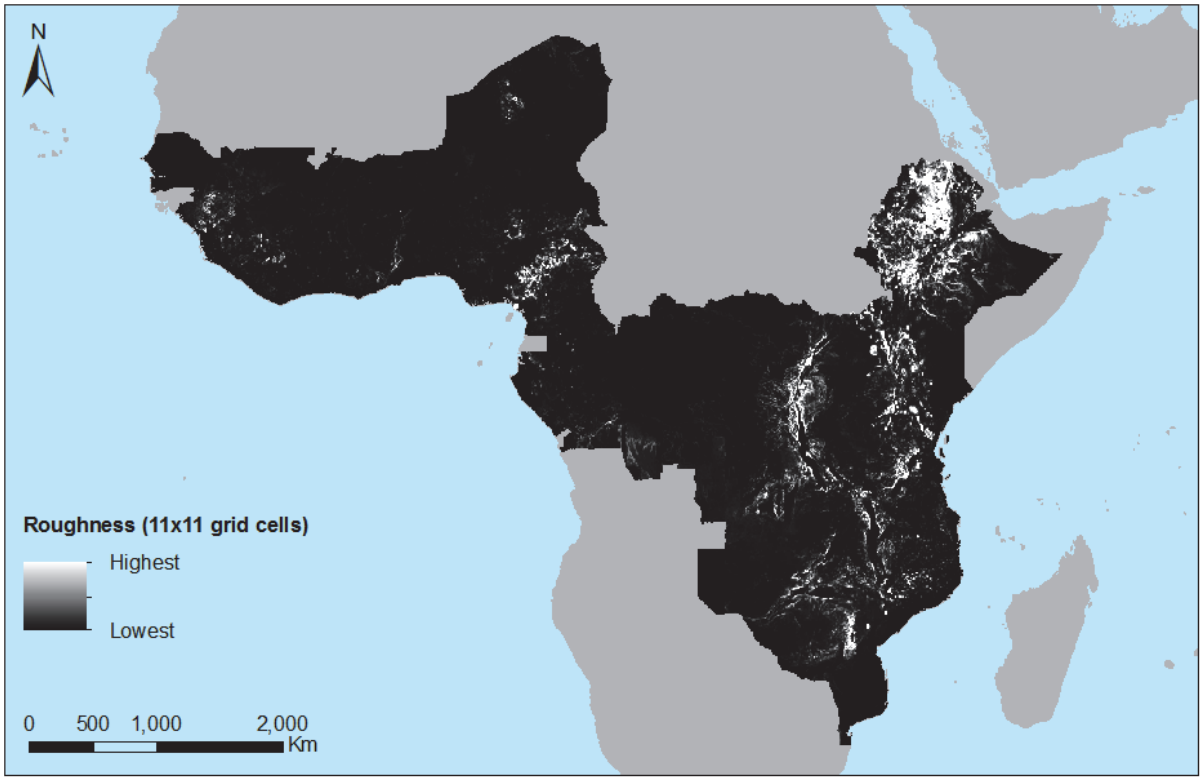


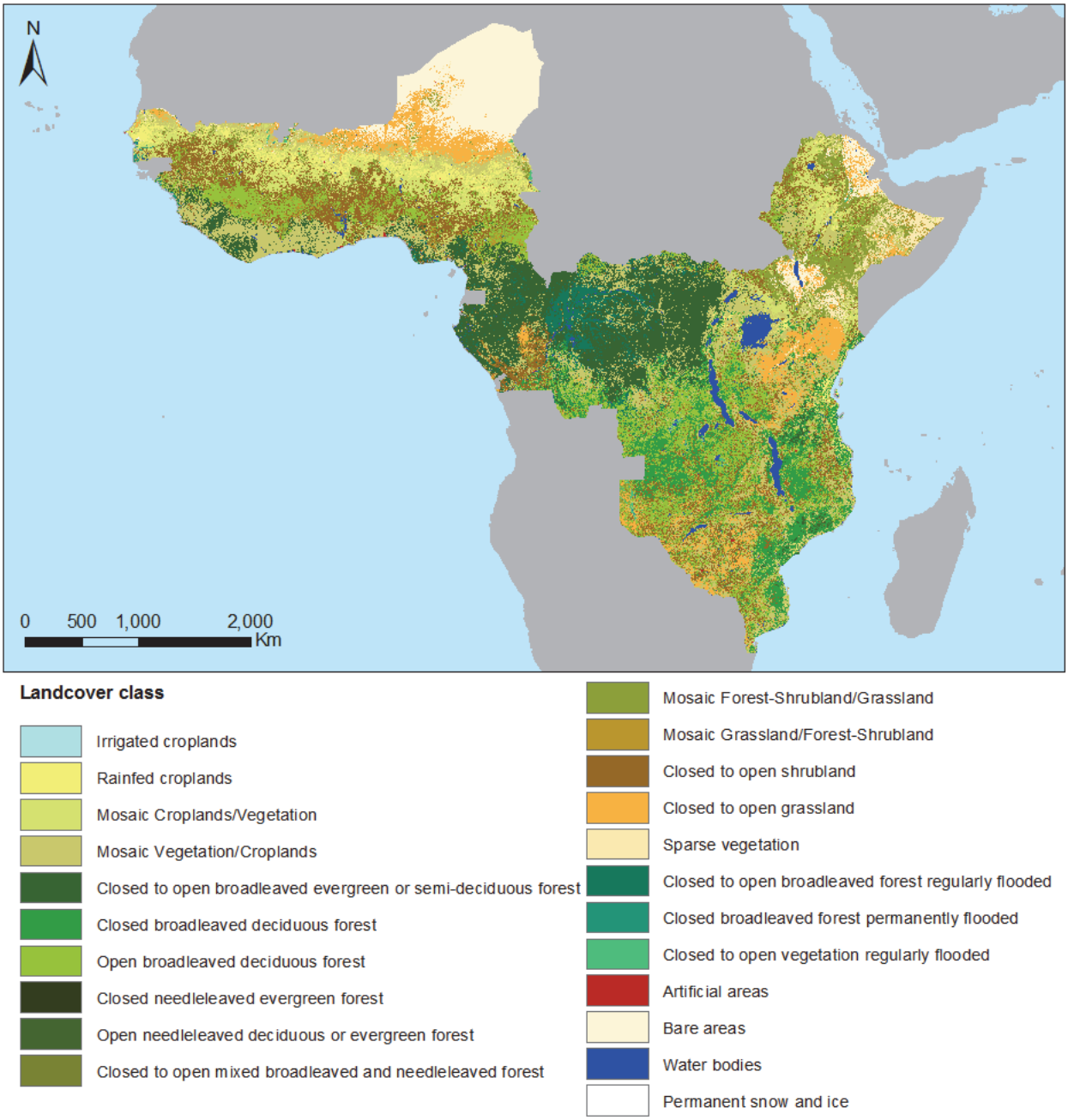
Figure A.15e. Topographic roughness calculated using neighborhood window of 11x11 grid cells



Landcover

The ESA 2009 GlobCover dataset (http://due.esrin.esa.int/page_globcover.php) was produced using Envisat MERIS fine resolution data and an automatic pre-processing and classification chain to identify 22 classes compatible with the UN Land Cover Classification System (LCCS).

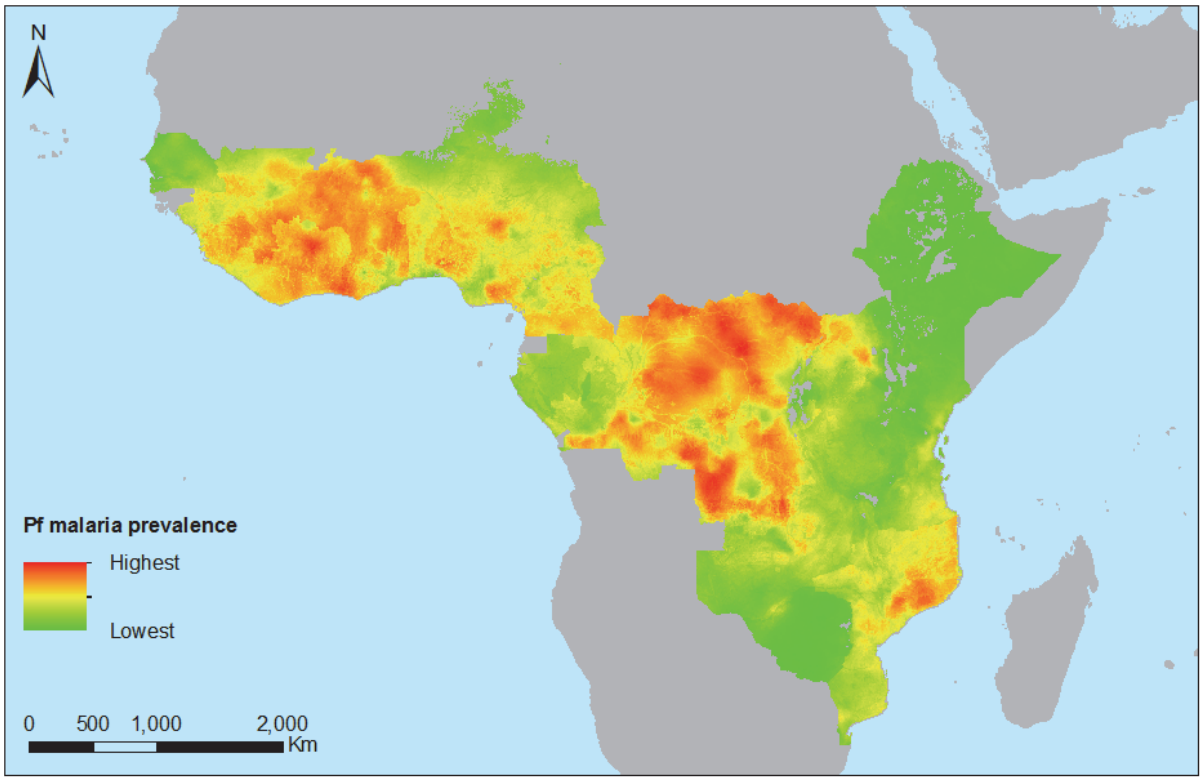
Figure A.16. MERIS-based landcover classification



Malaria prevalence

The Malaria Atlas Project "*Pf*PR2-10 in Africa 2000-2015" datasets (http://www.map.ox.ac.uk/) were produced using malaria prevalence survey data (1995-2014), along with temporally explicit intervention coverage, environmental and socio-demographic covariates, as input to a spatiotemporal Bayesian geostatistical model to estimate *Pf*PR trends (Bhatt et al. 2015). In this study, all "*Pf*PR2-10 in Africa 2000-2015" datasets were averaged across time.

Figure A.17. 15-year average *Plasmodium falciparum* prevalence distribution



Appendix B – Detailed Methods and Results

In this section of the appendix, we describe the complete spatial model discussed in Section 3.3. All the additional figures and tables referenced in Sections 3 and 4 are also presented.

B.1. Conditional Autoregressive (CAR) Model

The study region is partitioned into $n = 255$ areal units, otherwise known as the DHS regions. These regions are denoted by $\boldsymbol{\mathcal{S}} = \{\mathcal{S}_1, \dots, \mathcal{S}_n\}$. Let $\mathbf{Y} = (\mathrm{Y}_1, \dots, \mathrm{Y}_n)$ denote the vector of child mortality rates associated with each region. Also, each region, $\mathcal{S}_i$, is linked with a vector of covariates (including an intercept term and p covariates) which we denote by $\mathbf{x}_i^T = (1, \mathrm{x}_{i1}, \dots, \mathrm{x}_{ip})$. Considering a Gaussian distribution for the mortality rates, the spatial linear model is given by

$$\mathrm{Y}_i | \mu_i \sim N(\mathrm{y}_i | \mu_i, \sigma^2), \text{ for } i = 1, \dots, n,$$

$$\mu_i = \mathbf{x}_i^T \boldsymbol{\beta} + \phi_i. \tag{1}$$

In model (1), $\boldsymbol{\beta}_{(p+1)\times 1}$ denotes the regression coefficients, σ^2 is a scale parameter for the mortality rates and ϕ_i is a spatial random effect for region, $\mathcal{S}_i$.

The random effects $\boldsymbol{\phi} = (\phi_1, \dots, \phi_n)$ model residual spatial autocorrelation in $\mathbf{Y}$. In models of the type given in (1), these random effects $\boldsymbol{\phi}$ are typically characterized using a conditional autoregressive (CAR) model. A number of *global* CAR models have been proposed in the statistical literature for this purpose. These include the intrinsic CAR and Besag-York-Mollie (BYM) models (Besag, York, and Mollié 1991) and their alternatives (Leroux, Lei, and Breslow 2000; Stern and Cressie 1999). However, based on a recent review study by (Lee 2013), we adopted the CAR model by Leroux et al. for our analysis. This CAR prior is given by

$$\phi_i | \boldsymbol{\phi}_{-i} \sim N\left(\frac{\rho \sum_{j=1}^{n} w_{ij}\phi_j}{\rho \sum_{j=1}^{n} w_{ij} + 1 - \rho}, \frac{\tau^2}{\rho \sum_{j=1}^{n} w_{ij} + 1 - \rho}\right),$$

where ρ is a spatial autocorrelation parameter and τ^2 is a variance parameter. The $n \times n$ matrix of the w_{ij}'s (W) characterizes the neighborhood structure of the regions; this induces spatial autocorrelation in $\boldsymbol{\phi}$. A binary specification is often used for W where $w_{ij} = 1$ if regions $(\mathcal{S}_i, \mathcal{S}_j)$ share a common border and is zero if they do not.

The model was estimated in a Bayesian framework that naturally takes care of uncertainty in predictions and parameter estimation. To complete our model specification, we placed the following priors on the parameters: $\beta_j \sim N(0,1000)$ $(j = 0, \dots, p)$, $\sigma^2 \sim \text{Uniform}(0, 1000)$, $\rho \sim \text{Uniform}(0, 1)$ and $\tau^2 \sim \text{Uniform}\ (0, 1000)$.

Table B.1. Top 10 models resulting from the model selection process. Models are ranked according to R^2.

Model	AIC	R^2	PR^2
stunting + malaria prevalence*birth interval + literacy*birth interval + literacy*access to a health facility + birth interval*temperature + literacy*night-time lights + birth interval*ethnic groups variety	2132	0.54	0.53
stunting + temperature + malaria prevalence*birth interval + literacy*birth interval + literacy*access to a health facility + literacy*night-time lights + birth interval*ethnic groups variety	2130	0.54	0.53
stunting + night-time lights + malaria prevalence*birth interval + literacy * birth interval + literacy * access to a health facility + birth interval *temperature + birth interval*ethnic groups variety	2130	0.54	0.53
stunting + malaria prevalence*birth interval + literacy*birth interval + literacy * access to a health facility + birth interval * temperature + literacy*economic activity + birth interval * ethnic groups variety	2132	0.54	0.53
stunting + temperature + night-time lights +malaria prevalence* birth interval + literacy* birth interval + literacy * access to a health facility + birth interval* ethnic groups variety	2128	0.54	0.53
stunting + temperature + malaria prevalence*birth interval + literacy* birth interval + literacy* access to a health facility + literacy *economic activity + birth interval * ethnic groups variety	2130	0.54	0.53
stunting + economic activity + malaria prevalence* birth interval + literacy* birth interval + literacy* access to a health facility + birth interval* temperature + birth interval * ethnic groups variety	2130	0.54	0.53
stunting + malaria prevalence*birth interval + literacy * access to a health facility + birth interval*temperature + literacy *sanitation + literacy*economic activity + birth interval*ethnic groups variety	2134	0.54	0.53
stunting + temperature + economic activity + malaria prevalence* birth interval + literacy* birth interval + literacy * access to a health facility + birth interval * ethnic groups variety	2128	0.54	0.53
stunting + population density + malaria prevalence * birth interval + literacy* birth interval + literacy* access to a health facility + birth interval*temperature + birth interval * ethnic groups variety	2130	0.54	0.53

Table B.2. Posterior means of the parameters of the spatial model without region-level interactions

Parameter	Mean	SD	Median	95% Credible interval
(Intercept)	33.4298	1.1676	33.4633	(31.1412, 35.7095)
stunting	0.3254	0.1201	0.3233	(0.0915, 0.5600)
Malaria prevalence	50.6905	7.5511	51.0639	(35.3227, 65.1751)
Birth interval	0.7751	0.1722	0.7664	(0.4497, 1.1245)
Literacy	0.0146	0.0863	0.0103	(-0.1439, 0.1952)
Access to a health facility	-0.1614	0.0660	-0.1601	(-0.2905, -0.0287)
Temperature	0.1103	0.0576	0.1080	(-0.0006, 0.2278)
Night-time lights	-0.2923	1.2151	-0.3366	(-2.6548, 2.0590)
Ethnic groups variety	-0.5791	0.4347	-0.5878	(-1.3887, 0.2862)
Malaria prevalence: Birth interval	2.0500	0.8843	2.0393	(0.2837, 3.7821)
Birth interval: Literacy	-0.0162	0.0078	-0.0163	(-0.0310, -0.0008)
Literacy: Access to a health facility	0.0049	0.0025	0.0050	(0.0000, 0.0097)
Birth interval: Temperature	-0.0008	0.0051	-0.0007	(-0.0107, 0.0094)
Literacy: Night-time lights	-0.0174	0.0342	-0.0181	(-0.0811, 0.0537)
Birth interval: Ethnic group variety	0.0960	0.0491	0.0954	(0.0028, 0.1934)
$\boldsymbol{\sigma^2}$	179.5640	90.8631	221.3824	(0.0218, 278.2934)
$\boldsymbol{\tau^2}$	102.9706	161.5442	0.0107	(0.0004, 468.1414)
$\boldsymbol{\rho}$	0.3974	0.2289	0.3627	(0.0413, 0.8792)