

©Copyright 2021

Jessica Godwin

# Subnational Estimation of Period Child Mortality in a Low and Middle Income Countries Context

Jessica Godwin

A dissertation  
submitted in partial fulfillment of the  
requirements for the degree of

Doctor of Philosophy

University of Washington

2021

Reading Committee:

Jon Wakefield, Chair

Emilio Zagheni

Tyler McCormick

Program Authorized to Offer Degree:

Department of Statistics

University of Washington

**Abstract**

Subnational Estimation of Period Child Mortality in a Low and Middle Income Countries  
Context

Jessica Godwin

Chair of the Supervisory Committee:  
Professor Jon Wakefield  
Statistics and Biostatistics

Child mortality is an important metric used in quantifying and monitoring the health of a population's children. Moreover, child mortality can be a key indicator of the overall health of a population, and is often used to quantify mortality at other ages. Over the past several decades, there have been huge global reductions in child mortality. However, child mortality remains large in many low and middle income countries (LMICs). The United Nations' Sustainable Development Goals (SDGs) call for a reduction of child mortality in the period 2015–2030. In particular, SDG 3.2 continues the global initiative to improve child mortality outcomes by calling for an end to preventable child deaths and reaching a target under-five mortality rate (U5MR) of 25 deaths before age 5 per 1000 births and 12 deaths per 1000 births in the first month of life by the year 2030. Many methods exist for estimation of national child mortality measures, but there is a growing desire for more and better methods for subnational estimation of child mortality. Improved methods for subnational estimation of U5MR can allow for a better understanding of the geographic differences of trends in U5MR and more targeted intervention for child mortality reduction. Though there are many ways to think about geographic variability within a country, in this thesis we focus on the administrative divisions of a country that have political and infrastructural meaning. The

national level of a country is referred to as Admin-0, the coarsest subnational administrative division is referred to as Admin-1, and the next coarsest as Admin-2. Many sources of child mortality data collect information for which Admin-1 estimates are reasonable, but this leads to small sample sizes at the Admin-2 level. The focus of this thesis is on subnational child mortality estimation at the Admin-2 level.

In this thesis, we provide a review of demographic and statistical methods for age-specific period child mortality, synthesize notation across fields, and develop two methods for subnational estimation of U5MR at the Admin-2 data in LMICs. We make use of child mortality data from household surveys throughout this thesis and discuss in detail how an individual's mortality information is used in various existing demographic and statistical methods for mortality estimation. The two methods we develop for subnational child mortality estimation in LMICs at the Admin-2 level take different approaches to accounting for the method of data collection in the household surveys and one method incorporates census data. We apply the method developed to incorporate census data separately to the countries of Kenya and Malawi. We find mixed evidence of improvements in estimation with the incorporation of census data, and note the method's shortcomings when small sample sizes result in many Admin-2 areas with no observed deaths. The limitations of this method, motivate the need for a method that can be applied more broadly to countries with sparser coverage of household survey data. We develop a method and reproducible, replicable pipeline for data acquisition, cleaning, estimation, and visualization. We apply this method to Admin-1 and Admin-2 levels in 22 LMICs, fitting separate models for each country and administrative division. The country-specific modeling and steps in the pipeline allow us to address the unique context of child mortality in each country if the generic base model is inappropriate, such as in countries with generalized HIV/AIDS epidemics, the genocide events in Rwanda, and Cyclone Nargis in Myanmar. The estimates for 22 LMICs have been published in col-

laboration with UNICEF and the UN Inter-agency Group for Child Mortality Estimation, have undergone review in consultations with country representatives, and are available at <https://childmortality.org>. This thesis improves on the current understanding in the processing and use of child mortality data in the literature and develops two new methods for estimation of child mortality in LMICs at the Admin-2 level. In particular, the development of a clearly-defined and replicable pipeline that allows for easy adaptation to address the unique mortality estimation needs of individual countries fills a previous gap in existing methodology. However, there are areas for future research and methodological improvement. Throughout the thesis we make note of ongoing work to improve existing aspects of the developed methods and make clear the issues that remain unaddressed.

# TABLE OF CONTENTS

	Page
List of Figures . . . . .	iv
List of Tables . . . . .	x
Chapter 1: Introduction . . . . .	1
1.1 Motivation . . . . .	1
1.2 Data . . . . .	5
1.3 Organization of Dissertation . . . . .	7
1.4 Methodological Contribution of Dissertation . . . . .	8
Chapter 2: Background . . . . .	12
2.1 Notation . . . . .	12
2.2 Demographic methods for period child mortality estimation . . . . .	15
2.3 Statistical methods for mortality estimation . . . . .	26
2.4 Survey statistics . . . . .	33
2.5 Bayesian spatiotemporal smoothing . . . . .	37
Chapter 3: A comparison of direct estimation methods of period child mortality with complex surveys . . . . .	42
3.1 Introduction . . . . .	42
3.2 Counting deaths and exposures for period mortality . . . . .	43
3.3 Estimation of period mortality . . . . .	50
3.4 A comparison using Senegal 2006 DHS . . . . .	54
3.5 Discussion . . . . .	92

Chapter 4:	Combining indirect and direct subnational child mortality estimates at the Admin-2 Level . . . . .	94
4.1	Introduction . . . . .	94
4.2	Methods . . . . .	97
4.3	Application in Kenya and Malawi . . . . .	112
4.4	Discussion . . . . .	123
Chapter 5:	Model-based estimation of subnational child mortality at the Admin-2 level with complex surveys . . . . .	127
5.1	Introduction . . . . .	127
5.2	Methods . . . . .	129
5.3	A betabinomial model for estimating U5MR . . . . .	134
5.4	Benchmarking to UN IGME national estimates . . . . .	136
5.5	Country-specific crisis adjustment . . . . .	137
5.6	Application to 22 countries . . . . .	139
5.7	Discussion . . . . .	145
Chapter 6:	Discussion and Future Work . . . . .	147
Appendix A:	Appendix to Chapter 4 . . . . .	161
A.1	Zero Adjustment details . . . . .	161
A.2	Bias in time . . . . .	163
A.3	Hyperparameters and Model parameters . . . . .	163
A.4	National results . . . . .	166
A.5	Period Results by Area: Kenya . . . . .	166
A.6	Yearly Results by Area: Kenya . . . . .	166
A.7	Period Results by Area: Malawi . . . . .	166
A.8	Yearly Results by Area: Malawi . . . . .	166
Appendix B:	Appendix to Chapter 5 . . . . .	169
B.1	Benin . . . . .	169
B.2	Burundi . . . . .	194
B.3	Ethiopia . . . . .	228

B.4	Ghana	254
B.5	Kenya	289
B.6	Lesotho	304
B.7	Liberia	312
B.8	Malawi	334
B.9	Mali	344
B.10	Myanmar	364
B.11	Namibia	386
B.12	Nigeria	417
B.13	Pakistan	430
B.14	Rwanda	436
B.15	Senegal	453
B.16	Sierra Leone	473
B.17	Tanzania	486
B.18	Togo	530
B.19	Uganda	548
B.20	Zambia	581
B.21	Zimbabwe	605

## LIST OF FIGURES

Figure Number	Page
1.1 Progress toward MDGs . . . . .	10
1.2 Administrative divisions . . . . .	11
2.1 The Lexis diagram: Individual . . . . .	17
2.2 The Lexis diagram: Population . . . . .	18
2.3 Period life table: Individuals . . . . .	21
3.1 Lexis vs. life table: Individual . . . . .	44
3.2 Lexis: three cohorts . . . . .	46
3.3 Comparison of deaths: 8 age bands . . . . .	58
3.4 Comparison of deaths: 5 age bands . . . . .	59
3.5 Comparison of deaths: 2 age bands . . . . .	60
3.6 Comparison of deaths: 1 age band . . . . .	61
3.7 Comparison of exposures: 8 age bands . . . . .	62
3.8 Comparison of exposures: 5 age bands . . . . .	63
3.9 Comparison of exposures: 2 age bands . . . . .	64
3.10 Comparison of exposures: 1 age band . . . . .	65
3.11 Comparison of personmonths: 8 age bands . . . . .	66
3.12 Comparison of personmonths: 5 age bands . . . . .	67
3.13 Comparison of personmonths: 2 age bands . . . . .	68
3.14 Comparison of personmonths: 1 age band . . . . .	69
3.15 Comparison of age-specific mortality: 8 age bands . . . . .	72
3.16 Comparison of age-specific mortality: 5 age bands . . . . .	73
3.17 Comparison of age-specific mortality: 2 age bands . . . . .	74
3.18 Comparison of age-specific mortality: 1 age band . . . . .	75
3.19 Comparison of age-specific mortality: 8 age bands . . . . .	76

3.20	Comparison of age-specific mortality: 5 age bands . . . . .	77
3.21	Comparison of age-specific mortality: 2 age bands . . . . .	78
3.22	Comparison of age-specific mortality: 1 age band . . . . .	79
3.23	Comparison of U5MR using E: 8 age bands . . . . .	80
3.24	Comparison of U5MR using E: 5 age bands . . . . .	81
3.25	Comparison of U5MR using E: 2 age bands . . . . .	82
3.26	Comparison of U5MR using E: 1 age bands . . . . .	83
3.27	Comparison of U5MR: 8 age bands . . . . .	84
3.28	Comparison of U5MR: 5 age bands . . . . .	85
3.29	Comparison of U5MR: 2 age bands . . . . .	86
3.30	Comparison of U5MR: 1 age bands . . . . .	87
3.31	Comparison of U5MR, discrete time: 8 age bands . . . . .	88
3.32	Comparison of U5MR, discrete time: 5 age bands . . . . .	89
3.33	Comparison of U5MR, discrete time: 2 age bands . . . . .	90
3.34	Comparison of U5MR, discrete time: 1 age bands . . . . .	91
4.1	Admin-2 by Admin-1: Kenya and Malawi . . . . .	95
4.2	Zeroes Adjustment: KDHS 2003, 1980–1984 . . . . .	102
4.3	Zeroes Adjustment Type: KDHS 2003, 1980–1984 . . . . .	103
4.4	Zeroes Adjustment Map: KDHS 2003, 1980–1984 . . . . .	104
4.5	HIV Adjustments: Kenya, pt. 1 . . . . .	111
4.6	HIV Adjustments: Kenya, pt. 2 . . . . .	112
4.7	HIV Adjustments: Malawi . . . . .	113
4.8	Estimates: Nairobi, Kenya and Lilongwe, Malawi . . . . .	116
4.9	Estimates: Malawi, 2010–2014 . . . . .	117
4.10	Estimates: Kenya, 2010–2014 . . . . .	118
4.11	Estimates: Malawi, 1980–1984 to 2010–2014 . . . . .	119
4.12	Estimates: Kenya, 1980–1984 to 2010–2014 . . . . .	120
5.1	Myanmar areas affected by Cyclone Nargis . . . . .	138
5.2	Subnational variability by country, 2019 . . . . .	140
5.3	Subnational percent decline by country, 1990–2019 . . . . .	141

5.4	Myanmar U5MR during Cyclone Nargis, a comparison with IHME . . . . .	143
5.5	Myanmar expected deaths during Cyclone Nargis, a comparison with IHME	144
A.1	Zero Adjustments: direct vs. smoothed . . . . .	162
A.2	Bias by time prior to survey: Kenya . . . . .	163
A.3	Bias terms for SBH data sources: Kenya and Malawi . . . . .	165
A.4	National results: Kenya . . . . .	167
A.5	National Results: Malawi . . . . .	168
B.1	Admin-1 Names: Benin . . . . .	172
B.2	Admin-1 by space: Benin . . . . .	173
B.3	Uncertainty across Admin-1 areas: Benin , 2019 . . . . .	174
B.4	Admin-2 Names: Benin . . . . .	179
B.5	Admin-2 by space: Benin . . . . .	180
B.6	Uncertainty across Admin-2 areas– Benin , 2019 . . . . .	181
B.7	Admin-1 Names: Burundi . . . . .	197
B.8	Admin-1 by space: Burundi . . . . .	198
B.9	Uncertainty across Admin-1 areas: Burundi , 2019 . . . . .	199
B.10	Admin-2 Names: Burundi . . . . .	205
B.11	Admin-2 by space: Burundi . . . . .	206
B.12	Uncertainty across Admin-2 areas– Burundi , 2019 . . . . .	207
B.13	Admin-1 Names: Ethiopia . . . . .	231
B.14	Admin-1 by space: Ethiopia . . . . .	232
B.15	Uncertainty across Admin-1 areas: Ethiopia , 2019 . . . . .	233
B.16	Admin-2 Names: Ethiopia . . . . .	238
B.17	Admin-2 by space: Ethiopia . . . . .	239
B.18	Uncertainty across Admin-2 areas– Ethiopia , 2019 . . . . .	240
B.19	Admin-1 Names: Ghana . . . . .	257
B.20	Admin-1 by space: Ghana . . . . .	258
B.21	Uncertainty across Admin-1 areas: Ghana , 2019 . . . . .	259
B.22	Admin-2 Names: Ghana . . . . .	263
B.23	Admin-2 by space: Ghana . . . . .	264

B.24 Uncertainty across Admin-2 areas– Ghana , 2019	265
B.25 Admin-1 Names: Kenya	292
B.26 Admin-1 by space: Kenya	293
B.27 Uncertainty across Admin-1 areas: Kenya , 2019	294
B.28 Admin-1 Names: Lesotho	307
B.29 Admin-1 by space: Lesotho	308
B.30 Uncertainty across Admin-1 areas: Lesotho , 2019	309
B.31 Admin-1 Names: Liberia	313
B.32 Admin-1 by space: Liberia	314
B.33 Uncertainty across Admin-1 areas: Liberia , 2019	315
B.34 Admin-2 Names: Liberia	320
B.35 Admin-2 by space: Liberia	321
B.36 Uncertainty across Admin-2 areas– Liberia , 2019	322
B.37 Admin-2 Names: Malawi	337
B.38 Admin-2 by space: Malawi	338
B.39 Uncertainty across Admin-2 areas– Malawi , 2019	339
B.40 Admin-1 Names: Mali	347
B.41 Admin-1 by space: Mali	348
B.42 Uncertainty across Admin-1 areas: Mali , 2019	349
B.43 Admin-2 Names: Mali	353
B.44 Admin-2 by space: Mali	354
B.45 Uncertainty across Admin-2 areas– Mali , 2019	355
B.46 Admin-1 Names: Myanmar	366
B.47 Admin-1 by space: Myanmar	367
B.48 Uncertainty across Admin-1 areas: Myanmar , 2019	368
B.49 Admin-2 Names: Myanmar	373
B.50 Admin-2 by space: Myanmar	374
B.51 Uncertainty across Admin-2 areas– Myanmar , 2019	375
B.52 Admin-1 Names: Namibia	389
B.53 Admin-1 by space: Namibia	390
B.54 Uncertainty across Admin-1 areas: Namibia , 2019	391

B.55 Admin-2 Names: Namibia . . . . .	396
B.56 Admin-2 by space: Namibia . . . . .	397
B.57 Uncertainty across Admin-2 areas– Namibia , 2019 . . . . .	398
B.58 Admin-1 Names: Nigeria . . . . .	420
B.59 Admin-1 by space: Nigeria . . . . .	421
B.60 Uncertainty across Admin-1 areas: Nigeria , 2019 . . . . .	422
B.61 Admin-1 Names: Pakistan . . . . .	431
B.62 Admin-1 by space: Pakistan . . . . .	432
B.63 Uncertainty across Admin-1 areas: Pakistan , 2019 . . . . .	433
B.64 Admin-1 Names: Rwanda . . . . .	439
B.65 Admin-1 by space: Rwanda . . . . .	440
B.66 Uncertainty across Admin-1 areas: Rwanda , 2019 . . . . .	441
B.67 Admin-2 Names: Rwanda . . . . .	445
B.68 Admin-2 by space: Rwanda . . . . .	446
B.69 Uncertainty across Admin-2 areas– Rwanda , 2019 . . . . .	447
B.70 Admin-1 Names: Senegal . . . . .	456
B.71 Admin-1 by space: Senegal . . . . .	457
B.72 Uncertainty across Admin-1 areas: Senegal , 2019 . . . . .	458
B.73 Admin-2 Names: Senegal . . . . .	463
B.74 Admin-2 by space: Senegal . . . . .	464
B.75 Uncertainty across Admin-2 areas– Senegal , 2019 . . . . .	465
B.76 Admin-1 Names: Sierra Leone . . . . .	475
B.77 Admin-1 by space: Sierra Leone . . . . .	476
B.78 Uncertainty across Admin-1 areas: Sierra Leone , 2019 . . . . .	477
B.79 Admin-2 Names: Sierra Leone . . . . .	481
B.80 Admin-2 by space: Sierra Leone . . . . .	482
B.81 Uncertainty across Admin-2 areas– Sierra Leone , 2019 . . . . .	483
B.82 Admin-1 Names: Tanzania . . . . .	489
B.83 Admin-1 by space: Tanzania . . . . .	490
B.84 Uncertainty across Admin-1 areas: Tanzania , 2019 . . . . .	491
B.85 Admin-2 Names: Tanzania . . . . .	499

B.86 Admin-2 by space: Tanzania	500
B.87 Uncertainty across Admin-2 areas– Tanzania , 2019	501
B.88 Admin-1 Names: Togo	533
B.89 Admin-1 by space: Togo	534
B.90 Uncertainty across Admin-1 areas: Togo , 2019	535
B.91 Admin-2 Names: Togo	539
B.92 Admin-2 by space: Togo	540
B.93 Uncertainty across Admin-2 areas– Togo , 2019	541
B.94 Admin-1 Names: Uganda	551
B.95 Admin-1 by space: Uganda	552
B.96 Uncertainty across Admin-1 areas: Uganda , 2019	553
B.97 Admin-2 Names: Uganda	556
B.98 Admin-2 by space: Uganda	557
B.99 Uncertainty across Admin-2 areas– Uganda , 2019	558
B.100Admin-1 Names: Zambia	584
B.101Admin-1 by space: Zambia	585
B.102Uncertainty across Admin-1 areas: Zambia , 2019	586
B.103Admin-2 Names: Zambia	590
B.104Admin-2 by space: Zambia	591
B.105Uncertainty across Admin-2 areas– Zambia , 2019	592
B.106Admin-1 Names: Zimbabwe	608
B.107Admin-1 by space: Zimbabwe	609
B.108Uncertainty across Admin-1 areas: Zimbabwe , 2019	610
B.109Admin-2 Names: Zimbabwe	614
B.110Admin-2 by space: Zimbabwe	615
B.111Uncertainty across Admin-2 areas– Zimbabwe , 2019	616

## LIST OF TABLES

Table Number	Page
2.1 Period life table: Population . . . . .	23
4.1 Data sources and types: Kenya and Malawi . . . . .	96
4.2 Validation: Malawi leave out area-time . . . . .	122
4.3 Validation: Kenya, leave out area-time . . . . .	123
4.4 Validation: Malawi, leave out MW2015 . . . . .	124
4.5 Validation: Kenya, leave out KDHS2014 . . . . .	125
5.1 Surveys, administrative divisions, models by country . . . . .	131
A.1 Distribution of proportion of spatiotemporal variation . . . . .	164
B.1 Data summary for Benin. . . . .	170
B.2 Data summary for Burundi. . . . .	195
B.3 Data summary for Ethiopia. . . . .	229
B.4 Data summary for Ghana. . . . .	255
B.5 Data summary for Kenya. . . . .	290
B.6 Data summary for Lesotho. . . . .	305
B.7 Data summary for Liberia. . . . .	312
B.8 Data summary for Malawi. . . . .	335
B.9 Data summary for Mali. . . . .	345
B.10 Data summary for Myanmar. . . . .	365
B.11 Data summary for Namibia. . . . .	387
B.12 Data summary for Nigeria. . . . .	418
B.13 Data summary for Pakistan. . . . .	430
B.14 Data summary for Rwanda. . . . .	437
B.15 Data summary for Senegal. . . . .	454

B.16 Data summary for Sierra Leone. . . . .	474
B.17 Data summary for Tanzania. . . . .	487
B.18 Data summary for Togo. . . . .	531
B.19 Data summary for Uganda. . . . .	549
B.20 Data summary for Zambia. . . . .	582
B.21 Data summary for Zimbabwe. . . . .	606

## ACKNOWLEDGMENTS

Thank you to my advisor, Professor Jon Wakefield. Thank you for your mentorship and thank you for every opportunity. Thank you for creating an environment where research is fun and enjoyable and collaborative. Thank you, most importantly, for supporting me and allowing me to achieve my goals in the face of extreme personal hardship.

Thank you to my dissertation committee, Professors Jon Wakefield, Emilio Zagheni, Tyler McCormick and Mark Ellis. Thank you for the feedback and constructive criticism that has elevated my work. I would like to thank Emilio Zagheni and the Max Planck Institute for Demographic Research, especially, for the opportunity and the funding to study and work in Rostock.

Thank you to Center for Studies in Demography & Ecology and Professor Sara Curran for the constant mentorship, support, and opportunities. I would not be the scholar I am today without the training I received from CSDE or the funding provided that has allowed me moments in time to focus solely on my research.

Thank you to the advisor of my Master's thesis, Professor Nedret Billor. Thank you for opening up opportunities for me I did not know were possible. War eagle!

Thank you to my colleagues, especially those in STAB and at MPIDR, for helping me grow my confidence as a researcher, demographer, and statistician. Thank you for all the support and friendship you've all provided me over the years.

Thank you to Cafe Solstice Capitol Hill and the Bison family. Thank you for providing a space to complete nearly all of my degree requirements while at the University of Washington. Thank you for making Seattle a home to me, for being an emergency contact when I did not

have one, for being a place of refuge in some of the hardest moments of my life. Thank you for doing the same for so many. I miss you all every day.

Thank you to my family, given and chosen, and my friends for your constant support. Thank you for being all the good things in life and for reminding me of those things when I needed it most. Thank you for making all of this worth it.

## DEDICATION

to my dads, I wish you were here to see me cross the finish line

## Chapter 1

# INTRODUCTION

### **1.1 Motivation**

Child mortality is an important metric used in quantifying and monitoring the health of a population's children. Moreover, child mortality can be a key indicator of the overall health of a population, and is often used to quantify mortality at other ages (Clark, 2019; Wilmoth et al., 2012). In 1990, the United Nations' Millennium Development Goals (MDGs) Target 4.A set a goal of a two-thirds reduction in child mortality between 1990 and 2015 (UN and UN DESA, 2015, p. 32). Globally, the *under-five mortality rate* (U5MR), or the number of deaths before age 5 per 1000 births was halved between 1990 and 2015 (UN and UN DESA, 2015, p. 5). While every region of the globe saw large declines over the period of the MDGs initiative, few regions achieved the two-thirds target (UN and UN DESA, 2015, p. 32). Following the MDGs, the UN announced the Sustainable Development Goals (SDGs) for the period 2015–2030. In particular, SDG 3.2 continues the global initiative to improve child mortality outcomes by calling for an end to preventable child deaths and reaching a target of 25 deaths before age 5 per 1000 births and 12 deaths per 1000 births in the first month of life by the year 2030 (UN and UN DESA-Div. for SDGs, 2015, p. 18).

The top left Figure 1.1 shows the change in the global distribution of posterior median estimates of national U5MR (UN IGME, 2020; Alkema and New, 2014) from the beginning of the MDGs to 2019 – four years into the SDGs. The yellow region indicates the proportion of countries meeting SDG 3.2 with less than 25 deaths per 1000 births before age 5. As of

2019, the proportion of countries achieving SDG 3.2 was 62%, up from 35% in 1990 (UN IGME, 2020).

At the start of the MDGs, the UN defined regions of sub-Saharan Africa and Southern Asia were the only regions of the world with estimated median U5MRs above 100 at 179 and 126 deaths per 1000 births, respectively (UN and UN DESA, 2015, p. 32). Despite huge improvements, sub-Saharan Africa still had the highest estimated U5MR (86) at the end of the MDGs initiative followed by Oceania (51) and Southern Asia (50). In addition to the risk of death before age 5 remaining high in sub-Saharan Africa and Southern Asia as compared to the rest of the globe, the two regions are projected to have the largest populations under age 5 through the remainder of the SDG period (UN and UN DESA, 2015, p. 33). Figure 1.1 shows the remarkable progress of sub-Saharan Africa (top right) and Southern Asia (bottom left) toward child mortality reduction goals. However, when compared to the global distribution in the top left panel, the relatively high burden of child mortality in these two regions is obvious. Most notably, the proportion of countries in sub-Saharan Africa meeting SDG 3.2 has barely changed since 1990.

Sub-Saharan Africa has both the largest number of people in the world both receiving and in need of antiretroviral treatment (ART) for HIV/AIDs (WHO, 2010, p. 53). The vertical white line on the top right panel containing only countries in sub-Saharan Africa indicates the beginning of the UN commitment to the scale-up of treatment and ART for HIV/AIDS in low and middle income countries (LMICs) in hopes of reaching universal coverage by 2010 (UN, 2006). Though some countries began to implement ART interventions as early as 2000, the commitment to scale-up ART coverage and treatment had a huge impact on both mortality due to HIV/AIDs and mother-to-child transmission in the latter half of the decade (WHO, 2010, p. 59-61). The steepest slopes of increase for categories of lower U5MR in Figure 1.1 occur in sub-Saharan Africa over this period of time.

To assess a country's progress toward the MDGs or SDGs, one must be able to count or

estimate the number of deaths before age 5 per 1000 births. Ideally this information comes in the form of dates of birth and either dates of death or ages at death for individuals in the population. Many countries collect this information via the country's vital registration (VR) systems using birth and death certificates. If every birth and death is observed, VR data will describe the true mortality burden of a population. However, even the best VR systems in the world do not reach full coverage of death registration, and VR systems in many LMICs may not cover larger proportions of a population (Mathers et al., 2005). Where VR data is insufficient, household surveys often collect information about child mortality via a mother's full birth history (FBH), or the birth and death date of each of their children (Hill et al., 2012). While the type of information contained in FBH data is the same as VR data, FBH data arising from surveys are based on a sample of the population and, therefore, provide a less precise picture of a population's mortality burden. Censuses and some household surveys in LMICs collect a different type of information on child mortality, a mother's summary birth history (SBH). Instead of reporting the date of birth and age at death of each child, SBH questionnaires ask women of reproductive age to report their age and, often, time since first birth (TSFB), the number of children they have had, or parity, and the number of those children who have died (Hill et al., 2012). Though not the most ideal information on child mortality, SBH data indirectly provide some information on timing of births and deaths in the relationship between the calendar year of data collection, mother's parity, and either age or TSFB. When collected via a census, SBH data sources also cover much more of a population than data arising from a survey. Section 1.2 summarizes the two sources of FBH data and one source of SBH data that will be used in this thesis.

Populations where the number of children expected to die before their 5<sup>th</sup> birthday are relatively high, tend to also have higher mortality burdens at other ages as well (Wilmoth et al., 2012). Unfortunately, countries where child mortality remains high today are likely to also be countries that cannot rely on VR data to record comprehensive individual mortality

information for the population (Hill et al., 2012). While the major multi-national household surveys that measure child mortality provide information about national child mortality and some information about the subnational location of respondents, the number of observations per survey becomes smaller as we concern ourselves with finer levels of geographic division (or any other subgrouping) within a country. SDG 17.18 calls for improvements in data systems in less and least developed countries to provide demographic and health information for subgroups of the population including, “income, gender, age, race, ethnicity, migratory status, disability, geographic location and other characteristics relevant in national contexts” (UN and UN DESA-Div. for SDGs, 2015, p. 17). Additionally, understanding geographic variability in mortality levels and trends can allow for focused resource allocation and intervention on the communities in most need.

Though there are many ways to think about geographic variability within a country, subnational estimation is a catch-all term that can refer to a continuously varying spatial surface within the country or small area estimation in discrete partitions of the nation. Of particular interest for discrete subnational estimation are the administrative divisions of a country that have political and infrastructural meaning. Throughout this dissertation we will refer to the national level as Admin-0, the coarsest administrative division as Admin-1, and the next coarsest administrative division as Admin-2, Figure 1.2 shows these three definitions for the country of Zimbabwe. This dissertation focuses on the synthesis and advancement of statistical and demographic methods for estimation of child mortality from FBH or SBH, with the majority of work focusing on estimating subnational U5MR in at the Admin-1 and Admin-2 levels in the situation in which survey sample sizes are relatively small in each area. Methods of child mortality estimation developed in this thesis are applied to countries in the regions of sub-Saharan Africa and Southern and South-Eastern Asia, where child mortality remains relatively high, but the methods developed can be applied in similar contexts.

## 1.2 Data

### 1.2.1 Demographic and Health Surveys

The Demographic and Health Surveys (DHS) began in 1984 as an endeavor to collect demographic information via household surveys in countries without sufficient VR data (ICF International, 2021a). DHS surveys are funded by United States Agency for International Development (USAID) and often implemented in collaboration with a country's national statistical office (DHS, 2021). Since their inception, DHS has implemented over 400 surveys in over 90 countries in the world (DHS, 2021).

Most DHS surveys follow a similar implementation that relies on partitioning the population into geographic regions called enumeration areas (EAs) such that the number of households in each EA is approximately equal (MEASURE DHS, 2012, p. 5). Before EAs are selected from the sampling frame, they are stratified according to region within the country and to whether they are urban or rural (MEASURE DHS, 2012, p. 7). EAs are then sampled from within each strata, typically with probability proportional to size (PPS) sampling where the size of the EA is measured in number of households. Once EAs are selected within strata, all households within selected EAs are enumerated by the survey team and then households are selected from this list (MEASURE DHS, 2012, p. 2). Within each selected household, all women aged 15–49 who slept in the house the night before are eligible to complete the survey (MEASURE DHS, 2012, p. 15). The questionnaire answered by women includes FBH information for each child they've had, in addition to many more detailed questions about child health for those children born in the 5 years preceding the survey. Nonresponse is quite low in DHS surveys, but is higher amongst individuals with higher levels of education or economic security, those residing in urban areas, and men (MEASURE DHS, 2012, p. 24). In addition to the FBH collected by DHS surveys, this dissertation makes use of the GPS locations of sampled EAs and also maps of subnational divisions of countries

provided by the DHS (DHS, 1990–2021).

### *1.2.2 Multiple Indicator Cluster Surveys*

The Multiple Indicator Cluster Surveys (MICS) were first implemented by the United Nations Children’s Fund (UNICEF) in 1994 as a faster, shorter, and more cost-effective household survey than carried out by the DHS to monitor progress toward international goals in child health (UNICEF, 2015, p. 14-15). Since then, 346 MICS surveys have been carried out in 118 countries (UNICEF, 2021). Many of these surveys follow a similar household selection scheme to that of the DHS; however, other sampling schemes are used when MICS surveys are designed to target specific subgroups or subnational areas (UNICEF, 2015, p. 32-33).

MICS surveys initially collected child mortality information from mothers in the form of SBH information. The 4<sup>th</sup> wave of MICS surveys (MICS4) included the first MICS surveys that collected FBH information, which became standard in MICS5 (UNICEF, 2015, p. 87-91). Chapter 4 uses MICS surveys from Kenya and Malawi spanning the 2<sup>nd</sup> to 5<sup>th</sup> waves, some of which contain SBH data and some of which contain FBH data. While there are many MICS, it wasn’t until MICS6, the most recently completed, that countries were given the option to collect and make available GPS locations of sampled EAs (UNICEF, 2000–2004). Surveys currently in planning are part of the 7<sup>th</sup> wave. Chapter 5 presents methods for subnational child mortality estimation using FBH data from household surveys where GPS locations of EAs are available. Currently, the results presented in that chapter only contain DHS surveys, but future work will incorporate the MICS6 surveys with available GPS data.

### *1.2.3 IPUMS-International*

National censuses have collected SBH information from women since the beginning of the 20<sup>th</sup> century (Hill, 2013b, p. 159). The first methods to use the information contained in census SBH data to estimate population fertility or child mortality were not developed until

the 1960s (Brass, 1964; Brass and Coale, 1968). At the time, countries outside the United States only published or made census data available in aggregate tables (Ruggles et al., 2015). In the late 1990s, individual-level microdata stored on paper was at increasing risk of being lost (Ruggles, 2014). A push was made to expand an ongoing project to make United States census microdata available called the Integrated Public Use Microdata Series (IPUMS) to incorporate microdata from other countries (McCaa, 2013). Over the next two decades IPUMS worked with national statistical agencies to digitize and disseminate international census microdata (Ruggles et al., 2015). To date, IPUMS-International has digitized microdata from over 100 countries in the world (Minnesota Population Center, 2021). A typical IPUMS dataset uses a 10% sample of individuals to anonymize from each census (Ruggles et al., 2015; Ruggles, 2014). Chapter 4 of this thesis uses women’s SBH data and subnational geographic information from IPUMS-International samples for the countries of Kenya and Malawi (Minnesota Population Center, 2019).

### ***1.3 Organization of Dissertation***

The rest of this thesis is organized as follows. Chapter 2 provides an introduction to the demographic and statistical notation, vocabulary, and methods used in Chapters 4–5. Chapter 3 compares demographic and statistical methods for child mortality measures estimated from DHS FBH data with applications to the Senegal 2006 DHS. Chapter 4 and Chapter 5 both explore subnational estimation of child mortality over time in the absence of VR data. The two chapters differ in the underlying statistical methods used to account for the sampling scheme of household survey and in whether the mortality data for individual children in a subnational region of a country at a specific point is aggregated and input into the model at the level of the EA or at the subnational region. Chapter 4 develops new models for estimating subnational mortality over time using both FBH and SBH data, and applies the developed methods to data from Malawi and Kenya. Chapter 5 develops subnational

child mortality estimation methods under a different modeling paradigm than the previous chapter, using only FBH data, and applies the method to 22 LMICs. Finally, Chapter 6 provides a discussion of the work and outlines areas of future research.

#### ***1.4 Methodological Contribution of Dissertation***

After introducing and synthesizing the notation of necessary demographic and statistical concepts in Chapter 2, Chapter 3 compares demographic and statistical methods for period child mortality estimation from FBH data arising from household surveys carried out in LMICs. Different organizations use different approaches to calculate rates of child mortality. In this chapter we compare and contrast these methods and, then, examine the implications of applying these methods using data from the Senegal DHS carried out in 2006. Ultimately, the goals and contributions of this thesis center on expanding existing methods for subnational estimation of child mortality over time using household surveys. The methods described in Chapters 4 and 5 take different statistical approaches to account for the survey sampling scheme of FBH data but are similar in the methods of spatiotemporal estimation used. Chapter 4 extends the spatiotemporal smoothing methods of Mercer et al. (2015) for estimation of child mortality across discrete subnational areas to allow for synthesis over different data sources and at different temporal resolutions. Application to a combination of FBH and SBH data from the countries of Kenya and Malawi provide mixed evidence of the level of benefit of incorporating SBH data. Additionally, current methods cannot be used when the sample sizes in each subnational area and period of time are very small.

The smoothing methods of Mercer et al. (2015) and the temporal smoothing methods introduced in Chapter 4, were previously applied to FBH data from 35 countries to obtain Admin-1 level estimates in Li et al. (2019). Chapter 5 describes the methods development and reproducible computational pipeline that was devised to produce the official UN Admin-2 estimates of U5MR for 22 countries in sub-Saharan Africa, Southern Asia, and South-

Eastern Asia. This work was done in collaboration with a multi-national team of colleagues at the University of Washington and the United Nations' Inter-agency Group for Child Mortality Estimation (UN IGME). After country consultations conducted by UNICEF, these results have been published by UN IGME and UNICEF and can be found at <https://childmortality.org> (2020). The work pipeline include steps to: decide which countries and DHS surveys for which reasonable estimates can be made, acquiring and cleaning these data – including resolving geographic anomalies, software development, methods development, model choice, data analyses, and final reporting of results in meaningful and understandable ways.

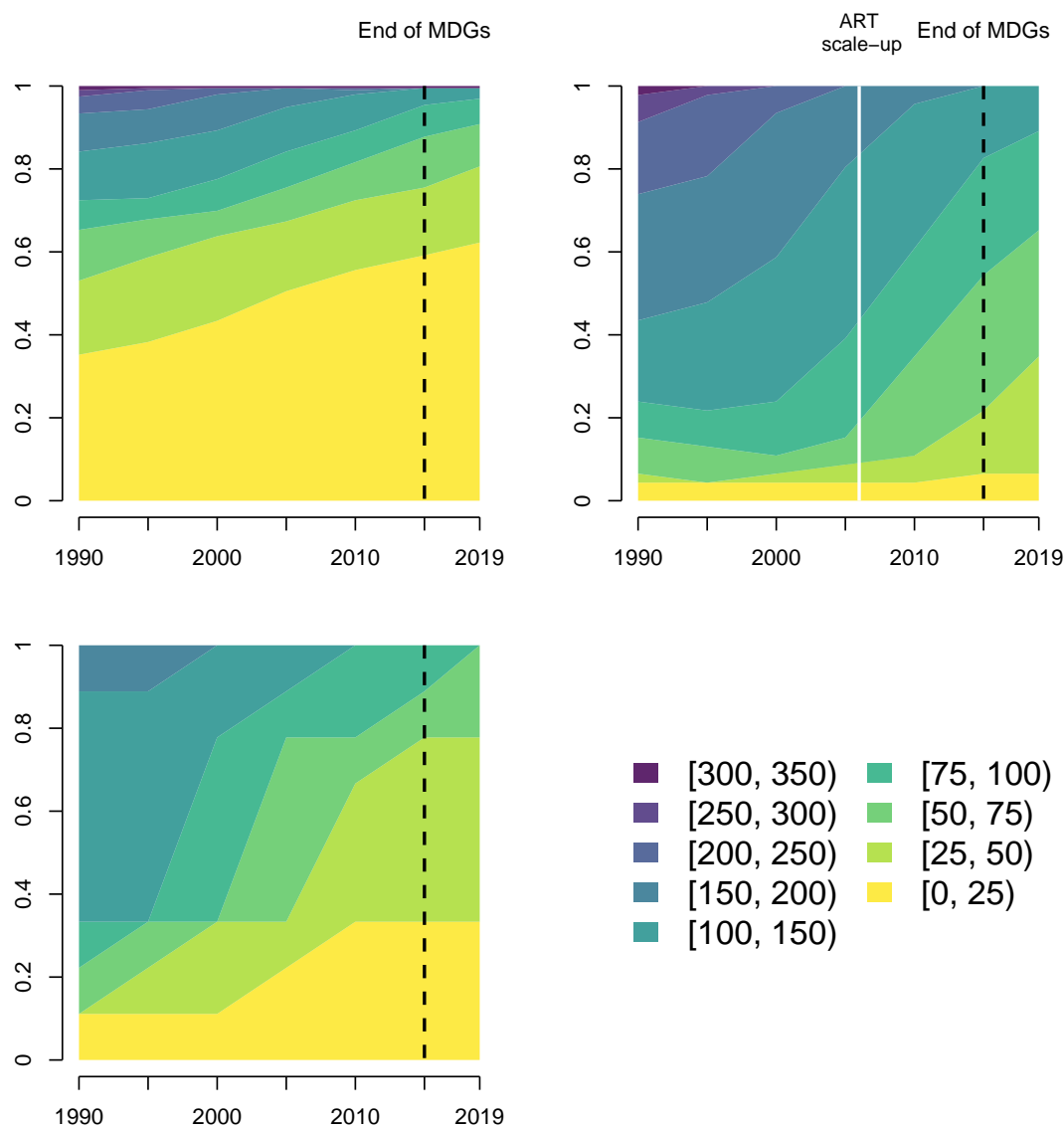


Figure 1.1: Change in distribution over the years 1990–2019 of the proportion of countries whose estimated posterior median number of deaths before age 5 per 1000 live births lie within the ranges that correspond to each color (UN IGME, 2020). The yellow portion of each plot represents countries who have achieved SDG 3.2. The vertical dashed line indicates the year 2015, and the transition from the MDGs to the SDGs. **Top Left:** All countries. **Top Right:** Countries in sub-Saharan Africa. **Bottom Left:** Countries in Southern Asia.

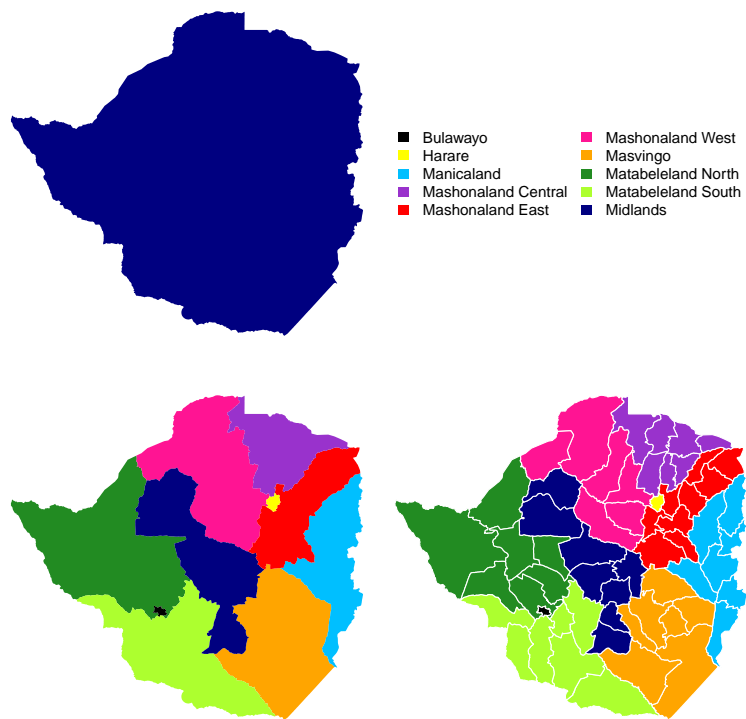


Figure 1.2: From left to right the Admin-0, Admin-1, and Admin-2 administrative divisions of Zimbabwe.

## Chapter 2

### BACKGROUND

#### 2.1 Notation

Throughout this dissertation we will default to the canonical demographic notation for age-specific mortality rates and probabilities. We will denote probabilities of death by age  $x + n$  conditional on survival to age  $x$  as follows:

$${}_nq_x = \Pr(\text{death before reaching age } x + n | \text{survival to age } x). \quad (2.1)$$

We will denote the mortality rates between ages  $x$  and  $x + n$  as

$${}_nM_x = \frac{\text{Number of deaths in age range } [x, x + n)}{\text{Number of units of person-time lived in age range } [x, x + n)}. \quad (2.2)$$

Note that  ${}_nM_x$  will refer to the observed mortality rate for age group  $[x, x + n)$ , and, when appropriate,  ${}_nm_x$  will refer to the underlying true age-specific mortality rate in a hypothetical infinite population. Both age-specific probabilities and age-specific rates can be calculated for a pre-specified birth *cohort* or for a pre-specified *period* of calendar time. A cohort mortality measure for a particular age band,  $[x, x + n)$ , will describe the mortality experiences of only people that fall within the cohort defined by shared calendar year of birth. A period mortality measure for the age band  $[x, x + n)$  will describe the mortality experiences of anyone who was that age over a period of calendar time regardless of birth cohort. Taken together, a set of age-specific cohort mortality measures over many age bands reflect the lived mortality experiences of a single subgroup of a population that move through those age bands together over calendar time. On the other hand, a set of age-specific period mortality measures reflect

the mortality experiences of an entire population during a single period of calendar time by partitioning the population into their respective age bands during the period.

This dissertation focuses exclusively on age-specific period measures of mortality. In particular, the remainder of this dissertation will focus on measures of child mortality up to five years of age. Typically risk of death for children is at its highest in the first month of life and decreases with age until adolescence and young adulthood, regardless of the overall level of mortality in a population (Lawn et al., 2005). To that end, there are three canonical indicators of child mortality we will refer to throughout this thesis. The *neonatal mortality rate* (NMR) is the probability of death within the first month of life, the *infant mortality rate* (IMR) refers to the probability of death in the first year of life, and the *under-five mortality rate* (U5MR) is the probability of death within the first five years of live. Though all of these commonly used measures contain the word “rate” they are actually probabilities, usually defined as the number of deaths before the respective age divided by the number of births in the population. Often the literature will use IMR and  ${}_1q_0$  and U5MR and  ${}_5q_0$  interchangeably, counting ages in years. However, unless otherwise noted, throughout this dissertation we will measure age in months,  $m = 0, 1, 2, \dots$ , so that  ${}_1q_0$ ,  ${}_{12}q_0$ , and  ${}_{60}q_0$  refer to NMR, IMR, and U5MR respectively. It is common (Population Division, 2011; Rutstein et al., 2006; Elkasabi, 2019) to partition the first 60 months of life into a smaller number of age bands  $[a_1, a_2)$ , where  $0 \leq a_1 < a_2 \leq 60$ , instead of focusing on probabilities  ${}_1q_m$  for  $m = 0, \dots, 59$ . Unless otherwise noted we partition the first 60 months of life into six age-bands below, following previous authors (Li et al., 2019; Wakefield et al., 2019; Mercer et al., 2015):

$$a[m] = \begin{cases} 1, & 0 \leq m < 1 \\ 2, & 1 \leq m < 12 \\ 3, & 12 \leq m < 24 \\ 4, & 24 \leq m < 36 \\ 5, & 36 \leq m < 48 \\ 6, & 48 \leq m < 60 \end{cases}. \quad (2.3)$$

When estimating period child mortality from FBH data collected via complex household surveys, we will follow conventional notation from survey statistics and refer to child  $k$  from cluster  $c$  within strata  $h$ . For a calendar period of interest,  $t$ , FBH data from child  $k$  will be first expanded into set of binary indicator for each month of life starting at birth,  $m = 0$ , until, either, the age month in which death occurred or the month of observation if the child is alive at the time of survey. The observations corresponding to  $m \in [0, 60)$  are retained. These binary indicators are further collapsed into the number of months of life in age group  $a[m]$  child  $k$  begins alive,  $n_{a[m],tk}$ , and a binary indicator,  $y_{a[m],tk}$ , that takes the value 1 if child  $k$  dies before the end of age band  $a[m]$  and 0 otherwise. For any age group  $a[m]$  child  $k$  does not reach before death or time of observation,  $n_{a[m],tk} = 0$  and no observations are contributed. The retrospective nature of FBH data means that multiple surveys can contribute data to the estimation of a country's child mortality in a particular period of time; where necessary, surveys within a country are indexed by  $s$ . When subnational mortality is the indicator of interest,  $i$  will index administrative units at either Admin-1 or Admin-2 level.

The rest of this chapter will first introduce demographic (Section 2.2) and statistical (Section 2.3) methods for estimating child mortality in a single period of calendar time. These sections will frequently refer to an example population observed over the calendar time  $[t_0, t_3)$ , partitioned by discrete time points  $\{t_0, t_1, t_2, t_3\}$ , comprised of exactly three

cohorts. The three cohorts are denoted by the left-endpoint of their cohort-defining period of calendar time, e.g., members of Cohort  $t_0$  are born in the interval  $[t_0, t_1)$ . The interval  $[a_0, a_3)$  is partitioned into discrete age groups for which measures of period mortality by the ages  $\{a_0, a_1, a_2, a_3\}$ . Where this example is not used, notation will revert to the generic notation using age  $[x, x + n)$ . Section 2.4 will cover statistical methods for inference with complex survey data. Finally, Section 2.5 introduces Bayesian methods that will be used to smooth child mortality estimates over space and time.

## 2.2 Demographic methods for period child mortality estimation

*Direct estimation* is a demographic term that refers to estimation of measures of a population's mortality based on information about individuals' birth dates and death dates or ages at death (Hill, 2013a). On the other hand, *indirect estimation* describes demographic methods for mortality estimation from data that do not contain information regarding timing of each individual's birth and death (Hill, 2013b; Preston et al., 2001). In many nations, vital registration systems provide the necessary information for direct estimation in the form of birth and death certificates. However, the focus of this dissertation is on the estimation of child mortality in countries that rely on household surveys like DHS and MICS to collect individual's demographic and health information. FBH data collected from household surveys allow for use of direct estimation techniques. National censuses and some household surveys collect summary birth histories from mothers, which require indirect estimation methods. Section 2.2.1 and Section 2.2.2 describe direct and indirect mortality estimation methods, respectively.

### 2.2.1 Direct estimation

This subsection introduces two different demographic frameworks for summarizing the age-specific mortality experiences of both individuals and populations. Both frameworks sum-

marize the observed mortality experience of an individual using line segments called *lifelines* that extend from time of birth to time of death or, if an individual is alive, to time of observation. However, the two frameworks differ in how they an individual's lifeline and in how they display summarize the collective mortality experiences of whole cohorts or populations. As with many traditional demographic methods, those introduced in this subsection to describe the age-specific mortality schedule of a finite population, are deterministic. When age-specific mortality measures for a population are estimated from a sample of the population, these methods provide no uncertainty for these estimates.

### *Using the Lexis diagram*

One approach to direct estimation of period mortality involves careful consideration of the relationship between the period of calendar time for which an estimate of child mortality in a population is desired and the ages and birth cohorts of each child in the population. The Lexis diagram is graphical tool for displaying mortality data as a function of age, period, and cohort (Preston et al., 2001, Ch. 2). The generic Lexis diagram uses the horizontal axis to represent calendar time, the vertical axis to represent age, and cohorts pass through the two societal measures of time diagonally. Typically, a Lexis diagram measures age and time in the same units though this is not necessary.

Figure 2.1 uses a Lexis diagram to summarize the mortality information of individuals in an example population. The color of the line segment is blue for the individual in Cohort  $t_0$ , red for the individual in Cohort  $t_1$ , and yellow for the individual in Cohort  $t_2$ . The lifelines for each individual begin at the point  $(t^*, a_0)$ , where  $t^*$  is the time each individual reaches age  $a_0$ . The lifelines increase (with a slope of 1, if age and time are measured in the same units) until the point  $(t^* + a^*, a^*)$ , where  $a^*$  is the age at death of the individual. Age at death is further highlighted with the dashed lines. Despite the two individuals in blue and red being a part of separate cohorts, we can see they both die between ages  $[a_1, a_2)$  during

the period of interest,  $[t_2, t_3)$ , highlighted in grey.

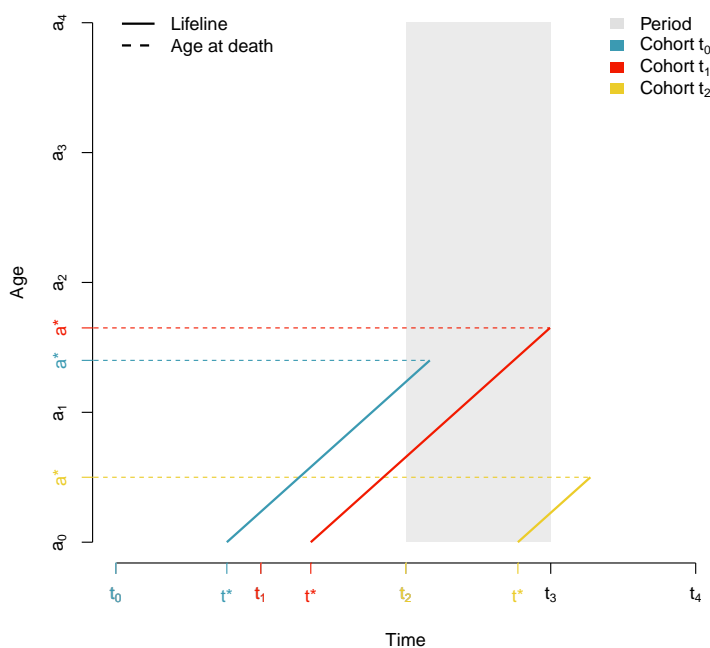


Figure 2.1: Each solid line is a lifeline for a single individual, that starts when the individual reaches ages  $a_0$  at time  $t^*$  and increases diagonally until age at death,  $a^*$ . The color of the life line indicates the cohort of the individual—  $[t_0, t_1)$  (blue),  $[t_1, t_2)$  (red), and  $[t_2, t_3)$  (yellow). The grey rectangle indicates the period of interest  $[t_2, t_3)$ .

Alternatively, a Lexis diagram can summarize the collective mortality experiences of entire cohorts, as in Figure 2.2. Using the intersection of horizontal lines representing units of age, vertical lines representing units of calendar time, and diagonal lines representing cohorts born at each unit of calendar time, the Lexis diagram is divided into triangles. To summarize collective population mortality information, total numbers of deaths,  ${}_nD_x$ , can be displayed in the proper age-period-cohort triangle, e.g., as in Figure 2.2. The number of

people from each cohort to reach age  $x$  is denoted  $B_x$ , so when  $x = 0$  this number equals the initial size of the respective birth cohort.

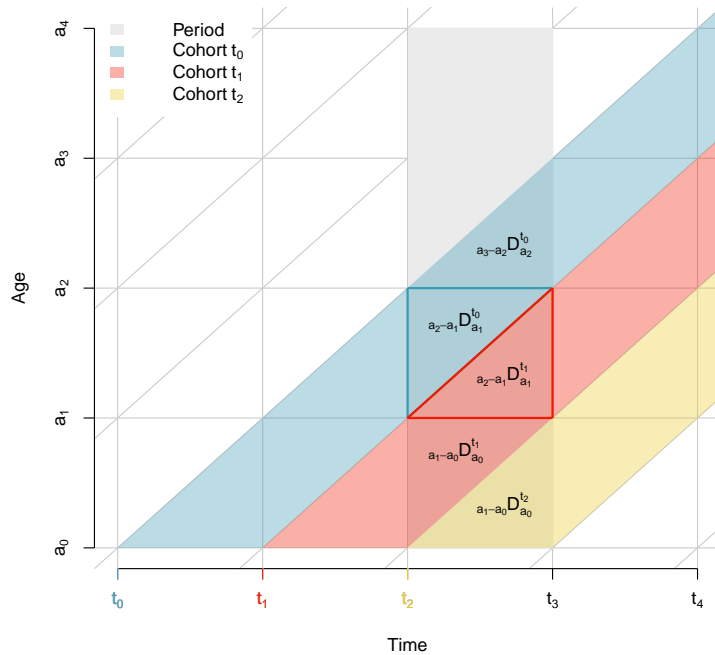


Figure 2.2: Each parallelogram represents the collective mortality experiences of an entire birth cohort. Numbers of deaths,  $D$ , denote observed deaths in each age band (subscript) and birth cohort (superscript) triangle during the period of interest. The color of the life line indicates the cohort of the individual—  $[t_0, t_1)$  (blue),  $[t_1, t_2)$  (red), and  $[t_2, t_3)$  (yellow). The grey rectangle indicates the period of interest  $[t_2, t_3)$ .

If a population consists of exactly one cohort, age-specific the observed probability of death defined in (2.1) can be calculated as

$${}_nq_x = \frac{{}_nD_x}{B_x}. \quad (2.4)$$

When a population consists of more than one cohort, the calculation is slightly more complicated. In our example population in Figure 2.2 the translucent blue, red, and yellow regions between the diagonal cohort lines represent the regions of the Lexis diagram in which the lifelines of individuals of a population comprised of Cohorts  $t_0$ ,  $t_1$ , and  $t_2$  could be drawn respectively. The region defined by the intersection of age group  $[a_1, a_2)$  and period of interest  $[t_2, t_3)$  contains the mortality experience of some individuals born in  $t_0$  (the triangle outlined in blue) and people born in  $t_1$  (the triangle outlined in red).

Following Chapter 2 of Preston et al. (2001), for the period  $[t_2, t_3)$ ,

$$\begin{aligned}
 {}_{a_2-a_1}q_{a_1} &= \Pr(a^* < a_2 | a^* \geq a_1) \\
 &= \underbrace{\Pr(a^* < a_2 \cap t^* \in [t_0, t_1) | a^* \geq a_1)}_{\text{Contribution from Cohort } t_0} \\
 &\quad + \underbrace{\Pr(a^* < a_2 \cap t^* \in [t_1, t_2) | a^* \geq a_1)}_{\text{Contribution from Cohort } t_1} \\
 &= \Pr(a^* < a_2 | a^* \geq a_1, t^* \in [t_0, t_1)) \times \Pr(t^* \in [t_0, t_1) | a^* \geq a_1) \\
 &\quad + \Pr(a^* < a_2 | a^* \geq a_1, t^* \in [t_1, t_2)) \times \Pr(t^* \in [t_1, t_2) | a^* \geq a_1) \\
 &= \frac{{}_{a_2-a_1}D_{a_1}^{t_0}}{B_{a_1}^{t_0}} \times \frac{B_{a_1}^{t_0}}{B_{a_1}^{t_0} + B_{a_1}^{t_1}} + \frac{{}_{a_2-a_1}D_{a_1}^{t_1}}{B_{a_1}^{t_1}} \times \frac{B_{a_1}^{t_1}}{B_{a_1}^{t_0} + B_{a_1}^{t_1}} \tag{2.5}
 \end{aligned}$$

$$= \frac{{}_{a_2-a_1}D_{a_1}^{t_0} + {}_{a_2-a_1}D_{a_1}^{t_1}}{B_{a_1}^{t_0} + B_{a_1}^{t_1}} \tag{2.6}$$

### *Using the life table*

The *period life table* is another demographic tool commonly used to quantify the underlying age-specific probabilities in a population at a given point in calendar time. Contrary to the Lexis diagram approach described in Section 2.2.1, a life table approach assumes all members of a cohort are born at the same moment in calendar time, reducing the temporal dimension of the problem by one (Preston et al., 2001, Ch. 3). Graphically, these assumptions allow one

to consider lifelines that grow in length only along a single temporal dimension, as opposed to diagonally through age and time according to their cohort designation, as in the Lexis diagram.

Figure 2.3 shows the lifelines for the three individuals in the left panel of Figure 2.1 using the same cohort-specific color scheme. The horizontal axis above each set of lifelines indicate the age axis for the individual's cohort. The translucent lifelines show the actual time of birth and death for each individual. The solid lifelines drawn are under the assumption that all members of Cohort  $t_0$  are born at time  $t_0$ , all members of Cohort  $t_1$  are born at time  $t_1$ , and all members of Cohort  $t_2$  are born at time  $t_2$ . Though Figure 2.1 clearly shows that the individuals represented in blue and red died between ages  $[a_1, a_2)$  during the period  $[t_2, t_3)$ , Figure 2.3 shows that the assumptions used to collapse the temporal dimension lead us to incorrectly assign the death of the individual in Cohort  $t_0$  to the period  $[t_1, t_2)$ .

A period life table can be used to summarize the collective mortality experience of a population during a specific moment in time. When age and time are measured in the same units, each age-specific period mortality measure is informed by a single cohort, in direct contrast to the Lexis diagram framework. If we compare the age axes each cohort in Figure 2.3 to the period of interest  $[t_2, t_3)$  shaded in grey, we can see visually that Cohort  $t_0$  is contributing to period age-specific mortality for the age group  $[a_2, a_3)$ , Cohort  $t_1$  contributes to the previous age band  $[a_1, a_2)$ , and Cohort  $t_2$  contributes to the youngest ages  $[a_0, a_1)$ . Chapter 3 of this dissertation will further explore the differences between summarizing population mortality experiences with these assumptions or with the Lexis diagram.

A life table summarizes the age-specific mortality experiences of an entire population, or collection of individual lifelines like those in Figure 2.3, by describing the expected life course of a *hypothetical cohort* born today if they were to pass through the population's observed mortality in each age group until they die out (Preston et al., 2001, Ch. 3). Table 2.1 displays age-specific mortality information over  $[t_2, t_3)$  for the same population as the Lexis diagram

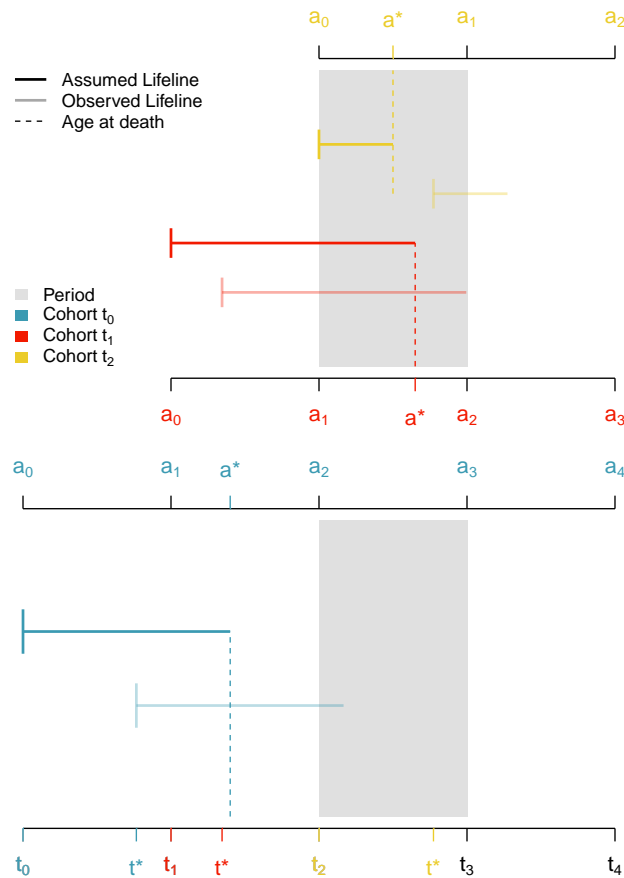


Figure 2.3: Lifelines for the same individuals from Figure 2.1 plotted horizontally in calendar time, represented by the bottom axis, where color represents the individual's cohort. Translucent lifelines connecting points  $t^*$  and  $t^* + a^*$  represent each individual's true calendar times at birth and death. Solid lifelines show the lifeline for each individual under the assumption they were born at the beginning of their cohort-defining period. Axes above each set of lifelines represent the age of the cohort indicated by color at each point in calendar time. The period of interest is represented by the solid grey rectangle.

in the right panel of Figure 2.1 under the assumptions of cohort birth-timing displayed in Figure 2.3. The color of the first three rows of the first column of Table 2.1 reflect the sole cohort whose individuals inform age-specific mortality in each row for the period  $[t_2, t_3)$ . Following traditional demographic notation, the column names of Table 2.1 denote the following hypothetical cohort quantities

- $x$ : the left-inclusive beginning of the age band represented by each row,
- $l_x$ : the number of people to reach age  $x$
- ${}_n d_x$ : the number of deaths in age band  $[x, x + n)$ ,
- ${}_n q_x$ : the probability of death by age  $x + n$ , given survival to age  $x$
- ${}_n L_x$ : the units of person-time lived between ages  $x$  and  $x + n$ ,
- ${}_n m_x$ : the mortality rate in age band  $[x, x + n)$ .

Note, this is not an exhaustive list of all possible columns a period life table may contain. Some of the most important mathematical relationships between columns of a life table include:

$$l_{x+n} = l_x - {}_n d_x,$$

$${}_n q_x = {}_n d_x \times l_x,$$

and

$${}_n m_x = {}_n d_x \times {}_n L_x.$$

The starting size of the hypothetical cohort represented in a period life table,  $l_0$ , is called the *radix*. The radix of a life table may be chosen to reflect the actual size of an existing population, but is often chosen to be a multiple of 10 for ease of calculation and interpretation.

$x$	$l_x$	${}_n d_x$	${}_n q_x$	${}_n L_x$	${}_n m_x$
$a_0$	$l_0$	${}_{a_1-a_0} D_{a_0}^{t_2}$	$\frac{a_1-a_0 D_{a_0}^{t_2}}{B_{a_0}^{t_2}}$	$(a_1 - a_0) \times (B_{a_0}^{t_2} - \frac{1}{2} D_{a_0}^{t_2})$	$\frac{a_1-a_0 d_{a_0}}{a_1-a_0 L_{a_0}}$
$a_1$	$l_0 - {}_{a_1-a_0} d_{a_0}$	${}_{a_2-a_1} D_{a_1}^{t_1}$	$\frac{a_2-a_1 D_{a_1}^{t_1}}{B_{a_1}^{t_1}}$	$(a_2 - a_1) \times (B_{a_1}^{t_1} - \frac{1}{2} D_{a_1}^{t_1})$	$\frac{a_2-a_1 d_{a_1}}{a_2-a_1 L_{a_1}}$
$a_2$	$l_{a_1} - {}_{a_2-a_1} d_{a_1}$	${}_{a_3-a_2} D_{a_2}^{t_0}$	$\frac{a_3-a_2 D_{a_2}^{t_0}}{B_{a_2}^{t_0}}$	$(a_3 - a_2) \times (B_{a_2}^{t_0} - \frac{1}{2} D_{a_2}^{t_0})$	$\frac{a_3-a_2 d_{a_2}}{a_3-a_2 L_{a_2}}$
$a_4$	$l_\infty$	$l_\infty$	1		

Table 2.1: Life table summarizing the age-specific mortality during the period  $[t_2, t_3)$  under the assumption that each cohort's members are all born at the beginning of the cohort-defining interval. Column 1 denotes age group. The color of the first three entries reflect the cohort contributing to the age group during this period. Columns 2 and 3 contain the number of individuals that survive to the start of each row's age group and the number of those individuals who die, respectively. Column 4 contains the probability of death before reaching the end of the age group, conditional on surviving to the start. Column 5 column contains the person-time lived in the age group. The final column contains the mortality rate in the age group.

When the exact age at death,  $x^*$ , is known for every individual, the person-time lived can be calculated

$${}_n L_x = \underbrace{n \times (l_x - {}_n d_x)}_{x^* \geq x+n} + \underbrace{\sum_{k=1}^{n d_x} (x_k^* - x)}_{x^* \in [x, x+n)}.$$

In many cases, however, the person-time cannot be directly observed and the contribution to  ${}_n L_x$  for those who die in  $[x, x + n)$  is made based on assumptions about the average age at death in the age interval, denoted  ${}_n a_x$  (Preston et al., 2001, Ch. 3). Though not included in the simple example in Table 2.1,  ${}_n a_x$  will be discussed in further detail in Chapter 3. The last row of any life table represents all remaining members of the hypothetical cohort alive at the start of the last age group, denoted  $l_\infty$ . By construction,  $l_\infty$  is also equal to the number

of deaths in the last age group and the probability of death for this age group is 1. Though death of each individual of a population surviving the start of the last age group is certain, it has not always occurred by the end of the period being described. The bottom row of Table 2.1 leaves the entries  ${}_nL_x$  and  ${}_nm_x$  empty for this reason. Using Table 2.1, one might calculate an alternative estimate to that in (2.6),

$$\begin{aligned} {}_{a_2-a_1}q_{a_1} &= \Pr(a^* < a_2 | a^* \geq a_1) \\ &= \frac{{}_{a_2-a_1}D_{a_1}^{t_1}}{B_{a_1}^{t_1}}. \end{aligned}$$

### 2.2.2 Indirect estimation of period child mortality

This section describes an indirect estimation technique for child mortality that relies on summary birth histories collected by censuses and some MICS surveys called *the Brass method*. SBH data include only the age of the mother, the number of children ever born (CEB) to her, and the number of those children that ever died (CED). They contain no information on birth date or age at death. A number of different methods have been described to analyze summary birth history data (Brass, 1964; Rajaratnam et al., 2010; Hill et al., 2015; Brady and Hill, 2017; Burstein et al., 2018; Wilson and Wakefield, 2020). The Brass method is the most commonly-used method to analyze SBH data (Hill et al., 1983; Preston et al., 2001), and has the most straightforward implementation. The original method used the proportion of children ever died to children ever born in five-year age groups of mothers to make an estimate of  ${}_60q_0$ , the U5MR (Brass, 1964).

Specifically, mothers ages are broken down into five-year bands, 15 – 19, . . . , 45 – 49. The proportion of children who have died to mothers in age group  $a$ ,  $d_a$ , is first calculated. This proportion is a function of the underlying probabilities of giving birth at different ages and the child mortalities by age of the population over the relevant time period. Brass equated the proportions of dead children,  $d_a$ , to different mortalities depending on the associated

age group: the dead proportion in the 15–19 group corresponds to death within the first year of life, the 20–24 group to death in the first two years of life, and the 30–34 age group corresponds to death in the first five years of life (Preston et al., 2001, p. 228). For the age groups that do not directly equate to U5MR (i.e., all but the 30–34 year age group), a life table can be leveraged to convert to U5MR. Extensions have been made to account for parity of women by age groups in the sample and to attach a reference date of the estimate, a year in time prior to the census or survey (Hill and Trussell, 1977; Coale and Trussell, 1977).

The proportion  $d_a$  approximates death before age  $x$  for some age  $x$  that depends on  $a$ . This basic relationship will change as a function of the reproductive histories of the women who supplied the data. Details on these histories are unavailable, and instead information on the parities in the relevant population are used in an empirical model that adjusts the mortality estimate. Specifically, for each age group, the following adjustment is used,

$${}_xq_0 = d_a \left( b_{1a} + b_{2a} \frac{P_1}{P_2} + b_{3a} \frac{P_2}{P_3} \right),$$

where  $P_1, P_2, P_3$  are average parities of mothers in age groups 15–19, 20–24 and 25–29, respectively, that provide some empirical estimates of birth timing. The coefficients  $b_{1a}, b_{2a}, b_{3a}$  were estimated via simulation using model life tables for fertility and mortality schedules (Trussell, 1975; Hill and Trussell, 1977). Different coefficients are used in different settings, and we use the Coale and Demeny (C-D) North life table coefficients in our analysis, which is the version that is relevant for Kenya and Malawi (Coale and Demeny, 1966). These coefficients are available in Table 47 of Hill et al. (1983).

The final step is to assign the estimate to a relevant period in the past, and another set of simulated coefficients,  $c_{1a}, c_{2a}, c_{3a}$ , are used in conjunction with the parity ratios to estimate a reference date (Coale and Trussell, 1977),

$$t = c_{1a} + c_{2a} \frac{P_1}{P_2} + c_{3a} \frac{P_2}{P_3}.$$

Thus, the age group 30–34 directly produces an estimate of U5MR in year  $t$ ,  ${}_{60}q_{0,t}$ .

### 2.3 Statistical methods for mortality estimation

The direct estimation methods described in Section 2.2.1 are deterministic. When the deaths and births in a population are (nearly) fully observed, this is (nearly) appropriate. However, if observations of child mortality events contain measurement error in birth or death dates or data come from a sample much smaller than the population size, statistical methods must be used to quantify the uncertainty in period child mortality measures. Classic deterministic demographic methods first developed for populations with VR data and infrastructure will be inappropriate for many LMICs unless adapted to account for smaller sample sizes and methods of data collection in those countries. This section covers some statistical methods and assumptions that are used to estimate age-specific mortality for a single period in time, under the assumptions of simple random sampling (SRS).

#### 2.3.1 Binary outcomes: finite and infinite populations

Returning to the notation used to describe our example population in Section 2.2, let  $t_k^*$  be the measured calendar time of birth or reaching age  $a_0$  for individual  $k$  and  $a_k^*$  be the age at death if individual  $k$  has died at the time of observation. We will focus again on estimation of  ${}_{a_2-a_1}q_{a_1}$ , or the probability of death for an individual who is alive at  $t_2$ , whose age at death,  $a^*$ , lies in  $[a_1, a_2)$ , and time at death,  $t^* + a^*$ , occurs before  $t_3$ , during period  $[t_2, t_3)$ . Suppose the population contains  $N$  individuals such that  $t_2 \leq t_k^* + a_k^*$  and  $t_2 \in [t_k^* + a_1, t_k^* + a_2)$ . Then, define

$$y_k = \begin{cases} 1, & a_1 \leq a_k^* < a_2, \\ 0, & \text{otherwise.} \end{cases} \quad (2.7)$$

for  $k = 1, \dots, N$  so that

$$D = \sum_{k=1}^N y_k \quad (2.8)$$

will represent the total number of deaths in age group  $[a_1, a_2)$  during period  $[t_2, t_3)$  in the population. If we select  $n$  individuals from the population via SRS, then the number of deaths in the sample is

$$d = \sum_{k=1}^n y_k. \quad (2.9)$$

Statistical inference requires assumptions about the probability distribution of  $d$ , and there are two natural choices. Point estimation of the target of inference,  ${}_{a_2-a_1}q_{a_1}$ , is identical under either assumption, but variance estimation is not. Under the assumption  $d \sim \text{Binomial}(n, {}_{a_2-a_1}q_{a_1})$ ,

$${}_{a_2-a_1}\hat{q}_{a_1} = \frac{d}{n} \quad (2.10)$$

$$\widehat{\text{Var}}({}_{a_2-a_1}\hat{q}_{a_1}) = \frac{{}_{a_2-a_1}\hat{q}_{a_1}(1 - {}_{a_2-a_1}\hat{q}_{a_1})}{n} \quad (2.11)$$

If one assumes, instead, that  $d \sim \text{HyperGeometric}(N, D, n)$  where  ${}_{a_2-a_1}q_{a_1} = \frac{D}{N}$ ,

$${}_{a_2-a_1}\hat{q}_{a_1} = \frac{d}{n} \quad (2.12)$$

$$\widehat{\text{Var}}({}_{a_2-a_1}\hat{q}_{a_1}) = \frac{{}_{a_2-a_1}\hat{q}_{a_1}(1 - {}_{a_2-a_1}\hat{q}_{a_1})}{n} \times \left(1 - \frac{n}{N}\right). \quad (2.13)$$

The variance estimator is constructed so that as the sample size grows to cover the population, or  $n \rightarrow N$ , the uncertainty in the estimate decreases until it reaches 0 when the population has been fully observed.

Differences in distributional assumptions almost always imply differences in statistical inference from sampled data, regardless of context or application. The difference in binomial and hypergeometric distributions can be described using the examples of flipping coins and drawing from an urn. In the case of flipping a coin, the target of inference represents the true, unknown and unknowable, probability of success in an infinite number of trials, i.e. the probability of seeing heads if the same coin were flipped in exactly the same way and in exactly the same conditions an infinite number of times or assuming each coin flip's

probability of seeing heads is drawn independently from a distribution with a common mean. In the case of drawing from an urn that contains  $N$  balls,  $D$  of which are red and  $N - D$  of which are not red, the target of inference represents the true, possibly knowable, proportion of red balls in the urn. In this case, the target of inference can only be interpreted as probability of success in a single trial, i.e. a single red ball is drawn from the urn when it is full and no more are drawn from the remaining  $N - 1$ . If data arising from the  $n$  draws from the urn are assumed to follow a binomial distribution, the target of inference would be interpreted as the true probability of drawing a red ball from the urn in an infinite number of replicates of the experiment of drawing a single ball from the urn and replacing it before the next draw. Though this does not reflect the actual method of data collection, the consequences of assuming a binomial distribution over a hypergeometric distribution are small unless  $N$  is small or  $\frac{n}{N}$  is not close to 0.

Under a binomial distribution  ${}_{a_2-a_1}q_{a_1}$  is, then, the true underlying probability or *risk* of death between ages  $[a_1, a_2)$  of an individual who survives to age  $t_2$  who has already reached age  $a_1$  or will before  $t_3$  if this experiment were repeated an infinite number of times. In the context of age-specific mortality, this is analogous to assuming the outcomes for the  $n$  observations in our sample represent outcomes from an experiment that could, hypothetically, be repeated a countably infinite number of times. In other words, this assumes the  $n$  observations of  $y_k$  are a sample from a hypothetical *infinite population* or *superpopulation*. Under a hypergeometric distribution  ${}_{a_2-a_1}q_{a_1}$ , represents, instead, the true population *prevalence* of death in ages  $[a_1, a_2)$  during a period  $[t_2, t_3)$  amongst those alive at time  $t_2$  in a *finite population*. Alternatively, it might represent the probability of selecting a single record from  $N$  fully observed vital records of all individuals in the population who meet the age and period criteria for  $[a_1, a_2) \times [t_2, t_3)$  and finding the individual died between ages  $[a_1, a_2)$  in period  $[t_2, t_3)$ .

While these two methods of estimation of a population's mortality for a single age group

in a single time period provide estimates of uncertainty, unlike the demographic methods mentioned previously, they still leave something to be desired. In reality, births and deaths do not happen on the same time scale as the discrete scales by which we as a society define or collect information about those events. Choices of units used to discretize calendar time and, especially, age are largely informed by typical societal measures each of these things, but are also informed by the method in which a population records an observed birth or death. If births and deaths are observed through vital registration, units of time and age can be chosen according to the finest level of time records, usually days, but often, minutes. When births and deaths must be observed through survey interviews where responses are not informed by vital records, time will be discretized according to both the units reported by the respondent and, also, the units in which the survey instrument records time and age.

Time periods of interest are chosen in such a way that the period length is sufficient to observe enough deaths over the age group of interest to make meaningfully precise estimates. The time period must also be short enough to reasonably assume the individuals within the age group of interest experienced similar levels of risk throughout the period. Age groups of interest are chosen so that the length of the age band provides sufficient information and also so that, regardless of period, individuals in the population experience similar levels of risk of mortality across the age group of interest. Some of the most rapid changes in age-specific mortality occur as the risk of death falls from its height at birth until beginning to increase again in adolescence. The rate of decrease slows with time since birth, necessitating shorter age band lengths at ages closest to birth if the assumption of constant mortality in that age group is to be at all appropriate.

The definition of the main mortality indicators used to quantify child mortality in any population or to assess a country's progress toward international mortality goals – NMR, IMR, and U5MR – are all informed by a combination of knowledge about how human mortality changes with time since birth, the ways a population defines age, and the methods

of birth history collection in countries where child mortality remains high today. Chapter 3 will further explore the consequences of different choices in lengths of age groups or time periods, the units used to measure them, and demographic or statistical methods within a single age and time period combination.

Regardless of how age and time are measured in data, we know death before reaching age  $a + 1$  after survival to age  $a$  is not determined by an instantaneous the flip of a coin at the moment of birth that determines whether an individual will remain living or die the exact moment before reaching age  $a + 1$ . Even if we are selecting vital records at random from the fully observed population of records with no measurement error in units of time that more closely reflect continuous time, like days or minutes, there are remaining issues with summarizing the age-specific period mortality of a population using only a single binary outcome for each individual. As the lifelines in Figure 2.1 and Figure 2.3 show, not every individual in a population who dies in age group  $[a_1, a_2)$  during period  $[t_2, t_3)$  has the same time of birth,  $t^*$ , or the same age at death,  $a^*$ . The range of possible lengths of lifelines of individuals from a population that contribute to the mortality in age group  $[a_1, a_2)$  during time period  $[t_2, t_3)$  grows as the length of age band  $[a_1, a_2)$  increases or the units by which age and time are measured become larger. In this way, treating each individual's death outcome as simply binary becomes less appropriate as the amount of time each individual contributes to the the age band and period of interest before death or reaching age  $a_2$  or  $t_3$  can take on a larger range of values (Allison, 2014, Ch. 1).

### 2.3.2 Discrete time survival analysis

*Survival analysis* or *event history analysis* refers to statistical methods for analyzing data time to event data (Allison, 2014, Ch. 1). Instead of collapsing all the information contained in each individual's *event history*, information about the time of each individual was observed before the death or data collection event is also incorporated. Life tables like those described

in Section 2.2.1 are actually one of the oldest methods used to summarize a sample of survival data without fully reducing the amount of information contained in each event history to a single binary variable, and have been used for centuries (Preston et al., 2001; Allison, 2014, Ch. 1). As previously mentioned, life table methods are not statistical models and do not yield quantification of uncertainty. Statistical methods for time to event or failure data began to arise in the mid 20<sup>th</sup> century (Allison, 2014, Ch. 1). Most event history analysis treats time continuously, even if time is always measured discretely in data collection (Allison, 2014, Chs. 1 & 3). The choice between continuous and discrete methods is motivated by how close the discrete unit of time used in data collection is to continuous time and, also, by the computational requirements that might be necessary if time is treated discretely and the length of observation window is long (Allison, 1982; Carstensen, 2007). If an event history analysis will make use of time-varying covariates, continuous methods may not be ideal (Allison, 2014, Ch. 2). In this dissertation, each chapter will employ *discrete time survival analysis* to incorporate information about the length of time a child was alive during period of interest using months for units of time, the finest available in FBH data. The discrete time analysis methods also allow for the simultaneous estimation of the period hazard of mortality for all discrete age groups, regardless of length, by including an indicator of age group membership as a time-varying covariate.

Let  $\{a_0, a_1, \dots, a_A\}$  denote a discrete partition of the continuous age interval  $[a_0, a_A)$  and  $\{t_0, t_1, \dots, t_T\}$  denote a discrete partition of the continuous interval of calendar time  $[t_0, t_T)$ . For each individual,  $k$ , let  $t_k^*$  denote time age  $a_0$  is reached and  $a_k^*$  denote age at death. Let  $t$  denote the left-inclusive interval of calendar time for which measures of period mortality are desired. All individuals observed alive and having reached age  $m$  at the start of interval  $t$  make up the *risk set* (Allison, 2014, Ch. 2). For all individuals from the sample who belong

to the risk set, define

$$y_{m,tk} = \begin{cases} 1, & m \leq a_k^* < m + 1, \\ 0, & a_k^* > m + 1, \end{cases} \quad (2.14)$$

for  $m = a_0, a_0 + 1, \dots, a_1 - 1, a_1, a_1 + 1, \dots, a_A - 1, a_A$ . These binary indicators are then used to estimate the monthly *hazard rate*, or conditional probability of death within one month given survival to age  $m$ ,  ${}_1q_{m,t}$ . Despite the dependence between the binary outcomes for a single individual, the hazards can be modelled with logistic regression, as though each binary outcome is an independent Bernoulli trial, and estimated via maximum likelihood or quasilielihood methods (Allison, 1982). No methods for addressing dependence in the outcomes arising from a single individual are needed unless the event of interest can happen to an individual more than once, which is not an issue with mortality (Allison, 2014, Ch. 2). For age bands of interest denoted  $a[m]$  and defined in the manner of (2.3),

$$y_{m,tk} | {}_1q_{m,t} \sim \text{Bernoulli}({}_1q_{m,t})$$

$$\text{logit}({}_1q_{m,t}) = \log\left(\frac{{}_1q_{m,t}}{1 - {}_1q_{m,t}}\right) = \beta_{a[m],t}, \quad (2.15)$$

Where the age intervals are not of width equal to one month, this mean model specification implies that all age months in the same age interval have the same estimated hazard rate,

$${}_1\hat{q}_{m,t} = \text{expit}\left(\hat{\beta}_{a[m],t}\right) = \frac{\exp(\hat{\beta}_{a[m],t})}{1 + \exp(\hat{\beta}_{a[m],t})}. \quad (2.16)$$

When estimates are desired for a left-inclusive age interval beginning with age  $a$  of length  $n_a$ , they can be calculated from the monthly hazard rates as

$${}_{n_a}\hat{q}_{a,t} = 1 - (1 - {}_{n_a}\hat{q}_{a,t}) = 1 - \prod_{m=a}^{a+n_a-1} (1 - {}_1\hat{q}_{m,t}), \quad (2.17)$$

for any  $a$  measured in months, not just  $\{a_1, \dots, a_A\}$ .

## 2.4 Survey statistics

All methods introduced in Section 2.3 are appropriate for observations from SRS data. However, SRS sampling can be impractical when populations are large, geographically vast, an outcome of interest is rare, or no *sampling frame* of all the individuals in a population exists to choose from. Sampling from a frame containing a list of groups of individuals to be sampled is called cluster sampling, and is often employed to minimize cost while maximizing sample size. Even if clusters are chosen from the sampling frame via SRS, outcomes from individuals within the same sampled cluster or primary sampling unit (PSU) are dependent. In the case of DHS and MICS surveys, the sampling frame is a list of EAs, not children or mothers. The regions used to stratify EAs geographically are typically Admin-1 divisions within the country. EAs within each geographic strata are then further stratified by urban or rural designation. The EAs are not sampled independently within strata but with probability proportional to size sampling (PPS). Once an EA has been sampled, a full enumeration of households in the cluster is made often in the form of a simple hand drawn map. Households are selected as secondary sampling units (SSU) according to a sampling scheme called *systematic sampling* where one or more starting points on the sample frame is chosen and every  $r^{th}$  unit after is sampled. The typical DHS or MICS sampling scheme described here is, then, a *stratified, two-stage cluster* design. If the dependence between observations from the same PSU or SSU is not accounted for, the variance of estimators will be underestimated. On the other hand, if stratification is ignored and strata are related to the outcome, one will miss out on the precision gains that arise from stratification. Additionally, ignoring stratification will result in bias. Methods to account for data arising from sampling schemes more complex than SRS, fall into two categories. *Design-based methods* consider the sample as coming from a finite population whereas *model-based methods* assume the samples come from an infinite population.

Design-based methods treat an individual's sampled outcome as fixed and statistical

estimates account for the probability an individual from the population would make it into the sample under replication of the survey. Design-based methods require knowledge of  $\pi_k$ , the *sampling probability* individual  $k$  is sampled, and also,  $\pi_{k,k'}$ , the *joint sampling probability* that individuals  $k$  and  $k'$  are both sampled. A *sampling weight*,  $w_k = \frac{1}{\pi_k}$ , is then calculated for each individual  $k$  and, if there is no nonresponse, can be interpreted as the number of individuals in the population individual  $k$ 's response can be thought to represent. In the DHS, the sampling weights are not exactly the probability of the individual being sampled for interview, but are adjusted after survey implementation to account for survey nonresponse.

Consider individual binary outcomes,  $y_k$ , and sampling weights,  $w_k$ , from a probability sample of  $n$  individuals from a population of size  $N$  containing  $T$  individuals for whom  $Y_k = 1$ , letting capital letters denote population quantities. Typical targets of inference are the population total,  $T$ , and the population proportion or prevalence,  $P = \frac{T}{N}$ . The commonly used design-based estimator of the population total (Horvitz and Thompson, 1952) is

$$\hat{T} = \sum_{k=1}^n w_k y_k, \quad (2.18)$$

$$\widehat{Var}(\hat{T}) = \sum_{k,k'} \frac{y_k y_{k'}}{\pi_k \pi_{k'}} - \frac{y_k y_{k'}}{\pi_{kk'}}. \quad (2.19)$$

The Horvitz-Thompson estimator of the population prevalence or proportion is

$$\hat{P} = \frac{\sum_{k=1}^n w_k y_k}{N} \quad (2.20)$$

$$\approx \frac{\sum_{k=1}^n w_k y_k}{\hat{N}} = \frac{\sum_{k=1}^n w_k y_k}{\sum_{k=1}^n w_k}, \quad (2.21)$$

where  $N$  is estimated via the sum of the sampling weights when it is unknown (Hájek, 1971). If  $N$  is known, as would be the case when a sampling frame lists a population in full,

$$\widehat{Var}(\hat{P}) = \frac{\widehat{Var}(\hat{T})}{N^2}. \quad (2.22)$$

In the case where  $N$  is estimated variance estimation is not straightforward, as  $\hat{P}$  is a nonlinear function of estimators of two population totals. If we define the population values of an indicator variable  $X_k = \mathbb{1}_{k \in \text{pop}}$  to be 1 for each individual of the population, we can estimate population size,  $N = T_X$ , by

$$\hat{N} = \hat{T}_X = \sum_{k=1}^n w_k x_k, \quad (2.23)$$

so that the estimator in (2.21) can be written as a  $\hat{P} = \frac{\hat{T}_Y}{\hat{T}_X}$ . Design-based estimators of this form are also called *ratio estimators*. If a target of inference is a nonlinear function of population totals, like the ratio estimator  $\hat{P}$ , the variance can be approximated using a method called *linearization* that approximates the nonlinear function of population totals with a linear function comprised of terms from a Taylor series expansion (Woodruff, 1971; Binder, 1983).

When making estimates for a subgroup within a population, as will be done with administrative areas throughout this dissertation, the size of the the number of children or births in area  $i$ ,  $N_i$ , is often unknown. One drawback to linearization methods for variance estimation is the computational cost as the expansions would need to be derived and coded up by each practitioner. Additionally, when sample sizes are small, variances are large and variance estimates are unstable. (Lohr, 2010, p. 369). In this dissertation, we avoid the former issue by making use of the `survey` package in `R` which performs the computational aspects of linearization under the hood (Lumley, 2004, 2010). The latter issue is commonly encountered in *small area estimation* and the increase in the severity of this problem as the level of subnational estimation goes from Admin-1 down to Admin-2 motivates much of the difference in methods discussed in Chapter 4 and Chapter 5 of this dissertation. The instability of the design-based variance estimator in the presence of small samples also motivates the use of Bayesian smoothing methods introduced in the next section and used in the method developed by Mercer et al. (2015) extended in Chapter 4.

Under stratified sampling,  $n_h$  individuals are sampled from the  $N_h$  people in each strata  $h = 1, \dots, H$ , such that  $\sum_{h=1}^H N_h = N$ . The Horvitz-Thompson estimator of the total becomes,

$$\hat{T} = \sum_{h=1}^H \hat{T}_h = \sum_{h=1}^H \sum_{k=1}^{n_h} w_{hk} y_{hk}, \quad (2.24)$$

$$\widehat{Var}(\hat{T}) = \sum_{h=1}^H \widehat{Var}(\hat{T}_h) = \sum_{h=1}^H \sum_{k,k'} \frac{y_{hk} y_{hk'}}{\pi_{hk} \pi_{hk'}} - \frac{y_{hk} y_{hk'}}{\pi_{hkk'}}, \quad (2.25)$$

where  $w_{hk}$  is the sampling weight for individual  $k$  in strata  $h$ . Under one-stage cluster sampling,  $C$  clusters or PSUs are chosen from a sampling frame. All individuals  $N_c$  within each cluster  $c$  are sampled, so  $w_{ck} = w_c = 1/\pi_c$  is the sampling weight for individual  $k$  in cluster  $c$ . Variance estimators account for within cluster dependence.

*Multi-stage sampling* refers to surveys where not every individual within sampled PSUs are observed. DHS and MICS sample PSUs from strata, SSUs from PSUs, and then all women in a selected SSU are eligible for interview. In reality nonresponse means SSUs are not fully observed in these surveys either, and no total at any stage of sampling is fully observed as in one-stage sampling. The Horvitz-Thompson estimator of the population total takes a similar form again, summing over estimates of strata totals. However within strata totals the nested dependence of PSUs, SSUs, and subsequent sampling units within each strata are considered.

When estimates of population prevalence,  $\hat{P}$  are desired and population sizes at various stages of sampling are estimated, as in estimation of child mortality from FBH data in this dissertation, the variance estimation becomes even more complicated as linearization methods and approximations are needed. Though we do not cover model-based methods in as much detail as the Horvitz-Thompson estimators, in general, fixed effects are used to account for stratification and random effects for each PSU can be used to account for dependence in observations with them (Lohr, 2010, Ch. 11).

The Horvitz-Thompson estimators described here can be used to estimate FBH data that arise from complex surveys, if the FBHs are reduced to single binary indicators as in Section 2.3.1. The linearization methods used in variance estimation for population means and totals can also be used to estimate generalized linear regression models with a pseudolikelihood method where each observation’s contribution to the score function is weighted by the sampling weight (Binder, 1983). Thus, the discrete time survival model defined in can be combined with the pseudolikelihood method to estimate age-specific hazards from survey data (Mercer et al., 2015). This approach yields design-consistent estimates of  $\text{logit}({}_1\hat{q}_m) = \hat{\beta}_{a[m]}$  as in (2.16), for each age band,  $a[m]$  along with the estimated  $(A - 1) \times (A - 1)$  variance-covariance matrix  $\hat{\Sigma}$  of the estimated vector  $\hat{\beta}$  using a sandwich estimator (Binder, 1983). To make estimates for any arbitrary age interval  $[a, a + n_a)$  as in (2.17), the delta method can be used to estimate the variance for  ${}_{n_a}\hat{q}_a$  by transforming  $\hat{\Sigma}$  (Mercer et al., 2015). This method of age-specific mortality estimation will be used for estimation of national child mortality measures in Chapter 3 and Admin-2 child mortality measures in Chapter 4.

## 2.5 Bayesian spatiotemporal smoothing

Though many surveys can contribute to age-specific mortality in an admin area  $i$  and time period  $t$ , as the level of administrative division becomes finer and the number of years in  $t$  becomes smaller, the variance of the design-based age-specific mortality estimator will grow. Bayesian hierarchical models that can be used to exploit the the dependence between estimates at near by points in time and/or space to smooth the design-based estimates of  ${}_{n_a}\hat{q}_a$  are introduced in this section (Rue and Held, 2005; Mercer et al., 2015). The fast computation of Bayesian hierarchical models with integrated nested Laplace approximations implemented in R via INLA allow these methods to be fit easily (Rue et al., 2009; Martins et al., 2013).

### 2.5.1 Temporal smoothing models

Let  $y_t$  be an observed outcome at time  $t$ , and  $g(\cdot)$  be a link function for a generalized linear model with  $\eta_t$  is the mean linear predictor at time  $t$ , i.e.  $E[y_t] = g(\eta_t)$ . If we have prior information that  $y_t$  are dependent, one might specify the mean model below

$$\eta_t = \mu + \gamma_t, \quad (2.26)$$

where  $\mu$  is a fixed intercept and  $\gamma_t$  is a random effect that will describe the trend in time net of  $\mu$ . The prior for the  $\gamma$  terms can be chosen to reflect different assumptions about dependence in the  $y_t$ 's across time. A random walk of order 1 (RW1) assumes the realization of the process at time  $t + 1$  is only informed by that at  $t$ , and penalizes differences between terms in adjacent time points (Rue and Held, 2005). The prior specification for the  $1 \times T$  vector  $\boldsymbol{\gamma}$  of elements  $\gamma_t$ , can be written

$$\pi(\boldsymbol{\gamma}|\kappa_\gamma) \propto \kappa_\gamma^{(T-1)/2} \exp \left[ -\frac{\kappa_\gamma}{2} \sum_{t=1}^{T-1} (\gamma_t - \gamma_{t+1})^2 \right] = \kappa_\gamma^{(T-1)/2} \exp \left[ -\frac{\kappa_\gamma}{2} \boldsymbol{\gamma}^\top \mathbf{Q}_\gamma \boldsymbol{\gamma} \right],$$

where  $\mathbf{Q}_\gamma$  specifies the second order dependency between time points and has rank  $T - 1$ , and makes the prior improper (Rue and Held, 2005). A random walk of order 2 (RW2) assumes the realization of the process at time  $t + 2$  depends upon those at both times  $t + 1$  and  $t$ , and penalizes large second differences. The improper prior on second differences is

$$\pi(\boldsymbol{\gamma}|\kappa_\gamma) \propto \kappa_\gamma^{(T-2)/2} \exp \left[ -\frac{\kappa_\gamma}{2} \sum_{t=1}^{T-2} (\gamma_t - 2\gamma_{t+1} + \gamma_{t+2})^2 \right] = \kappa_\gamma^{(T-2)/2} \exp \left[ -\frac{\kappa_\gamma}{2} \boldsymbol{\gamma}^\top \mathbf{Q}_\gamma \boldsymbol{\gamma} \right],$$

where  $\mathbf{Q}_\gamma$  specifies the second order dependency between time points and has rank  $T - 2$  (Rue and Held, 2005).

Similar to the RW1, an autoregressive process of order 1 (AR(1)), assumes that realizations of a process at time  $t + 1$  depend only on that at time  $t$ . However, instead of assuming the conditional mean of the process at time  $t + 1$ ,  $E[\gamma_{t+1}|\gamma_t]$  is equal to the realization of the

process at time  $t$ ,  $\gamma_t$ , it assumes the mean at time  $t + 1$  is a product of the *autocorrelation* parameter,  $\rho$ , and the previous realization of the process. Formally,

$$\begin{aligned}\gamma_t &= \rho\gamma_{t-1} + \varepsilon_t, \\ \gamma_1 &\sim \text{N}(0, \kappa_\gamma^{-1}),\end{aligned}$$

for  $t = 2, \dots, T$  where  $\rho \in (0, 1)$ ,  $\varepsilon_t \sim \text{N}(0, \tau^{-1})$ , and  $\kappa_\gamma = \tau(1 - \rho^2)$ . The prior can be written in matrix notation as

$$\pi(\gamma|\kappa_\gamma) \propto \kappa_\gamma^{(T-1)/2} \exp \left[ -\frac{\kappa_\gamma}{2} \sum_{t=1}^{T-1} (\gamma_{t+1} - \rho\gamma_t)^2 \right] = \kappa_\gamma^{(T-1)/2} \exp \left[ -\frac{\kappa_\gamma}{2} \gamma^T \mathbf{Q}_\gamma \gamma \right].$$

When predicting the value of a process at a time point that has not been observed, a RW1 prior will project forward the same mean as the previous time period with added uncertainty. An AR(1) process will project forward as the product of  $\rho$  and the previous mean,  $y_t$ . However, a RW2 prior will project forward the difference between the previous two time points to get an estimate in the next time point, allowing for projections to continue previously observed increasing or decreasing trends.

### 2.5.2 Spatial smoothing models

Similar methods can be used for observations from admin areas. Denote the data by  $y_i$  for  $i = 1, \dots, I$ , and  $g(\cdot)$  be a link function for a generalized linear model, i.e.  $E[y_i] = g(\eta_i)$  with

$$\eta_i = \mu + \phi_i, \tag{2.27}$$

is the linear predictor for area  $i$ . If the prior chosen for the spatial random effects is  $\phi_i \stackrel{\text{iid}}{\sim} \text{N}(0, \kappa_\phi^{-1})$ , there is no implicit assumption about the dependence between outcomes in areas that are nearby. The Besag, York and Mollié (BYM) (1991) model instead defines the mean model for an area using the sum of two latent terms as follows,

$$\eta_i = \mu + \theta_i + \phi_i. \tag{2.28}$$

The BYM model places the same IID prior on the vector of  $\phi_i$  and  $\theta_i$  is an intrinsic conditional autoregressive (ICAR) spatial random effect. The improper prior for  $\theta_i$  encourages the borrowing of information from area  $i$ 's neighbors which in this dissertation we define as those areas that share a common boundary. The conditional density is

$$\theta_i | \theta_j, j \sim i, \kappa_\theta \sim \text{N} \left( \frac{1}{n_i} \sum_{j \sim i} \theta_j, \frac{1}{n_i \kappa_\theta} \right),$$

where  $j \sim i$  indexes the neighbors of area  $i$  and  $n_i$  are the number of such neighbors. The joint prior “distribution” is compactly written as

$$\pi(\theta | \kappa_\theta) \propto \kappa_\theta^{(n-1)/2} \exp \left[ -\frac{\kappa_\theta}{2} \theta^\text{T} \mathbf{Q}_\theta \theta \right],$$

where  $\mathbf{Q}_\theta$  encodes the neighbors.

In this dissertation, we adopt a reparameterization of the BYM which allows one to place a joint penalized-complexity (PC) prior on the traditional terms ICAR and IID terms of a BYM that penalizes deviations from a “null” model without random effects (Riebler et al., 2016; Simpson et al., 2017; Fuglstad et al., 2019). The BYM2 parameterization will be denoted with asterisks as follows

$$\eta_i = \mu + \theta_i^* + \phi_i^*. \quad (2.29)$$

The joint prior on the spatial terms has two parameters, one which represents the proportion of the marginal variance,  $r_{\text{BYM}}$ , which is attributed to the ICAR term instead of the IID term and one which represents the marginal precision,  $\kappa_{\text{BYM}}$ , of the weighted sum of  $\theta_i$  and  $\phi_i$ . In the resulting linear predictor,

$$\theta_i^* + \phi_i^* = \kappa_{\text{BYM}}^{-1/2} \left[ \sqrt{r_{\text{BYM}}} \theta_i + \sqrt{1 - r_{\text{BYM}}} \phi_i \right].$$

### 2.5.3 Spatiotemporal smoothing models

If spatiotemporal estimates are desired, the spatial and temporal terms above can be used for main smoothing terms, but a space-time interaction smoothing term,  $\delta_{it}$ , may still be

desired. The linear predictor is, then, of the form:

$$\eta_i = \mu + \theta_i^* + \phi_i^* + \alpha_t + \gamma_t + \delta_{it}. \quad (2.30)$$

Spatiotemporal smoothing terms can be classified into four types according to whether they assume dependence in space or time (Knorr-Held, 2000). A Type I interaction assumes no dependence across space or time in the interaction terms, i.e. that  $\delta_{it}$  are IID. Type II interactions assume the interactions are dependent in time, but there is no spatial structure to them. Type III interactions assume structure in space but not in time for  $\delta_{it}$ 's. Finally, as used in Chapter 4 and Chapter 5, a Type IV interaction assumes a different dependence model in time and in space. The prior for the vector of  $\delta_{it}$  is determined by a precision matrix that is the Kronecker product of the precision matrices of any of the temporal smoothing priors in Section 2.5.1 and an ICAR, i.e.  $\mathbf{Q}_\delta = \mathbf{Q}_\gamma \otimes \mathbf{Q}_\theta$ . We can write the joint prior “distribution” as,

$$\pi(\delta|\kappa_\delta) \propto \kappa_\delta^{(n-1)(T-2)/2} \exp \left[ -\frac{\kappa_\delta}{2} \delta^T \mathbf{Q}_\delta \delta \right]. \quad (2.31)$$

This form is improper, since the matrix  $\mathbf{Q}_\delta$  has rank  $(n-1)(T-2)$ .

## Chapter 3

# A COMPARISON OF DIRECT ESTIMATION METHODS OF PERIOD CHILD MORTALITY WITH COMPLEX SURVEYS

### *3.1 Introduction*

In Section 2.2.1 and Section 2.3, we introduced demographic and statistical methods for period age-specific mortality estimation in a single time period of interest. The Lexis diagram (Preston et al., Ch. 2) and life table (Preston et al., Ch. 3) methods are well-established and commonly used. They each take different approaches in their treatment of age, birth cohort, and calendar period as functions of time. As with many demographic methods, neither address uncertainty or measurement error. While Sections 2.3.1 and 2.3.2 both introduce statistical methods that allow for estimation of uncertainty, only discrete time survival analysis is comparable to the demographic methods in addressing the reality that individuals reach an age band of interest according to the timing of their birth and that deaths may occur at any point within a time period or age band. In the context of estimating child mortality from FBH data collected in household surveys in LMICs, quantification of uncertainty is important as we do not see more than a small percentage of the population. Additionally, different agencies who make estimates use different methods and quantification of uncertainty is important to compare these different estimates.

In this chapter we provide a more in depth comparison of the Lexis, life table, and discrete time analysis approaches to age-specific child mortality estimation. First, Section 3.2 describes the differences in how the methods count deaths and exposures for a time period and age group of interest. Section 3.3 discusses difference in estimators of age-specific mortality measures from each method. Finally, Section 3.4 presents a comparison of these

methods applied to the Senegal 2006 DHS FBH data. In particular, we compare variants of the methods used by UN IGME and DHS to the statistical discrete time methods used in Chapter 4 and Chapter 5.

### 3.2 *Counting deaths and exposures for period mortality*

Let us return to our example from Chapter 2 concerning age group of interest  $[a_1, a_2)$  and time period  $[t_2, t_3)$ . As Figure 3.1 (which is a combination of Figure 2.1 and Figure 2.3) makes clear, the Lexis approach and life table approaches will treat each of the three individuals differently. Both methods count 2 deaths in the period of interest represented in grey rectangles. Both methods count the individual from Cohort  $t_1$  as a death in  $[a_1, a_2) \times [t_2, t_3)$ . The Lexis diagram correctly includes the individual from Cohort  $t_0$  as a death in  $[a_1, a_2) \times [t_2, t_3)$ , while the life table method treats this individual's death correctly in age and incorrectly in period as  $[a_1, a_2) \times [t_1, t_2)$ . The individual from Cohort  $t_2$  who died between ages  $[a_0, a_1)$  but in time  $[t_3, t_4)$  receives some exposure in the correct period from both methods; however, the life table treats this death as if it happened during  $[t_2, t_3)$ . While it is hard to visually assess the difference in the lengths of time spent alive by individuals or *exposure* to the force of mortality, it is clear that the life table method does not include any exposure or time lived for the individual from Cohort  $t_0$ .

If we measured time continuously and there is no measurement error in our birth and death data, the Lexis diagram is the correct method for relating the three temporal dimensions of age, period, and cohort and displaying an individual's mortality information. Hence, there are no translucent lines in the left panel of Figure 3.1 as in the right panel. However, as we established throughout Chapter 2, age, period, and cohort are measured in discrete units both socially and in the collection of FBH data by DHS. A further complication is the censoring of children who have not reached age 60 months or died before the time of survey and the censoring of children whose mothers have died. Moreover, our purpose for using

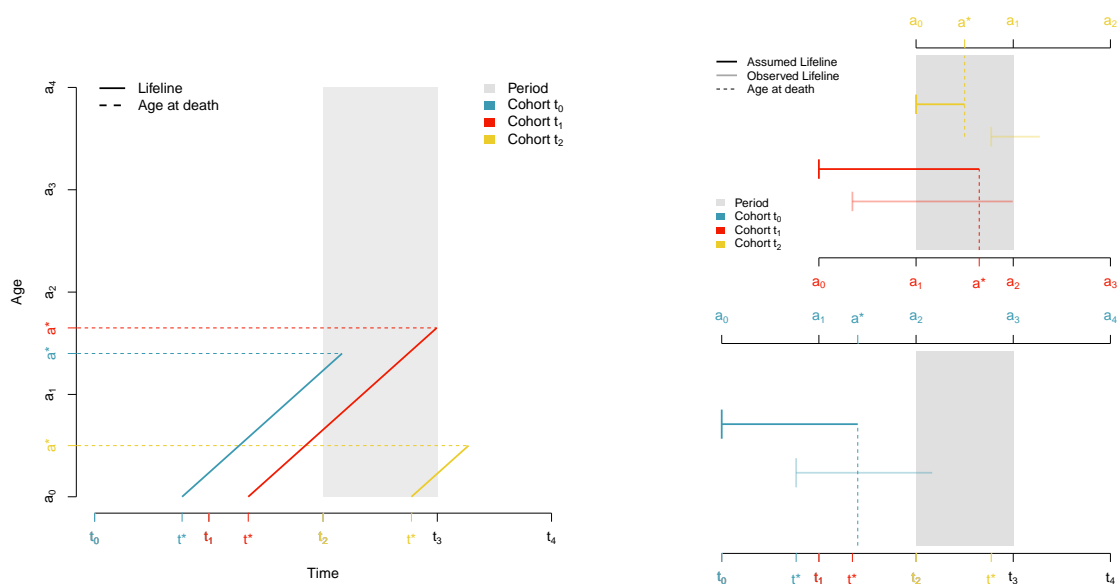


Figure 3.1: The color of the life line indicates the cohort of the individual—  $[t_0, t_1)$  (blue),  $[t_1, t_2)$  (red), and  $[t_2, t_3)$  (yellow). The grey rectangle indicates the period of interest  $[t_2, t_3)$ . **Left:** Each solid line is a lifeline for a single individual, that starts when then the individual reaches ages  $a_0$  at time  $t^*$  and increases diagonally until age at death,  $a^*$ . **Right:** Lifelines for the same individuals plotted horizontally in calendar time, represented by the bottom axis, where color represents the individual’s cohort. Translucent lifelines connecting points  $t^*$  and  $t^* + a^*$  represent each individual’s true calendar times at birth and death. Solid lifelines show the lifeline for each individual under the assumption they were born at the beginning of their cohort-defining period.

these methods is not to visualize or to quantify the mortality experiences of individuals, but a population. The accuracy of the aggregation of deaths in Figure 2.1 to the cohort-specific parallelograms in Figure 2.2 will depend upon the units at which time is measured in birth and death information as well as the length of age band and time period. If we were to observe a full population in the past with no measurement error and continuous measures of

time of birth and death, the Lexis diagram framework will accurately count the number of *observed* deaths in the *finite population* as a “synthesis of cohorts” (Preston et al., p. 35).

The methods used by UN IGME (Population Division, 2011) and DHS (Rutstein et al., 2006; Elkasabi, 2019) both use a Lexis diagram approach to counting deaths and exposures for mortality measures. However, they divide the intersection of a period of interest and age group of interest into a synthesis of contributions from 3 cohorts A (blue), B (yellow), C (red) as seen in Figure 3.2. These methods both count every death that occurs to a member of Cohort B, or those born between  $[t_2 - a_1, t_3 - a_2)$  in our example. Anyone in this cohort will reach age  $a_1$  between  $[t_2, t_3 - a_2 + a_1)$  and  $a_2$  in  $[t_2 - a_1 + a_2, t_3)$ , thus experiencing all exposure to mortality at ages  $[a_1, a_2)$  during  $[t_2, t_3)$ . For these individuals  $k$  in Cohort B, both methods define a binary indicator of death  $y_{a_1, t_2 k}$ . Formally,

$$y_{a_1, t_2 k} = \begin{cases} 1, & a_k^* \in [a_1, a_2) \text{ and } t_k^* \in [t_2 - a_1, t_3 - a_2) \\ 0.5, & a_k^* \in [a_1, a_2) \text{ and } t_k^* \in [t_2 - a_2, t_2 - a_1) \cup [t_3 - a_2, t_3 - a_1) \\ 0, & \text{otherwise.} \end{cases} \quad (3.1)$$

Finally the Lexis approaches used by UN IGME and DHS, treat individual deaths differently in the period prior to survey interviews. If an individual is a member of Cohort C and died in the period just before the survey, they are counted as a whole death, i.e.

$$y_{a_1, t_2 k} = \begin{cases} 1, & a_k^* \in [a_1, a_2) \text{ and } t_k^* \in [t_2 - a_1, t_3 - a_2) \cup [t_3 - a_2, t_3 - a_1) \\ 0.5, & a_k^* \in [a_1, a_2) \text{ and } t_k^* \in [t_2 - a_2, t_2 - a_1) \\ 0, & \text{otherwise.} \end{cases} \quad (3.2)$$

The Lexis methods calculate an exposure or denominator for each of these individuals,  $E_{a_1, t_2 k} = 1$ , regardless of death or length of  $[a_1, a_2)$ . For individuals in Cohort A, born

between  $[t_1 - a_2, t_1 - a_1)$ , and Cohort C, born between  $[t_2 - a_2, t_2 - a_1)$ , they treat each observed death and exposure as a half-death or exposure. Formally,

$$E_{a_1, t_2 k} = \begin{cases} 1, & a_k^* \geq a_1 \text{ and } t_k^* \in [t_2 - a_1, t_3 - a_2) \\ 0.5, & a_k^* \geq a_1 \text{ and } t_k^* \in [t_2 - a_2, t_2 - a_1) \cup [t_3 - a_2, t_3 - a_1) \\ 0, & \text{otherwise.} \end{cases} \quad (3.3)$$

Though half the parallelograms of Cohort A and C do, indeed, overlap with  $[a_1, a_2) \times [t_2, t_3)$ ,  $E_{a_1, t_2 k} \in \{0, 0.5, 1\}$  regardless of how many age groups for which one is making estimates or whether those groups have different lengths, e.g. one month for NMR or 12 months for IMR.

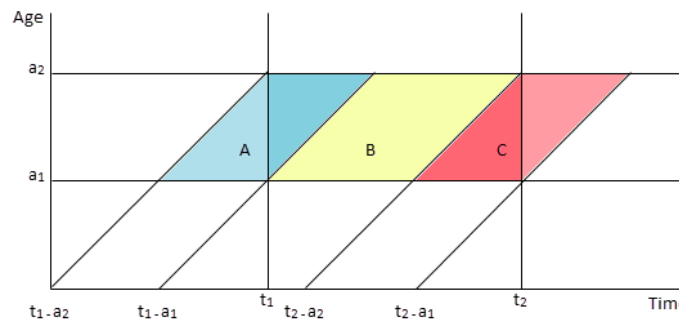


Figure 3.2: Visualization of the cohorts used to divide up mortality contributions to a period of interest and age group of interest (Rutstein et al., 2006).

Regardless of the precision with which time and age are measured in data, the life table method of counting deaths will depend on the consequences of the assumptions made to treat time as one-dimensional. Even if we were to observe a full population in the past with no measurement error and continuous measures of time of birth and death, the life table method is still a *model* used to describe the mortality experiences of a *hypothetical cohort* or a *synthetic cohort* of unborn individuals (Preston et al., 2001, p. 42), reminiscent of a

statistical *superpopulation* approach.

Recall the definitions of the life table columns introduced in Section 2.2.1. In our example, the parameter  ${}_{a_2-a_1}d_{a_1}$  represents the number of *expected* deaths in  $[a_1, a_2)$  to a hypothetical cohort of size  $l_0$  born (or reaching  $a_0$ ) at time  $t_0$ ,  $l_{a_1}$  of whom survived to  $t_1$  or age  $a_1$ . When one is given  ${}_{a_2-a_1}q_{a_1}$ , this can be calculated as the expectation of a binomial random variable with  $l_{a_1}$  trials or more general collection of  $l_{a_1}$  binary indicators of death,  ${}_{a_2-a_1}d_{a_1} = l_{a_1} \times {}_{a_2-a_1}q_{a_1}$ . When  ${}_{a_2-a_1}q_{a_1}$  is not known, we must estimate it from the number of deaths in our data. Assuming all individuals in a cohort are born at the same time at the beginning of the cohort-definition period, implies that only a single cohort contributes to the mortality estimation for any age and time period of interest as seen in the right panel of Figure 3.1 and Table 2.1. So, a life table approach in our toy example would define

$$y_{a_1, t_2 k} = \begin{cases} 1, & a_k^* \in [a_1, a_2) \text{ and } t_k^* \in [t_1, t_2) \\ 0, & \text{otherwise.} \end{cases} \quad (3.4)$$

Note, that the cohort-defining period for the lifetable approach includes individuals from both Cohorts A and Cohort C as defined in the Lexis diagram used by IGME and DHS in Figure 3.2.

Following, Section 2.2.1 one could calculate the exposures by defining the following

$${}_{a_2-a_1}L_{a_1, t_2 k} = \begin{cases} 1, & a_k^* \in [a_0, a_1) \cup [a_2, a_A) \text{ and } t_k^* \in [t_1, t_2), \\ a_k^* - a_1, & a_k^* \in [a_1, a_2) \text{ and } t_k^* \in [t_1, t_2), \\ 0, & \text{otherwise,} \end{cases} \quad (3.5)$$

and calculating,

$${}_{a_2-a_1}L_{a_1, t_2} = \sum_k {}_{a_2-a_1}L_{a_1, t_2 k} \quad (3.6)$$

However, when time is discretized, we may not know the exact time of death for child  $k$ ,  $a_k^*$ , exactly for each individual  $k$ . The potential magnitude of measurement error of  $a_k^*$  for each individual grows as the unit of time measurement increases from months to years or multi-year age groups. When exact ages at death are not known for each individual or data are tabulated, one must make assumptions about the average time spent alive in the age interval,  ${}_n a_x$ . With a value or estimate of  ${}_n a_x$ , one can approximate exposures in a life table, e.g. returning to the example in (3.6),

$$\begin{aligned} {}_{a_2-a_1} L_{a_1, t_2} &= \sum_{k: a_k^* \geq a_2} {}_{a_2-a_1} L_{a_1, t_2 k} + \sum_{k: a_k^* \in [a_1, a_2)} {}_{a_2-a_1} L_{a_1, t_2 k} \\ &\approx \sum_{k: a_k^* \geq a_2} {}_{a_2-a_1} L_{a_1, t_2 k} + {}_{a_2-a_1} a_{a_1, t_2}. \end{aligned} \quad (3.7)$$

The simplest assumption is  ${}_n a_x = \frac{x+n}{2}$ , equivalent to the expected value of time spent alive under a uniform distribution over  $[x, x+n)$ . This assumption does not reflect reality, and the violations of this assumption worsen when the force of mortality varies widely within  $[x, x+n)$ . This can be partially alleviated by using finer units of time, but always remains an issue for the last open-ended age interval  $[x, \infty)$ . It also remains an issue for the first age interval of a life table when it represents times just after birth, as the force of human mortality experiences some of its steepest declines following birth in the interval  $[0, 1)$  whether units of time represent months (NMR), years (IMR), or 5 year periods (U5MR). More complicated demographic methods exist for estimating  ${}_1 a_0$  from data or making more complicated mathematical assumptions about  ${}_1 a_0$  (Coale and Demeny, 1966; ?). An assumption of constant  ${}_1 a_x$  for later months of life contributing to IMR and U5MR using monthly measurements of time, as in the DHS, is not wholly inappropriate. However, the force of mortality decreases rapidly between birth and age one month, making the assumption less appropriate for estimation of NMR.

The assumptions in the life table that lead to temporal dimension reductions are actually implicit assumptions about the way population fertility and mortality changes over time. If each cohort contains the same number of births and age-specific mortality does not change over time, then this will not result in a biased estimate of period age-specific mortality. However, if there are big differences in cohort size and cohort age-specific mortality, the experience of the single cohort used to estimate mortality in a particular age band may not be a suitable proxy for the reality that deaths in an age band and period occur to individuals born at different times in different cohorts. The Lexis diagram makes implicit assumptions about fertility and mortality changes as well, by assuming that exactly half of the deaths and half of the exposures from Cohorts A and C occur to either side of the cohorts' parallelogram. This will reflect reality if exactly half of all the deaths and time lived in the age band of interest in each cohort fall within the period of interest, i.e. if births occur uniformly in time across the cohort-defining period and deaths occur uniformly across the age group of interest. Other assumptions may be made about the way the three cohorts combine to contribute to period age-specific mortality estimation, Carstensen (2007) discuss these in detail. Though, the Lexis assumptions do not have implications for the way age-specific mortality is allowed to change between cohorts.

The discrete time survival analysis methods introduced in Section 2.3.2 do allow for assumptions about changes in the hazard of death over time, but make no assumptions about or considerations of the cohort mortality or fertility like demographic methods. Despite this, the statistical assumptions of independent trials in a binomial process and discrete time survival analysis closely connect to the life table methods. The discrete time survival analysis methods in Section 2.3.2, define

$$y_{a_1, t_2 k} = \begin{cases} 1, & a_k^* \in [a_1, a_2) \text{ and } t_k^* + a_k^* \in [t_2, t_3) \\ 0, & \text{otherwise,} \end{cases} \quad (3.8)$$

and

$$E_{a_1, t_2 k} = \begin{cases} 1, & a_k^* \geq a_1 \text{ and } t_k^* + a_1 \in [t_2, t_3) \\ 0, & \text{otherwise,} \end{cases} \quad (3.9)$$

so that  $n_{a_1, t_2} = \sum_k E_{a_1, t_2 k}$  is the size of the risk set for  $[a_1, a_2)$  in  $[t_2, t_3)$ . Under the assumption that all members of a cohort are born at the beginning of their cohort-defining interval and since  ${}_a a_{a_1} = 1$ , the life table and discrete time survival methods for age bands of length  $a_2 - a_1 = 1$  will count deaths and exposures the same way. Then,  ${}_{a_2 - a_1} L_{a_1} = l_{a_1}$  and the hypothetical cohort of the model life table aligns with a binomial model where  $l_{a_1}$  is the number of independent trials. The target of inference is underlying *hazard* of death in a superpopulation, in contrast to the target of inference of *prevalence* of death in Lexis diagram and finite population approaches. When life table approaches are used on data for a fully observed population,  ${}_{a_2 - a_1} L_{a_1}$  will be the same as the sum of exposures  $E_{a_1, t_2 k}$  in (3.3) for individuals  $k$  belonging to the individuals from Cohorts A and B that also fall into the cohort contributing to life table quantities for  $[a_1, a_2)$ . Note, for consistency across the methods we will refer to assigned exposure values for DHS data using  $E$ , leaving  $L$  to refer to corresponding hypothetical life table values.

### 3.3 Estimation of period mortality

After converting each individual's birth history information into deaths and exposures for the discrete ages and time periods of interest, there are also many choices and methods of estimation of period age-specific mortality. In the Lexis diagram framework, age-specific period mortality can be estimated as the prevalence of deaths in a finite population or the rate of deaths per units of person-time spent alive in the age interval by each individual. Though what is being counted in the Lexis estimate is prevalence of deaths, demographic notation refers to this finite population quantity in lowercase as  ${}_n q_x$ . The mortality rate

from a Lexis diagram or for a finite population is denoted  ${}_nM_x$ . In the life table framework the underlying hazard of death before age  $x + n$  given survival to age  $x$ ,  ${}_nq_x$ , in an infinite hypothetical cohort of individuals, or  ${}_nm_x$ , the mortality rate in the hypothetical cohort. Though the distinction is made for mortality rates, there is no standard uppercase notation for observed prevalence of death in a finite population. The statistical approaches introduced in Section 2.3.1-2.3.2 only described the target of inference of  ${}_nq_x$ , the underlying hazard of mortality in a hypothetical cohort or superpopulation of people all reaching age  $x$  at the same time when using discrete survival methods and binomial outcomes or the true prevalence of death in a finite population. However, mortality rates can be estimated statistically as the simple ratio of deaths to total observed person-time in one's sample or using a Poisson likelihood or the quasiliikelihood version with a log link function. Throughout the rest of this chapter we will use  ${}_nq_x$  interchangeably across methods as is standard, though we will see many quantities commonly denoted  ${}_nq_x$  are representing different metrics.

Under simple random sampling, the age-specific period probability of death can be estimated using formulas in Chapter 2. Continuing with our example we can estimate mortality in  $[a_1, a_2)$  during  $[t_2, t_3)$ , using the Lexis frameworks in (3.1) and (3.3) to estimate the observed prevalence of death,

$${}_{a_2-a_1}\widehat{q}_{a_1,t_2} = \frac{\sum_{k=1}^n y_{a_1,t_2k}}{\sum_{k=1}^n E_{a_1,t_2k}} \quad (3.10)$$

$$= \frac{\sum_{k=1}^n \mathbb{1}_{a_k^* \in [a_1, a_2)} (\mathbb{1}_{k \in B} + 0.5 \times \mathbb{1}_{k \in A \text{ or } C})}{\sum_{k=1}^n (\mathbb{1}_{k \in B} + 0.5 \times \mathbb{1}_{k \in A \text{ or } C})}, \quad (3.11)$$

where the indicators are for membership Cohorts A, B, or C and death in  $[a_1, a_2)$ . To estimate the observed age-specific mortality rate for a population from the Lexis diagram,

$${}_{a_2-a_1}\widehat{M}_{a_1,t_2} = \frac{\sum_{k=1}^n y_{a_1,t_2k}}{\sum_{k=1}^n E_{a_1,t_2k} \times (a_2 - a_1)} \quad (3.12)$$

$$= \frac{\sum_{k=1}^n \mathbb{1}_{a_k^* \in [a_1, a_2]} (\mathbb{1}_{k \in B} + 0.5 \times \mathbb{1}_{k \in A \text{ or } C})}{\sum_{k=1}^n (a_2 - a_1) \times (\mathbb{1}_{k \in B} + 0.5 \times \mathbb{1}_{k \in A \text{ or } C})}. \quad (3.13)$$

In a life table framework, we can use (3.4) and (3.7) to estimate the age-specific mortality rate in a hypothetical cohort as

$${}_{a_2-a_1}\widehat{m}_{a_1,t_2} = \frac{\sum_{k=1}^n y_{a_1,t_2k}}{\sum_{k=1}^n {}_{a_2-a_1}L_{a_1,t_2k}} = \frac{\sum_{k=1}^n y_{a_1,t_2k}}{{}_{a_2-a_1}L_{a_1,t_2}} \quad (3.14)$$

$$\approx \frac{\sum_{k=1}^n y_{a_1,t_2k}}{{}_{a_2-a_1}\widehat{L}_{a_1,t_2}}, \quad (3.15)$$

where  ${}_{a_2-a_1}\widehat{L}_{a_1,t_2}$  uses assumptions about  ${}_{a_2-a_1}a_{a_1}, t_2$ . To make estimates of the underlying probability of death, the life table methods often make use of the relationship between mortality rates and probabilities. Following ? (Ch. 3preston:2001),

$${}_nq_x = \frac{n \times_n m_x}{1 + {}_n m_x (n - {}_n a_x)}. \quad (3.16)$$

In our example, then,

$${}_{a_2-a_1}\widehat{q}_{a_1,t_2} = \frac{(a_2 - a_1) \times {}_{a_2-a_1}\widehat{m}_{a_1,t_2}}{1 + {}_{a_2-a_1}\widehat{m}_{a_1,t_2} \times [(a_2 - a_1) - {}_{a_2-a_1}a_{a_1,t_2}]}. \quad (3.17)$$

If the length of the age group of interest is a single unit of time, then the only term governing the difference in the mortality rate and the probability of death is the average age at death parameter. It is easy to see by rewriting (3.16), that  ${}_1q_x \rightarrow {}_1m_x$  as  ${}_1a_x \rightarrow 1$ . Given the life table methods already assume every person in a cohort is born at the simultaneously beginning of the single unit of time,  ${}_1a_x \approx 1$  implies all deaths occur simultaneously in the last moment before the surviving cohort members turn age  $x + 1$ .

Using the discrete time survival methods to calculate deaths and exposures as in (3.8) and (3.9), proportions could be used to estimate age-specific mortality as

$${}_{a_2-a_1}\widehat{q}_{a_1,t_2} = \frac{\sum_{k=1}^n y_{a_1,t_2k}}{n_{a_1,t_2}} \quad (3.18)$$

and a binomial assumption on  $y_{a_1,t_2k}$  can be used to estimate uncertainty as in (2.13). The mortality rate can be estimated as

$${}_{a_2-a_1}\widehat{m}_{a_1,t_2} = \frac{\sum_{k=1}^n y_{a_1,t_2k}}{\sum_{k=1}^n E_{a_1,t_2k} [(a_k^* \vee a_2) - a_1]}. \quad (3.19)$$

When time is measured in months  $m$ , for  $m \in [a_1, a_2)$ , the discrete time survival methods will use the following model

$$y_{a_1[m],t_2k} | {}_1q_{a_1[m],t_2} \sim \text{Bernoulli}(1, {}_1q_{a_1[m],t_2}) \quad (3.20)$$

$$\text{logit}({}_1q_{a_1[m],t_2}) = \beta_{a_1[m],t_2}. \quad (3.21)$$

Following Section 2.3.2 maximum likelihood or quasilielihood estimation can be used to estimate  ${}_1\widehat{q}_{a_1[m],t_2} = \text{expit}(\widehat{\beta}_{a_1[m],t_2})$  from a simple random sample so that,

$${}_{a_2-a_1}\widehat{q}_{a_1,t_2} = 1 - \prod_{m=a_1}^{a_2-1} (1 - {}_1\widehat{q}_{a_1[m],t_2}), \quad (3.22)$$

and standard errors can be estimated via the delta method.

When using FBH data from DHS surveys, the above methods do not reflect the complex sampling design or survey implementation. Each estimator above can be extended, easily, to design-based estimators. Following the Lexis framework used by IGME and DHS,

$${}_{a_2-a_1}\widehat{q}_{a_1,t_2} = \frac{\sum_{k=1}^n w_k y_{a_1,t_2k}}{\sum_{k=1}^n w_k E_{a_1,t_2k}} \quad (3.23)$$

$$= \frac{\sum_{k=1}^n w_k \mathbb{1}_{a_k^* \in [a_1, a_2)} (\mathbb{1}_{k \in B} + 0.5 \times \mathbb{1}_{k \in A \text{ or } C})}{\sum_{k=1}^n w_k (\mathbb{1}_{k \in B} + 0.5 \times \mathbb{1}_{k \in A \text{ or } C})}, \quad (3.24)$$

where  $w_k$  is the design weight for individual  $k$ . Note, for Cohorts A and C the assignment of half-deaths and half-exposures provide can be interpreted in the design-based framework of sampling from a finite population. Under the assumptions of the Lexis framework that half of all deaths and exposures to each cohort occur in  $[a_1, a_2) \times [t_2, t_3)$ , we can rewrite (3.24) as something that looks like a more traditional Horvitz-Thompson estimator

$${}_{a_2-a_1}\hat{q}_{a_1,t_2} = \frac{\sum_{k=1}^n w_k^* y_{a_1,t_2k}^*}{\sum_{k=1}^n w_k^*}, \quad (3.25)$$

where  $w_k^* = w_k \times E_{a_1,t_2k}$  and  $y_{a_1,t_2k}^*$  is always equal to either 1 or 0. So when interpreting a weight  $w_k$  as the average number of people in the finite population individual  $k$  represents, the number of people each person from Cohorts A and C represents is reduced by half.

In the next section, we discuss briefly some nuances with DHS FBH data before comparing the Lexis methods with the discrete time survival methods used in Chapters 4 and 5. In the context of LMICs where FBH data informing child mortality is from a small sample of the full population, we are interested both in comparisons of point estimates of mortality measures and quantifying uncertainty. Using the Senegal 2006 DHS, we compare these methods for age-specific mortality between ages 0 and 60 months and, specifically, look at how differences between the methods change when the interval  $[0,60)$  is discretized into different numbers and sizes of age groups and when time is measured in units of months or years.

### **3.4 A comparison using Senegal 2006 DHS**

In this section, we compare the Lexis methods of IGME and DHS, both with and without the adjustment used for the period just before survey administration, to discrete time methods that more closely resemble life table methods. We consider three slight variations in the discrete time methods. The age at death is reported in a child's DHS birth history in months for age at death  $a^* \in [0, 24)$  months. Though there are a few exceptions, ages at death reported in DHS surveys for age at death  $a^* \in [24, 60)$  are reported in years, so that

$a^* \in [24, 36)$  is reported as 24 months,  $a^* \in [36, 48)$  is reported as 36 months, and  $a^* \in [48, 60)$  is reported as 48 months. So, we look at three ways of counting exposures in application to DHS data when the observed age at death for an individual  $k$ ,  $a_k^* \equiv 0 \pmod{12}$ :

1. **Unadjusted:** Assume  $a_k^*$  is correct, e.g. deaths marked age  $a = 48$  occur before reaching  $a = 49$
2. **Midpoint:** Assume the real age at death is  $a_k^* + 5$ , e.g. deaths marked age  $a = 48$  occur at the midpoint of the year, age  $a = 53$  months.
3. **Uniform:** Assume the real age at death is  $a_k^* + U_k$ , where  $U_k \sim \text{Uniform}(0, 11)$  as in the Ecuador Reproductive Survey (Population Division, 2011).

The following subsections present results for counting of deaths and exposures, as well as estimation of period-mortality for the period [2000, 2005). Not only do we compare across the two Lexis variants and three discrete time approaches outlined above, we compare across several partitions of the age interval [0,60). We always count time in months and consider the following:

1. One age band, [0, 60)
2. Two age bands, [0, 12), [12, 60)
3. Five age bands, [0, 12), [12, 24), [24, 36), [36, 48), [48, 60), and
4. Eight age bands, [0, 1), [1, 3), [3, 6), [6, 12), [12, 24), [24, 36), [36, 48), [48, 60).

### 3.4.1 Results: deaths and exposures

In this section, we compare the number of deaths and measures of exposure counted by method for the period [2000, 2005). We now examine how the different methods estimate deaths and exposures – both as a proportion of the age band the child survived and as number of personmonths. Figures 3.3–3.6 show the number of deaths counted by each of Cohorts A, B, and C. The colors of the bars represent the method variants. Different blues representing life table or discrete time survival methods that differ in treatment of age at death above 24 months, and yellow represent using an adjustment for the timing of interview or not, respectively. Figures 3.3–3.5 show similar results for counting of deaths, regardless of the number of age bands. The Lexis methods and discrete time survival methods treat all individuals in Cohort B equivalently for 2, 5, 8 age bands, i.e. (3.1), (3.2), and (3.4) are equal for Cohort B. Cohort B is always the largest cohort, so the equivalency of these counts means most children’s birth histories are being treated the same by each method. However, we see a standard pattern for Cohorts A and C: Lexis methods count more deaths in Cohort A and life table methods count more deaths in Cohort C. When one chooses a single age band, Cohort B is undefined, and the other cohorts are larger in size. Figure 3.6 shows that the number of deaths counted for Cohorts A and C are less different than when the age interval is partitioned further; however, we lose the nice finding that deaths are counted exactly the same across all methods for Cohort B.

When it comes to counting exposures, things are a bit less clear across methods. Figures 3.7–3.10 show the proportion of the age band each method counts. Figures 3.11–3.14 show the number of personmonths counted in each age band by method. When counting exposure as the proportion of the age band each child was alive for, there are less discrepancies in Cohorts A and C than when counting deaths. However, exposure counts for Cohort B are not equivalent as with death counts. Figures 3.7–3.9 all see a slightly larger number of exposure counts for Lexis methods in Cohort A. There is mixed evidence of a generally

larger number of exposures counted for Cohort C in life table methods, and the situation is certainly not as clear as the patterns in death counts. Though the differences are small, we see in Figures 3.7 and 3.8 the expected larger counts of exposures for life table methods using the midpoint or uniform adjustment for deaths at older ages. Though the adjustment for the date of interview in the Lexis methods, by definition, counts more exposures for deaths in Cohort C, this is barely visible in the results. This is to be expected, as the difference in exposure only applies to children who died in Cohort C and not all children in Cohort C.

When we begin to compare personmonths, the results are similar to those in Figures 3.7–3.10. The discrepancies are less consequential, however. Noting the height of of the vertical axes in Figures 3.11–3.14, though there are discrepancies across methods they make less practical difference than the differences in counts of deaths. Perhaps the most notable difference, is when we consider a single age band in Figure 3.14. Departing from the patterns we saw with deaths for Cohorts A and C, the Lexis methods count more personmonths for both Cohorts A and C. Combining this with the results in Figure 3.6, we see that the denominators would remain constant by method for either cohort though the Lexis methods count more deaths for Cohort A and the life table methods count more deaths for Cohort C. This discrepancy in the denominators between methods is also larger in number than in Figure 3.10.

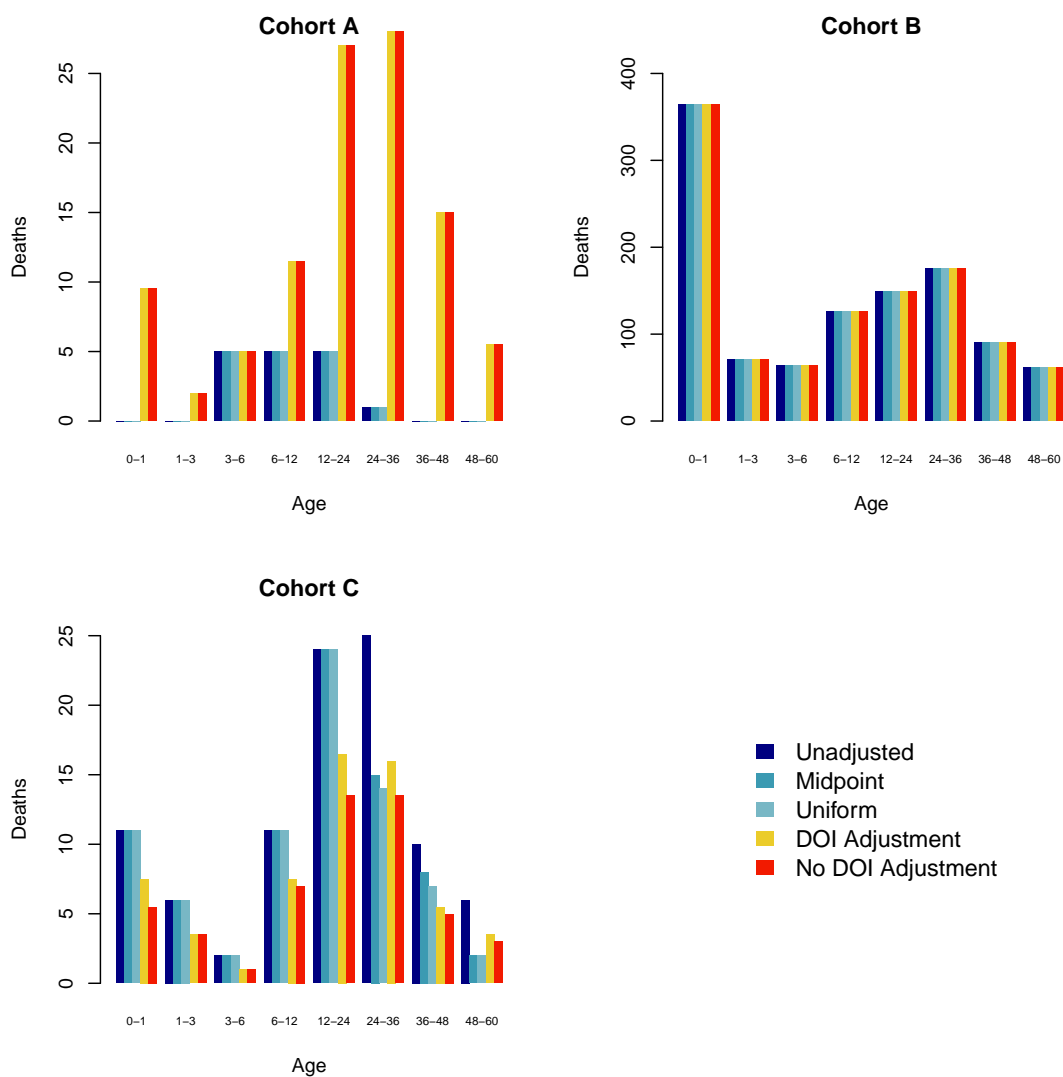


Figure 3.3: Comparison of deaths: 8 age bands. Death counts are compared across Cohorts A, B, C. Different blues represent life table methods that differ only in the way the treat the reported age at death for ages greater than or equal to 24. Red and yellow represent Lexis methods used by UN IGME and the DHS, the only difference indicating whether an adjustment for death counts in the period prior to the timing of the survey or date of interview (DOI) in months is made for Cohort C.

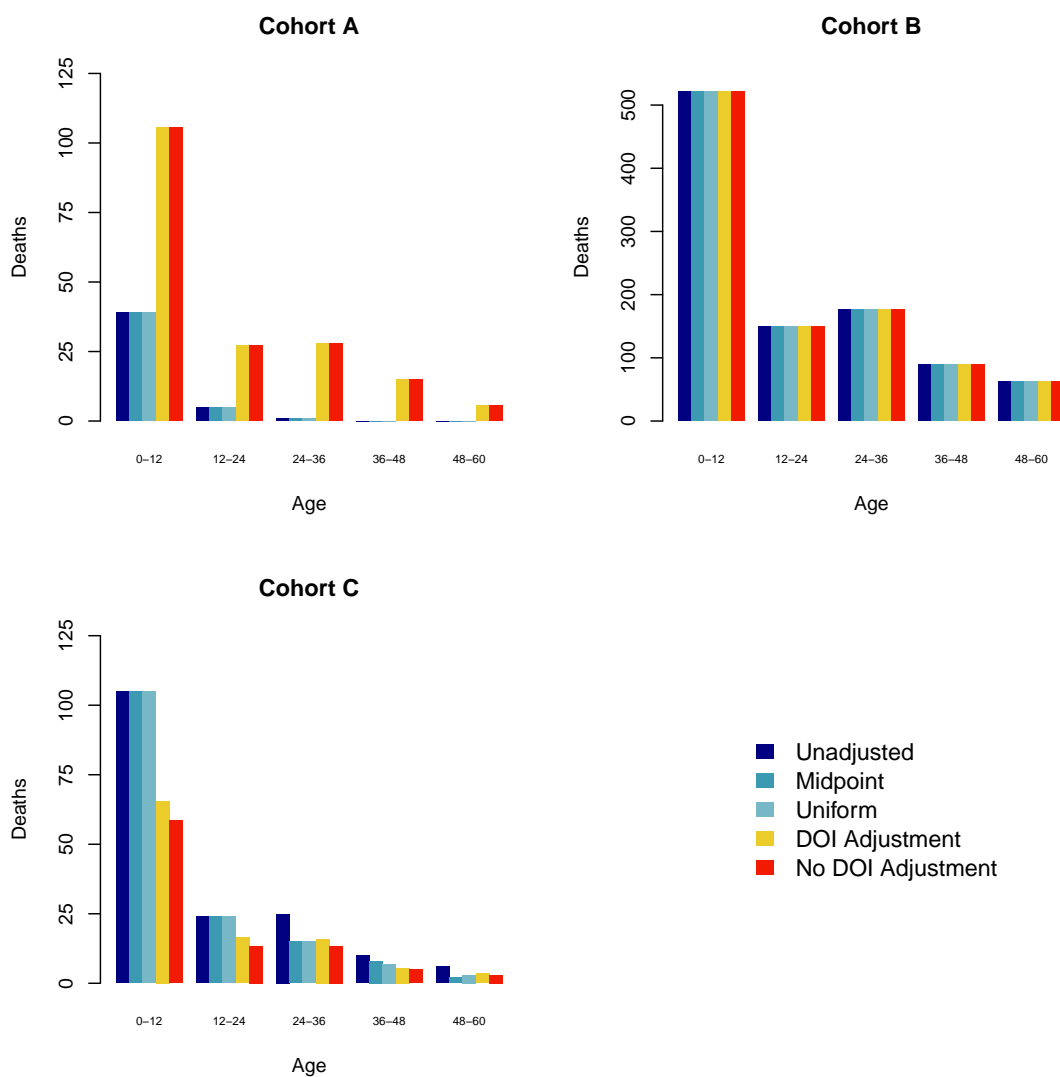


Figure 3.4: Comparison of deaths: 5 age bands. Comparison of deaths: 8 age bands. Death counts are compared across Cohorts A, B, C. Different blues represent life table methods that differ only in the way the treat the reported age at death for ages greater than or equal to 24. Red and yellow represent Lexis methods used by UN IGME and the DHS, the only difference indicating whether an adjustment for death counts in the period prior to the timing of the survey or date of interview (DOI) in months is made for Cohort C.

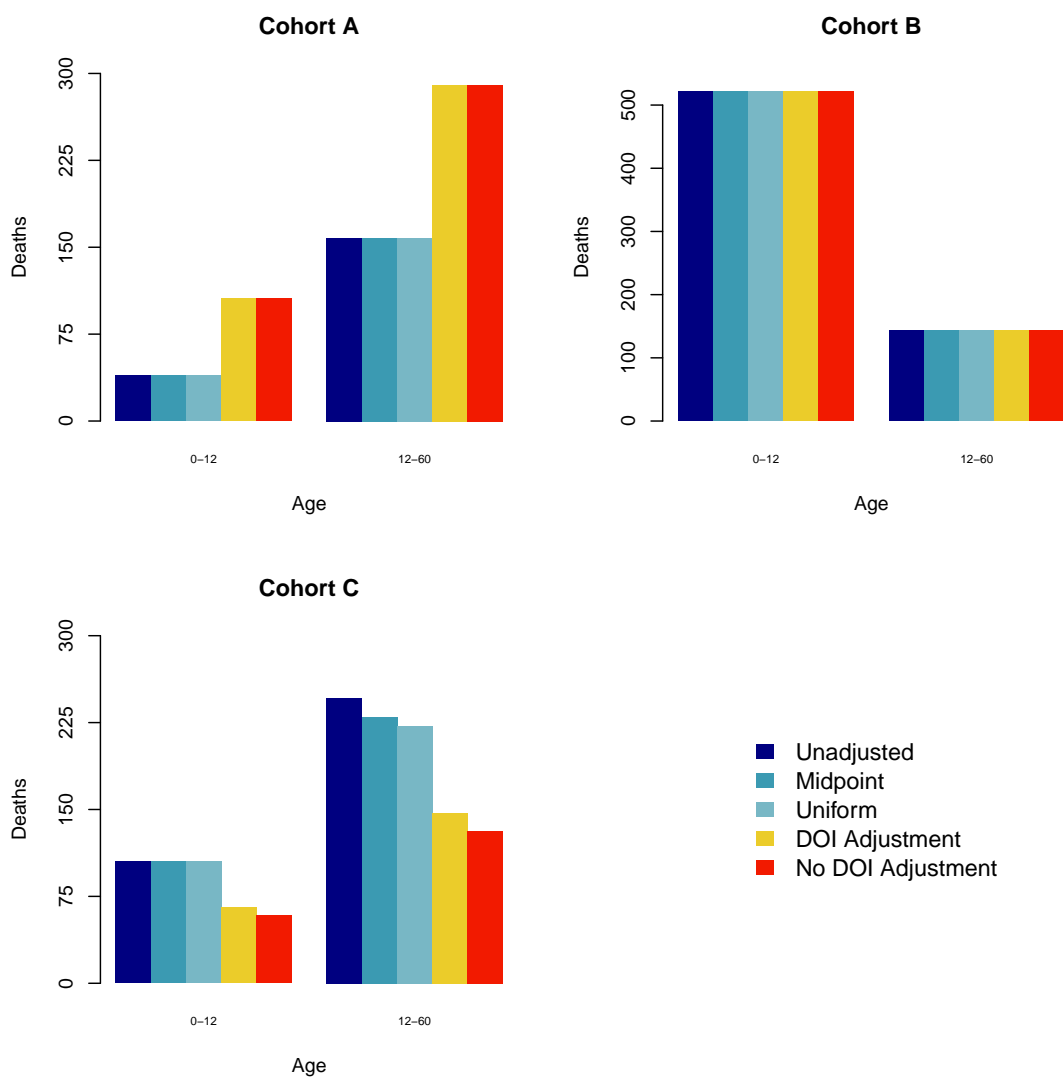


Figure 3.5: Comparison of deaths: 2 age bands. Comparison of deaths: 8 age bands. Death counts are compared across Cohorts A, B, C. Different blues represent life table methods that differ only in the way the treat the reported age at death for ages greater than or equal to 24. Red and yellow represent Lexis methods used by UN IGME and the DHS, the only difference indicating whether an adjustment for death counts in the period prior to the timing of the survey or date of interview (DOI) in months is made for Cohort C.

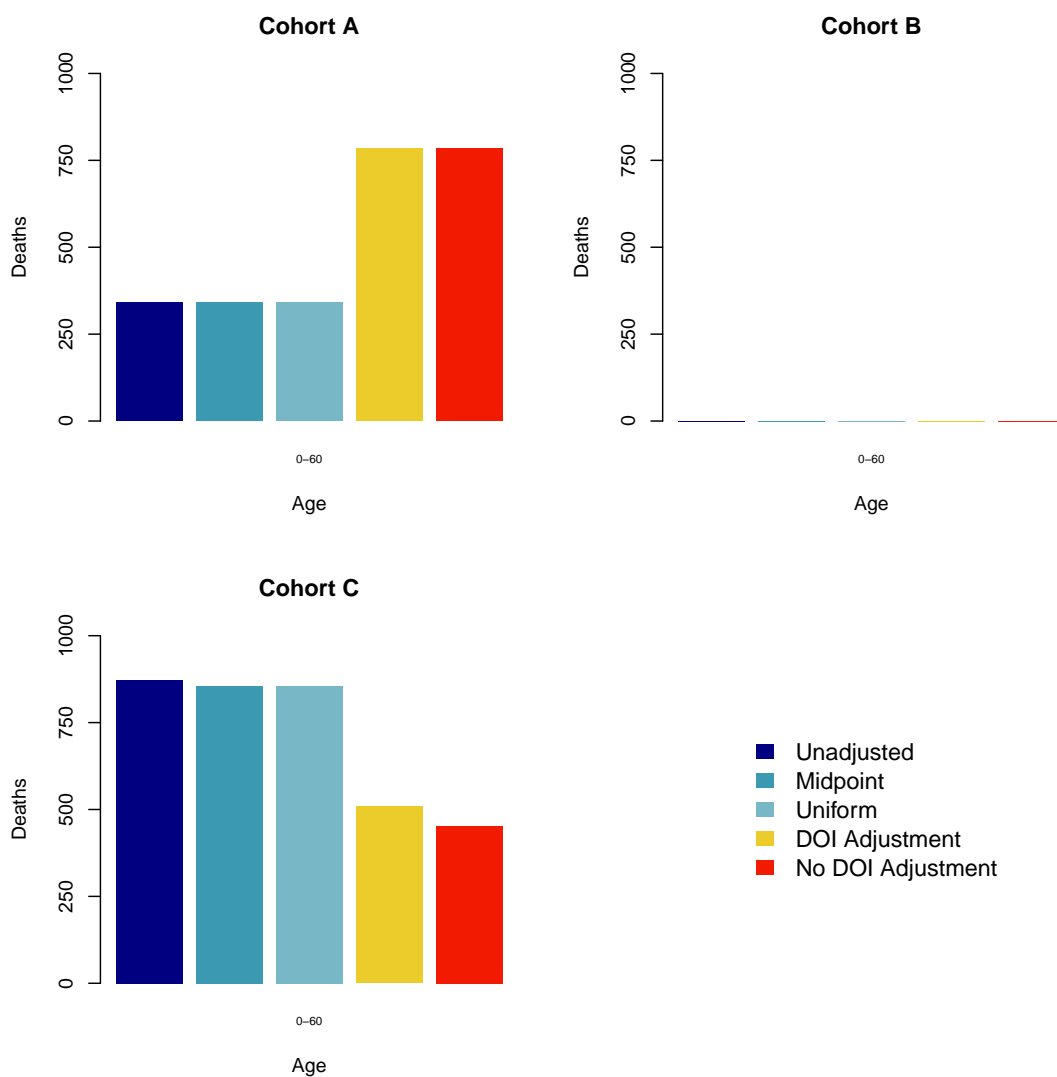


Figure 3.6: Comparison of deaths: 1 age band. Comparison of deaths: 8 age bands. Death counts are compared across Cohorts A, B, C. Different blues represent life table methods that differ only in the way the treat the reported age at death for ages greater than or equal to 24. Red and yellow represent Lexis methods used by UN IGME and the DHS, the only difference indicating whether an adjustment for death counts in the period prior to the timing of the survey or date of interview (DOI) in months is made for Cohort C.

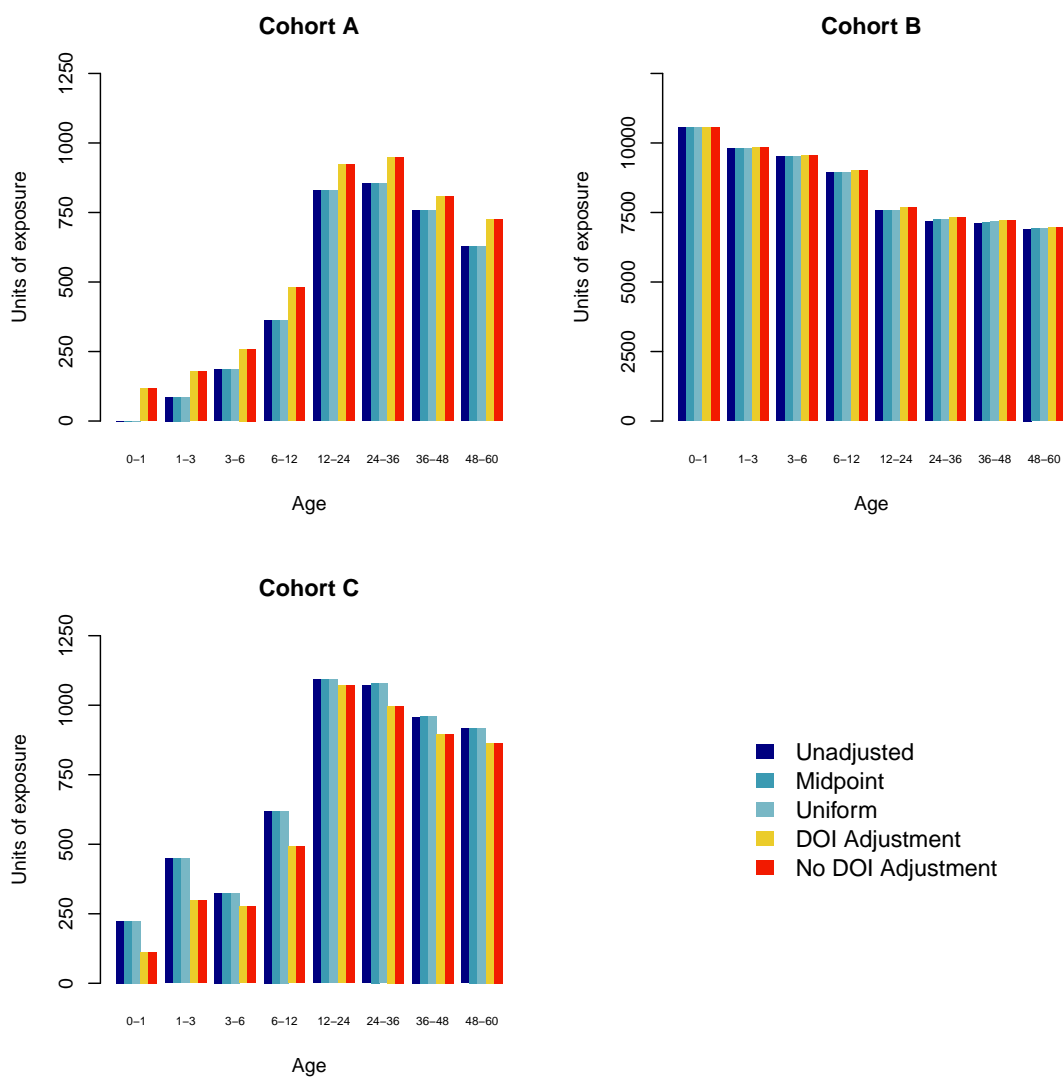


Figure 3.7: Comparison of exposures: 8 age bands. Comparison of deaths: 8 age bands. Death counts are compared across Cohorts A, B, C. Different blues represent life table methods that differ only in the way they treat the reported age at death for ages greater than or equal to 24. Red and yellow represent Lexis methods used by UN IGME and the DHS, the only difference indicating whether an adjustment for death counts in the period prior to the timing of the survey or date of interview (DOI) in months is made for Cohort C.

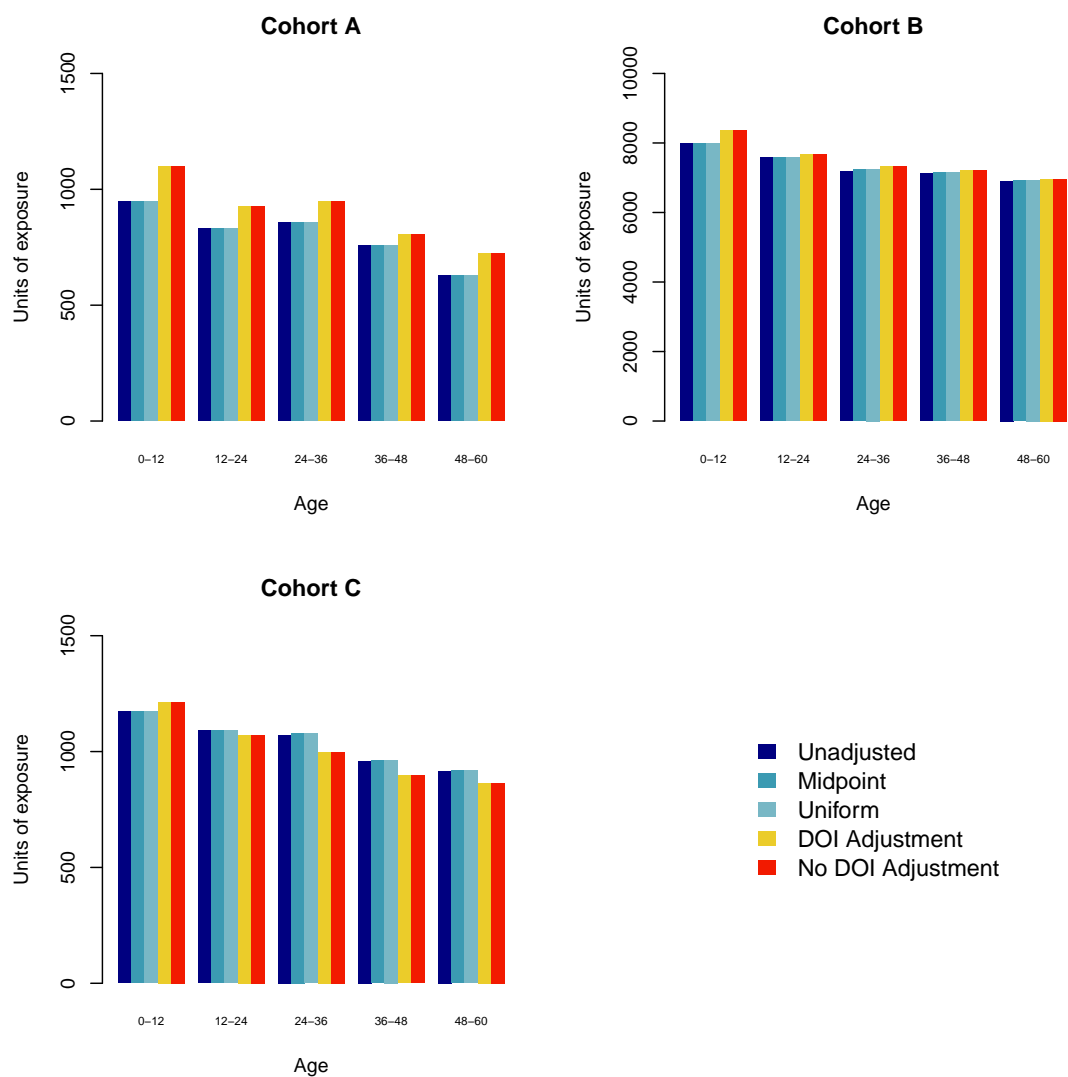


Figure 3.8: Comparison of exposures: 2 age bands. Comparison of deaths: 8 age bands. Death counts are compared across Cohorts A, B, C. Different blues represent life table methods that differ only in the way they treat the reported age at death for ages greater than or equal to 24. Red and yellow represent Lexis methods used by UN IGME and the DHS, the only difference indicating whether an adjustment for death counts in the period prior to the timing of the survey or date of interview (DOI) in months is made for Cohort C.

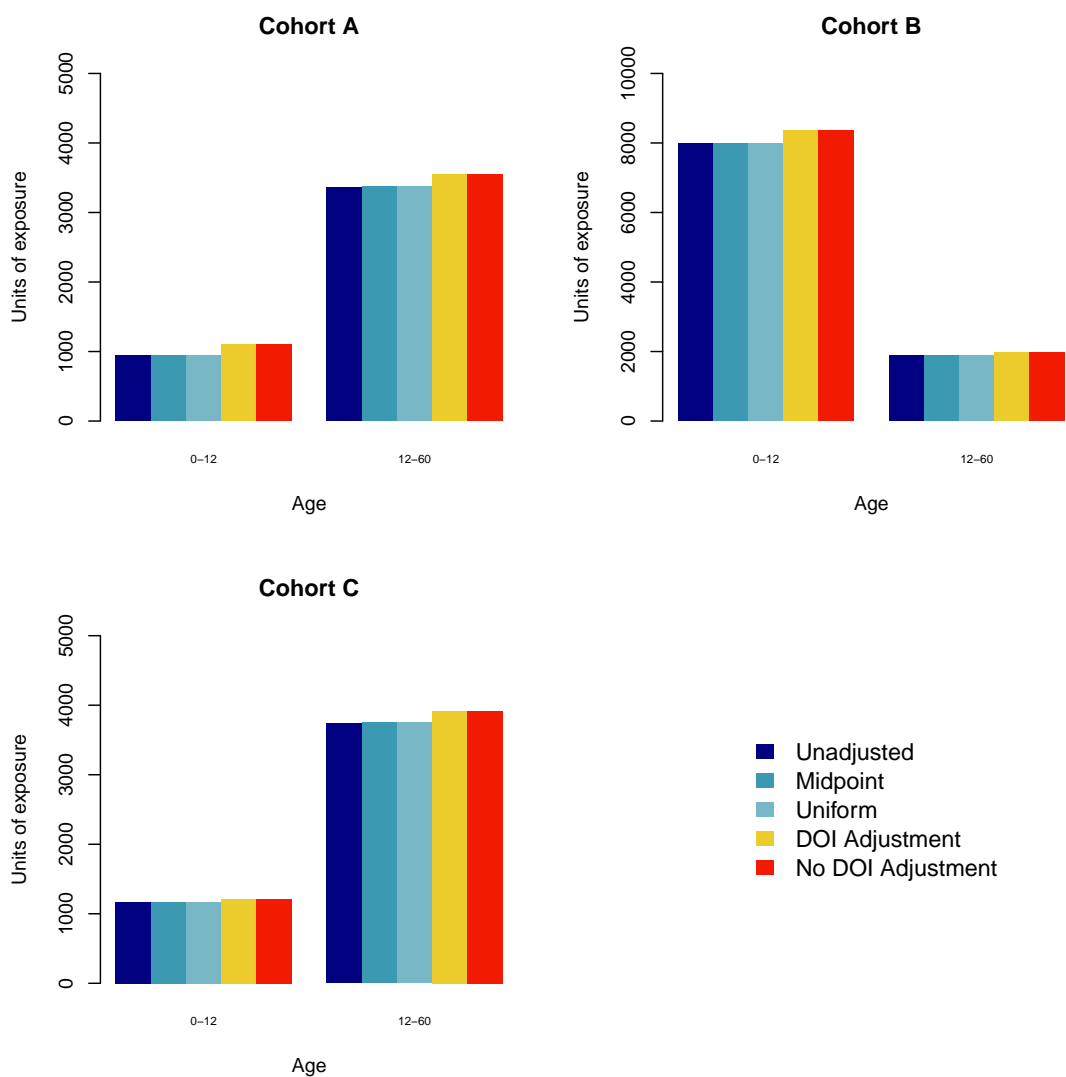


Figure 3.9: Comparison of exposures: 2 age bands. Comparison of deaths: 8 age bands. Death counts are compared across Cohorts A, B, C. Different blues represent life table methods that differ only in the way they treat the reported age at death for ages greater than or equal to 24. Red and yellow represent Lexis methods used by UN IGME and the DHS, the only difference indicating whether an adjustment for death counts in the period prior to the timing of the survey or date of interview (DOI) in months is made for Cohort C.

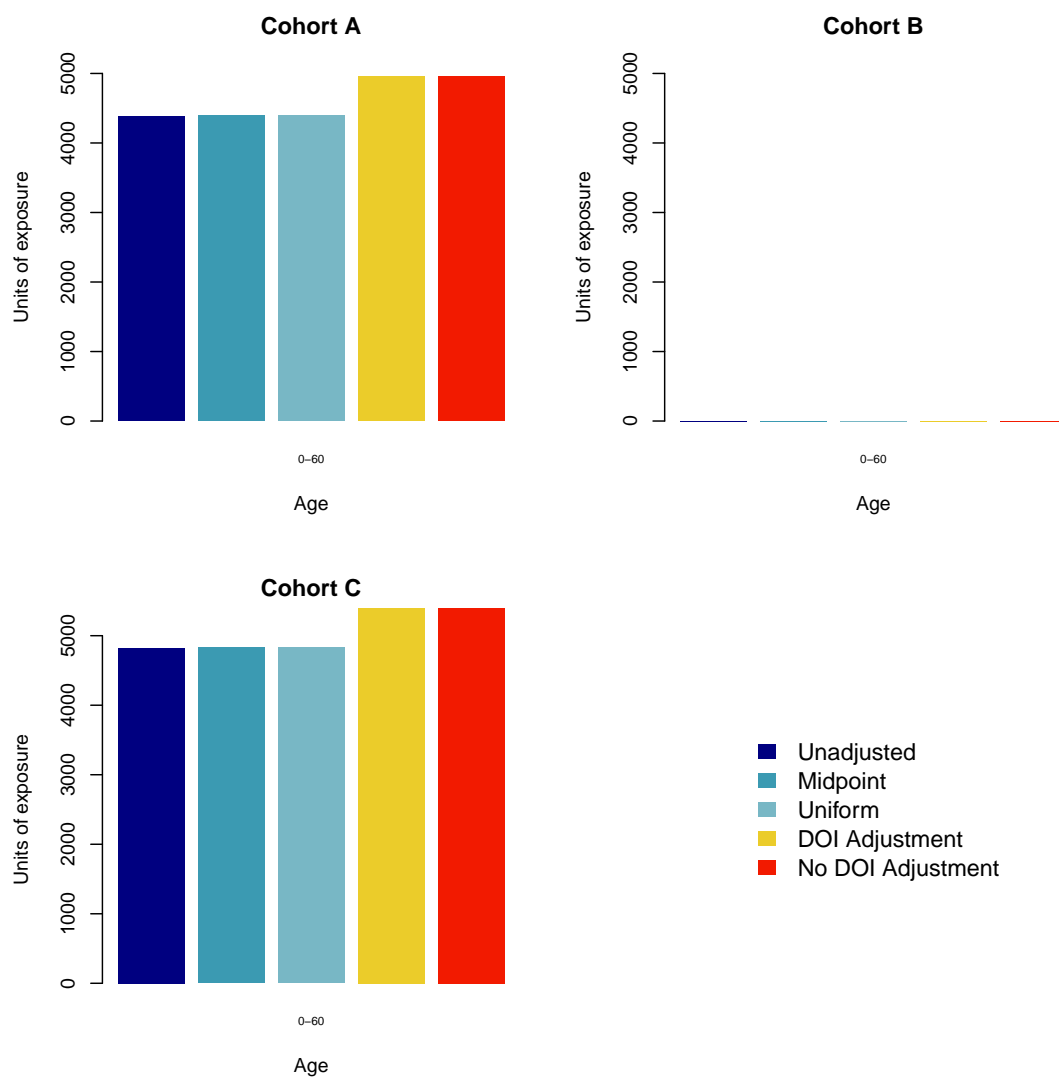


Figure 3.10: Comparison of exposures: 1 age bands. Comparison of deaths: 8 age bands. Death counts are compared across Cohorts A, B, C. Different blues represent life table methods that differ only in the way they treat the reported age at death for ages greater than or equal to 24. Red and yellow represent Lexis methods used by UN IGME and the DHS, the only difference indicating whether an adjustment for death counts in the period prior to the timing of the survey or date of interview (DOI) in months is made for Cohort C.

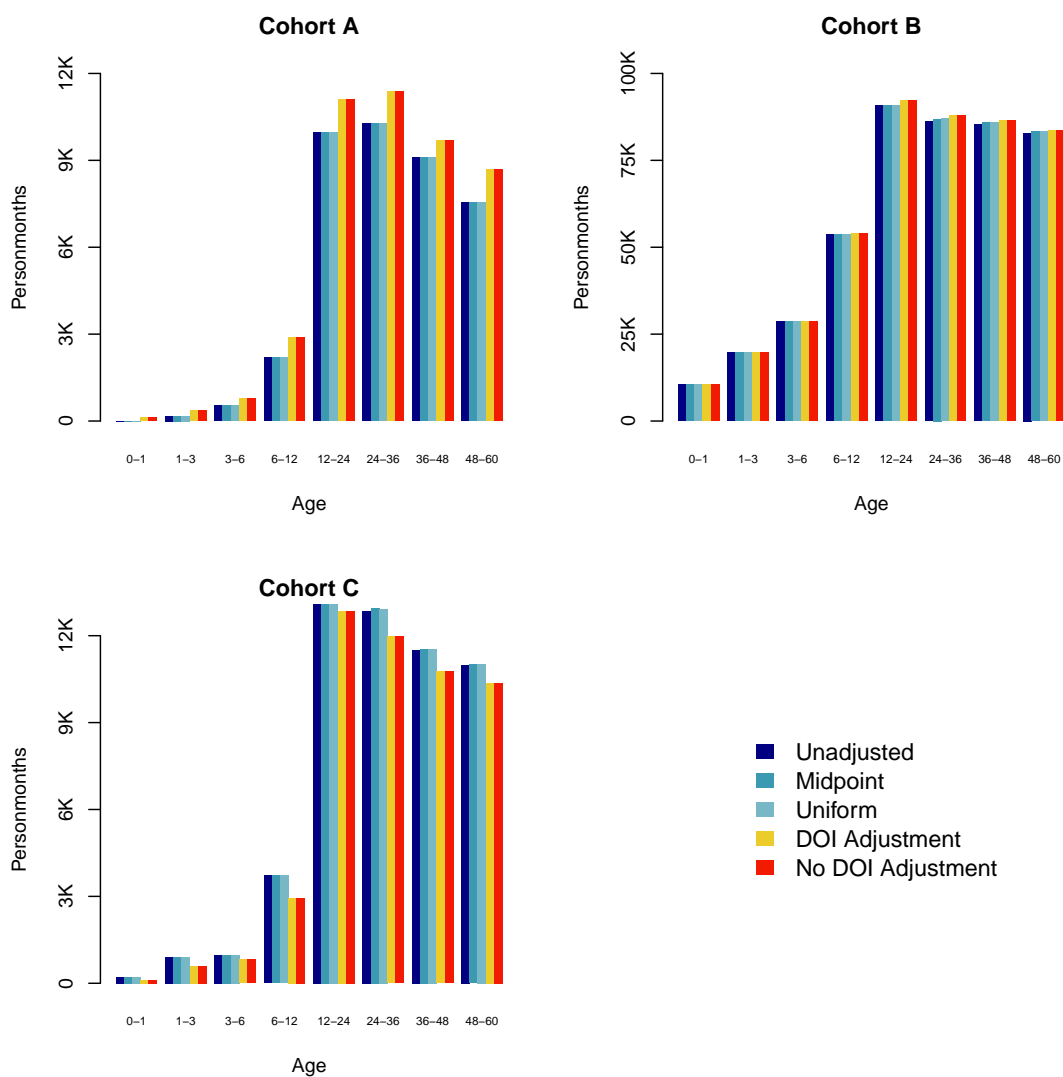


Figure 3.11: Comparison of personmonths: 8 age bands. Comparison of deaths: 8 age bands. Death counts are compared across Cohorts A, B, C. Different blues represent life table methods that differ only in the way they treat the reported age at death for ages greater than or equal to 24. Red and yellow represent Lexis methods used by UN IGME and the DHS, the only difference indicating whether an adjustment for death counts in the period prior to the timing of the survey or date of interview (DOI) in months is made for Cohort C.

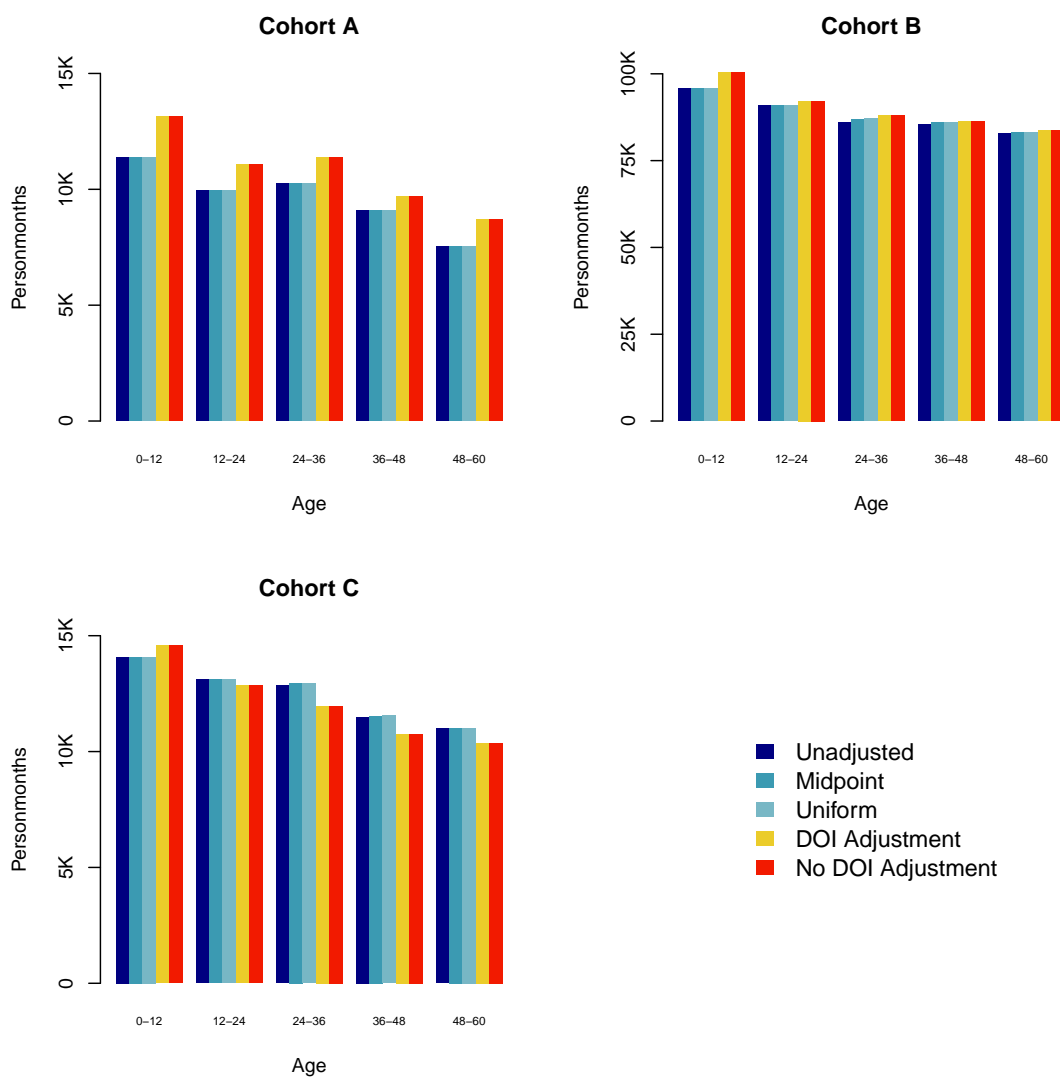


Figure 3.12: Comparison of personmonths: 2 age bands. Comparison of deaths: 8 age bands. Death counts are compared across Cohorts A, B, C. Different blues represent life table methods that differ only in the way they treat the reported age at death for ages greater than or equal to 24. Red and yellow represent Lexis methods used by UN IGME and the DHS, the only difference indicating whether an adjustment for death counts in the period prior to the timing of the survey or date of interview (DOI) in months is made for Cohort C.

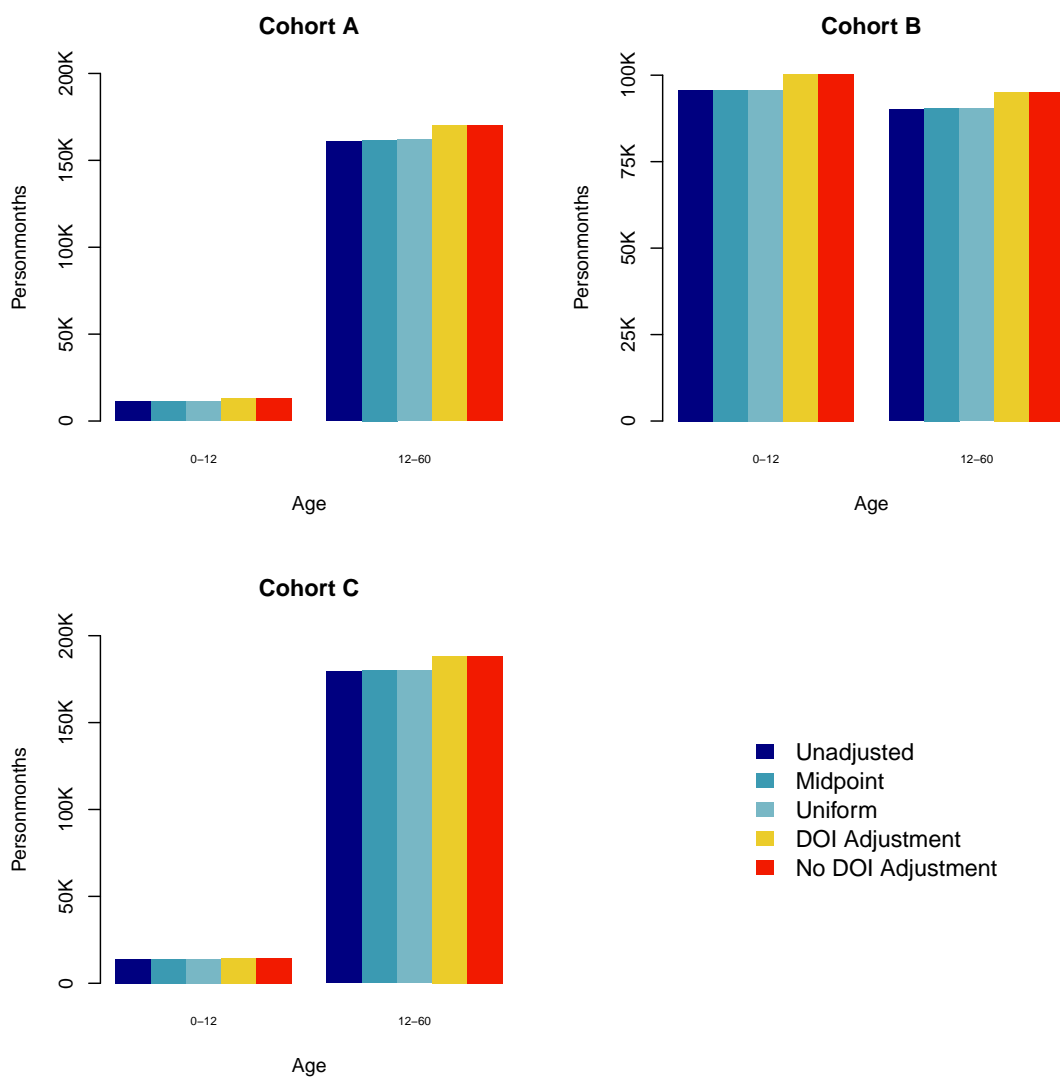


Figure 3.13: Comparison of personmonths: 2 age bands. Comparison of deaths: 8 age bands. Death counts are compared across Cohorts A, B, C. Different blues represent life table methods that differ only in the way they treat the reported age at death for ages greater than or equal to 24. Red and yellow represent Lexis methods used by UN IGME and the DHS, the only difference indicating whether an adjustment for death counts in the period prior to the timing of the survey or date of interview (DOI) in months is made for Cohort C.

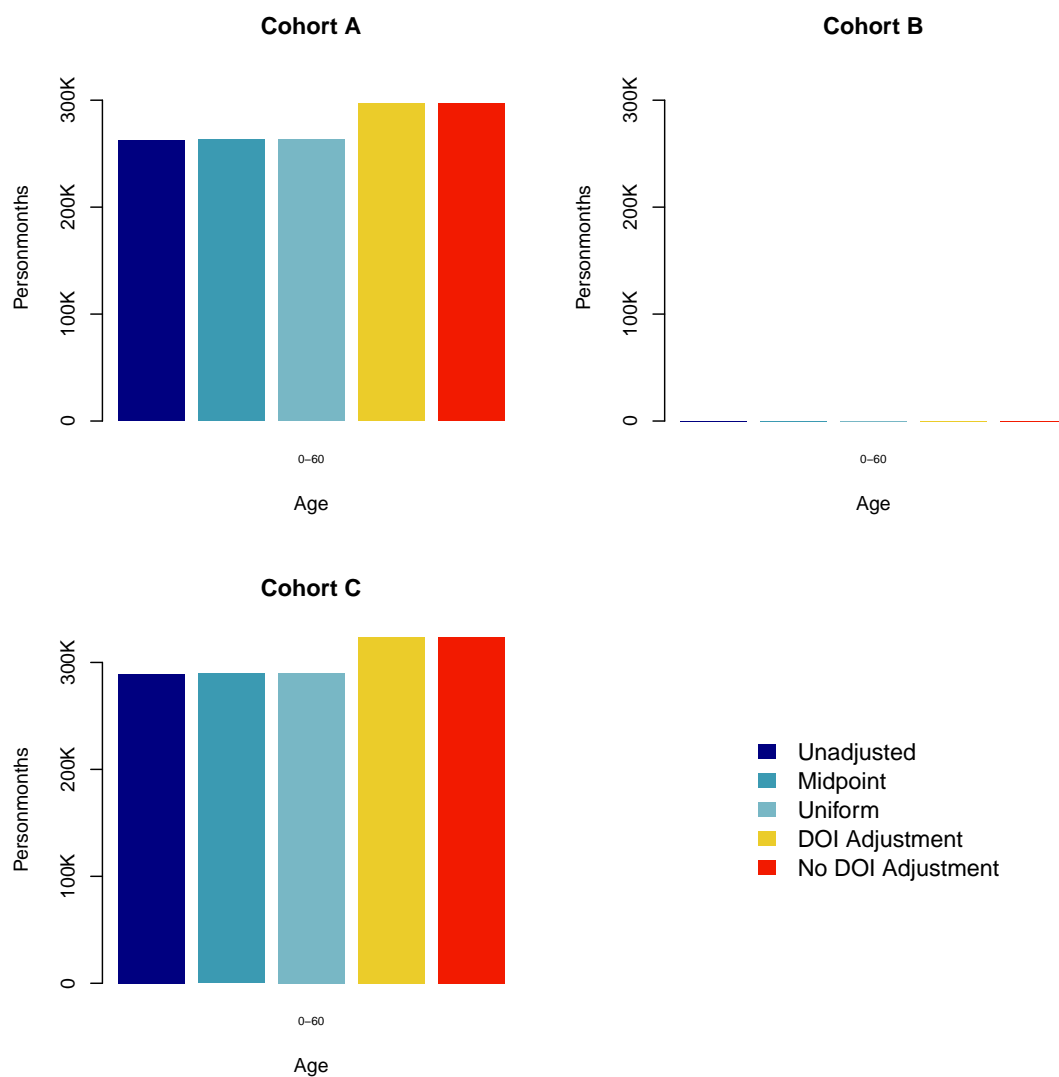


Figure 3.14: Comparison of personmonths: 1 age bands. Comparison of deaths: 8 age bands. Death counts are compared across Cohorts A, B, C. Different blues represent life table methods that differ only in the way they treat the reported age at death for ages greater than or equal to 24. Red and yellow represent Lexis methods used by UN IGME and the DHS, the only difference indicating whether an adjustment for death counts in the period prior to the timing of the survey or date of interview (DOI) in months is made for Cohort C.

### 3.4.2 Results: child mortality

First, we compare two different Horvitz-Thompson estimators of mortality in each age group. Figures 3.15–3.18 show the ratio of deaths to exposures as presented in Figures 3.7–3.10. Figures 3.19–3.22 show the ratio of deaths to exposures as presented in Figures 3.11–3.14. Despite the noted differences in counting deaths and exposures by each method, once uncertainty is accounted for, there does not appear to be many important differences in estimators of age band mortality. The differences in the two Lexis method estimates and the general correspondence of the estimates when the interview date is accounted for in the life table methods, indicate that this adjustment is important. When we consider only one age band, we see our first indication that perhaps the estimate from the Lexis methods without a date of interview adjustment is biased. One more notable result in Figures 3.15 and 3.16, is a clear consensus from all methods that age-specific mortality is not monotonically decreasing across age bands. Finally, we see more clearly the effects of not adjusting in the life table methods for the misreporting of age at death in older ages, as the dark blue estimates are a bit higher than the other life table methods. However, again, the uncertainty makes this difference less important. The monthly hazards calculated by dividing by the number of personmonths instead of exposures in Figures 3.19–3.22 show the same results.

Finally, we compare estimates of  ${}_{60}q_0$  using the two different age band measures of mortality in Figures 3.15–3.18 and Figures 3.19–3.22 aggregated up to U5MR to those estimated via logistic regression and discrete time survival methods. Figures 3.23–3.26 use the age specific mortality in Figures 3.15–3.18. Figures 3.27–3.30 use the age specific mortality in Figures 3.19–3.22. Finally, Figures 3.31–3.34 use discrete time survival analysis. In general, once we aggregate up to U5MR, there is no statistically significant difference in the estimates. So, we will only discuss the notable differences. All of the differences and issues with one age band in the age band specific mortality measures, exist here as well (as there is no aggregation). Not adjusting for the interview date in Lexis methods, yields an estimate that

appears to be quite off from any other. The most notable issue, is with the discrete time survival analysis with 8 age bands. When we separate out the first month of life, we cannot fit a logistic regression model as there are some individuals for whom  $y = 1$  and  $n = 0.5$  when the date of interview is adjusted for. It is interesting that the date of interview adjustment seems to bridge the gap between Lexis and life table methods, but is not helpful if one wants to use a survival analysis approach.

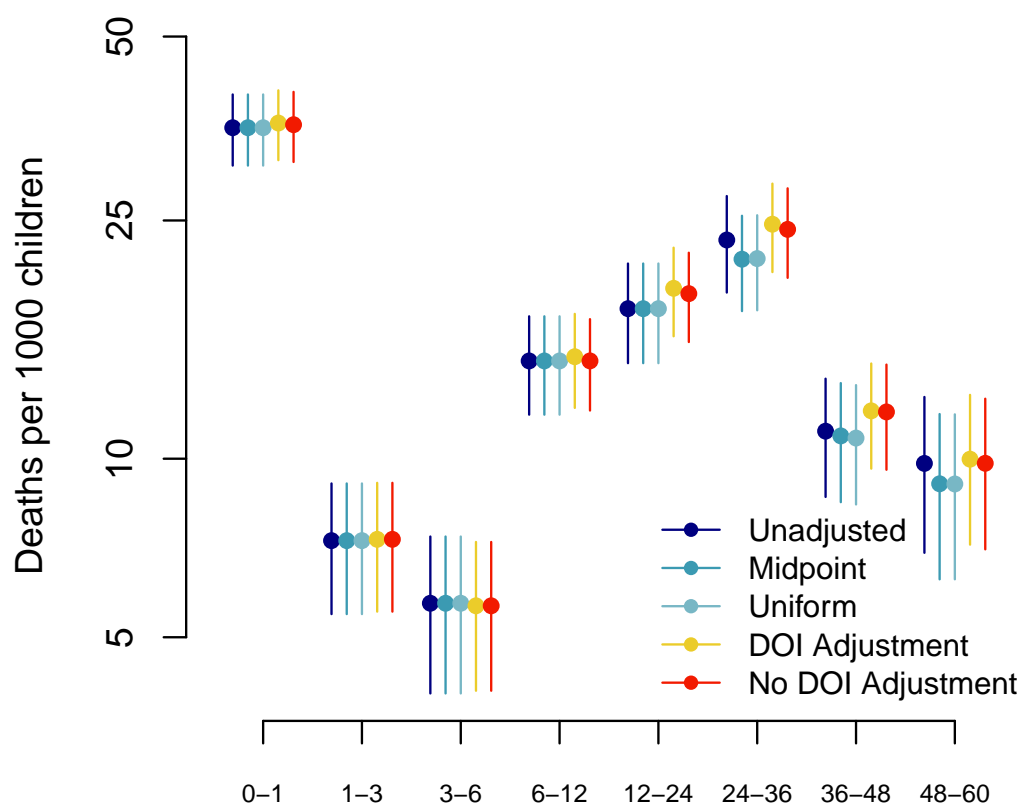


Figure 3.15: Comparison of Horvitz-Thompson estimates of deaths divided by exposures with 8 age bands. Comparison of deaths: 8 age bands. Death counts are compared across Cohorts A, B, C. Different blues represent life table methods that differ only in the way the treat the reported age at death for ages greater than or equal to 24. Red and yellow represent Lexis methods used by UN IGME and the DHS, the only difference indicating whether an adjustment for death counts in the period prior to the timing of the survey or date of interview (DOI) in months is made for Cohort C.

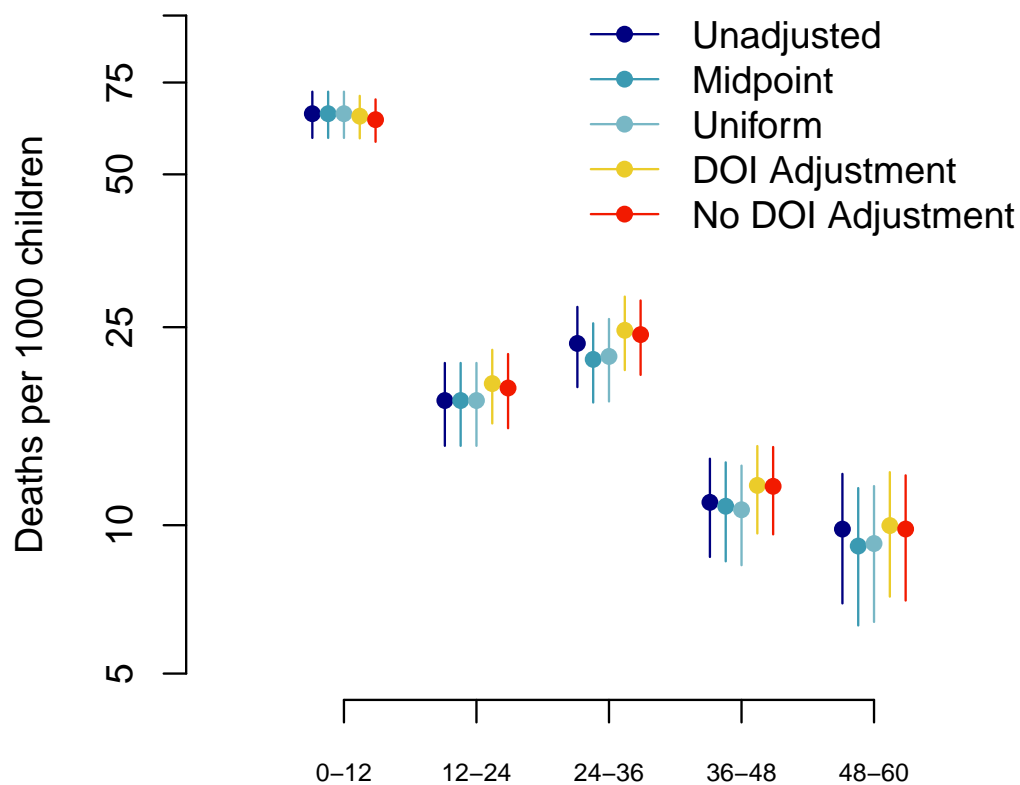


Figure 3.16: Comparison of Horvitz-Thompson estimates of deaths divided by exposures with 5 age bands. Comparison of deaths: 8 age bands. Death counts are compared across Cohorts A, B, C. Different blues represent life table methods that differ only in the way the treat the reported age at death for ages greater than or equal to 24. Red and yellow represent Lexis methods used by UN IGME and the DHS, the only difference indicating whether an adjustment for death counts in the period prior to the timing of the survey or date of interview (DOI) in months is made for Cohort C.

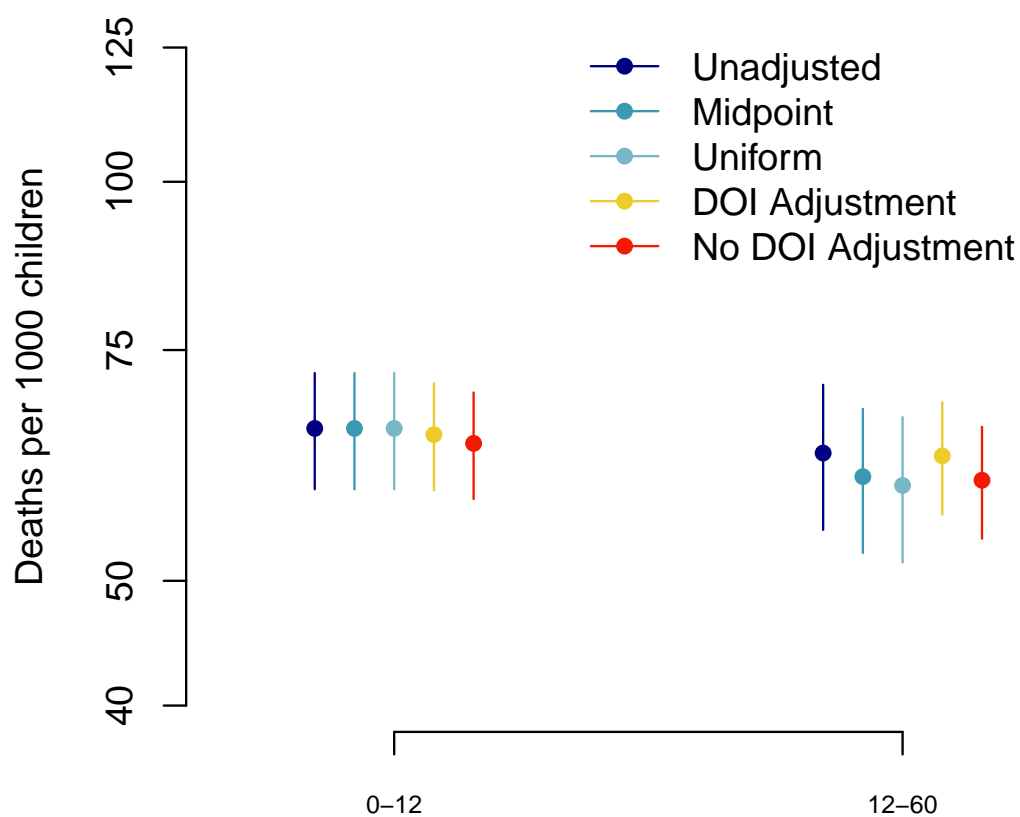


Figure 3.17: Comparison of Horvitz-Thompson estimates of deaths divided by exposures with 2 age bands. Comparison of deaths: 8 age bands. Death counts are compared across Cohorts A, B, C. Different blues represent life table methods that differ only in the way the treat the reported age at death for ages greater than or equal to 24. Red and yellow represent Lexis methods used by UN IGME and the DHS, the only difference indicating whether an adjustment for death counts in the period prior to the timing of the survey or date of interview (DOI) in months is made for Cohort C.

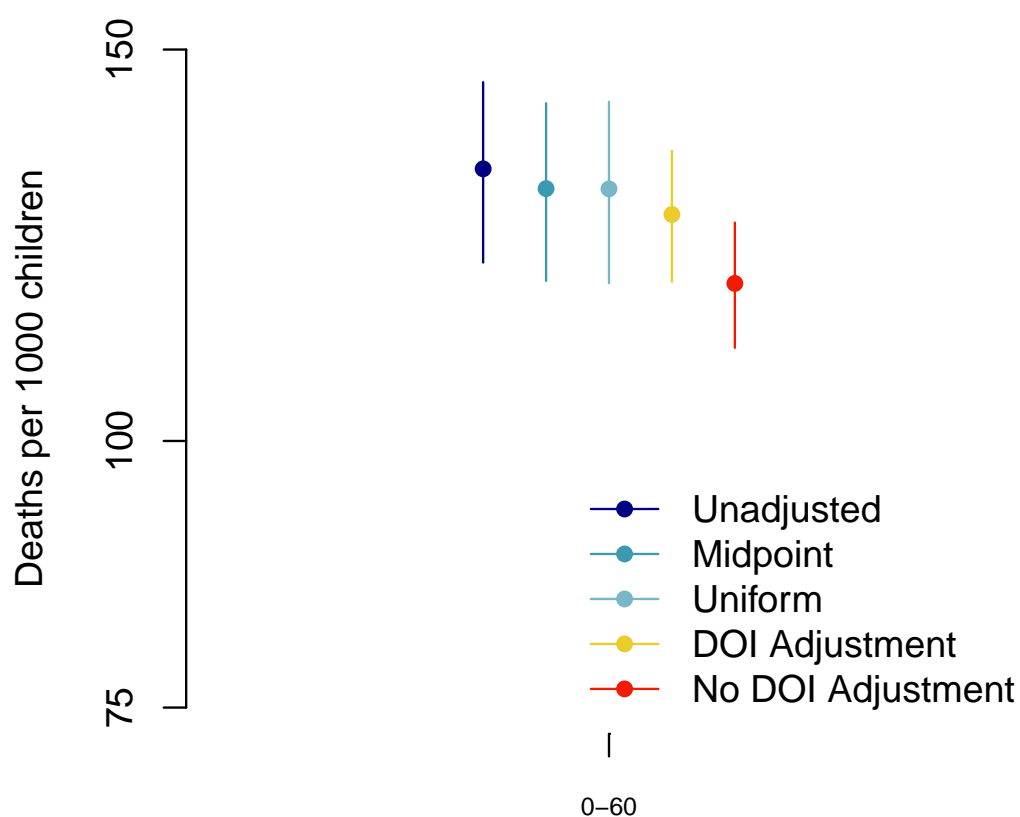


Figure 3.18: Comparison of Horvitz-Thompson estimates of deaths divided by exposures with 1 age band. Comparison of deaths: 8 age bands. Death counts are compared across Cohorts A, B, C. Different blues represent life table methods that differ only in the way the treat the reported age at death for ages greater than or equal to 24. Red and yellow represent Lexis methods used by UN IGME and the DHS, the only difference indicating whether an adjustment for death counts in the period prior to the timing of the survey or date of interview (DOI) in months is made for Cohort C.

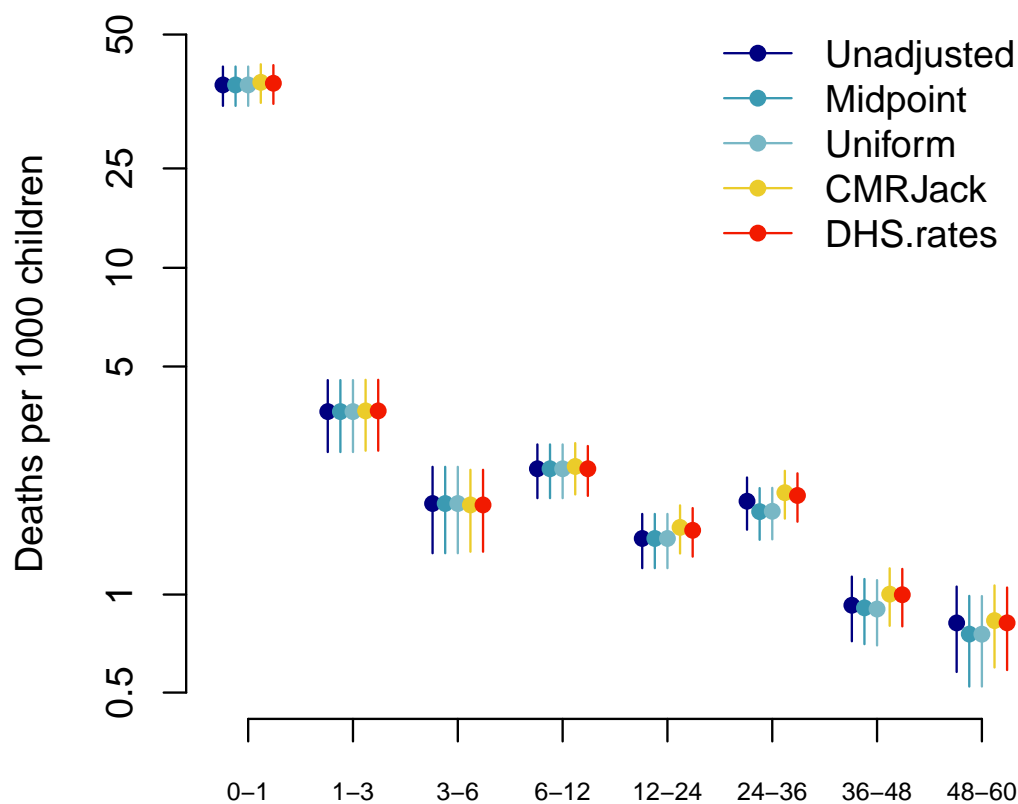


Figure 3.19: Comparison of Horvitz-Thompson estimates of deaths divided by personmonths with 8 age bands. Comparison of deaths: 8 age bands. Death counts are compared across Cohorts A, B, C. Different blues represent life table methods that differ only in the way the treat the reported age at death for ages greater than or equal to 24. Red and yellow represent Lexis methods used by UN IGME and the DHS, the only difference indicating whether an adjustment for death counts in the period prior to the timing of the survey or date of interview (DOI) in months is made for Cohort C.

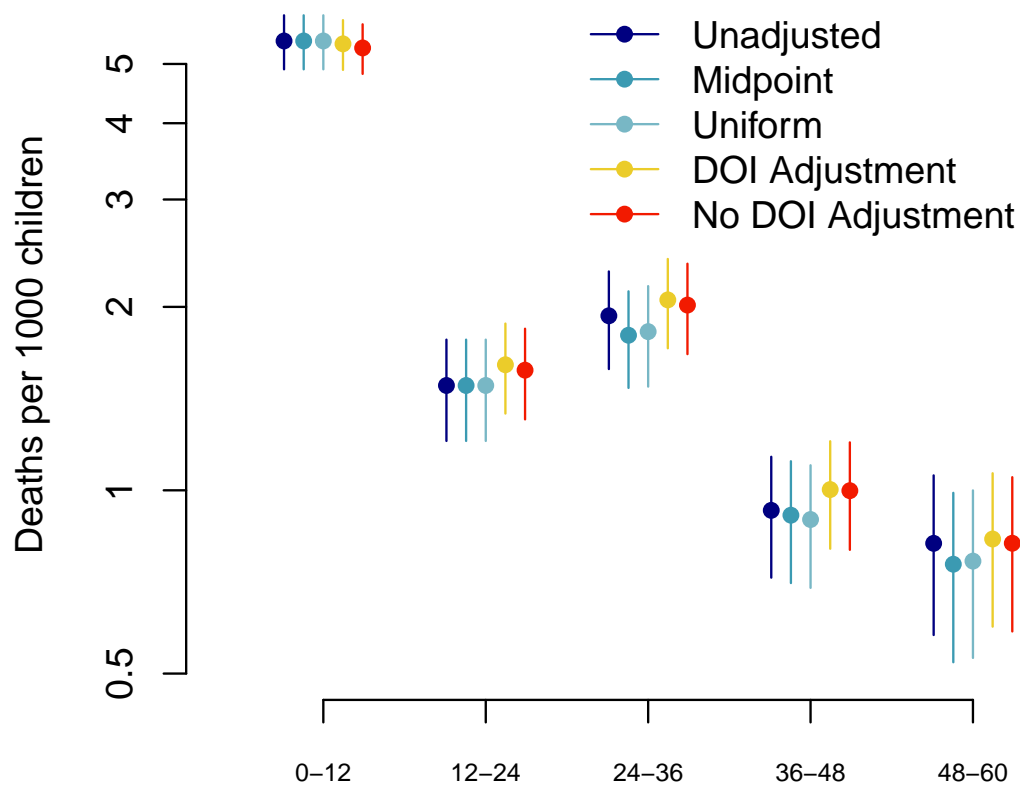


Figure 3.20: Comparison of Horvitz-Thompson estimates of deaths divided by personmonths with 5 age bands. Comparison of deaths: 8 age bands. Death counts are compared across Cohorts A, B, C. Different blues represent life table methods that differ only in the way the treat the reported age at death for ages greater than or equal to 24. Red and yellow represent Lexis methods used by UN IGME and the DHS, the only difference indicating whether an adjustment for death counts in the period prior to the timing of the survey or date of interview (DOI) in months is made for Cohort C.

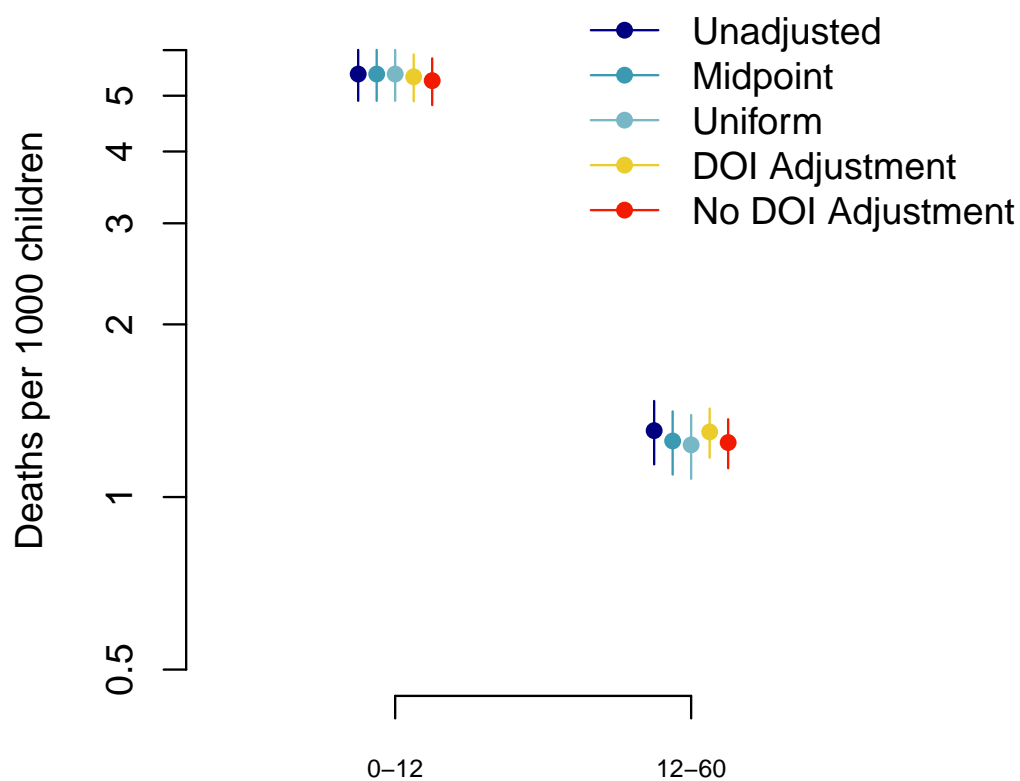


Figure 3.21: Comparison of Horvitz-Thompson estimates of deaths divided by personmonths with 2 age bands. Comparison of deaths: 8 age bands. Death counts are compared across Cohorts A, B, C. Different blues represent life table methods that differ only in the way the treat the reported age at death for ages greater than or equal to 24. Red and yellow represent Lexis methods used by UN IGME and the DHS, the only difference indicating whether an adjustment for death counts in the period prior to the timing of the survey or date of interview (DOI) in months is made for Cohort C.

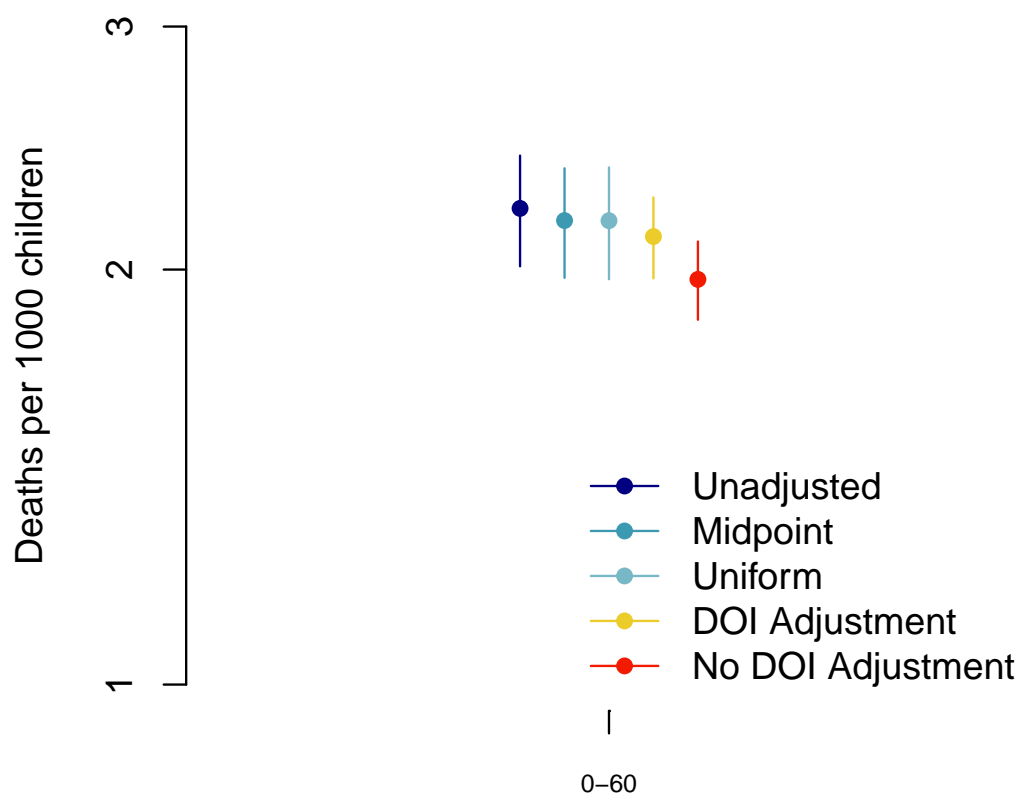


Figure 3.22: Comparison of Horvitz-Thompson estimates of deaths divided by personmonths with 1 age band. Comparison of deaths: 8 age bands. Death counts are compared across Cohorts A, B, C. Different blues represent life table methods that differ only in the way the treat the reported age at death for ages greater than or equal to 24. Red and yellow represent Lexis methods used by UN IGME and the DHS, the only difference indicating whether an adjustment for death counts in the period prior to the timing of the survey or date of interview (DOI) in months is made for Cohort C.

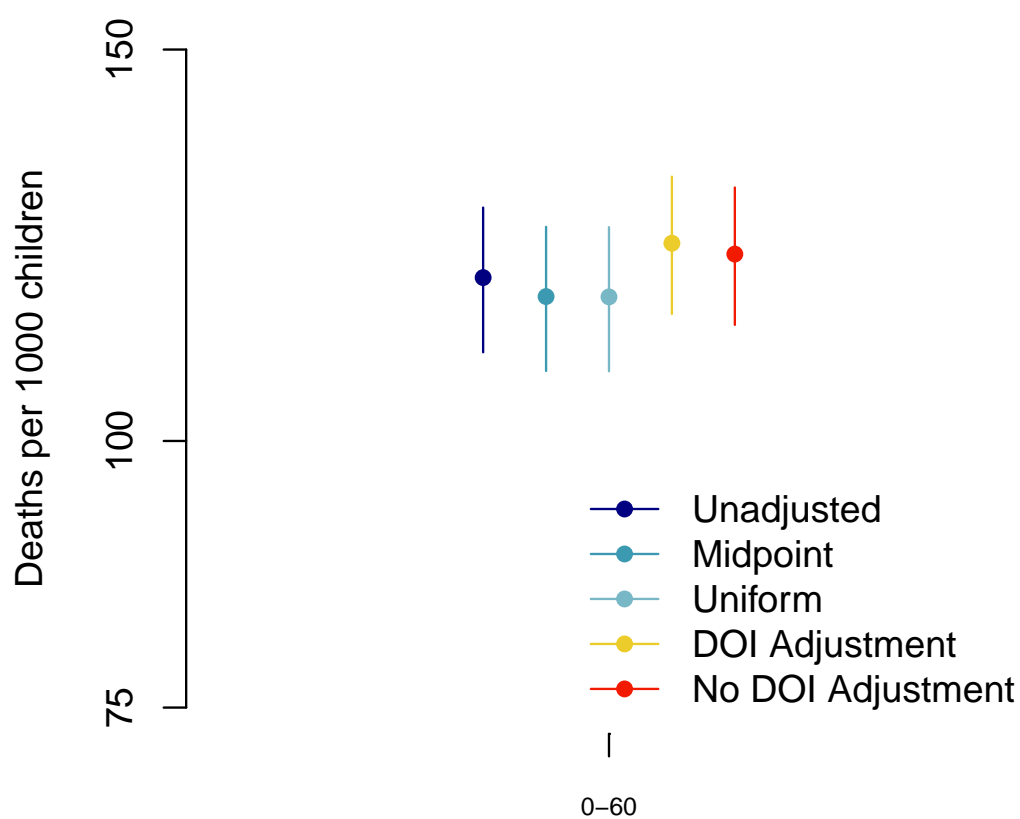


Figure 3.23: Comparison of U5MR estimates using age-band specific measures of mortality of deaths divided by exposure with 8 age bands. Comparison of deaths: 8 age bands. Death counts are compared across Cohorts A, B, C. Different blues represent life table methods that differ only in the way the treat the reported age at death for ages greater than or equal to 24. Red and yellow represent Lexis methods used by UN IGME and the DHS, the only difference indicating whether an adjustment for death counts in the period prior to the timing of the survey or date of interview (DOI) in months is made for Cohort C.

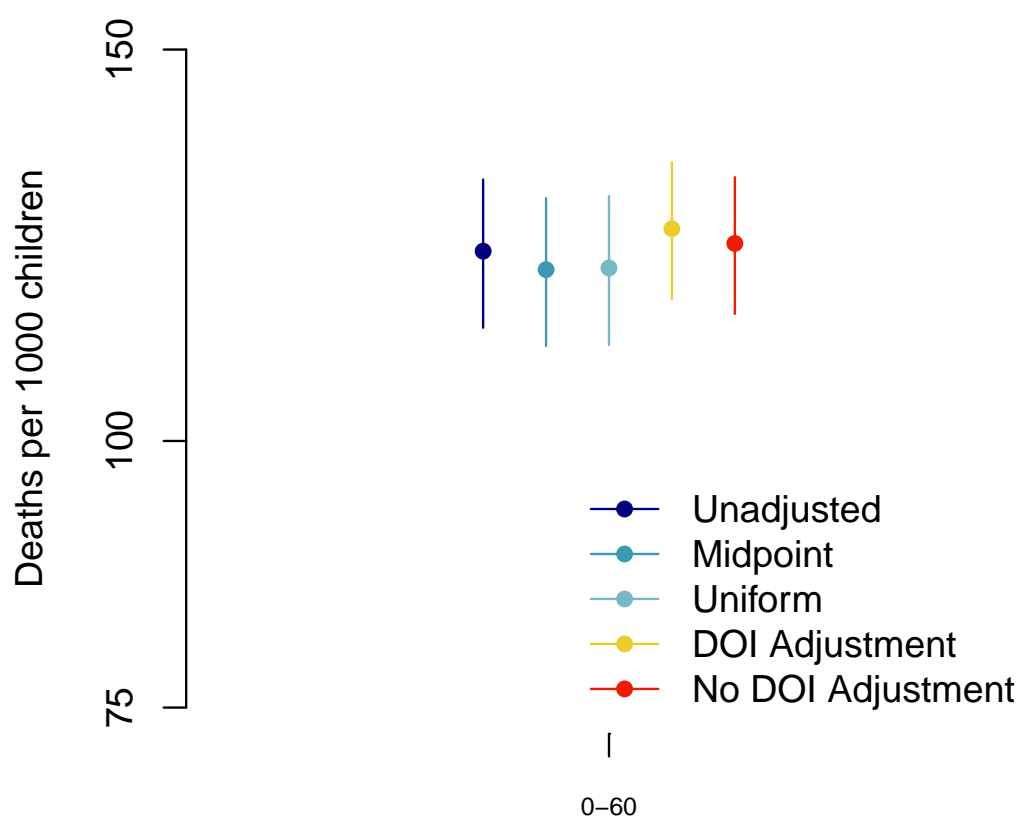


Figure 3.24: Comparison of U5MR estimates using age-band specific measures of mortality of deaths divided by exposure with 5 age bands. Comparison of deaths: 8 age bands. Death counts are compared across Cohorts A, B, C. Different blues represent life table methods that differ only in the way the treat the reported age at death for ages greater than or equal to 24. Red and yellow represent Lexis methods used by UN IGME and the DHS, the only difference indicating whether an adjustment for death counts in the period prior to the timing of the survey or date of interview (DOI) in months is made for Cohort C.

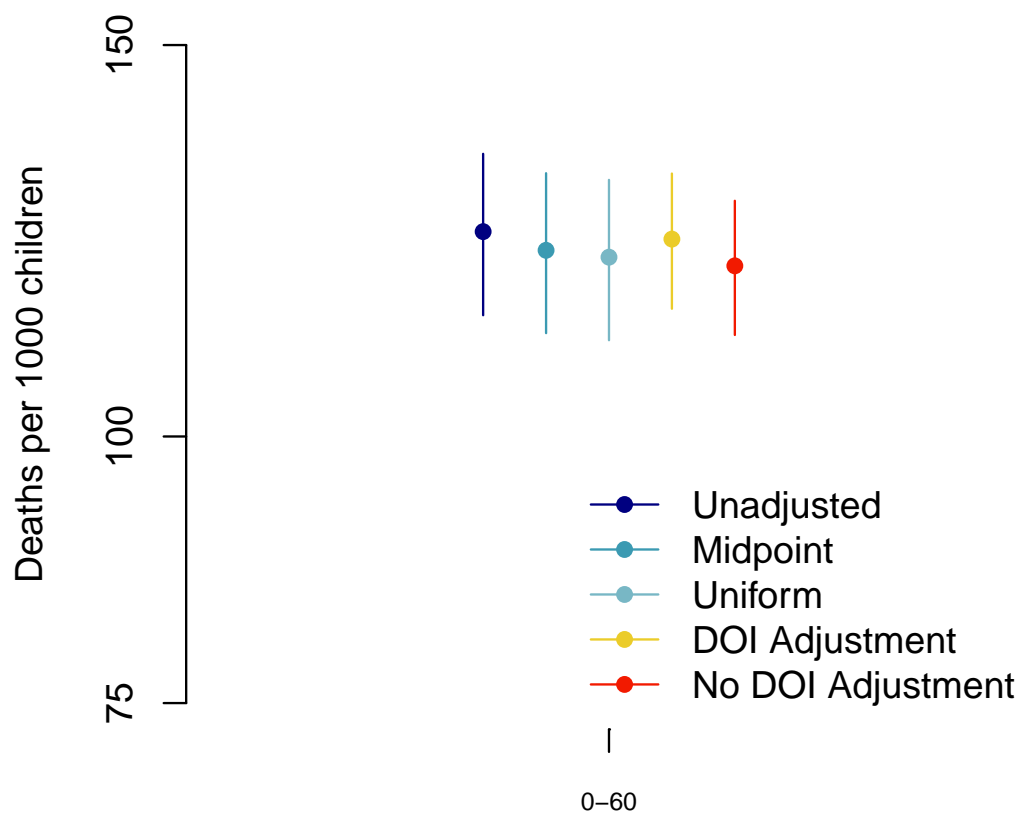


Figure 3.25: Comparison of U5MR estimates using age-band specific measures of mortality of deaths divided by exposure with 2 age bands.

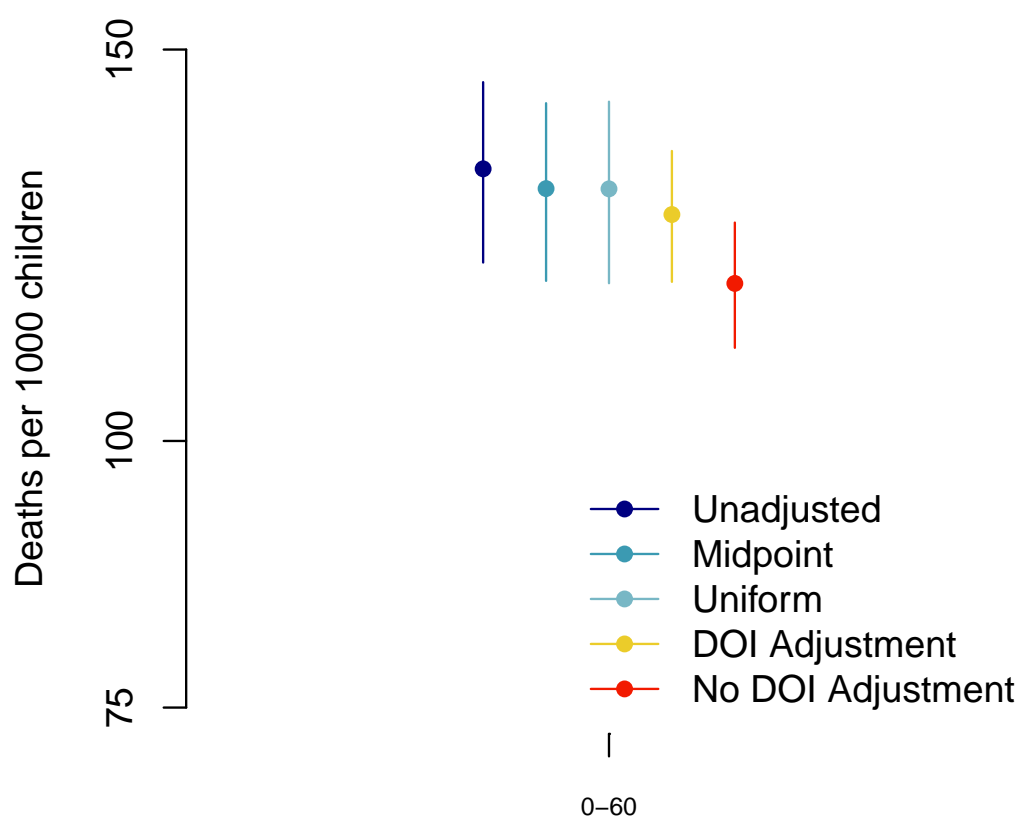


Figure 3.26: Comparison of U5MR estimates using age-band specific measures of mortality of deaths divided by exposure with 1 age band. Comparison of deaths: 8 age bands. Death counts are compared across Cohorts A, B, C. Different blues represent life table methods that differ only in the way the treat the reported age at death for ages greater than or equal to 24. Red and yellow represent Lexis methods used by UN IGME and the DHS, the only difference indicating whether an adjustment for death counts in the period prior to the timing of the survey or date of interview (DOI) in months is made for Cohort C.

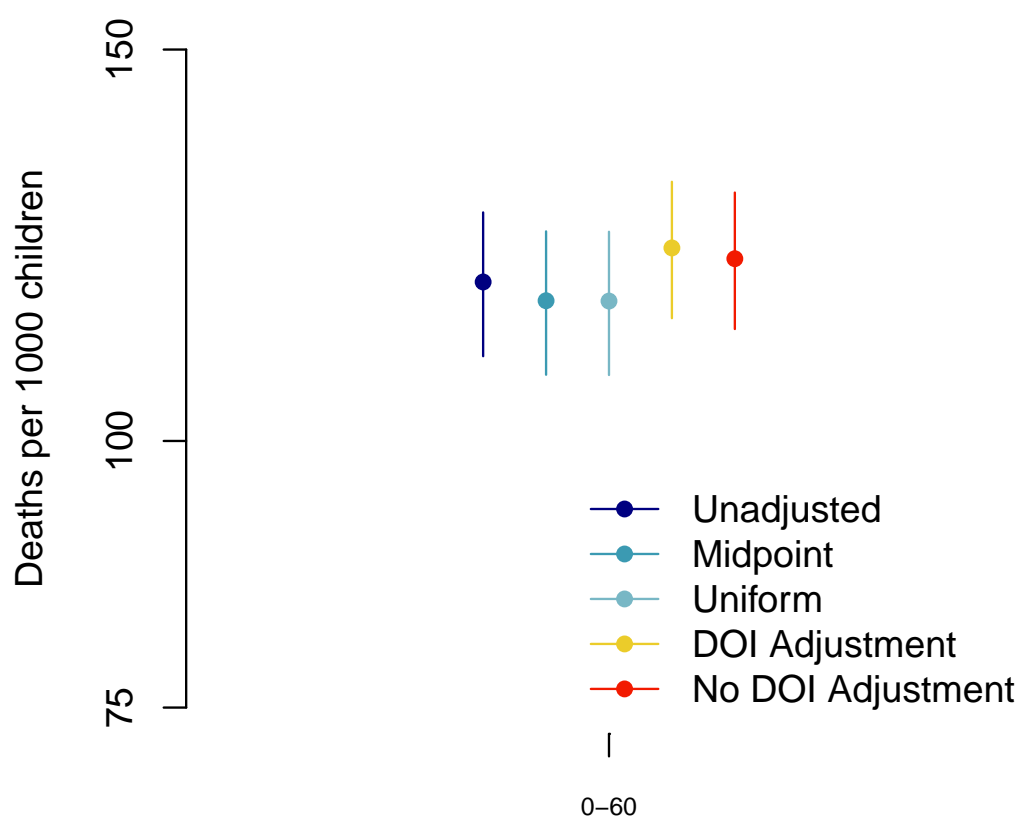


Figure 3.27: Comparison of U5MR estimates using age-band specific measures of mortality of deaths divided by personmonths with 8 age bands. Comparison of deaths: 8 age bands. Death counts are compared across Cohorts A, B, C. Different blues represent life table methods that differ only in the way the treat the reported age at death for ages greater than or equal to 24. Red and yellow represent Lexis methods used by UN IGME and the DHS, the only difference indicating whether an adjustment for death counts in the period prior to the timing of the survey or date of interview (DOI) in months is made for Cohort C.

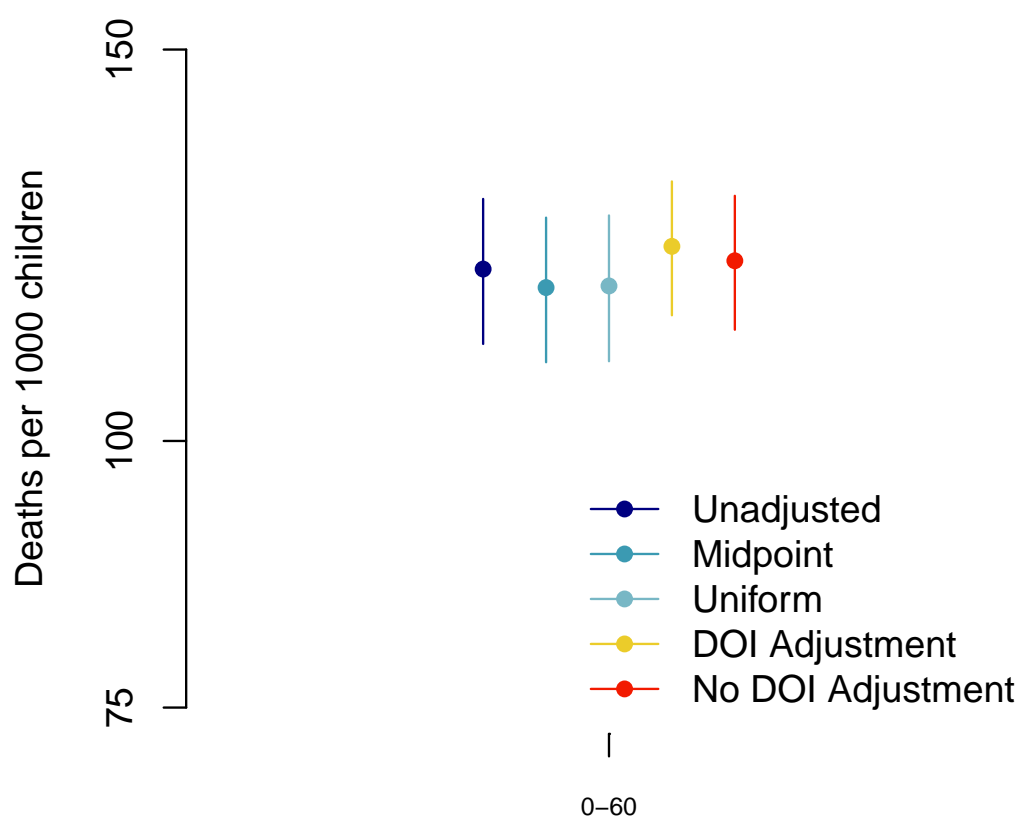


Figure 3.28: Comparison of U5MR estimates using age-band specific measures of mortality of deaths divided by personmonths with 5 age bands. Comparison of deaths: 8 age bands. Death counts are compared across Cohorts A, B, C. Different blues represent life table methods that differ only in the way the treat the reported age at death for ages greater than or equal to 24. Red and yellow represent Lexis methods used by UN IGME and the DHS, the only difference indicating whether an adjustment for death counts in the period prior to the timing of the survey or date of interview (DOI) in months is made for Cohort C.

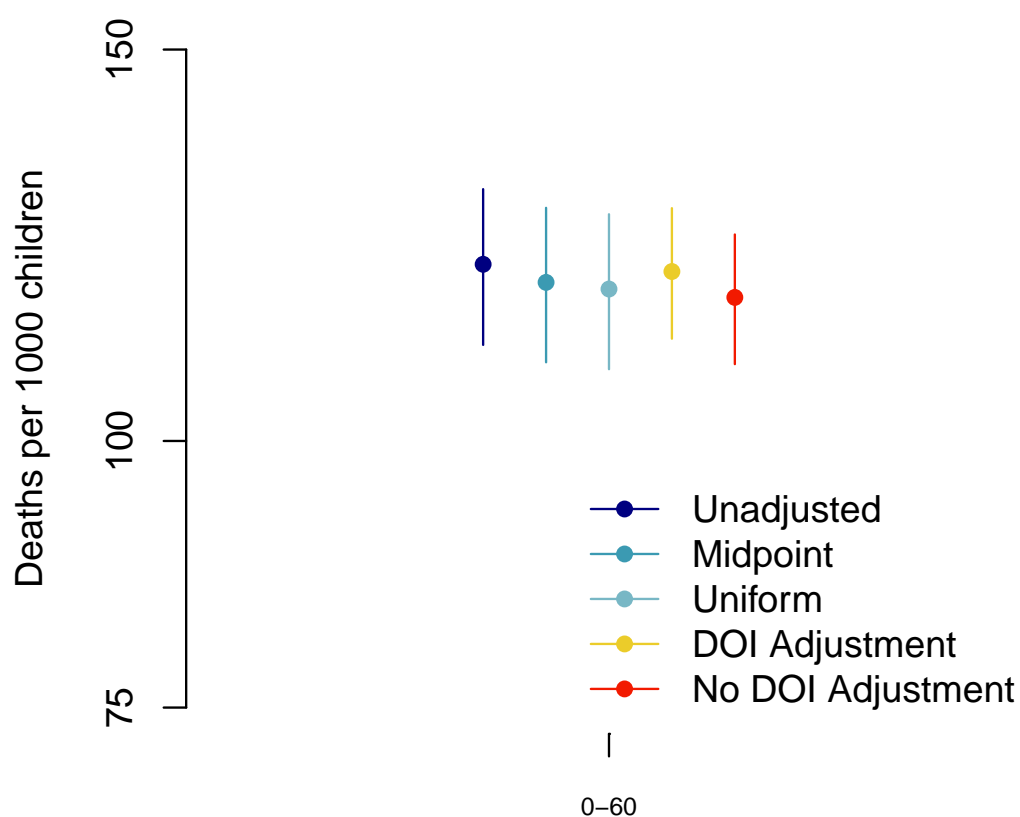


Figure 3.29: Comparison of U5MR estimates using age-band specific measures of mortality of deaths divided by personmonths with 2 age bands. Comparison of deaths: 8 age bands. Death counts are compared across Cohorts A, B, C. Different blues represent life table methods that differ only in the way the treat the reported age at death for ages greater than or equal to 24. Red and yellow represent Lexis methods used by UN IGME and the DHS, the only difference indicating whether an adjustment for death counts in the period prior to the timing of the survey or date of interview (DOI) in months is made for Cohort C.

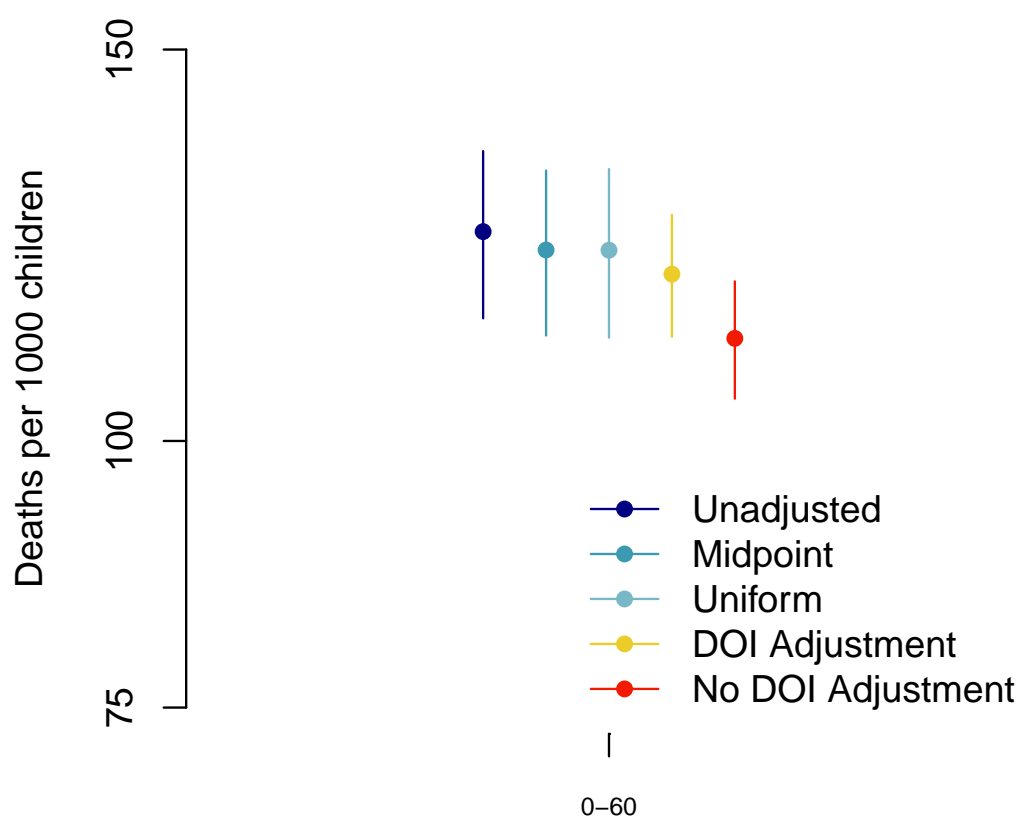


Figure 3.30: Comparison of U5MR estimates using age-band specific measures of mortality of deaths divided by personmonths with 1 age band. Comparison of deaths: 8 age bands. Death counts are compared across Cohorts A, B, C. Different blues represent life table methods that differ only in the way the treat the reported age at death for ages greater than or equal to 24. Red and yellow represent Lexis methods used by UN IGME and the DHS, the only difference indicating whether an adjustment for death counts in the period prior to the timing of the survey or date of interview (DOI) in months is made for Cohort C.

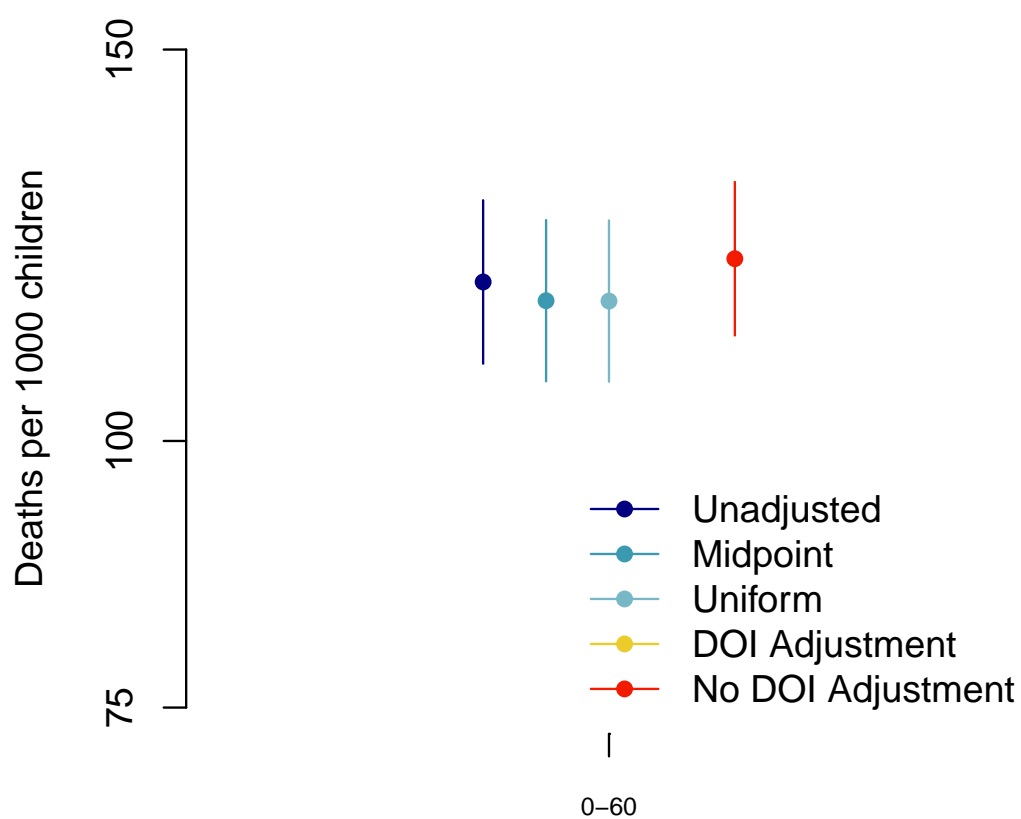


Figure 3.31: Comparison of U5MR estimates using discrete time survival analysis with 8 age bands. Comparison of deaths: 8 age bands. Death counts are compared across Cohorts A, B, C. Different blues represent life table methods that differ only in the way they treat the reported age at death for ages greater than or equal to 24. Red and yellow represent Lexis methods used by UN IGME and the DHS, the only difference indicating whether an adjustment for death counts in the period prior to the timing of the survey or date of interview (DOI) in months is made for Cohort C.

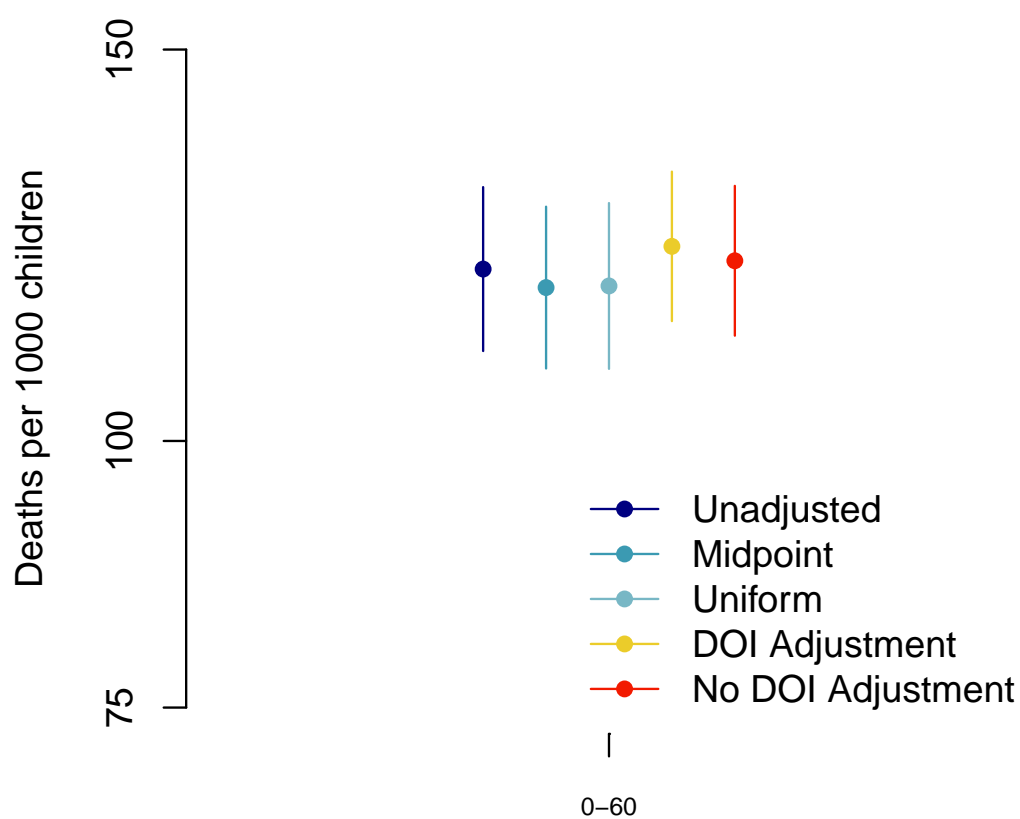


Figure 3.32: Comparison of U5MR estimates using discrete time survival analysis with 5 age bands. Comparison of deaths: 8 age bands. Death counts are compared across Cohorts A, B, C. Different blues represent life table methods that differ only in the way the treat the reported age at death for ages greater than or equal to 24. Red and yellow represent Lexis methods used by UN IGME and the DHS, the only difference indicating whether an adjustment for death counts in the period prior to the timing of the survey or date of interview (DOI) in months is made for Cohort C.

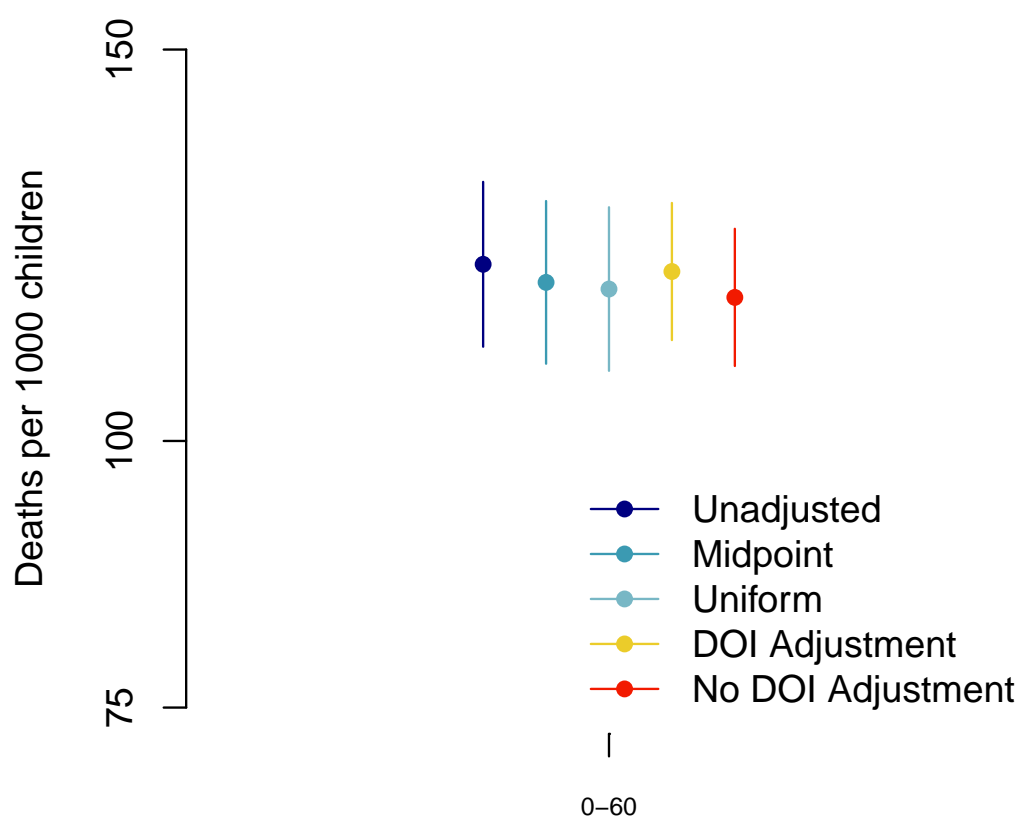


Figure 3.33: Comparison of U5MR estimates using discrete time survival analysis with 2 age bands. Comparison of deaths: 8 age bands. Death counts are compared across Cohorts A, B, C. Different blues represent life table methods that differ only in the way the treat the reported age at death for ages greater than or equal to 24. Red and yellow represent Lexis methods used by UN IGME and the DHS, the only difference indicating whether an adjustment for death counts in the period prior to the timing of the survey or date of interview (DOI) in months is made for Cohort C.

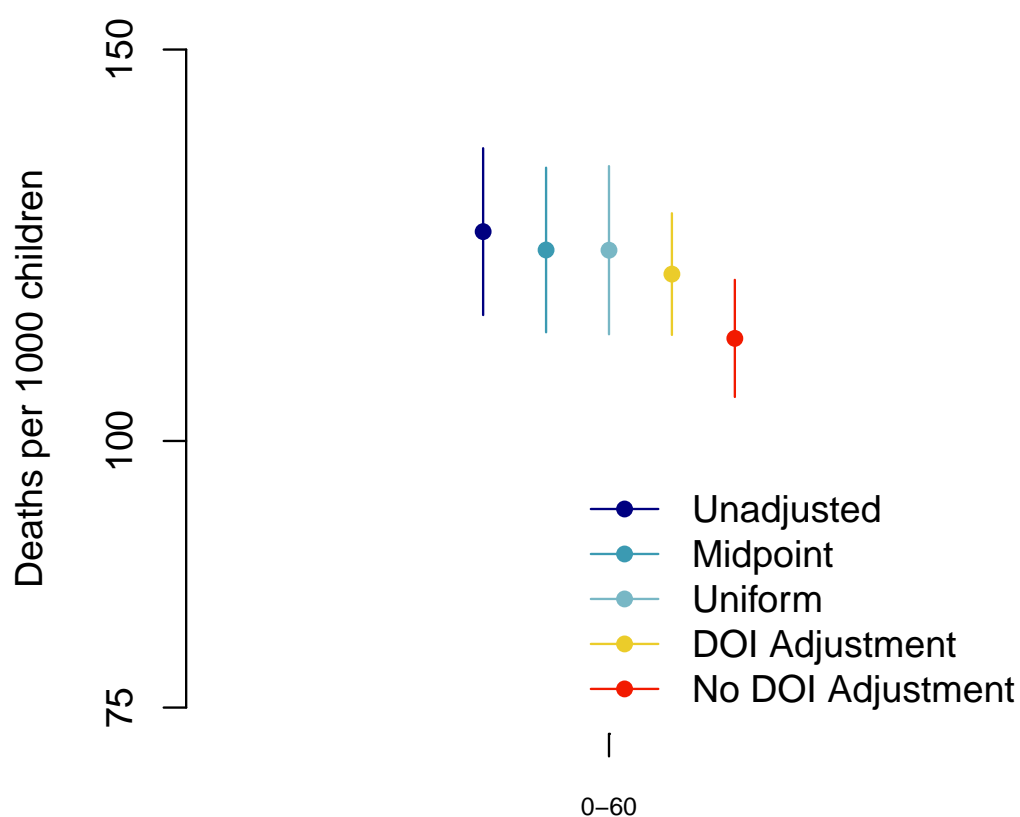


Figure 3.34: Comparison of U5MR estimates using discrete time survival analysis with 1 age bands. Comparison of deaths: 8 age bands. Death counts are compared across Cohorts A, B, C. Different blues represent life table methods that differ only in the way they treat the reported age at death for ages greater than or equal to 24. Red and yellow represent Lexis methods used by UN IGME and the DHS, the only difference indicating whether an adjustment for death counts in the period prior to the timing of the survey or date of interview (DOI) in months is made for Cohort C.

### 3.5 Discussion

In this chapter we have fully synthesized the statistical and demographic methods for age-specific period mortality. We connect life table and discrete time survival methods, and compare them to Lexis methods. We implement the Lexis methods used by UN IGME and the DHS, and compare them to the methods we use throughout the rest of this thesis. As expected, there are differences in counts of deaths, exposures and estimates of mortality across these methods. Though it may be of concern that the same child's birth history is treated differently by each method, we see that in the context of FBH data from household surveys, uncertainty renders many of these differences inconsequential. We find that Cohort B is treated equivalently in death counts regardless of method, but no method counts exposures equivalently. We also find that Cohort B does not exist when using a single age band of equivalent length to the period of interest. When the Lexis methods are used and the date of interview adjustment is made, we find more consistency with life table methods. The expected benefits of partitioning the age interval into bands, when estimating mortality are obvious. If we want to separate out the most unique month of life (the first month) this method of counting deaths and exposures cannot be used. Finally, we must note that when using estimating age band mortality as a rate, the aggregation up to U5MR is mathematically inappropriate. While the point estimates and their uncertainties do not indicate any huge differences from estimating U5MR via monthly hazards, there are no mathematical properties of rates that allow for the aggregation up to U5MR as in the formula in (3.22). Careful consideration of the target of interest for a particular analysis is important when choosing these methods. The Lexis methods used by UN IGME and DHS can still be considered estimates of prevalence of mortality for an individual age band in a finite population, as indicated by (3.25). They can be considered estimates of age-specific mortality rates in a finite population. However, when age bands of unequal length are taken together, the units of half or whole exposures mean different things in each age band. The discrete survival

analysis and life table approaches are more appropriate as estimators of the underlying hazard of death, therefore the aggregation across consecutive single age bands is mathematically motivated. Under certain assumptions about the relationship between the underlying hazard of mortality or mortality rates and in application to real data there may be little numerical differences between measures, despite the differing targets of interest. Finally, this chapter provides important context for comparison of estimates for child mortality for a single population. The Lexis diagram and life table approaches can be thought of as models. They each rely on differing assumptions that simplify reality to inform quantification and estimation of mortality in a single population. Even in a fully observed population where mortality is observed without censoring or measurement error, each method would produce a different number of deaths, different exposure counts, and different summaries of the age-specific mortality.

Future work will consider looking at the way these methods compare in subnational divisions of the country and also by urban/rural designation. We expect the decrease in sample size will lead to more unstable estimates, but that larger uncertainties will yield statistically indistinguishable methods. Most importantly, a simulation study will be conducted to understand the consequences of method choice, unit of time measurement, and age band choice when the truth is known. There are many issues that will complicate a simulation study, such as choosing whether to simulate birth histories for each child using continuous time or discrete time and how age-specific mortalities will be applied as simulated children age.

## Chapter 4

# COMBINING INDIRECT AND DIRECT SUBNATIONAL CHILD MORTALITY ESTIMATES AT THE ADMIN-2 LEVEL

### 4.1 Introduction

Currently, the United Nations Inter-agency group for Child Mortality Estimation (UN IGME) use the Bayesian B-spline bias reduction model for estimation of child mortality (Alkema and New, 2014). However, this method is used for national estimation and is not designed to deal with within-country variability. The model cannot incorporate surveys that are not carried out at the at the national or Admin-0 level. Mercer et al. (2015) developed a discrete space-time smoothing method to produce areal estimates that account for the complex survey design from FBH data. Li et al. (2019) apply a similar method to 35 countries at the spatial resolution of the Admin-1 area in work supported by UN IGME. Li et al. (2019) use only DHS FBH data, and do not incorporate other surveys or any SBH data, whereas Mercer et al. (2015) use additional information from demographic surveillance sites. We extend these methods to incorporate MICS surveys and census data. Moreover, we apply this method to make estimates at a finer administrative level, the 47 counties of Kenya and 28 districts of Malawi. Though Kenya redefined their Admin-1 level to be 47 counties, it was originally defined to be 8 regions. Figure 4.1 shows the Admin-1 regions by color with the counties or districts of each country outlined with white borders. Most household surveys in Kenya were stratified at the 8 region level, not the 47 counties. In Malawi, most surveys were stratified at the level of 3 regions. From this point on we will refer to the 47 counties and the 28 districts as Admin-2 level.

We analyze the combination of DHS, MICS and IPUMS census samples from Kenya

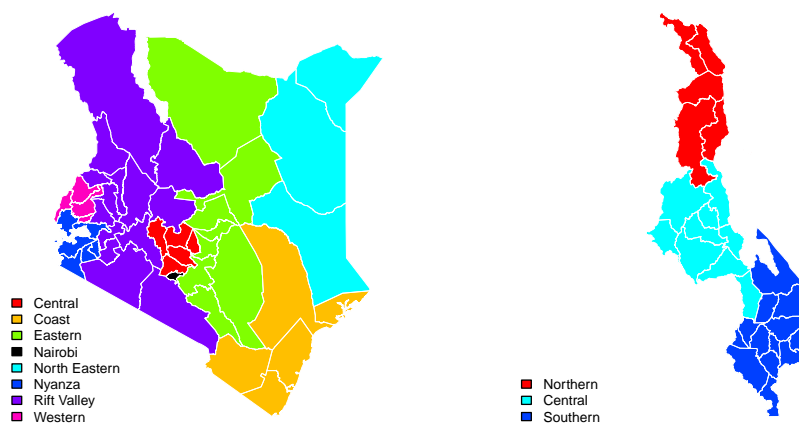


Figure 4.1: Colors show the 8 Admin-1 areas of Kenya and the white borders outline the 47 Admin-2 areas (Left) and the 3 Admin-1 areas of Malawi with the white borders outlining the 28 Admin-2 areas (Right).

and Malawi to make estimates of U5MR for Admin-2 areas over the period 1980-2014. As Table 4.1 shows, in time periods where IPUMS census samples are used, they comprise an overwhelming proportion of births contributing to that data although they do not provide birth date and death date information. Making estimates at the Admin-2 level leads to complications of unstable estimates for many small areas due to small sample sizes. We propose an adjustment to estimates for these areas prior to smoothing to make them more stable while still accounting for the complex survey design. Our model is fit on the yearly scale in a manner that allows us to produce yearly estimates and also estimates for periods of any year length. This allows us to incorporate both SBH estimates, which are made for a single year, and FBH estimates, which are made for five-year periods as the data become too sparse at finer levels of aggregation, into a space-time smoothing model.

The rest of this chapter is organized as follows. First, Section 4.2 begins by introduc-

Country	Source	Type	Year	Coverage	% births						
					80–84	85–89	90–94	95–99	00–04	05–09	10–14
Kenya	DHS	FBH	2003	National*	0.5	36.5	0.7	21.1	0.4	0.0	0.0
			2008	National	0.2	21.1	0.5	17.1	0.4	16.3	0.0
			2014	National	0.0	26.9	1.0	42.2	1.2	62.0	78.3
	MICS	SBH	2000	National*	0.0	0.0	2.1	3.8	0.0	0.0	0.0
			2011	1 Admin-1	0.0	12.2	0.3	11.0	0.3	14.4	12.9
		2013	3 Admin-2	0.0	3.3	0.1	4.8	0.1	7.3	8.9	
	IPUMS	SBH	1989	National	99.3	0.0	0.0	0.0	0.0	0.0	0.0
			1999	National	0.0	0.0	95.3	0.0	0.0	0.0	0.0
			2009	National	0.0	0.0	0.0	0.0	97.7	0.0	0.0
Malawi	DHS	FBH	2000	National	1.1	25.8	1.9	17.1	1.1	0.0	0.0
			2004	National	0.6	15.8	1.3	13.0	1.7	0.0	0.0
			2010	National	0.4	17.1	1.6	18.0	2.8	27.5	23.1
			2015	National	0.0	5.5	0.8	11.7	2.1	23.1	38.3
	MICS	FBH	2006	National	0.9	26.9	2.4	25.9	3.8	23.5	0.0
			2014	National	0.0	8.8	1.1	14.3	2.4	25.9	38.6
	IPUMS	SBH	1987	National	96.8	0.0	0.0	0.0	0.0	0.0	0.0
			1998	National	0.0	0.0	91.0	0.0	0.0	0.0	0.0
			2008	National	0.0	0.0	0.0	0.0	86.1	0.0	0.0

Table 4.1: List of all data sources and types for Kenya and Malawi. Columns 6-12 show the percentage of births each data source contributes to each five-year period. \*denotes this survey was carried out nationally, and does not have observations in all Admin-2 areas.

ing the method of spatiotemporal child mortality estimation extended in this chapter (Section 4.2.1) and addressing an adjustment needed for surveys carried out in countries with high prevalence of HIV/AIDs in the general population (Section 4.2.5). The rest of Section 4.2 covers methodological developments and extensions in this work necessary to make

estimates at the Admin-2 level – including adjusting for areas and time periods with no deaths in Section 4.2.2 and to incorporate SBH data including modeling at different temporal scales in Section 4.2.4 and bias adjustment in Section A.2. Section 4.3.1 presents results from application to data from Kenya and Malawi and, also, results from two validation exercises. Finally, Section 4.4 contains a discussion of methods developed, limitations, and future work.

## 4.2 Methods

### 4.2.1 Review of Smoothed Direct Estimation

Recalling notation from Chapter 2, particularly in Sections 2.1, 2.3.2 and 2.4, we will let  ${}_{60}q_{0,its}$  denote the probability of death before age 5 in years or 60 in months in area  $i$  and time period  $t$  collected by survey  $s$ . We will use the function  $a[m]$  defined in (2.3) and repeated below to partition the age interval from  $[0, 60)$  months into 6 distinct age groups:

$$a[m] = \begin{cases} 1, & 0 \leq m < 1 \\ 2, & 1 \leq m < 12 \\ 3, & 12 \leq m < 24 \\ 4, & 24 \leq m < 36 \\ 5, & 36 \leq m < 48 \\ 6, & 48 \leq m < 60. \end{cases}$$

The method extended in this chapter, begins estimating a discrete hazards model for each area and time period of interest each survey contains observations (Mercer et al., 2015). For each child  $k$ , their birth record is expanded into a single binary observation for each month of life the child began that indicates whether or not the child dies in that month. Here a child who dies in the first month of life would contribute a single observation of value 1, and

a child who survives until their 5<sup>th</sup> birthday contributes 60 observations all of which equal to 0. Then, assuming the monthly hazard is the same for each month in  $a[m]$ , define the model as

$$y_{m,itsk} | 1q_{m,its} \sim \text{Bernoulli}(1q_{m,its})$$

$$\text{logit}(1q_{m,its}) = \log\left(\frac{1q_{m,its}}{1 - 1q_{m,its}}\right) = \beta_{a[m],its}. \quad (4.1)$$

A design-weighted pseudolikelihood method (Binder, 1983) yields design-consistent estimates of the monthly hazard for each age band,  $1\hat{q}_{m,its} = \text{expit}(\hat{\beta}_{a[m],its})$ , along with the variance-covariance matrix of the age-specific intercepts,  $\hat{\Sigma}_{its}$ . The model is fitted in R using the `survey` package (Lumley, 2004).

From these estimates one can derive an estimate of  $60q_{0,its}$ ,

$$60\hat{q}_{0,its} = 1 - \prod_{m=0}^{59} (1 - 1\hat{q}_{m,its}). \quad (4.2)$$

A transformation via the delta method of  $\hat{\Sigma}_{its}$  results in  $\hat{V}_{its}$ , the design-based variance of  $\text{logit}(60\hat{q}_{0,its})$  (Mercer et al., 2015). Let  $y_{its} = \text{logit}(60\hat{q}_{0,its})$ . To get smoothed estimates, define the likelihood as the asymptotic distribution of the estimator,

$$y_{its} | \eta_{its} \sim N(\eta_{its}, \hat{V}_{its}). \quad (4.3)$$

As in Section 2.5.2 then decompose the mean of the log odds of the probability as

$$\eta_{its} = \mu + \alpha_t + \gamma_t + \theta_i + \phi_i + \delta_{it} + \nu_s, \quad (4.4)$$

where  $\mu$  is an overall mean, the terms  $\nu_s$  are IID random effects with a zero-mean normal priors with precision  $\kappa_\nu$  to estimate the bias of estimates from data source  $s$ , relative to  $\mu$ . Together,  $\gamma_t$  and  $\alpha_t$  comprise the temporal components of the model. The  $\alpha_t$  are IID terms that follow a zero-mean normal prior with precision  $\kappa_\alpha$  which allow for random shocks

in time. The temporal trend is modeled using terms  $\gamma_t$  which have a RW2 prior placed on them. Following the BYM model,  $\phi_i$  is an IID spatial effect with a zero-mean normal prior with precision,  $\tau_\phi$ , and  $\theta_i$  is an ICAR spatial random effect. Finally, the space-time interaction,  $\delta_{it}$ , is a Type IV interaction as defined by Knorr-Held (2000) with a precision matrix that is a cross between a RW2 and an ICAR. The model is fit via INLA using the R-INLA package (Rue et al., 2009; Martins et al., 2013), which is very fast. After fitting the model, one obtains smoothed direct estimates and credible intervals via estimation of the posterior distribution for

$${}_{60}q_{0,it} = \text{expit}(\mu + \alpha_t + \gamma_t + \theta_i + \phi_i + \delta_{it}). \quad (4.5)$$

The complete distributions are available, but we can summarize using, for example, the 2.5%, 50% and 97.5% quantiles. Note that terms  $\nu_s$  are not included in the predictions in (4.5).

#### 4.2.2 Zeroes Adjustment

Most household surveys are geographically stratified at an administration level higher than Admin-2, thus sample sizes can be small or nonexistent in some Admin-2 areas. If the data are sparse, we may not observe any deaths in such an area, leading to a direct estimate of  ${}_{60}\widehat{q}_{0,its} = 0$ . In these cases, and in cases where all the children in a cluster have the same response (all deaths or all non-deaths), the variance estimate is 0. These cases are problematic for our smoothing model as  $\text{logit}({}_{60}\widehat{q}_{0,its})$  is undefined and we require a (non-zero) design based variance. In the case of no data in an area and time period, we treat these as missing values, which can be dealt with accordingly within a Bayesian framework. In the aforementioned other cases, we cannot go this route, as they are informative, albeit weakly.

Some authors (Ha et al., 2014; Sugawara et al., 2018) have chosen to take as likelihood the sampling distribution of the estimator on the original scale (the U5MR in our setting)

while smoothing on a different scale, in an approach labeled *area-level unmatched sampling and linking models* (the linking model is the second stage of the hierarchy, (4.4) in our case). This approach removes the numerical problems with the estimated probability being 0 or 1, though the asymptotic normal approximation to the sampling distribution is likely to be inaccurate in these cases. In Ha et al. (2014, Section 3.2) the design variance is written as the product of the variance of the estimator for a binomial response and the design effect. The former depends on the unknown proportion, and this is estimated using a synthetic estimator that uses data from a larger region, the design effect is also estimated from a larger region. Generalized variance function estimation (Wolter, Ch. 7) provides a general approach to alleviating problems of variance instability.

We propose a model-based solution for areas where we have computational issues; we preprocess such data to replace the problematic estimates with a reasonable value. We stress that this should only be carried out for a small proportion of areas and time periods, perhaps not much more than 5%. In a problematic area labeled  $i$ , we take the set of its first order neighbors  $j \sim i$ , to constitute a larger area. For the problematic areas and its neighbors we fit a beta binomial model that accounts for the complex survey design by including an urban/rural strata fixed effect and a scale parameter that allows for clustering. We let  $i$  index areas,  $h$  urban/rural strata, and  $c$  clusters within strata. For a time period  $t$  and a survey  $s$ ,

$$\begin{aligned} Z_{its}^{(hc)} | {}_{60}q_{0,its}^{(hc)} &\sim \text{BetaBinomial}(n_{its}^{(hc)}, {}_{60}q_{0,its}^{(hc)}, d), \\ \text{logit} \left( {}_{60}q_{0,its}^{(hc)} \right) &= \mu + \beta 1(h = 2) + e_i, \\ e_i &\sim N(0, \sigma_e^2) \end{aligned}$$

where  $Z_{its}^{(hc)}$  is the number of deaths,  $n_{its}^{(hc)}$  is the number of children in cluster  $c$  for urban and rural strata,  $h = 1, 2$ , and  $d$  is an overdispersion parameter. This model can be fitted in R with INLA; we use relatively uninformative priors on  $\beta$ ,  $\sigma_e^2$  and  $d$ .

An estimate of the aggregate U5MR in area  $i$ , time period  $t$  and for source  $s$  is,

$${}_{60}q_{0,its} = p_{its} \times \text{expit}(\mu + e_i) + (1 - p_{its}) \times \text{expit}(\mu + \beta + e_i), \quad (4.6)$$

where  $p_{its}$  and  $1 - p_{its}$  are the proportions of births that are urban and rural, respectively. In the DHS sampling manuals, the numbers of clusters in urban and rural strata in the sampling frame is reported by area – usually this is by Admin1, but it is Admin2 for some surveys, including the ones we use in Kenya and Malawi.

Within the INLA implementation in R there is a function `INLA.posterior.marginal`, that allows one to draw samples  $[\beta_h^{(j)}, e_i^{(j)}]$  for  $j = 1, \dots, J$ , from the approximate joint posterior distribution. These can be converted to samples  ${}_{60}q_{0,its}^{(j)}$  using (4.6) and we use the posterior mean and posterior variance of these samples for  $y_{its}$  and  $\widehat{V}_{its}$ , respectively that we require for the smoothing model. Let  $y_{its}^{(j)} = \text{logit}({}_{60}q_{0,its}^{(j)})$ , then

$$y_{its} = \frac{1}{J} \sum_{j=1}^J y_{its}^{(j)},$$

and

$$\widehat{V}_{its} = \frac{1}{J-1} \sum_{j=1}^J (y_{its}^{(j)} - y_{its})^2.$$

We show results of the adjustment method for the 2003 Kenya DHS, as it has more areas and time periods that need to be adjusted. Malawi, on the other hand, has only a handful of areas/time periods that need adjustments. Figure 4.2 shows a side by side comparison of estimates of logit U5MR in all areas in Kenya before (Left) and after (Right) adjustment. Note, the estimates on the right-hand side and the corresponding variances used to create the confidence intervals are those later included in the smoothing model. Areas to be adjusted are in red in both panels. We see three types of estimates to be adjusted on the left. Lamu, Embu, Laikipia, Trans-Nzoia, Kajiado and Samburu all have an estimate of  ${}_{60}\widehat{q}_{0,i,80-84,2003} = 0$  and  $\widehat{V}_{i,80-84,2003} = 0$ . Isiolo has no estimate, as no clusters in the 2003 DHS were sampled in that area. Lastly, Marsabit and West Pokot have  ${}_{60}\widehat{q}_{0,i,80-84,2003} \neq 0$

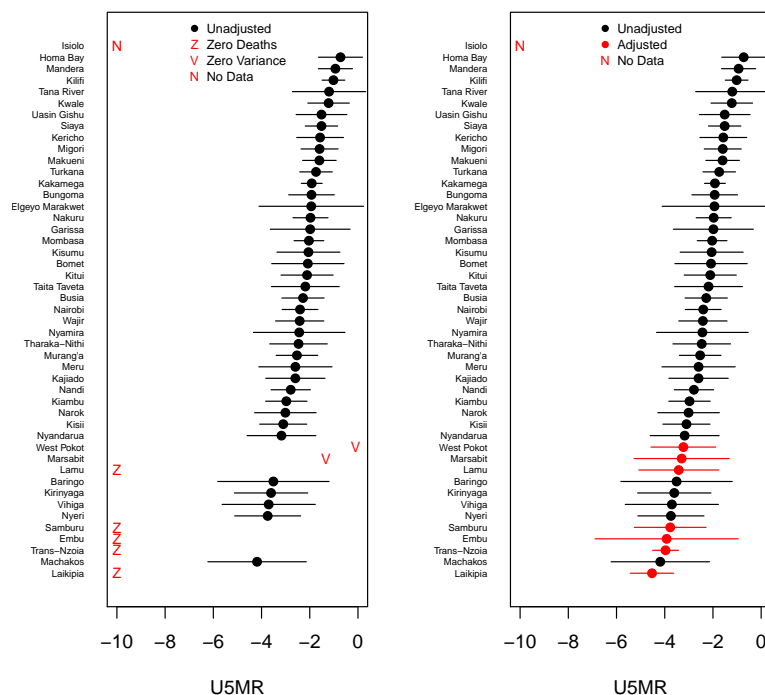


Figure 4.2: Left: The direct estimates on the logit scale for counties in Kenya from the 2003 DHS for 1980–1984. A red Z indicates an area with no observed deaths, whose logit U5MR is undefined. A red V indicates an area where deaths are observed, but estimated variances are zero. A red N indicates an area that contains no sampled clusters. Right: The adjustment comparisons on the logit scale.

and  $\widehat{V}_{i,80-84,2003} = 0$ . In Marsabit there are 4 sampled households, in West Pokot there are 5. In every household either all the children die or all the children survive, which leads to a variance estimate of 0. Figure 4.3 shows the location and types of adjustments on a map. Our adjusted estimates  ${}_{60}\widehat{q}_{0,its}$  are pulled away from 0 and informed by the data in surrounding areas. The average 95% CI width for adjusted areas is 2.01, which is nearly

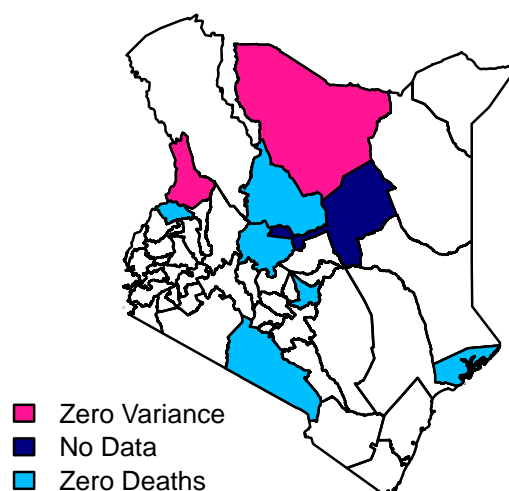


Figure 4.3: A map showing the areas whose estimates need adjusting from the KDHS 2003 for the period 1980–1984. Colors distinguish the types of problematic areas. Light blue areas have zero observed deaths. Pink areas have observed deaths, but zero variance. The navy area has no sampled clusters in the 2003 survey. We do not make adjustments for this area as it can be included as an NA in INLA.

identical to the average width of unadjusted areas, 2.03.

Figure 4.4 shows smoothed results on a map for the two ways of treating the observed

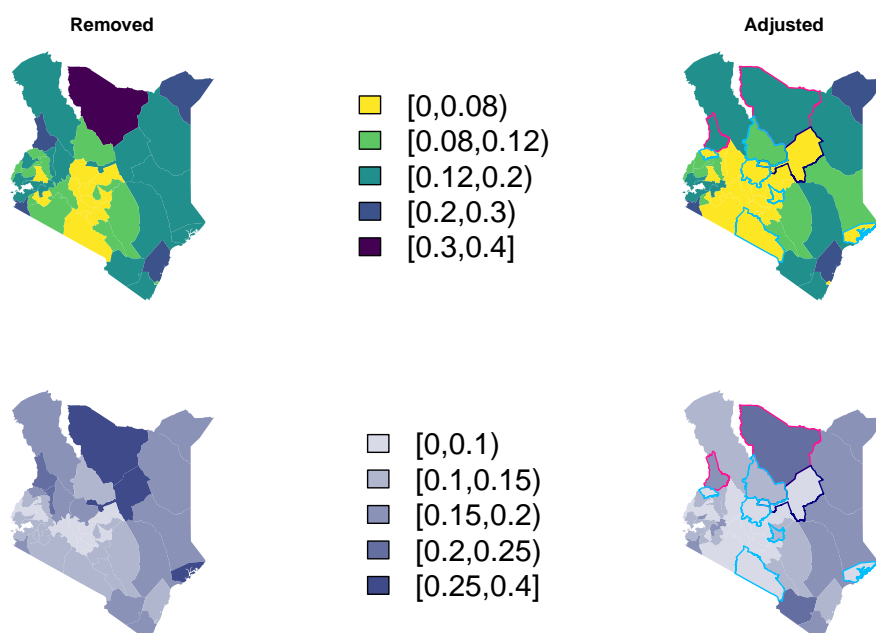


Figure 4.4: **Bottom:** Smoothed estimates for the period 1980–1984 when we throw out observed zeroes (Left) or adjust them (Right). **Bottom:** Width of 95% CI interval estimates for the period 1980–1984 when we throw out observed zeroes (Left) or adjust them (Right). Colored boundaries correspond to colors in Figure 4.3: pink are areas with observed deaths, but a zero variance estimate, light blue are areas with no observed deaths, and the navy outline is an area with no data.

zeroes for the period 1980–1984. Recall the pink areas, Marsabit and West Pokot, in Figure 4.3. These places had small sample sizes, with non-zero estimates of  ${}_{60}\widehat{q}_{0,i,80-84,2003}$  and zero estimates of  $\widehat{V}_{i,80-84,2003}$ . In the left column of Figure 4.4, we see that when we throw out these observations, we get incredibly wide CIs. However, when we use the adjusted method we still estimate relatively high mortality, which we would expect since we observed deaths in this area, but our uncertainty decreases. In areas where we observe no deaths,

we estimate smaller U5MRs when we use adjusted estimates instead of simply removing the data from the area. We also see shorter CI widths. Lastly, including the adjusted data allows us to make more precise estimates in Isiolo (outlined in navy) than when we throw out these Admin-2 area's data. In Isiolo, we have no observations whatsoever from the 2003 DHS. However, the adjusted data from Isiolo's neighbors that either had no observed deaths or an estimated variance of zero changes both the median estimate and the precision of the estimated U5MR in Isiolo after smoothing.

#### 4.2.3 Jackknife variance estimates for Brass estimates

Our smoothing model requires a point estimate and a variance estimate, but traditional Brass methods and variants do not provide uncertainty estimates. We use a jackknife to estimate  $\widehat{V}_{its}$  for SBH data (Pedersen and Liu, 2012). For the MICS surveys that provide SBH data, we must acknowledge the complex design, and we use a stratified cluster jackknife that removes clusters within strata. Let  $y^{(hc)} = \text{logit} \left( {}_{60}\widehat{q}_{0,its}^{(hc)} \right)$  be the estimate computed via the Brass method when the data from cluster  $c$  in strata  $h$  are removed for  $h = 1, \dots, H$  strata, with  $c = 1, \dots, n_{its}^{(h)}$  clusters in area  $i$ , at time  $t$  in strata  $h$  for study  $s$ . The estimate of the logit of U5MR for strata  $h$ , for  $i, t, s$ , is,

$$y_{its}^{(h)} = \frac{1}{n_{its}^{(h)}} \sum_{c=1}^{n_{its}^{(h)}} y_{its}^{(hc)},$$

where  $y_{its}^{(hc)}$  is the Brass estimate obtained when cluster  $c$  in strata  $h$  is removed. The stratified-cluster jackknife variance is,

$$\widehat{V}_{its} = \sum_{h=1}^H \frac{n_{its}^{(h)} - 1}{n_{its}^{(h)}} \sum_{c=1}^{n_{its}^{(h)}} (y_{its}^{(hc)} - y_{its}^{(h)})^2.$$

As each cluster-wise deletion must result in an estimate  $y_{its}^{(hc)}$ , we cannot calculate a measure of uncertainty for area, time, study combinations in which there are less than two clusters

in each strata. Moreover, if deleting a cluster  $c$  within a strata  $h$  results in the observed proportion of deaths being zero, then  $y_{its}^{(hc)}$  is undefined and we cannot compute the jackknife estimate of the variance for that area, time, study combination.

For the census samples we use a jackknife that removes women, as there are no design complications. There are, however, a large number of women in any given census and so we delete groups of women (Pedersen and Liu, 2012). First, we select the size of groups to delete as  $d = 30$ . Then, for each census we calculate the number of groups as  $n_d = N/d$  (in practice we round), where  $N$  is the number of women in the IPUMS census sample. Let  $y_{its}^{(k)} = \text{logit}\left(60\widehat{q}_{0its}^{(k)}\right)$  denote the estimate computed via the Brass method when data from women in group  $k$  are removed. Then,

$$\widehat{V}_{its} = \frac{n_d - 1}{n_d} \sum_{k=1}^{n_d} (y_{its}^{(k)} - y_{its})^2,$$

where  $y_{its} = \frac{1}{n_d} \sum_{k=1}^{n_d} y_{its}^{(k)}$ .

#### 4.2.4 Spatiotemporal models on two time scales

Both the FBH and SBH data produce a data pair  $\left[ \text{logit}(60\widehat{q}_{0,its}), \widehat{V}_{its} \right]$ , i.e. a point estimate and a variance, upon which a normal likelihood can be based. For the FBH data  $t$  refers to a five-year period, with  $t = 1980\text{--}1984, \dots, 2010\text{--}2014$ . For the SBH data,  $t$  refers to a single year (the reference year) between 1980 and 2014. To incorporate both 1-year and 5-year estimates in the direct smoothing model Mercer et al. (2015) we specify all time-related random effects on the yearly scale, which means that we have to aggregate over 5-year periods for the FBH data. Recall the model defined by (4.3) and (4.4) includes a RW2, IID effects for each period and a Type-IV interaction term for all area and time period combinations. We reformulate the model for each of these effects.

We still include a RW2 term,  $\gamma_t$ , but the random walk is defined on the yearly scale and effects for periods are averages of the yearly effects within that period. Let  $t = 1, \dots, n_T$

index 1-year time points,  $t^* = 1, \dots, n_{T^*}$  index 5-year periods. We then collect together the smoothing random effects on the 1-yearly and 5-yearly scales as  $\gamma = [ \gamma_T^T \quad \gamma_{T^*}^T ]^T$  where  $\gamma_T^T$  is a vector of length  $n_T$  and  $\gamma_{T^*}$  is a vector of length  $n_{T^*}$ . We define

$$\gamma_{T^*} = \mathbf{A}\gamma_T, \quad (4.7)$$

where  $\mathbf{A}$  is an  $n_{T^*} \times n_T$  matrix that averages each consecutive set of  $n_T/n_{T^*}$  points. The coarsened parameters  $\gamma_{T^*}$  are a deterministic function of the yearly versions  $\gamma_T$  but both are needed to define the model, for the FBH and SBH data, respectively. In our case,  $n_T = 35$  and  $n_{T^*} = 7$ , so that  $\gamma_{T^*}$  is a  $7 \times 1$  vector of five-year averages. The  $\text{RW2}(\kappa_\gamma)$  prior on  $\gamma_T$  can be written

$$\pi(\gamma_T | \kappa_\gamma) \propto \kappa_\gamma^{(n_T-2)/2} \exp \left[ -\frac{\kappa_\gamma}{2} \gamma_T^T \mathbf{Q}_\gamma \gamma_T \right],$$

where  $\kappa_\gamma \mathbf{Q}_\gamma$  is the precision matrix of the random walk. We practically implement the relationship (4.7) in INLA via a soft constraint in which we assume  $\gamma_{T^*} | \gamma_T \sim \text{N}(\mathbf{A}\gamma_T, \tau_\gamma^{-1} \mathbf{I}_{n_{T^*}})$ , where  $\mathbf{I}_{n_{T^*}}$  is the identity matrix and  $\tau_\gamma = 10^6$ . Written out in full,

$$\pi(\gamma_{T^*} | \gamma_T) \propto \tau_\gamma^{n_{T^*}/2} \exp \left[ -\frac{\tau_\gamma}{2} (\gamma_{T^*} - \mathbf{A}\gamma_T)^T \mathbf{I}_{n_{T^*}} (\gamma_{T^*} - \mathbf{A}\gamma_T) \right]. \quad (4.8)$$

Then, the joint prior on  $\gamma$  can be expressed as,

$$\pi(\gamma) = \pi(\gamma_T) \pi(\gamma_{T^*} | \gamma_T) \propto \exp \left[ -\frac{1}{2} \gamma^T \mathbf{Q}_\gamma^\dagger \gamma \right],$$

where

$$\mathbf{Q}_\gamma^\dagger = \begin{bmatrix} \kappa_\gamma \mathbf{Q}_\gamma + \tau_\gamma \mathbf{A}^T \mathbf{A} & -\tau_\gamma \mathbf{A}^T \\ -\tau_\gamma \mathbf{A} & \tau_\gamma \mathbf{I}_{n_{T^*}} \end{bmatrix}.$$

We can define the IID model similarly, where  $\alpha = [ \alpha_{n_T}^T \quad \alpha_{n_{T^*}}^T ]^T$  is the stacked vector of yearly IID effects and period IID effects and

$$\pi(\alpha_{n_T}) = (2\pi)^{-n_{n_T}/2} \kappa_\alpha^{n_{n_T}/2} \exp \left[ -\frac{\kappa_\alpha}{2} \alpha_{n_T}^T \mathbf{I}_{n_T} \alpha_{n_T} \right].$$

Then,  $\pi(\alpha_{T^*}|\alpha_T)$  is defined analogously to (4.8), and

$$\pi(\alpha) = \pi(\alpha_T)\pi(\alpha_{T^*}|\alpha_T) \propto \exp\left[-\frac{1}{2}\alpha^T \mathbf{Q}_\alpha^\dagger \alpha\right],$$

where

$$\mathbf{Q}_\alpha^\dagger = \begin{bmatrix} \kappa_\alpha \mathbf{I}_{n_T} + \tau_\alpha \mathbf{A}^T \mathbf{A} & -\tau_\alpha \mathbf{A}^T \\ -\tau_\alpha \mathbf{A} & \tau_\alpha \mathbf{I}_{n_{T^*}} \end{bmatrix}.$$

We again define a space-time interaction model using a Type IV interaction as described by Knorr-Held (2000) To specify the prior for all space-time interactions, first define  $\delta = [\delta_T^T \delta_{T^*}^T]^T$  where  $\delta_T$  is the  $(n \times n_T) \times 1$  is the vector of yearly interactions, and  $\delta_{T^*}$  is the  $(n \times n_{T^*}) \times 1$  vector of interactions defined with respect to periods, and  $n$  is the number of areas. Then,

$$\pi(\delta_T) \propto \exp\left[-\frac{\kappa_\delta}{2}\delta_T^T \mathbf{Q}_\delta \delta_T\right],$$

where the dependency structure is defined as  $\mathbf{Q}_\delta = \mathbf{Q}_\gamma \otimes \mathbf{Q}_\theta$  and  $\mathbf{Q}_\gamma$  and  $\mathbf{Q}_\theta$  specify the dependency structures of the RW2 and ICAR models respectively. Then  $\delta_{T^*} = \mathbf{B}\delta_T$ , where  $\mathbf{B}$  is an  $(n \times n_{T^*}) \times (n \times n_T)$  matrix that consists of  $n$  copies of  $\mathbf{A}$  on the diagonal and zeroes elsewhere. Then,

$$\pi(\delta) = \pi(\delta_T)\pi(\delta_{T^*}|\delta_T) \propto \exp\left[-\frac{1}{2}\delta^T \mathbf{Q}_\delta^\dagger \delta\right],$$

where

$$\mathbf{Q}_\delta^\dagger = \begin{bmatrix} \kappa_\delta \mathbf{Q}_\delta + \tau_\delta \mathbf{B}^T \mathbf{B} & -\tau_\delta \mathbf{B}^T \\ -\tau_\delta \mathbf{B} & \tau_\delta \mathbf{I}_{n \times n_{T^*}} \end{bmatrix}.$$

These specifications of the time effects allow us to smooth in time and incorporate data on both yearly and period scales. This also allows us to easily make estimates and predictions on both yearly and period scales.

#### 4.2.5 HIV Adjustment

One major source of selection bias in household surveys that collect child mortality information from mothers, is maternal mortality. Mothers who have died before the time of survey cannot be surveyed, and in many cases children who have lost their mothers also have increased risk of child mortality. Before the scale up of ART in LMICs, mother-to-child transmission led to a higher mortality burden from HIV/AIDs for children (WHO, 2010, p. 59-61). If household surveys are implemented after a population has experienced high HIV/AIDs mortality, mothers who died due to HIV/AIDs cannot be sampled. Moreover, their children, who were more likely to be HIV-positive, and, therefore, more likely to have died by age 5, can never be observed via the sampling method. In this work we make adjustments for the nonignorable missingness of mothers using a method that produces estimates of the proportion of deaths of children born to women who have died due to HIV prior to a survey,  $HIV_{its}$  (Walker et al., 2012). The method takes as inputs the number of women of reproductive age, the number of HIV-positive women of reproductive age, the number of births and the HIV incidence in children. We take the latter information from Spectrum (Stover et al., 2012), a simulation software used to simulate the course of the HIV/AIDS epidemic in a country and calculates many demographic measures of the population. To adjust our direct-estimates, we divide by a correction factor

$${}_{60}\widehat{q}_{its}^{HIV} = \frac{{}_{60}\widehat{q}_{its}}{1 - \% \text{ mothers missing}} = \frac{{}_{60}\widehat{q}_{its}}{1 - HIV_{its}}, \quad (4.9)$$

and adjustments to  $\widehat{V}_{its}$  are applied via the delta method. HIV adjustments used for each Admin-1 area and each region in Kenya. While we make estimates of child mortality on the Admin-2 level, we must make the same HIV adjustments for all Admin-2 areas within an Admin-1 area, as that is the finest level for which Spectrum (UNAIDS, 2014) makes estimates of the necessary inputs. Figures 4.5 and 4.6 show the HIV adjustments by area for each of the 8 Kenya regions. For Malawi, we only have national HIV adjustments and

they are pictured in Figure 4.7. The estimates reported in these plots is (1 - % children dead born to mothers missing) due to HIV/AIDS plotted against years before the survey. For FBH estimates in five-year periods the point estimate plotted represents the adjustment for the five year period before the survey starting with that year. For SBH estimates, we use adjustments for single years before the survey.

#### 4.2.6 Space-time smoothing model

In this section we discuss two difficult practical issues. The first is that each data source has its own idiosyncrasies in design and data collection, which leads to a unique set of potential biases – estimates resulting from SBH data in particular may be subject to bias (Hill et al., 2015), and we don't want them to have undue influence on our smoothed estimates. The second difficulty is that even though the IPUMS data are a 10% sample, the census data can dominate in terms of information available. Due to the large discrepancies in sample size between census data and survey data and the different methods used to analyze them, we modify the likelihood. The previous version of the model in equations (4.3) and (4.4) had a normal likelihood, and a random effect for source type. We remove the latter term for the overall estimates, which is consistent with believing that the study types are exchangeable, and no more or less likely to be susceptible to bias. We now change this specification to have fixed effects associated with the SBH data, to allow for the possibility of bias. The inclusion of these bias correction terms corrects for a data source having consistently low or high estimates and these fixed effects are assumed constant across time and space. The new model is,

$$\text{logit}(\widehat{q}_{0,its}) \sim N\left(\eta_{its}, \widehat{V}_{its}\right), \quad (4.10)$$

$$\eta_{its} = \mu + \beta_s \mathbf{1}(s \text{ is SBH}) + \alpha_t + \gamma_t + \theta_i^* + \phi_i^* + \delta_{it}, \quad (4.11)$$

where individual  $\beta_s$  are fixed bias correction terms for each data source  $s$  that provides SBH data. This allows these data sources to contribute to understanding temporal, spatial and

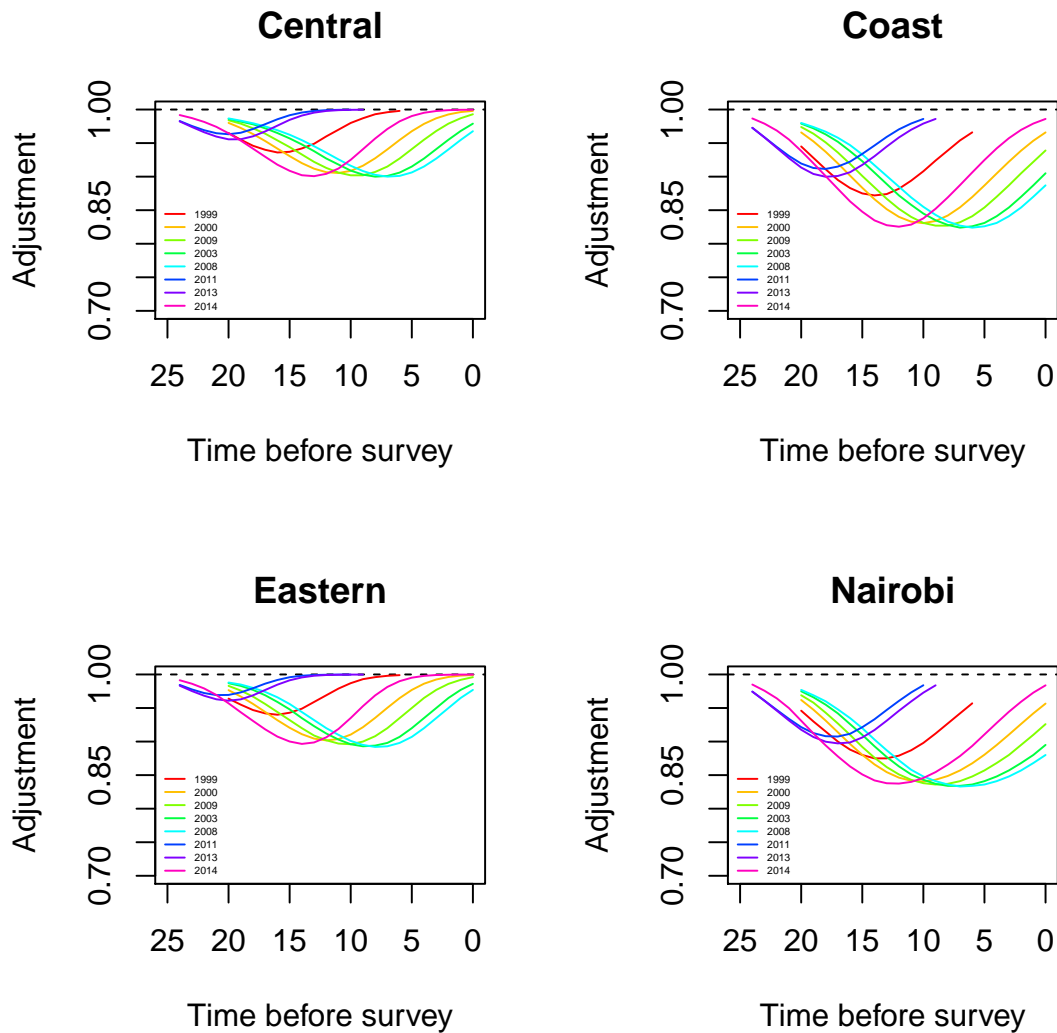


Figure 4.5: HIV adjustments for Central, Coast, Eastern regions and Nairobi.

spatio-temporal variability, but not the absolute level. This model specification is simple, but makes the strong assumption that the bias is constant across time and space.

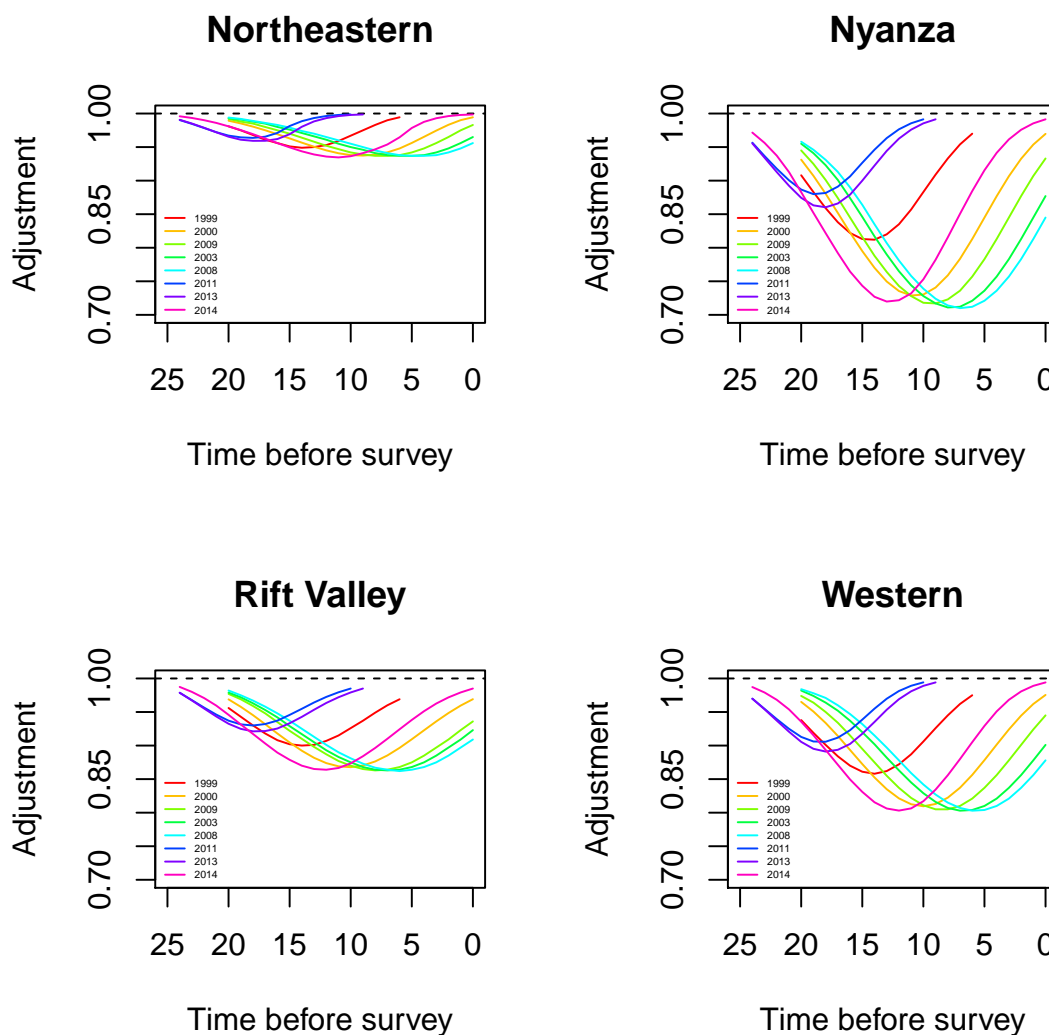


Figure 4.6: HIV adjustments for Northeastern, Nyanza, Rift Valley and Western regions.

### 4.3 Application in Kenya and Malawi

We apply the methods detailed in Section 4.2 to the data found in Table 4.1. The Kenya DHS in 2014 and the Malawi DHS in both 2010 and 2015 were stratified at the Admin-2 level

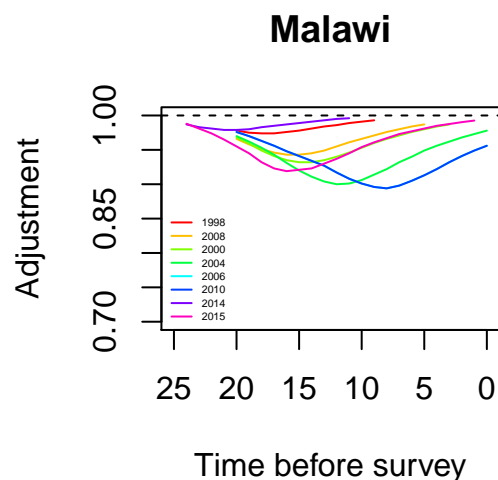


Figure 4.7: HIV adjustments for Malawi.

crossed with urban/rural, yielding sufficient sample sizes in all areas. However, the other DHS were stratified at the Admin-1 level. Figure 4.1 shows the difference in granularity of these administrative levels. Using jittered GPS locations of clusters sampled in the DHS stratified at the Admin-1 level, we can assign clusters to the appropriate Admin-2 area. For MICS surveys and IPUMS census samples, we do not have GPS locations, but we have place names. Usually, these place names are at an even finer granularity than Admin-2 (the census) or are no longer in use (older MICS and census). In these cases we query the Google Maps API to get a GPS location for the center associated with the name and then assign it to the appropriate Admin-2 area. Modeling at the subnational level also allows us to incorporate MICS surveys that were carried out with subnational coverage. In Kenya, this includes the 2011 MICS carried out in Nyanza (Admin-1) and three 2013 MICS carried out in Bungoma, Kakamega and Turkana (Admin-2).

First, we process all data according to their type, using the survey-weighted discrete

hazards survival method for the FBH data and the Brass method and jackknife for the SBH data. Then, we apply the adjustment for the areas that have no observed zeroes. These areas account for 51 of the 987 DHS area, time, survey points (5.2%) in Kenya and just 14 of the 784 points (1.8%) in Malawi. There are no MICS FBH estimates that need adjustment. Once we have converted all of our data sources into  $\left[ \text{logit}(\widehat{{}_{60}q_{0,its}^{\text{HIV}}}), \widehat{V}_{its}^{\text{HIV}} \right]$  pairs, we smooth these estimates across space and time using the model in (4.11) from Section 4.2.6. The posterior estimate of the SBH biases can be seen in Appendix A, Section A.2. To summarize, there is far more bias for the Kenya SBH data than for the Malawi data. To get estimates for an area  $i$  and time  $t$ , whether  $t$  is a year or a period, we take draws from the joint posterior. to compute  ${}_{60}q_{0,it}$ . Fitting the model in INLA using 8GB for Malawi took 421 seconds with 28 Admin-2 areas. For Kenya’s 47 Admin-2 areas it took 1072 seconds, which is much longer but still a relatively short amount of time for such a complex Bayesian model.

#### 4.3.1 Results

In this section we review results from the space-time smoothing model using the normal likelihood, (4.11). Figure 4.8 shows the data used in estimation and the smoothed estimates with 95% intervals for years 1980–2014 for two areas, Nairobi county in Kenya (Left) and Lilongwe district in Malawi (Right). Point sizes of the FBH and SBH estimates are weighted relative to the median 95% CI for any estimate in that Admin-2 area. Estimates from surveys and censuses with 95% CI widths larger than the median width in that Admin-2 area are made smaller using the ratio of the median width to the point’s CI width. Points whose 95% CI widths are below the median width for that county are made larger using the ratio of the median width to that point’s 95% CI width. Thus, larger points represent more precise estimates with narrower CIs and have a larger influence on the smoothing model than smaller points. In Nairobi, the increased mortality due to the HIV/AIDS epidemic is evident in the period between 1985 and 2005. Though U5MR in Nairobi has declined since the peak of the

epidemic, its current U5MR is roughly the same as it was before the epidemic began. The additional information from the SBH estimates gives us a slightly different estimate of the time period of increased U5MR due to the HIV/AIDS crisis. Moreover, the rate at which U5MR has been declining in Nairobi has slowed in recent years. Addition of SBH data gives us shorter CI widths historically, and even slightly shorter PIs for the period 2015-2019. For Lilongwe, Figure 4.8 shows a large decline in child mortality over the period analyzed. The inclusion of SBH data provides us with expected precision gains, but only changes the U5MR trend in Lilongwe before 1990. Compared to Kenya, the Malawi SBH estimates are more consistent with the FBH estimates over time. Thus, in general, we see smaller differences in trends estimated with and without SBH data. Similar results for all Admin-2 areas in Malawi and Kenya can be found in Sections A.5–A.8 of Appendix A. National results for both Malawi and Kenya can be found in Appendix A, Section A.4.

Figure 4.9 shows posterior median U5MR estimates in Malawi by district with and without SBH data for the period 2010–2014 on the top row. We see two Admin-2 areas in the Southern region (Neno and Mulanje) and one area in the Central region (Mchinji) with consistently high median estimates of U5MR regardless of the inclusion or exclusion of SBH data. Figure 4.10 shows the same set of results for Kenya. Three areas stand out as having particularly high U5MR relative to other Admin-2 areas: one in Rift Valley province (Turkana) and the other two in Nyanza province (Migori and Homa Bay). While both the minimum and maximum of the U5MR scale in Kenya is lower than in Malawi, the interval widths in Kenya tend to be larger. For both countries, the addition of SBH data leads to more Admin-2 areas with smaller 95% interval widths than estimates using only FBH data. This can be seen in the presence of more lightly shaded areas in the bottom right panels than in the bottom left panels of Figures 4.9 and 4.10.

Figure 4.11 shows district level estimates in five-year periods from 1980–1984 to 2010–2014 and projections to 2015–2019 using all the data. The huge decline in U5MR over this

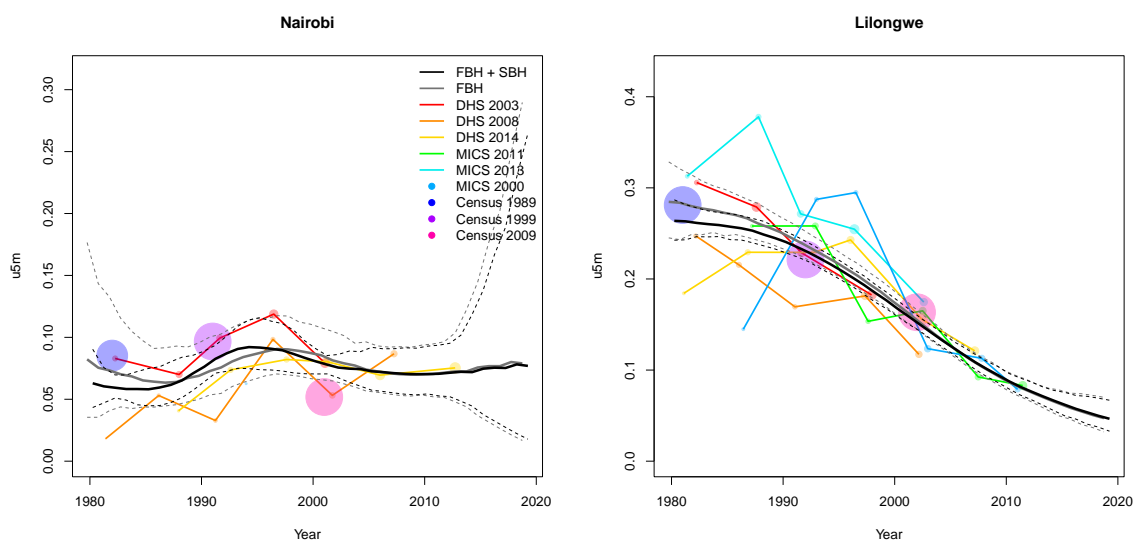


Figure 4.8: Survey and census estimates overlaid with smoothed estimates and intervals. Estimates connected by lines are direct estimates from either DHS or MICS; single points are Brass estimates from MICS or the census. For the estimates, the size of the point is weighted by the width of 95% CI. Larger points imply a shorter width and more precise estimates. Left: Results for Nairobi county in Kenya. Right: Results for Lilongwe district in Malawi.

period is easily seen. The relatively higher mortality of the Central region is also apparent relative to the Northern and Southern regions. Throughout the period of time analyzed, the Southern region appears to have more within region variability across Admin-2 areas than either the Northern or Central regions. Though there is still subnational variability in Malawi today, it is easier seen via the scale used in Figure 4.9 as all districts have relatively low U5MR compared to earlier periods. Relative subnational variation remains, however.

In Figure 4.12 we see a similar pattern of broad U5MR decline across all 47 counties in Kenya between the 1980–1984 and the 2010–2014 periods. However, the impact of the

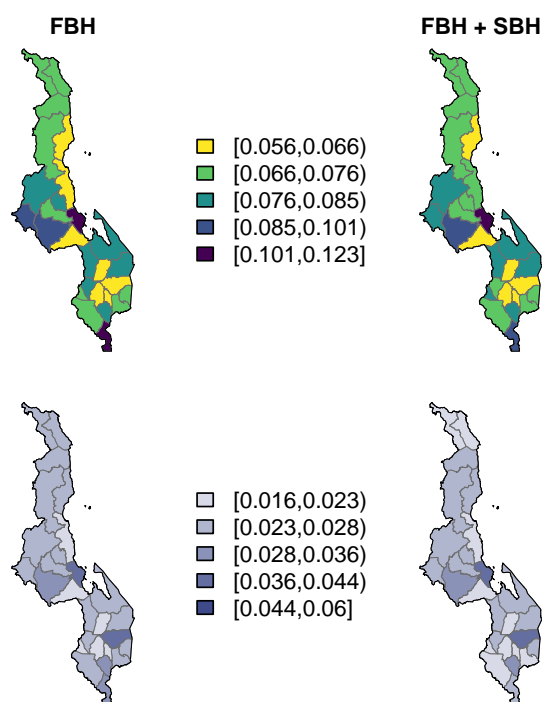


Figure 4.9: Results for Malawi by area for the period 2010–2014. **Top:** Smoothed estimates by area, with only FBH data (Eq 4.3) (Left) and FBH and SBH data (Right) (4.11). **Bottom:** 95% interval widths by area with FBH data only (Left) and FBH and SBH data (Right).

HIV/AIDS epidemic on U5MR in Kenya is apparent in the top row of Figure 4.12 as many areas underwent a clear rise in child mortality during this period, followed by a decline after the turn of the century. Consistently, counties in the Kenyan region of Nyanza experience higher relative mortality when compared to those in other regions. Though the highest experienced U5MRs in both Kenya and Malawi are similar, the lowest experienced and projected U5MRs in Kenya are a bit lower than in Malawi. Comparing the two figures, it is also evident that the high U5MR in the Nyanza region of Kenya lasted much later than in any region which previously experienced these high U5MRs in Malawi.

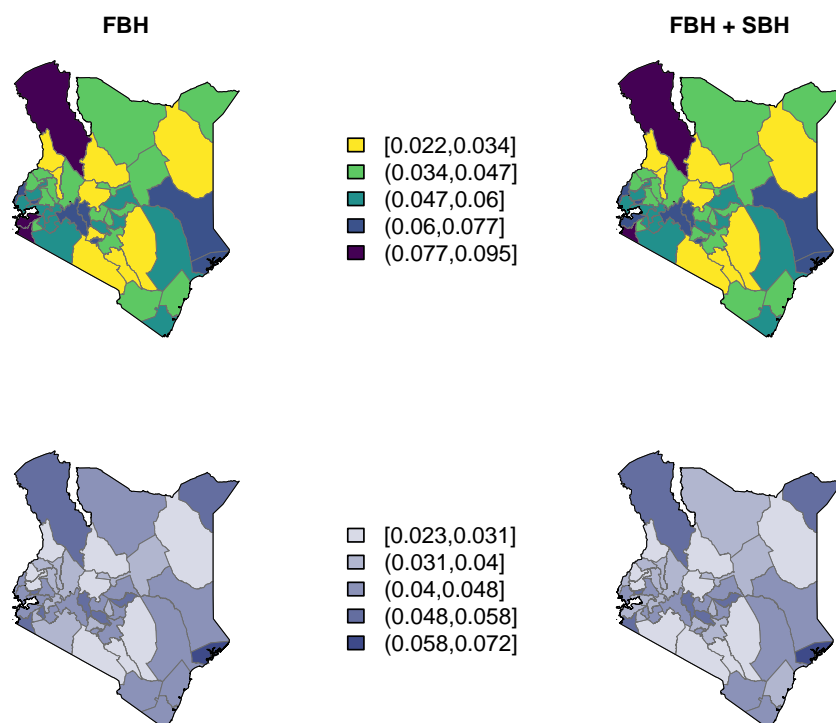


Figure 4.10: Results for Kenya by area for the period 2010–2014. **Top:** Smoothed estimates by area, with only FBH data (Eq 4.3) (Left) and FBH and SBH data (Right) (4.11). **Bottom:** 95% interval widths by area with FBH data only (Left) and FBH and SBH data (Right).

#### 4.3.2 Validation

As an exercise in model validation, cycling through all areas and periods from 1990–1994 to 2010–2014, we leave out data from an area  $i$  and a time period  $t$ , fit our model and then make estimates of  ${}_{60}q_{0,it}$ . We assess the performance by computing precision-weighted versions of standard model assessment statistics. Let  $\hat{y}_{it}$  be the smoothed median estimate of  $\text{logit}({}_{60}q_{0,it})$  and  $\hat{V}_{it}$  be its variance when data from area  $i$  and period  $t$  are omitted. We

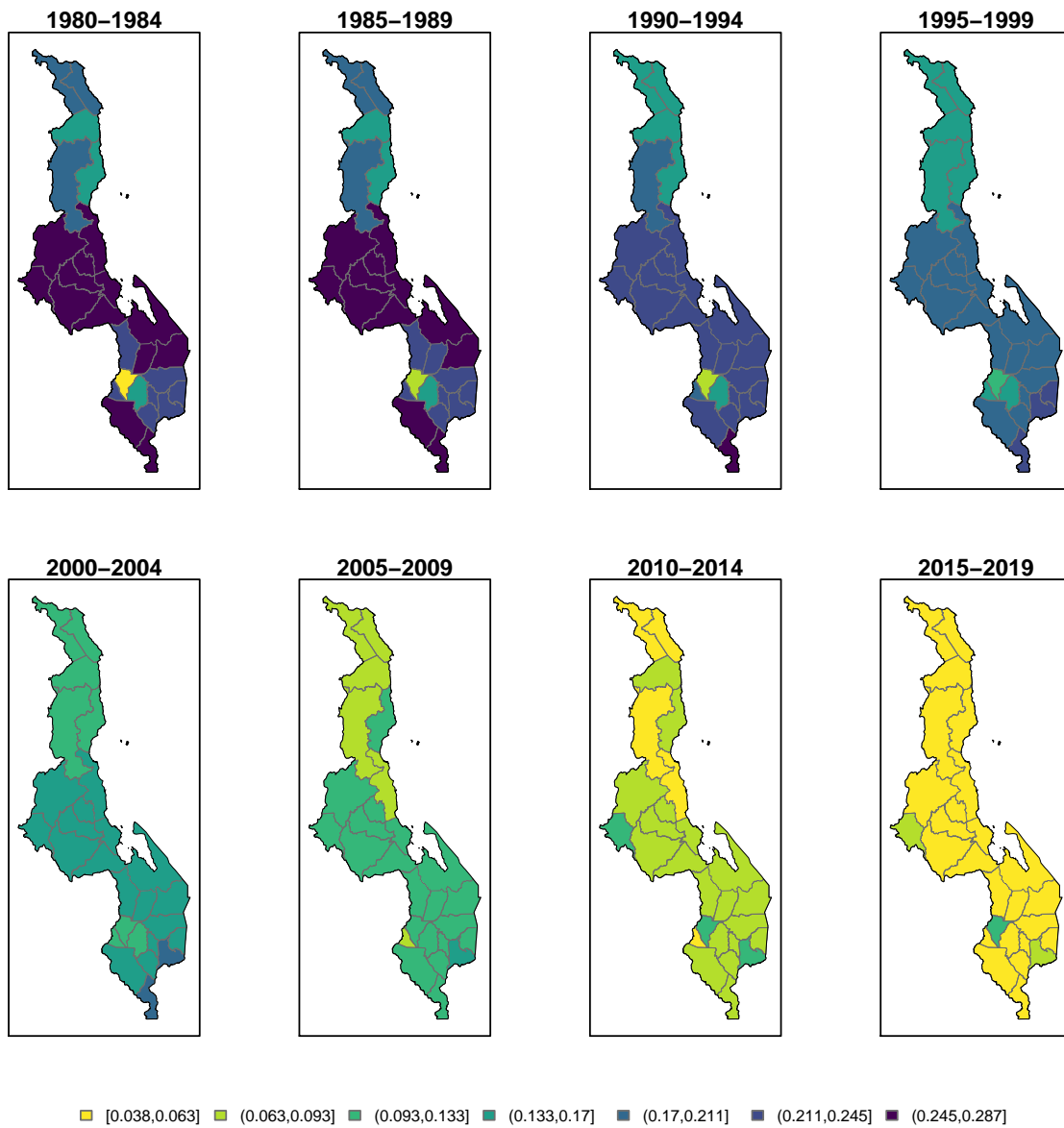


Figure 4.11: Median U5MR estimates for five-year periods in districts in Malawi, using all data sources, from 1980–1984 to 2010–2014, with projections to 2015–2019.

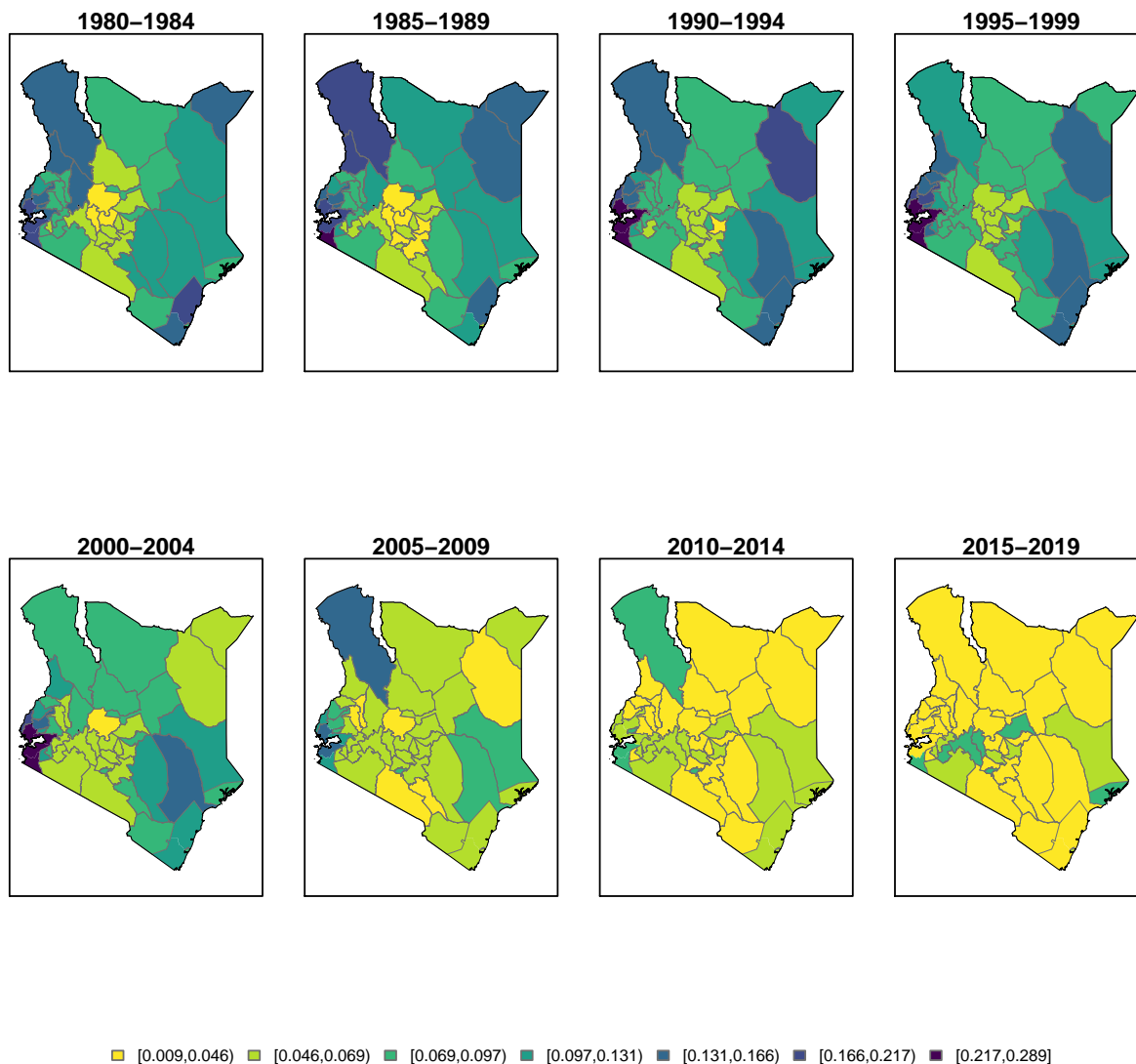


Figure 4.12: Median U5MR estimates for five-year periods in counties in Kenya, using all data sources, from 1980–1984 to 2010–2014, with projections to 2015–2019.

take

$$\hat{y}_{it} = \sum_{s \in F_{it}} w_{its} \hat{y}_{its} = \sum_{s \in F_{it}} \frac{\hat{V}_{its}^{-1}}{\sum_{s \in F_{it}} \hat{V}_{its}^{-1}} \hat{y}_{its},$$

and

$$\hat{V}_{it} = \frac{1}{\sum_{s \in F_{it}} \hat{V}_{its}^{-1}},$$

where  $s \in F_{it}$  indicates  $s$  is an FBH data source for area  $i$  and time  $t$ , as measures of truth and variance, respectively, for each area and time period. Given that the estimates can have a wide range of precisions, we weight the estimate of the “truth” by its relative precision. For the MSE, then,

$$\text{MSE}_t = \sum_{i=1}^n w_{it} (y_{it} - \hat{y}_{it})^2,$$

where  $n$  is the total number of areas and  $w_{it} = \frac{\hat{V}_{it}^{-1}}{\sum_{i=1}^I \hat{V}_{it}^{-1}}$ . Similarly, we can define the mean absolute error (MAE) as

$$\text{MAE}_t = \sum_{i=1}^n w_{it} |y_{it} - \hat{y}_{it}|,$$

and the mean absolute percent error (MAPE) as,

$$\text{MAPE}_t = \sum_{i=1}^n w_{it} \frac{|60q_{0,it} - 60\hat{q}_{0,it}|}{60\hat{q}_{0,it}}.$$

Tables 4.2 and 4.3 compare the above validation metrics between a model with only FBH data sources included and a model with both FBH and SBH data sources included. Reference dates for SBH estimates vary across space, but tend to contribute to a single five-year period. For example, in Malawi, all Brass estimates from the 1987 census have reference dates between years 1980-1981, all estimates from the 1998 census lie between years 1991-1992, and all estimates from the 2008 census lie between years 2001-2002. This means the effect of incorporating SBH data into the model will have different effects in different time periods. As a result, we see that neither model out performs the other in all time periods. In all periods with SBH estimates, Table 4.2 shows the MSE is smaller for the model with

both data types. For Kenya, we see the addition of SBH data decreases the MAE and MAPE for all periods with SBH data (1990–1994, 2000–2004) as well as in most other time periods. Comparing Sections A.5–A.8 of Appendix A shows much more between data source variability within Admin-2 area in Kenya than in Malawi. There is more estimated bias in the SBH data sources from Kenya than Malawi (see Figure A.3). These factors all contribute to give larger estimates of validation metrics across the board in Kenya as compared to Malawi.

Year	FBH				FBH + SBH			
	Bias	MSE	MAE	MAPE	Bias	MSE	MAE	MAPE
1990–1994	<b>0.001</b>	0.030	0.107	0.096	0.003	<b>0.024</b>	<b>0.089</b>	<b>0.078</b>
1995–1999	<b>0.001</b>	<b>0.041</b>	<b>0.108</b>	<b>0.083</b>	-0.005	0.043	0.113	0.086
2000–2004	<b>0.003</b>	0.015	0.087	0.077	0.004	<b>0.009</b>	<b>0.069</b>	<b>0.060</b>
2005–2009	<b>-0.009</b>	0.022	0.122	<b>0.100</b>	-0.013	<b>0.020</b>	<b>0.117</b>	0.103
2010–2014	0.007	0.058	<b>0.190</b>	<b>0.177</b>	<b>0.006</b>	<b>0.055</b>	0.194	0.181
<b>Mean</b>	<b>0.000</b>	0.033	0.121	0.106	-0.001	<b>0.030</b>	<b>0.116</b>	<b>0.102</b>

Table 4.2: Assessment measures for Malawi by period comparing the FBH + SBH smoothing model (columns 6-9) and the FBH smoothing model (columns 2-5) when leaving out data from each area and period. For each period, “best” values in each metric are bolded.

We perform an additional validation exercise by fitting the model to all data except the most recent DHS stratified at the Admin-2 level. Both the Kenya 2014 DHS and the Malawi 2015 DHS were stratified at a finer geographic level, and there are decent sample sizes in all areas and most time periods going back to 1985–1989. In this case, we use the direct estimate from these removed DHS surveys as the “truth” in our validation metrics. We calculate the same metrics as in the first validation exercise. In the scenario where we leave out the 2015 DHS in Malawi, the advantage to using both data sources is much

Year	FBH				FBH + SBH			
	Bias	MSE	MAE	MAPE	Bias	MSE	MAE	MAPE
1990–1994	0.091	0.203	0.318	0.311	<b>0.038</b>	<b>0.189</b>	<b>0.288</b>	<b>0.251</b>
1995–1999	<b>0.074</b>	<b>0.119</b>	<b>0.268</b>	<b>0.255</b>	0.101	0.136	0.287	0.279
2000–2004	0.105	0.179	0.304	0.314	<b>0.061</b>	<b>0.152</b>	<b>0.277</b>	<b>0.291</b>
2005–2009	0.120	0.288	0.398	0.480	<b>0.092</b>	<b>0.260</b>	<b>0.373</b>	<b>0.436</b>
2010–2014	<b>0.028</b>	0.393	0.480	<b>0.477</b>	0.055	<b>0.390</b>	<b>0.474</b>	<b>0.477</b>
<b>Mean</b>	0.084	0.236	0.354	0.368	<b>0.069</b>	<b>0.225</b>	<b>0.340</b>	<b>0.347</b>

Table 4.3: Assessment measures for Kenya by period comparing the FBH + SBH smoothing model (columns 6-9) and the FBH smoothing model (columns 2-5) when leaving out data from each area and period. For each period, “best” values in each metric are bolded.

clearer. Table 4.4 shows that in each of the six time periods assessed in this exercise, the MSE is much lower when we incorporate SBH data. In nearly all time periods the bias is also smaller with the incorporation of SBH data. In Kenya, Table 4.5 shows less evidence of improvements in estimates, but we do see decreases in MSE and MAPE in at least every time period with SBH data contributions. Moreover, the gains in MSE and MAPE tend to be larger in magnitude than in the cross-validation exercise. This is likely due to the fact that we remove much more data when removing an entire DHS survey (see Table 4.1) than when leaving out data for a single area and time period.

#### 4.4 Discussion

We have extended the previous work of Mercer et al. (2015) in several ways to incorporate data at a finer (yearly) time-scale and to make estimates at a finer (Admin-2 level) geographic scale. We have proposed a method for making more stable estimates of U5MR in a small area

Year	FBH				FBH + SBH			
	Bias	MSE	MAE	MAPE	Bias	MSE	MAE	MAPE
1985–1989	<b>0.014</b>	0.382	0.418	0.447	0.015	<b>0.328</b>	<b>0.388</b>	<b>0.387</b>
1990–1994	-0.310	0.202	0.390	0.273	<b>-0.237</b>	<b>0.124</b>	<b>0.305</b>	<b>0.217</b>
1995–1999	-0.353	0.234	0.431	0.308	<b>-0.177</b>	<b>0.105</b>	<b>0.274</b>	<b>0.210</b>
2000–2004	-0.370	0.237	0.427	0.302	<b>-0.103</b>	<b>0.077</b>	<b>0.224</b>	<b>0.181</b>
2005–2009	-0.447	0.274	0.463	0.320	<b>-0.127</b>	<b>0.065</b>	<b>0.191</b>	<b>0.153</b>
2010–2014	-0.493	0.298	0.493	0.347	<b>-0.165</b>	<b>0.081</b>	<b>0.217</b>	<b>0.174</b>
<b>Mean</b>	-0.327	0.271	0.437	0.333	<b>-0.132</b>	<b>0.130</b>	<b>0.267</b>	<b>0.220</b>

Table 4.4: Assessment measures for Malawi by period comparing the FBH + SBH smoothing model (columns 6-9) and the FBH smoothing model (columns 2-5) when leaving out the Malawi 2015 DHS and treating estimates from this survey as the “truth”. For each period, “best” values in each metric are bolded. (Note: results here exclude the district of Likoma. Removal of the 2015 DHS removes our only FBH data source with usable estimates for Likoma. Methods such as those included in Section 4.2.2 cannot be used as Likoma is an island, and is quite distinct from areas on the mainland, making a spatial smoothing model inappropriate.)

with small sample sizes by borrowing information from neighboring areas while accounting for the complex survey design. We applied this method to the 47 counties of Kenya and the 28 districts of Malawi. The estimates account for the selection bias that arises from mothers dying prematurely from HIV/AIDS, by estimating a time series of the proportion of missing children of women who have died due to the epidemic in years before the survey. However, the method does not estimate uncertainty in that proportion, which is a source of uncertainty we would like to include in future work.

Year	FBH				FBH + SBH			
	Bias	MSE	MAE	MAPE	Bias	MSE	MAE	MAPE
1985–1989	<b>-0.295</b>	<b>0.820</b>	<b>0.668</b>	0.521	-0.334	0.982	0.706	<b>0.488</b>
1990–1994	0.096	0.322	0.437	0.481	<b>-0.005</b>	<b>0.162</b>	<b>0.321</b>	<b>0.275</b>
1995–1999	0.064	0.321	0.412	0.463	<b>-0.017</b>	<b>0.170</b>	<b>0.308</b>	<b>0.255</b>
2000–2004	<b>-0.312</b>	0.379	0.484	0.373	-0.390	<b>0.332</b>	<b>0.462</b>	<b>0.320</b>
2005–2009	<b>-0.541</b>	<b>0.543</b>	<b>0.625</b>	0.425	-0.610	0.581	0.661	<b>0.428</b>
2010–2014	-0.353	0.361	<b>0.453</b>	<b>0.322</b>	<b>-0.343</b>	<b>0.335</b>	0.462	0.341
<b>Mean</b>	<b>-0.224</b>	0.458	0.513	0.431	-0.283	<b>0.426</b>	<b>0.487</b>	<b>0.351</b>

Table 4.5: Assessment measures for Kenya by period comparing the FBH + SBH smoothing model (columns 6-9) and the FBH smoothing model (columns 2-5) when leaving out the Kenya 2014 DHS and treating estimates from this survey as the “truth”. For each period, “best” values in each metric are bolded.

Alkema and New (2014) include a linear in time bias term in their model. We explored this idea with the Kenya data, but plots (See Appendix A, Section A.2) showed no indication of a systematic bias with time at the subnational level. While the incorporation of SBH data in this new method provide gains in precision, especially in the time periods the SBH data are specifically contributing to, comparing the results and validation exercises between Malawi and Kenya show that the magnitude of improvement depends on how the SBH estimates compare to the FBH estimates for a particular country. The use of SBH data is appealing, but in any one country the appropriateness of inclusion must be carefully determined as there are many sources of potential bias (Hill et al., 2015; Brady and Hill, 2017).

Existing methods for subnational estimation of child mortality in low and middle income countries fall into two main categories: discrete space and continuous space models. They

differ in whether the spatial structure of the data is modeled on the administrative area scale (e.g., counties, districts) or on the continuous scale using the GPS locations of sampled clusters. Previous estimates have been made for the probability of death before age 5 given survival to age 1,  ${}_{60}q_{12}$ , on an areal scale, using ecological covariate information (Pezzulo et al., 2016). These estimates only make use of the most recent DHS survey in each country, and, therefore, do not estimate time trends. As already mentioned, Li et al. (2019) use a discrete spatial model to obtain estimates over time at the Admin1 level for 35 African countries (with each country fitted separately). Continuous spatial models have also been used to estimate subnational child mortality (Golding et al., 2017; Burstein et al., 2019). Specifically, the SPDE approach (Lindgren et al., 2011) has been combined with an autoregressive temporal model to give estimates the  $5 \times 5$  km scale. In these approaches, SBH and FBH data are used along with covariate surfaces. No adjustment for the HIV/AIDS epidemic is made, and when data on the GPS of the clusters is unavailable but rather the areal polygon within which the cluster lies (which is often the case for SBH data), the deaths are randomly distributed within the polygon according to population density which is ad hoc. The limitations of allocating areal data to unknown locations have been illustrated (Marquez and Wakefield, 2020; Wilson and Wakefield, 2018). Also, the same space-time model is used for many contiguous counties, and the reasonableness of this assumption has not been addressed. Finally, the stacking procedure that is used for covariate modeling is not statistically legitimate (Wakefield et al., 2019).

In this chapter we have presented estimates for U5MR, but the discrete hazards model we use allows for estimating other measures of interest such as neonatal mortality and infant mortality. We did not incorporate covariate information yet though in principle we could include area-level variables within the models described in the chapter. Initial explorations of such models (in the context of HIV prevalence mapping) is encouraging (Wakefield et al., 2020).

## Chapter 5

# MODEL-BASED ESTIMATION OF SUBNATIONAL CHILD MORTALITY AT THE ADMIN-2 LEVEL WITH COMPLEX SURVEYS

### **5.1 Introduction**

In this chapter, we leave behind design-based inference to pursue a model-based solution for Admin-2 level estimates of U5MR from DHS FBH data in a larger number of countries that, unlike Malawi and Kenya, do not have surveys stratified at the Admin-2 level. If the amount of data that require adjustments like those introduced in Section 4.2.2 is more than a small percentage, the method introduced in Chapter 4 is inappropriate. When sample sizes are adequate, sampling weights are provided, and not all variables related to the sampling scheme are available for each individual in the sample, design-based methods will more adequately account for the sampling design. Model-based methods can be used when design-based methods are unstable, but rely on sampling variables to be present in the data for full adjustment. The DHS do not provide information on number of households in each cluster, so adjustment for PPS sampling is not possible. Depending on the complexity of the sampling scheme, there may also be limitations to the number of parameters that can be estimated with a useful amount of precision, and, therefore, the number of stages or strata in a sampling design for which the model can account. We present a model-based discrete time hazards model that attempts to account for as much of the sampling scheme as possible. Beyond the movement to a model-based framework, this method differs from that in Chapter 4 in that it they simultaneously estimate age-specific hazards and spatiotemporal smoothing terms.

Previously, the Institute for Health Metrics and Evaluation (IHME) have produced model-based subnational estimates for 99 countries (Burstein et al., 2019). Philosophically, our approach contrasts sharply with that of IHME as we model each country individually, instead of larger groups of countries simultaneously. This gives ample opportunity for spatial smoothing but allows country-specific models to be specified in a way that acknowledges the nuances within a country. In Sections 5.5 and 5.6.1, we show how simple, thoughtful deviations from the generic model produce more reasonable estimates without sacrificing reproducibility. IHME makes use of FBH, SBH, and VR data to varying degrees; however, reliable statistical methods have not been developed for certain data for SBH data, VR data and data with incomplete geographical information. For example, in some survey the locations of the clusters are not provided, rather only the administrative division of the mother is reported. IHME consider space as a continuous surface and use an ad hoc procedure (Golding et al., 2017; Utazi et al., 2019) to assign mothers with incomplete data to random locations within the area. This procedure does not account for the data generating mechanism that would produce areal level data (Marquez and Wakefield, 2020; Wilson and Wakefield, 2018). In this chapter, we consider space as a discrete surface partitioned according to administrative boundaries which allow for more meaningful interpretation and use of results and avoids increased computational demands of continuous spatial methods. We limit our data to complex household surveys providing FBH data where the GPS location of sampled locations is available, so that we can assign clusters to any partition of the country. Further discussion of and motivation for the modeling choices in this chapter are discussed in Section 5.2.2. All of our estimation is carried out using the `SUMMER` package (Li et al., 2020), which runs within the `R` programming environment, and begins with survey data files as they have been published. This allows our estimates to be reproduced step-by-step, which is not true of IHME's estimates.

## 5.2 Methods

### 5.2.1 Data

As mentioned, we restrict our focus in this chapter to DHS FBH data from 22 countries selected based on amount of and timing of the most recent data collected. The retrospective nature of the birth histories means that a survey provides information on child mortality going backward in time and not just in the year the survey was administered. These birth histories combined with the geographic information of sampled clusters and administrative divisions comprise the data we use to make estimates of U5MR. Most surveys provide only the Admin-1 location (the geographic level of stratification) of each child, not the Admin-2 location. GPS datasets are provided for select DHS that contain coordinates of EAs which have been jittered for privacy (Burgert et al., 2013). When these GPS data are available for a survey, we can assign the cluster locations to administrative boundaries of the shapefiles available in the Database of Global Administrative Areas (GADM) data repository (2019). We then assign all children from a sampled EA to the Admin-2 area the GPS coordinate of the cluster lies within. The jittering algorithm for DHS cluster locations guarantees that the jittered GPS points do not move outside the correct geographic level of stratification (usually Admin-1) (Burgert et al., 2013). There is no such guarantee made for lower levels of geography. There is also no guarantee the boundaries of GADM shapefiles and those used in the jittering of GPS locations are the same. Therefore these assignments suffer from the risk that the jittered coordinates lie outside the true Admin-2 area, but we hope this occurs rarely.

In order to make Admin-2 level estimates of U5MR for the period 1990-2019, we restrict our estimates to a subgroup of DHS surveys with available GPS data. We only make estimates for countries whose most recent survey was administered no earlier than 2013, so that our projection period of U5MR forward in time is kept to a minimum. Table 5.1 contains

the list of countries and surveys with sufficient data and used in subsequent estimation. The second and third columns contain the number of administrative units at Admin-1 and Admin-2 levels for which we make estimates. In cases where our number of units do not match those available from GADM, we were given different shapefiles for administrative boundaries directly from the countries themselves.

Other data used in analysis include UN IGME national U5MR estimates (UN IGME, 2020) which are used in benchmarking procedures and WorldPop population surfaces for ages 0–1 and 1–4 (Tatem, 2017; Stevens et al., 2015) used in crisis adjustments for the 2008 Cyclone Nargis in Myanmar and Rwandan genocide.

Country	No. Admin-1	No. Admin-2	No. Surveys	Years	Spacetime Model
Benin	12	76	3	1990–2017	RW1
Burundi	18	119	2	1990–2016	RW1
Ethiopia	11	79	4	1990–2016	AR1
Ghana	10	137	5	1990–2014	RW1
Kenya	47	–	3	1990–2014	RW1
Lesotho	10	–	3	1990–2014	RW1
Liberia	15	66	2	1990–2013	RW1
Malawi	3	28	4	1990–2015	RW1
Mali	9	50	5	1990–2018	RW1
Myanmar	15	63	1	1990–2015	RW1
Namibia	13	107	3	1990–2013	RW1
Nepal	7	77	4	1990–2016	RW1
Nigeria	37	–	5	1990–2018	RW1
Pakistan	8	32	2	1990–2017	RW1
Rwanda	5	30	4	1990–2014	RW1
Senegal	14	45	9	1990–2019	RW1
Sierra Leone	5	16	1	1990–2013	RW1
Tanzania	30	169	3	1990–2015	AR1
Togo	5	39	2	1990–2013	RW1
Uganda	4	135	4	1990–2016	RW1
Zambia	10	72	3	1990–2018	RW1
Zimbabwe	10	60	3	1990–2015	RW1

Table 5.1: Countries and DHS surveys for which subnational estimates of U5MR are produced in this chapter. Columns 2 and 3 contain the number of administrative units used to make estimates at the Admin-1 and Admin-2 levels respectively. (–) indicates estimates were not made for the particular country and administrative level combination.

### 5.2.2 Review of small area estimation methods of U5MR

Estimating U5MR at the Admin-2 level in countries reliant on household survey data is an endeavor in both small area estimation (SAE) which includes accounting for complex sampling schemes. In SAE, it is common to distinguish between *area-level* models (Fay and Herriot, 1979), where data are modeled at the small area level itself, and *unit-level* models (Battese et al., 1988). Previously, Li et al. (2019) made Admin-1 level estimates of U5MR for 35 countries using a design-based, area-level model which is an extension of the classic Fay-Herriot model (Fay and Herriot, 1979; Mercer et al., 2015). This method requires area-level, design-based estimates of U5MR and their corresponding variances which are then smoothed in space and time. When surveys are stratified at the Admin-1 level, sample sizes quickly become too small at finer levels of geographic division. Small sample sizes lead to design-based estimates of zero in many Admin-2 areas and extremely large variances in others, both of which lead to the area-level, design-based method of Li et al. (2019) being unsuitable at the Admin-2 level (Godwin and Wakefield, 2021).

To make reasonable estimates of U5MR at the Admin-2 level, we rely on a unit-level, model-based approach that models data on the PSU, or EA, level. This alleviates the small sample issues encountered in the design-based, area-level approach, but comes with a cost of larger datasets and increased computation. Previous work explored unit-level modeling of U5MR using smoothing models that consider space as continuous, allowing estimates to be aggregated up to the administrative division desired. However, estimates at any given point in space can have large uncertainty if there are no sampled points nearby (Wakefield et al., 2019). IHME use continuous spatial models to make predictions of U5MR and other child mortality measures, but fit models to large regional groups of countries at once instead of modeling data from each country separately (Burstein et al., 2019). The continuous spatial model requires an assumption that the distance between two points beyond which there is no spatial dependence is the same for all countries within that group. Given the

varying sizes of nations within these groups, this assumption seems unreasonable. Regional modeling also makes it difficult to address the unique circumstances informing child mortality in each country such as epidemics, natural disasters or crises. Moreover, Paige et al. (2020) compare continuous and discrete space model-based approaches to small area estimation in complex surveys and find that while continuous spatial models allow estimation of subnational variability at a smaller scale than administrative divisions, they are much more difficult to implement and come at a much higher computational cost than discrete space models which also perform well. The work also elucidates the importance of accounting for the dependency in outcomes between children in the same PSU in a model-based approach. Following Wakefield et al. (2020), we use a betabinomial likelihood to account for this within-cluster dependence and use a discrete spatial model for easier computation, implementation, and communication of results. The betabinomial model presented in this chapter does not directly account for within-household dependence in SSUs at the second stage of sampling but does account for within cluster dependence. Finally, the model does not attempt to account for either the urban/rural stratification or the PPS sampling at the first stage. Exploratory endeavors quickly found that using fixed effects for stratification required further work before appropriate methods could be implemented. As the number of admin areas  $I$  grows, the feasibility of estimated a fixed effect for each administrative strata decreases especially as data sparsity increases as  $I$  grows. If we restrict focus to fixed effects for the urban/rural strata, we can aggregate the urban and rural estimates of U5MR into a single estimate according to proportions of children under age 5 residing in urban and rural locations respectively. However, at Admin-2 level there is not enough information to estimate with precision these proportions as in (4.6) and other methods are needed. We proceed with ignoring urban/rural stratification in this model-based method in this thesis.

### 5.3 A betabinomial model for estimating U5MR

Our initial data come in the form of birth dates and death dates for each child born to a mother sampled by the DHS. Each of these records is expanded into monthly observations with a binary outcome indicating whether death occurs in the respective month of life in preparation for discrete time survival analysis with discrete hazards by age groups (Allison, 2014). For example, a child who survives to their 5<sup>th</sup> birthday will contribute 60 binary observations to our data and a child who dies within the first month of life will only contribute 1 observation. In order to speed up computation, we collapse the binary outcome data into binomial observations for each cluster  $c$  within area  $i$  in year  $t$  from survey  $s$  for age groups 0–1, 1–11, 12–23, 24–35, 36–47, 48–59 months. We fit a model that smooths over space and time to make estimates of the hazard of death within a month of life conditional on survival to the beginning of that month within each age band. Finally, we aggregate the monthly probabilities of death to calculate the desired measure, U5MR.

Formally, we let  $m = 0, 1, \dots, 59$  denote age in months, and define

$$a[m] = \begin{cases} 1, & 0 \leq m < 1 \\ 2, & 1 \leq m < 12 \\ 3, & 12 \leq m < 24 \\ 4, & 24 \leq m < 36 \\ 5, & 36 \leq m < 48 \\ 6, & 48 \leq m < 60 \end{cases} \quad a^*[m] = \begin{cases} 1, & 0 \leq m < 1 \\ 2, & 1 \leq m < 12 \\ 3, & 12 \leq m < 60 \end{cases} .$$

Our data come in the form of number of deaths observed,  $y_{m,i[c]ts}$ , in  $n_{m,i[c]ts}$  months of life in cluster  $c$  within area  $i$ , year  $t$  in agethmonth  $m$  from survey  $s$ . We fit

$$\begin{aligned}
y_{m,i[c]ts} | n_{m,i[c]ts}, {}_1q_{m,it} &\sim \text{BetaBinomial}(n_{m,i[c]ts}, {}_1q_{m,it}, d) \\
\eta_{m,i[c]ts} = \text{logit}({}_1q_{m,i[c]ts}) &= \mu_{a[m]} + \alpha_t + \gamma_{a^*[m]t} + \theta_i^* + \phi_i^* \\
&+ b_i t + \delta_{it} + \beta_s + \log(\text{HIV}_{i[c]t}),
\end{aligned} \tag{5.1}$$

where,  ${}_1q_{m,i[c]ts}$  is the monthly hazard of death for age  $m$  and  $d$  is an overdispersion parameter that allows the probabilities of death within clusters and surveys to vary around  ${}_1q_{m,it}$ . The mean model contains a fixed intercept for agebands,  $\mu_{a[m]}$ . First order temporal terms include a RW2 for each age group  $a^*[m]$ ,  $\gamma_{a^*[m]t}$ , and an IID random effect for temporal shocks,  $\alpha_t \sim N(0, \tau_\alpha)$ . The first order spatial terms taken together,  $\theta_i^* + \phi_i^*$ , follow a BYM2 prior (Riebler et al., 2016). Spacetime interactions are accounted for with two terms: random area-specific slopes,  $b_i$ , which are constrained to sum to zero and a Knorr-Held Type IV interaction terms,  $\delta_{it}$ , with proper sum-to-zero constraints for the spatial and temporal components (Knorr-Held, 2000). The Type IV prior for most countries is a ICAR $\times$ RW1, but there were a few of the countries for which found an ICAR $\times$ AR1 prior was more appropriate. The survey terms  $\beta_s$  are fixed effects constrained to sum to zero. Finally, a log-offset term,  $\log(\text{HIV}_{i[c]t})$ , is included in the model to account for the estimated proportion of mothers who could not be sampled due to death from HIV/AIDS. As mother and child HIV status are dependent, the children of these missing mothers are more likely to have been HIV positive and HIV positive children have increased risk of death before age 5 (WHO, 2010, p. 53). Ignoring this will lead to underestimates of U5MR in countries with generalized HIV/AIDS epidemics. We incorporate national offsets for Cameroon, Lesotho, Malawi, Namibia, Rwanda, Tanzania, and Uganda. We have subnational offsets for Kenya, Zambia, and Zimbabwe.  $\text{HIV}_{i[c]t}$  is taken to be 1 in countries where this is unnecessary.

#### 5.4 Benchmarking to UN IGME national estimates

UN IGME produce national yearly estimates of U5MR (UN IGME, 2020) using the B3 model (Alkema and New, 2014). These estimates are based on many sources of data that we cannot utilize for subnational estimation due to the unavailability of geographic information and/or a lack of reliable estimation methods available for modeling. We use a form of benchmarking to incorporate the additional national information contained in the UN IGME model. After the model in (5.1) is fit, taking posterior draws of the non-spatial terms, we can calculate posterior draws  $k = 1, \dots, 1000$  of a version of the national U5MR from our model in the following way:

$${}_{60}q_{0t}^{(k)} = 1 - \prod_{m=0}^{59} \left( 1 - \text{expit}(\mu_{a[m]}^{(k)} + \gamma_{a^*[m]t}^{(k)}) \right). \quad (5.2)$$

Taking the median of the national U5MR estimates,  ${}_{60}\widehat{q}_{0,t}$ , we calculate benchmarks,

$$\text{BENCH}_t = \frac{{}_{60}\widehat{q}_{0,t}}{\text{IGME}_t}, \quad (5.3)$$

for  $t = 1990, \dots, T^*$ , where  $T^*$  is the latest calendar year in which we have data. We then refit the model in (5.1), incorporating the benchmark into the log-offset term

$$\begin{aligned} y_{m,i[c]ts} | n_{m,i[c]ts}, q_{m,it} &\sim \text{BetaBinomial}(n_{m,i[c]ts}, q_{m,it}, d) \\ \eta_{m,i[c]ts} = \text{logit}(q_{m,i[c]t}) &= \mu_{a[m]} + \alpha_t + \gamma_{a^*[m]t} + \theta_i^* + \phi_i^* \\ &+ b_i t + \delta_{it} + \beta_s + \log(\text{HIV}_{i[c]t}) + \log(\text{BENCH}_t). \end{aligned} \quad (5.4)$$

Finally, we take draws from the joint posterior and calculate

$${}_{60}q_{0,it}^{B(k)} = 1 - \prod_{m=0}^{59} \left[ 1 - \text{expit}({}_1q_{m,it}^{B(k)}) \right] \quad (5.5)$$

$$= 1 - \prod_{m=0}^{59} \left[ 1 - \text{expit} \left( \mu_{a[m]}^{B(k)} + \gamma_{a^*[m]t}^{B(k)} + \theta_i^{B(k)} + \phi_i^{B(k)} + b_i^{B(k)}t + \delta_{it}^{B(k)} \right) \right], \quad (5.6)$$

for  $k = 1, \dots, 1000$ , where the superscript  $B$  indicate these parameters are estimated with the inclusion of a benchmarking term. Final estimates are taken to be the posterior median,  ${}_{60}\hat{q}_{0,it}^B$ , and quantiles can be calculated for intervals. Note, estimates of the national temporal shock term,  $\alpha_t$  is not included in prediction, nor are estimates of survey effects,  $\beta_s$ .

### 5.5 Country-specific crisis adjustment

In 2008, Cyclone Nargis struck the southwestern shore of Myanmar. It caused many fatalities, concentrated in the Ayeyarwady and Yangon Divisions (Tripartite Core Group, 2008). Figure 5.1 highlights the Admin-2 areas contained in these Admin-1 Divisions and displays the number of children under age 5 by Admin-2 area in the year 2008. Mortality shocks like Cyclone Nargis are hard to capture in a survey carried out after the event, as mothers who died in the event are removed from the sampling frame. The children of those mothers are more likely to have died during the shock event, leading to mortality underestimation in affected areas. The only DHS survey completed in Myanmar was carried out in 2015–2016, and the birth histories it contains do not reflect the child mortality burden of the cyclone. Our practice of estimating subnational mortality in each country separately allows us to speedily address the unique circumstances of Myanmar. To account for the mortality burden from Cyclone Nargis, we make use of national excess death estimates from the UN.

Departing slightly from the methodology described in Sections 5.3 and 5.4, we account for excess deaths after fitting the model in (5.1). First, we assume 20% of national excess under-five deaths are children under 1 year of age and the remaining 80% are children between ages 1 and 5. The national excess deaths by age group are then distributed to the Admin-2 areas

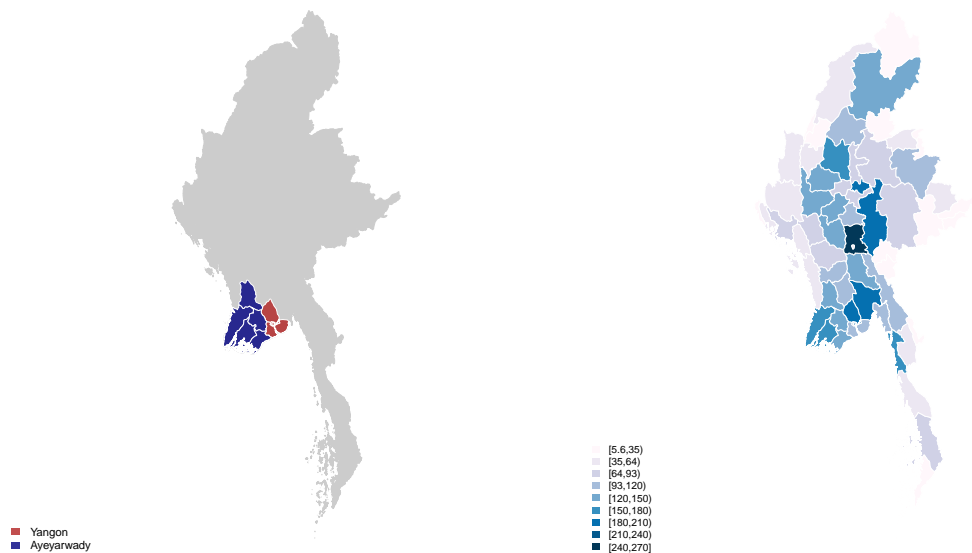


Figure 5.1: **Left:** Admin-2 areas (outlined in white) of Myanmar affected by Cyclone Nargis in the Admin-1 areas Yangon (red) and Ayeyarwady (blue). **Right:** Population under age 5 (in thousands) by Admin-2 areas in 2008.

according to the proportion of the total population by age group of Ayeyarwady and Yangon combined that resides in each Admin-2 area. Dividing excess deaths by population yields age-specific mortality rates which are converted, finally, to the probability of death in Cyclone Nargis for children under age 5,  ${}_{60}q_{0,it}^{ED}$  [CITE upon receipt from UNICEF](#). Taking posterior draws from the unbenchmarked betabinomial model, we calculate  ${}_{60}q_{0,it}^{ED,(k)} = {}_{60}q_{0,it}^{(k)} + {}_{60}q_{0,it}^{ED}$ . Note,  ${}_{60}q_{0,it}^{ED} = 0$  for all non-affected areas and years.

If we naively benchmark each area to the national estimate in 2008 from UN-IGME that has been crisis-adjusted, we would apply the excess deaths to all non-affected areas which is not correct. Instead, we calculate population-weighted national estimates of U5MR,

$${}_{60}q_{0,agg,t}^{ED,(k)} = \frac{\sum_{i=1}^I {}_{60}q_{0,it}^{ED,(k)} \times \text{POP}_{it}}{\sum_{i=1}^I \text{POP}_{it}}. \quad (5.7)$$

for years  $t = 1990, \dots, T^* = 2016$ . Benchmarks are then calculated as the ratio of the median national estimates in (5.7),  $\widehat{q}_{0agg,t}^{ED}$ , to the IGME national median estimate. We remove the benchmark  $BENCH_{2008}$  and fit a simple RW2 smoothing model to all other  $BENCH_t$  which allows us to estimate an appropriate benchmark in 2008 for areas that were not affected by the cyclone as well as those that were. Benchmarks are then projected forward in time to time  $T = 2019$ . The posterior medians,  $\overline{BENCH}_t$ , of the smoothed benchmarks in each year are then used to adjust posterior draws  $q_{0,it}^{ED,(k)}$  in each area. Summary measures of the distribution of  $q_{0,it}^{ED}$  can be calculated using quantiles.

We also make a post hoc crisis adjustment for the years of the Rwandan Civil War and the genocide of the Tutsi people during the years 1993–1999. Unlike the case of Cyclone Nargis in Myanmar, the increase in mortality due to the conflict occurred nationwide. We follow the benchmarking methodology described for Myanmar, but do not proceed with the smoothing step for  $BENCH_t$ . We attribute excess deaths across all areas of Rwanda according to the population size of children under age 5, before calculating benchmarks for each year. Initially, we performed a sensitivity analysis for Rwanda, by applying both the calculated benchmarks,  $BENCH_t$ , and posterior medians from a RW2 smoothing model,  $\overline{BENCH}_t$ . Not only did the Rwandan genocide impact the entire nation of Rwanda, it spanned multiple years with variability in conflict intensity over the period. We found that smoothing the benchmarks attenuated the benchmarks and final estimates in the earlier years of conflict known to be the most intense. As such the benchmarking issue pertaining to Cyclone Nargis are alleviated, and for Rwanda we apply the unsmoothed post hoc benchmarks  $BENCH_t$  for each year.

## 5.6 Application to 22 countries

Results for each of the 22 countries and their appropriate administrative divisions listed in Table 5.1 can be found in Appendix ???. Figure 5.2 presents subnational estimates of U5MR

for the year 2019 along with the IGME national median estimate. Countries are ordered by decreasing national U5MR. We see clearly the between country variability, the within country variability, and countries where there are a small number of subnational areas with outlying high child mortality. The majority of national and subnational areas presented here have not yet achieved the SDG 3.2 target of 25 deaths per 1000 births as of 2019.

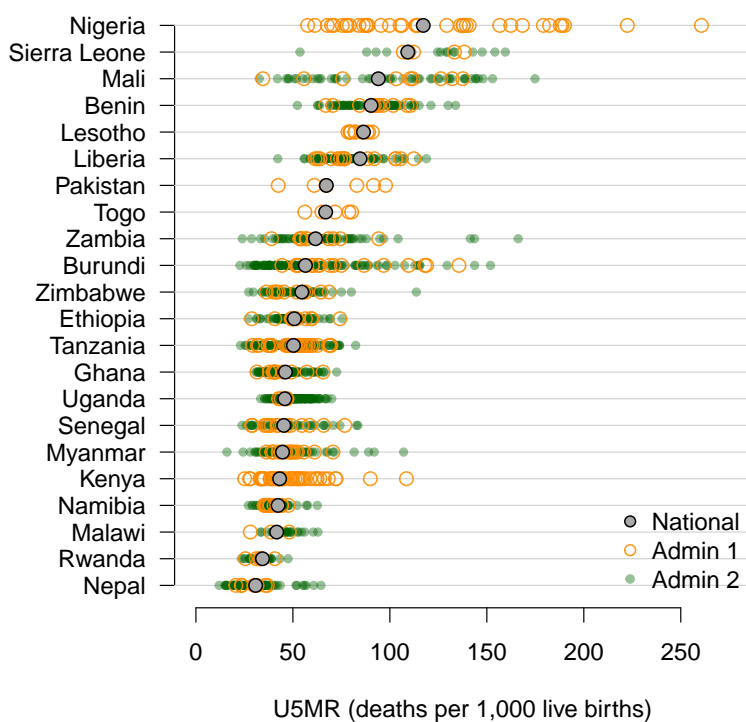


Figure 5.2: Median estimates of subnational U5MR compared to national estimates from the IGME B3 model for the year 2019.

Figure 5.3 shows the percentage decline of U5MR between the years 1990–2019 for national and subnational U5MR. The vast majority of countries and areas have seen at least some improvement in child mortality, but large between and within country variability in

mortality decline is evident. The countries of Burundi, Kenya, Namibia and Uganda, in particular, have skewed distributions of decline indicating more subnational areas have experienced less of a decline in U5MR less than that of the national improvement or those areas whose decline is larger than the national estimate.

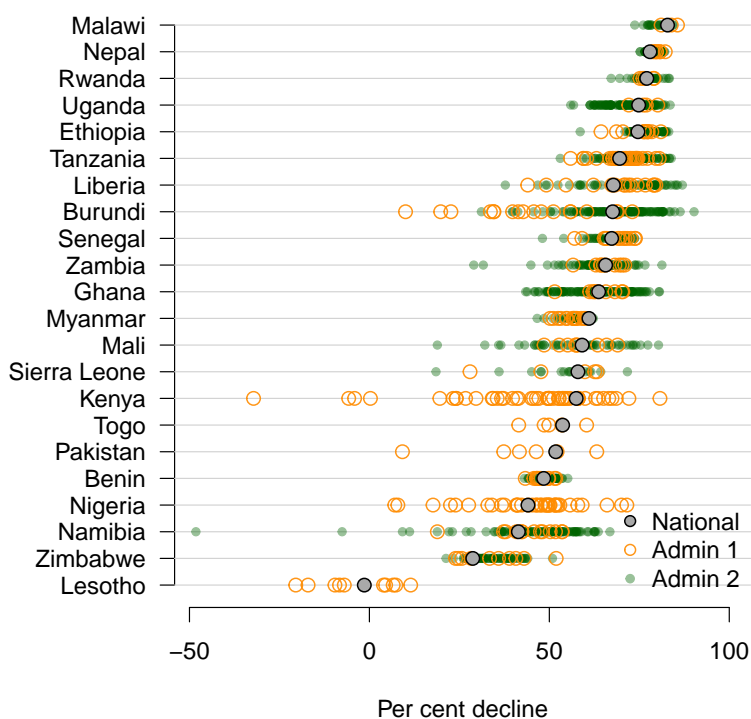


Figure 5.3: The percent decline of subnational U5MR compared to the national percent decline in U5MR as estimated from the IGME B3 model between 1990 and 2019.

### 5.6.1 Special Case: Myanmar

In this section, we will continue to focus on the special case of Myanmar and accounting for the geographically disparate increase in child mortality caused by Cyclone Nargis. Figure 5.4

compares subnational estimates of  ${}_{60}q_0$  using the method described in this thesis with those published by IHME (Burstein et al., 2019) in 2008. While the IHME estimates treat the cyclone as a national event, our estimates reflect the localized impact of Cyclone Nargis. We also find that the Admin-2 area Hkamti, Sagaing Division, in the northwest of Myanmar had relatively high U5MR during 2008, unrelated to the natural disaster. On the other hand, IHME’s subnational methodology for child mortality involves benchmarking to the national estimated of U5MR estimated in the Global Burden of Disease (Monasta et al., 2018). All subnational areas receive some inflation in estimated U5MR due to the benchmarking regardless of whether the cyclone affected that area. The affected areas of Ayeyarwady and Yangon receive too small of an adjustment. The contrast in estimates by method in Figure 5.4 visually illustrates the issue with making national adjustments for subnational crises described in Section 5.5.

Figure 5.5 compares the subnational distribution of estimated deaths by method, calculated as  ${}_{60}\widehat{q}_{0,i2008} \times \text{POP}_{i2008}$ . Whether comparing U5MR or expected number of deaths, the differences in estimates by method are quite large and the differences even vary by level of administrative division. Country-specific modeling allows us to make more realistic estimates of U5MR or numbers of deaths that reflect the geographic disparity in casualties caused by Cyclone Nargis.

Estimated temporal trends and the uncertainty of those trends by method and administrative area can be found in Appendix ???. These more clearly display the assumptions underlying a national adjustment for subnational crisis, and place the large number of excess deaths in Ayeyarwady and Yangon Divisions in context, relative to other years.

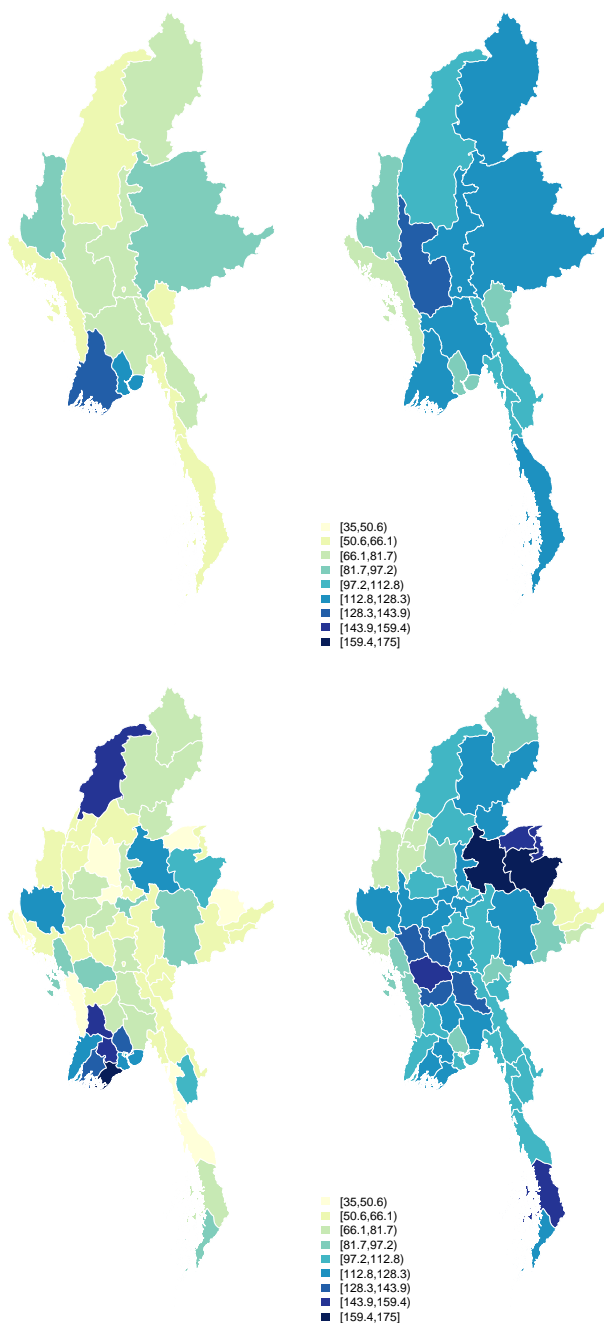


Figure 5.4: **Left:** Median estimates of  ${}_{60}q_0$  per 1000 children in 2008. **Right:** IHME estimates of  ${}_{60}q_0$  per 1000 children in 2008. **Top:** Admin-1 areas. **Bottom:** Admin-2 areas.

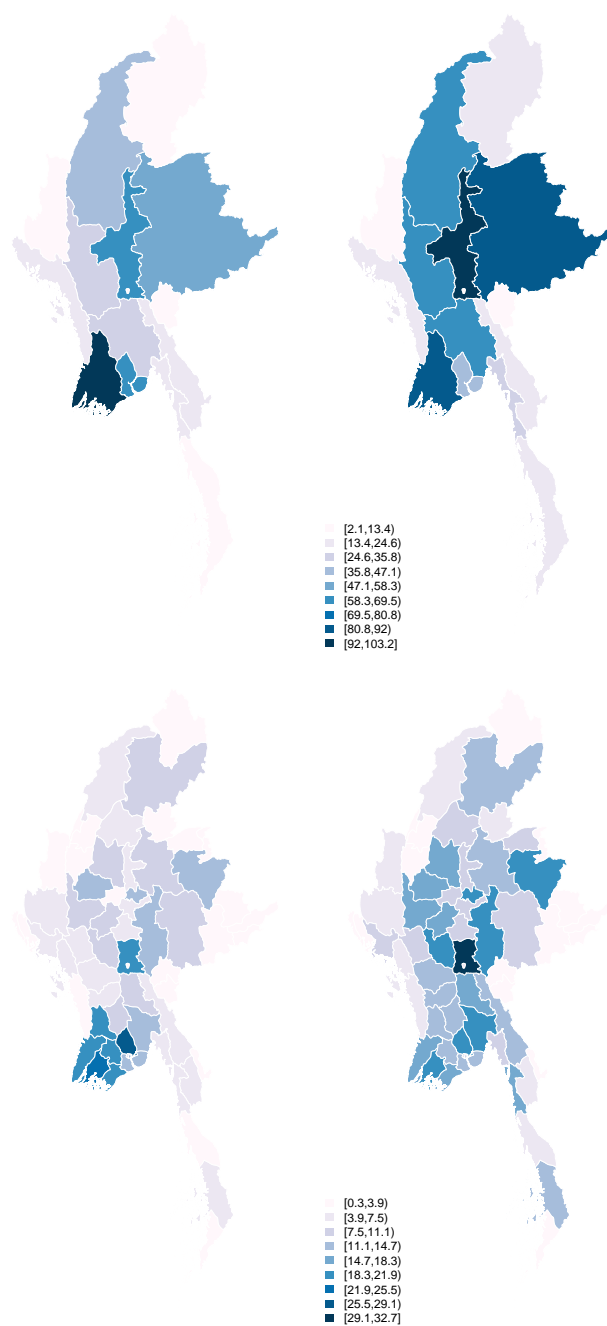


Figure 5.5: **Left:** Median estimates of number of deaths (in thousands) in 2008. **Right:** IHME estimates of number of deaths (in thousands) in 2008. **Top:** Admin-1 areas. **Bottom:** Admin-2 areas.

## 5.7 Discussion

We have developed methods for country-specific subnational estimation of child mortality from full birth histories gathered via complex household surveys in LMICs. Specifically, the methods described in this thesis allow for subnational estimation of U5MR at the Admin-2 level, even if the underlying survey data is stratified at the Admin-1 level. We applied this method to 22 countries for which the necessary data to implement this method exists. As shown in Figure 5.3 of the 22 countries in question, with the exception of Lesotho, has experienced a national decline in U5MR over the period of estimation from 1990 to 2019. Subnational estimates for each country indicate that both the level of and the decline in U5MR vary widely across administrative divisions within a country. The country-specific methodology allowed us to adjust for the impact of the generalized HIV/AIDs epidemics in 10 of the 22 countries as well as the Rwandan genocide and Cyclone Nargis in Myanmar. Where excess death estimates exist, crisis adjustments could easily be incorporated into future estimation for other events. We are continuing work to expand the number of countries for which we can estimate subnational U5MR. The implementation of our methodology via open source software using only accessible data fills a void of reproducible and transparent methods in the estimation of subnational child mortality in LMICs.

There are still many limitations and open problems in the development of better methods for child mortality estimation using complex household survey data. In particular, we have made no attempt to address age-heaping, the tendency of survey respondents to misreport the age at death of a child in months as a “typical” number (Lyons-Amos and Stones, 2017). For example, a mother whose child died at age 11 months or 13 months, may report the death as a death at 1 year of age. Additionally, we do not address issues related to migration before the time of the survey. We assume each child’s birth and death takes places in the location where the mother was interviewed. Though any mother may have previously migrated from another country, allowing this issue to affect national estimates of U5MR, internal migration

within a country is much more likely and has consequences for subnational estimation.

To more accurately reflect the stratification of the complex surveys, it would be more appropriate to make subnational estimates for each area by urban/rural classification and then aggregate these according to the proportion of the population under age 5 within each area that live in an urban or rural place of residence. Work is ongoing to correctly address the urban/rural stratification in a statistically sound manner. Our benchmarking to the IGME national estimates may alleviate the potential bias to some extent, but assumes the bias is constant across subnational areas.

Finally, there are many covariates—measured within the household surveys or otherwise—that could provide more accurate and precise estimation of child mortality in the presence of small sample sizes. Many statistical methods using covariates to better estimate an outcome with survey data require population totals of those covariates by administrative division that are rarely available for LMICs. For other covariate sources, such as raster surfaces, covariates are needed for each year of estimation as well as each location and proper aggregation from continuous to discrete space can be complex. Current work is focused on extending our methodology to incorporate covariates. Lastly, neither the re-estimation method of benchmarking nor the posthoc method are ideal. Better methods for simultaneous benchmarking and estimation that can be applied in all cases are under development.

Though we have presented estimates of under-five mortality in this chapter, the age-specific hazards allow for estimation of neonatal mortality, infant mortality, or the estimation of any of the former with differently defined age bands than those we have used.

## Chapter 6

### DISCUSSION AND FUTURE WORK

In this thesis we have provided a review of demographic and statistical methods for age-specific period child mortality, synthesized notation across fields, and developed two methods for subnational estimation of U5MR at the Admin-2 data in LMICs. In Chapter 3, we draw connections between demographic life table methods and statistical methods for discrete time survival analysis. We connect life table and discrete time survival methods to statistical notions of superpopulation or infinite population inference, and demographic Lexis diagram methods to statistical notions of finite population inference. We use these connections to compare discrete time survival methods used throughout the rest of this thesis to variants of Lexis methods used by UN IGME and DHS, with an application to FBH data from the Senegal 2006 DHS. While the nuances of and differences between Lexis and life table methods have been well studied in demographic literature, most attention has been paid to age-specific mortality estimation in populations with VR data or a close-to-fully-observed population. Comparing these methods to FBH data from a household survey allows one to see the comparison of these methods in a setting with sparser data where uncertainty of point estimates is sure to be larger. We break down in details how individual birth histories are treated differently by each method, but ultimately find the uncertainty of estimates of U5MR mean most differences in point estimates are statistically unimportant. We find that partitioning the first five years of life into any number of discrete bins that separate out the first month of life can, in some situations, lead to data for which discrete time survival analysis cannot be applied. We also find unique numeric instabilities when one fails to divide the first five years of life beyond the single interval of  $[0, 60)$  months.

We develop both design-based and model-based methods for subnational child mortality estimation in LMICs at the Admin-2 level. In Chapter 4, we extend previous design-based methods that take a finite population approach to estimation of period-specific U5MR for each area to allow of inclusion of additional child mortality information from SBH data sources. Spatiotemporal smoothing is performed on area-time estimates of child mortality arising from either an FBH or SBH data source. We apply these methods to the countries of Kenya and Malawi, and find that the inclusion of census SBH data does, indeed, improve precision for U5MR estimation in time periods where census estimates apply. However, we also observed inconsistency of SBH estimates across countries and across censuses within countries. Finally, the countries of Kenya and Malawi were chosen for the rare existence of DHS surveys whose stratification and sample size calculations were carried out at the Admin-2 level. Overall, we found mixed evidence of the appropriateness of application of this method to other countries, and saw a need to develop different methods for broader application.

In Chapter 5, we develop model-based methods for estimation of age-specific mortality at the Admin-2 level. Departing from the previous methodology, the model-based methods for survey data and spatiotemporal smoothing are performed simultaneously to model estimation. In contrast to existing model-based methods for subnational U5MR estimation, we estimate mortality in each country separately and do not allow spatial smoothing to be informed by adjacent areas from other nations. We apply the method to 22 countries via a completely reproducible pipeline. The country-specific modeling and steps in the pipeline, which include an array of exploratory analyses, allow us to address the unique context of child mortality in each country if the generic base model is inappropriate, such as in countries with generalized HIV/AIDs epidemics, the genocide events in Rwanda, and Cyclone Nargis in Myanmar. Despite the accommodation of country-specific context, the pipeline used to create the estimates is designed so that departures from the boiler plate method

are as minimal as possible and are also archived for replicability. There is ongoing work to improve the model-based methods introduced in this chapter to more appropriately account for urban/rural stratification, and to develop methods that allow for crisis adjustment and benchmarking to be performed simultaneously in model estimation.

In this thesis, we have addressed many issues in age-specific estimation of period child mortality in LMICs. However, there are many avenues of future research that remain open. Simulation studies can help us understand better the consequences of the choice of method, the choice of the unit of time measurement, and the choice of partition for the age interval  $[0, 60)$  when using FBH data. A more complex simulation study could also help further understanding of the consequences of these choices in a household survey and subnational context. There are many other known issues with FBH data this thesis does not address such as age-heaping, consequences of migration on subnational estimation when the current location of the interviewed mother is not the location of birth or death of one of their children, and censoring and selectivity issues arising from the retrospective nature of FBH data and the changes in cohort mortality and fertility across the broad 15–49 age group of interviewed women. Finally, a review and exploration of measures of model comparison and validation would be valuable. Though used Chapter 4, leave-one-out validation schemes are not obviously the correct choice for comparing U5MR estimates from survey data over time and space. As there is no way of knowing the true underlying U5MR in any of the areas for which estimates are made, comparisons are based on “truths” which are estimates and may have large uncertainty. Additionally, the spatiotemporal dependence of estimates and complex survey sampling schemes mean a standard leave-one-out method assuming independent sampling is inappropriate.

## BIBLIOGRAPHY

- Alkema, L. and New, J. (2014). Global estimation of child mortality using a Bayesian b-spline bias-reduction model. *The Annals of Applied Statistics*, pages 2122–2149.
- Allison, P. D. (1982). Discrete-time methods for the analysis of event histories. *Sociological methodology*, 13:61–98.
- Allison, P. D. (2014). *Event History and Survival Analysis, Second Edition*. SAGE publications.
- Battese, G. E., Harter, R. M., and Fuller, W. A. (1988). An error-components model for prediction of county crop areas using survey and satellite data. *Journal of the American Statistical Association*, 83:28–36.
- Besag, J., York, J., and Mollie, A. (1991). Bayesian image restoration with two applications in spatial statistics (with discussion). *Ann Inst Stat Math.*, 43:1–59.
- Binder, D. (1983). On the variances of asymptotically normal estimators from complex surveys. *International Statistical Review*, 51:279–292.
- Brady, E. and Hill, K. (2017). Testing survey-based methods for rapid monitoring of child mortality, with implications for summary birth history data. *PloS One*, 12:e0176366.
- Brass, W. (1964). Uses of census or survey data for the estimation of vital rates. Document No. E/CN.14/CAS.4/V57., United Nations Economic and Social Council, Addis Ababa: African Seminar on Vital Statistics. Available at <https://repository.uneca.org/ds2/stream/?#/documents/2ed26999-ee01-5bc9-b119-2167e6eaa246/page/1>.

- Brass, W. and Coale, A. (1968). Methods of analysis and estimation. In Brass, W., Coale, A., Demeny, P., Heisel, D., Romaniuk, A., and van de Walle, E., editors, *The Demography of Tropical Africa*. Princeton University Press, Princeton, NJ.
- Burgert, C. R., Colston, J., Roy, T., and Zachary, B. (2013). Geographic displacement procedure and georeferenced data release policy for the demographic and health surveys. DHS Spatial Analysis Reports 7, USAID and ICF International, Calverton, Maryland. Available at <https://dhsprogram.com/pubs/pdf/SAR7/SAR7.pdf>.
- Burstein, R., Henry, N. J., Collison, M. L., Marczak, L. B., Sligar, A., Watson, S., Marquez, N., Abbasalizad-Farhangi, M., Abbasi, M., Abd-Allah, F., et al. (2019). Mapping 123 million neonatal, infant and child deaths between 2000 and 2017. *Nature*, 574(7778):353–358.
- Burstein, R., Wang, H., Reiner Jr, R. C., and Hay, S. I. (2018). Development and validation of a new method for indirect estimation of neonatal, infant, and child mortality trends using summary birth histories. *PLoS Medicine*, 15:e1002687.
- Carstensen, B. (2007). Age–period–cohort models for the Lexis diagram. *Statistics in Medicine*, 26:3018–3045.
- Clark, S. J. (2019). A general age-specific mortality model with an example indexed by child mortality or both child and adult mortality. *Demography*, 56(3):1131–1159.
- Coale, A. and Demeny, P. (1966). Regional model life tables and stable populations.
- Coale, A. J. and Trussell, J. (1977). Annex I: estimating the time to which Brass estimates apply. *Population Bulletin of the United Nations*, pages 87–89.
- Database of Global Administrative Areas (GADM) (2019). Global administrative areas [shapefiles]. Available at <https://www.gadm.org>. Downloaded September 2019.

- Elkasabi, M. (2019). Calculating fertility and childhood mortality rates from survey data using the dhs rates r package. *PLoS One*, 14.
- Estimation, I. G. M. and Children's Fund (2020). Levels and Trends in Child Mortality . Report 2020, United Nations. Available at <https://childmortality.org/wp-content/uploads/2020/09/UNICEF-2020-Child-Mortality-Report.pdf>.
- Fay, R. E. and Herriot, R. A. (1979). Estimates of income for small places: an application of james-stein procedures to census data. *Journal of the American Statistical Association*, 74:269–277.
- Fuglstad, G.-A., Simpson, D., Lindgren, F., and Rue, H. (2019). Constructing priors that penalize the complexity of Gaussian random fields. *Journal of the American Statistical Association*, 114:445–452.
- General Assembly and Department of Economic & Social Affairs (2015). The Millennium Development Goals Report 2015. Report ST/DPI(058)/M646/2015, United Nations. Available at <https://digitallibrary.un.org/record/796052?ln=en>.
- General Assembly and Division for Sustainable Development Goals, D. E. S. A. (2015). Transforming Our World: the 2030 Agenda for Sustainable Development. General Assembly Document ST(02)/T77, United Nations. Available at <https://digitallibrary.un.org/record/1654217?ln=en>.
- General Assembly and Secretary-General (2006). Scaling up HIV prevention, treatment, care and support. General Assembly Document A/60/737, United Nations. Available at <https://digitallibrary.un.org/record/573075?ln=e>.
- Godwin, J. and Wakefield, J. (2021). Space-time modeling of child mortality at the admin-2 level in a low and middle income countries context. *Statistics in Medicine*, 40(7):1593–

- Golding, N., Burstein, R., Longbottom, J., Browne, A. J., Fullman, N., Osgood-Zimmerman, A., Earl, L., Bhatt, S., Cameron, E., Casey, D. C., et al. (2017). Mapping under-5 and neonatal mortality in africa, 2000–15: a baseline analysis for the sustainable development goals. *The Lancet*, 390(10108):2171–2182.
- Ha, N. S., Lahiri, P., and Parsons, V. (2014). Methods and results for small area estimation using smoking data from the 2008 national health interview survey. *Statistics in Medicine*, 33:3932–3945.
- Hájek, J. (1971). Discussion of an essay on the logical foundations of survey sampling, part i, by d. basu. *Foundations of statistical inference*, page 326.
- Hill, K. (2013a). Direct estimation of child mortality from birth histories. In Moultrie, T. A., Dorrington, R., Hill, A. G., Hill, K., Timæus, I., and Zaba, B., editors, *Tools for demographic estimation*. International Union for the Scientific Study of Population.
- Hill, K. (2013b). Indirect estimation of child mortality. In Moultrie, T. A., Dorrington, R., Hill, A. G., Hill, K., Timæus, I., and Zaba, B., editors, *Tools for demographic estimation*. International Union for the Scientific Study of Population.
- Hill, K., Brady, E., Zimmerman, L., Montana, L., Silva, R., and Amouzou, A. (2015). Monitoring change in child mortality through household surveys. *PLoS One*, 10.
- Hill, K. and Trussell, J. (1977). Further developments in indirect mortality estimation. *Population Studies*, 31:313–334.
- Hill, K., You, D., Inoue, M., and Oestergaard, M. Z. (2012). Child mortality estimation: accelerated progress in reducing global child mortality, 1990–2010. *PLoS Medicine*, 9.
- Hill, K., Zlotnik, H., and Trussell, J. (1983). *Demographic Estimation: A Manual on Indirect Techniques. Manual X*.

- Horvitz, D. G. and Thompson, D. J. (1952). A generalization of sampling without replacement from a finite universe. *Journal of the American statistical Association*, 47(260):663–685.
- ICF International (1990–2021a). Demographic and Health Surveys [Datasets]. Funded by USAID. Available at <https://dhsprogram.com>.
- ICF International (1990–2021b). The DHS Program Spatial Data Repository [Datasets]. Funded by USAID. Available at <https://spatialdata.dhsprogram.com/home/>.
- ICF International (2021c). The DHS Program website: Home. <http://www.dhsprogram.com>. [Accessed June 20, 2021].
- ICF International (2021d). The DHS Program website: Survey Process. <https://www.dhsprogram.com/Methodology/Survey-Process.cfm>. [Accessed June 20, 2021].
- Inter-agency Group for Child Mortality Estimation (2020). Levels and Trends in Child Mortality: Estimates. Available at <https://childmortality.org>.
- Knorr-Held, L. (2000). Bayesian modelling of inseparable space-time variation in disease risk. *Statistics in Medicine*, 19:2555–2567.
- Lawn, J. E., Cousens, S., Zupan, J., Team, L. N. S. S., et al. (2005). 4 million neonatal deaths: When? where? why? *The Lancet*, 365(9462):891–900.
- Li, Z., Hsiao, Y., Godwin, J., Martin, B. D., Wakefield, J., Clark, S. J., with support from the United Nations Inter-agency Group for Child Mortality Estimation, and its Technical Advisory Group (2019). Changes in the spatial distribution of the under-five mortality rate: Small-area analysis of 122 dhs surveys in 262 subregions of 35 countries in africa. *PloS One*, 14(1).

- Li, Z. R., Martin, B. D., Dong, T. Q., Fuglstad, G.-A., Godwin, J., Paige, J., Riebler, A., Clark, S., and Wakefield, J. (2020). Space-time smoothing of demographic and health indicators using the R package SUMMER. *arXiv preprint arXiv:2007.05117*.
- Lindgren, F., Rue, H., and Lindström, J. (2011). An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic differential equation approach (with discussion). *Journal of the Royal Statistical Society, Series B*, 73:423–498.
- Lohr, S. (2010). *Sampling: Design and Analysis, Second Edition*. Brooks/Cole Cengage Learning, Boston.
- Lumley, T. (2004). Analysis of complex survey samples. *Journal of Statistical Software*, 9:1–19.
- Lumley, T. (2010). *Complex Surveys: A Guide to Analysis using R*. John Wiley and Sons, Hoboken, Jersey.
- Lyons-Amos, M. and Stones, T. (2017). Trends in demographic and health survey data quality: an analysis of age heaping over time in 34 countries in sub saharan africa between 1987 and 2015. *BMC Research Notes*, 10(1):1–7.
- Marquez, N. and Wakefield, J. (2020). Harmonizing child mortality data at disparate geographic levels. *arXiv preprint arXiv:2002.00089*.
- Martins, T., Simpson, D., Lindgren, F., and Rue, H. (2013). Bayesian computing with INLA: new features. *Computational Statistics and Data Analysis*, 67:68–83.
- Mathers, C. D., Ma Fat, D., Inoue, M., Rao, C., and Lopez, A. D. (2005). Counting the dead and what they died from: an assessment of the global status of cause of death data. *Bulletin of the World Health Organization*, 83:171–177c.

- McCaa, R. (2013). The big census data revolution: IPUMS-International. trans-border access to decades of census samples for three-fourths of the world and more. *Revista de demografia historica*, 30(1):69.
- MEASURE DHS (2012). Demographic and Health Survey Sampling and Household Listing Manual. Demographic and Health Surveys Methodology, USAID and ICF International, Calverton, Maryland. Available at [https://dhsprogram.com/pubs/pdf/DHSM4/DHS6\\_Sampling\\_Manual\\_Sept2012\\_DHSM4.pdf](https://dhsprogram.com/pubs/pdf/DHSM4/DHS6_Sampling_Manual_Sept2012_DHSM4.pdf).
- Mercer, L., Wakefield, J., Pantazis, A., Lutambi, A., Masanja, H., and Clark, S. (2015). Space–time smoothing of complex survey data: Small area estimation for child mortality. *The Annals of Applied Statistics*, 9:1889–1905.
- Minnesota Population Center (2019). Integrated public use microdata series, international: Version 7.2 [datasets]. Available at <https://doi.org/10.18128/D020.V7.2>.
- Minnesota Population Center (2021). The IPUMS International website: Home. <https://international.ipums.org/international/index.shtml>. [Accessed June 20, 2021].
- Monasta, L., Ronfani, L., Gallus, S., Beghi, E., Giussani, G., Bosetti, C., Cortinovis, M., Bikbov, B., Perico, N., Remuzzi, G., et al. (2018). Global, regional, and national age-sex-specific mortality and life expectancy, 1950–2017. *The Lancet*, 392(10159):1684–1735.
- Paige, J., Fuglstad, G.-A., Riebler, A., and Wakefield, J. (2020). Design- and Model-Based Approaches to Small-Area Estimation in A Low- and Middle-Income Country Context: Comparisons and Recommendations. *Journal of Survey Statistics and Methodology*.
- Pedersen, J. and Liu, J. (2012). Child mortality estimation: Appropriate time periods for child mortality estimates from full birth histories. *PLoS Medicine*, 9:1001289.

- Pezzulo, C., Bird, T., Utazi, E., Sorichetta, A., Tatem, A., Yourkavitch, J., and Burgert-Brucker, C. (2016). Geospatial modeling of child mortality across 27 countries in sub-saharan africa. DHS Spatial Analysis Reports 13, USAID and ICF International, Rockville, Maryland. Available at <https://www.dhsprogram.com/pubs/pdf/SAR13/SAR13.pdf>.
- Population Division, D. E. S. A. (2011). Mortality estimates from major sample surveys: towards the design of a database for the monitoring of mortality levels and trends. Technical Report 2011/2, United Nations. Available at [https://www.un.org/development/desa/pd/sites/www.un.org.development.desa.pd/files/files/documents/2020/Jan/un\\_2011\\_techpaper2.pdf](https://www.un.org/development/desa/pd/sites/www.un.org.development.desa.pd/files/files/documents/2020/Jan/un_2011_techpaper2.pdf).
- Preston, S., Heuveline, P., and Guillot, M. (2001). *Demography: Measuring and Modeling Population Processes*. Oxford: Blackwell.
- Rajaratnam, J. K., Tran, L. N., Lopez, A. D., and Murray, C. J. (2010). Measuring under-five mortality: validation of new low-cost methods. *PLoS Medicine*, 7:e1000253.
- Riebler, A., Sørbye, S. H., Simpson, D., and Rue, H. (2016). An intuitive Bayesian spatial model for disease mapping that accounts for scaling. *Statistical Methods in Medical Research*, 25:1145–1165.
- Rue, H. and Held, L. (2005). *Gaussian Markov Random Fields: Theory and Applications*. CRC press.
- Rue, H., Martino, S., and Chopin, N. (2009). Approximate Bayesian inference for latent gaussian models by using integrated nested laplace approximations. *Journal of the Royal Statistical Society: Series B*, 71:319–392.
- Ruggles, S. (2014). Big microdata for population research. *Demography*, 51(1):287–297.

- Ruggles, S., McCaa, R., Sobek, M., and Cleveland, L. (2015). The IPUMS collaboration: integrating and disseminating the worlds population microdata. *Journal of Demographic Economics*, 81(2):203–216.
- Rutstein, S. O., Rojas, G., et al. (2006). Guide to dhs statistics. Technical report, Calverton, MD: ORC Macro. Available at [https://www.dhsprogram.com/pubs/pdf/DHSG1/Guide\\_to\\_DHS\\_Statistics\\_29Oct2012\\_DHSG1.pdf](https://www.dhsprogram.com/pubs/pdf/DHSG1/Guide_to_DHS_Statistics_29Oct2012_DHSG1.pdf).
- Simpson, D., Rue, H., Riebler, A., Martins, T., Sørbye, S., et al. (2017). Penalising model component complexity: A principled, practical approach to constructing priors. *Statistical Science*, 32:1–28.
- Stevens, F. R., Gaughan, A. E., Linard, C., and Tatem, A. J. (2015). Disaggregating census data for population mapping using random forests with remotely-sensed and ancillary data. *PloS one*, 10(2):e0107042.
- Stover, J., Brown, T., and Marston, M. (2012). Updates to the Spectrum/Estimation and Projection Package (EPP) model to estimate HIV trends for adults and children. *Sexually Transmitted Infections*, 88.
- Sugasawa, S., Kubokawa, T., and Rao, J. (2018). Small area estimation via unmatched sampling and linking models. *Test*, 27:407–427.
- Tatem, A. J. (2017). Worldpop, open data for spatial demography. *Scientific data*, 4(1):1–4.
- Tripartite Core Group (2008). Post-nargis. Periodic Review I. Available at <https://reliefweb.int/report/myanmar/myanmar-post-nargis-periodic-review-i>.
- Trussell, T. J. (1975). A re-estimation of the multiplying factors for the Brass technique for determining childhood survivorship rates. *Population Studies*, 29:97–107.
- UNAIDS (2014). *Quick Start Guide for Spectrum (Downloaded December 2015)*. UNAIDS.

- UNICEF (2015). *Monitoring the situation of children and women for 20 years: The Multiple Indicator Cluster Surveys (MICS) 1995–2015*. United Nations. Available at <https://mics.unicef.org/publications/reports-and-methodological-papers>.
- UNICEF (2021). The MICS website: About MICS. <https://mics.unicef.org/about>. [Accessed June 20, 2021].
- UNICEF–Statistics and Monitoring (2000–2014). Multiple Indicator Cluster Surveys (MICS) [Datasets]. Available at <https://mics.unicef.org/surveys>.
- Utazi, C., Thorley, J., Alegana, V., Ferrari, M., Nilsen, K., Takahashi, S., Metcalf, C., Lessler, J., and Tatem, A. (2019). A spatial regression model for the disaggregation of areal unit based data to high-resolution grids with application to vaccination coverage mapping. *Statistical methods in medical research*, 28(10-11):3226–3241.
- Wakefield, J., Fuglstad, G.-A., Riebler, A., Godwin, J., Wilson, K., and Clark, S. J. (2019). Estimating under-five mortality in space and time in a developing world context. *Statistical Methods in Medical Research*, 28:2614–2634.
- Wakefield, J., Okonek, T., and Pedersen, J. (2020). Small area estimation for disease prevalence mapping. *International Statistical Review*, 88(2):398–418.
- Walker, N., Hill, K., and Zhao, F. (2012). Child mortality estimation: methods used to adjust for bias due to aids in estimating trends in under-five mortality. *PLoS Medicine*, 9.
- Wilmoth, J., Zureick, S., Canudas-Romo, V., Inoue, M., and Sawyer, C. (2012). A flexible two-dimensional mortality model for use in indirect estimation. *Population Studies*, 66(1):1–28.
- Wilson, K. and Wakefield, J. (2018). Pointless spatial modeling. *Biostatistics*.

- Wilson, K. and Wakefield, J. (2020). Child mortality estimation incorporating summary birth history data. *Biometrics*.
- Wolter, K. (2007). *Introduction to Variance Estimation, Second Edition*. Springer Science and Business Media.
- Woodruff, R. S. (1971). A simple method for approximating the variance of a complicated estimate. *Journal of the American Statistical Association*, 66(334):411–414.
- World Health Organization, UNAIDS, and UNICEF (2010). Towards universal access: scaling up priority HIV/AIDS interventions in the health sector. Progress Report 2010, World Health Organization. Available at [http://apps.who.int/iris/bitstream/handle/10665/44443/9789241500395\\_eng.pdf?sequence=1](http://apps.who.int/iris/bitstream/handle/10665/44443/9789241500395_eng.pdf?sequence=1).

## Appendix A

## APPENDIX TO CHAPTER 4

**A.1 Zero Adjustment details**

A full approximation for the probability of death given residence in area  $A_i$  and given the data from  $A_i$ 's set of first-order neighbors,  $A^*$ :

$$\begin{aligned}
\Pr(\text{death}|A_i, A^*) &= \sum_{h=1}^H \sum_{c=1}^{n_h} \Pr(\text{death} \cap \text{strata } h \cap \text{cluster } c|A_i, A^*) \\
&= \sum_{h=1}^H \sum_{c=1}^{n_h} \Pr(\text{death}|\text{strata } h, \text{cluster } c, A_i, A^*) \times \\
&\quad \Pr(\text{strata } h|A_i) \times \Pr(\text{cluster } c|\text{strata } h, A_i) \\
&= \sum_{h=1}^H \sum_{c=1}^{n_h} \Pr(\text{death}|\text{strata } h, \text{cluster } c, A_i, A^*) \times \frac{N_{ih}}{N_i} \times \frac{N_{ihc}}{N_{ih}} \\
&= \sum_{h=1}^H \frac{N_{ih}}{N_i} \sum_{c=1}^{n_h} \frac{N_{ihc}}{N_{ih}} \Pr(\text{death}|\text{strata } h, \text{cluster } c, A_i, A^*) \\
&\approx \sum_{h=1}^H \frac{N_{ih}}{N_i} \sum_{c=1}^{n_h} \frac{\hat{N}_{ihc}}{\hat{N}_{ih}} \Pr(\text{death}|\text{strata } h, \text{cluster } c, A_i, A^*)
\end{aligned}$$

After fitting the model, we can draw samples from the approximated posterior of  $(\beta_h^{(j)}, \varepsilon_i^{(j)}, \varepsilon_c^{(j)})$ . From these, we get  ${}_{60}q_0^{(j)}$  and as follows to estimate posterior mean and variance,  $(\hat{y}, \hat{V})$ .

$$\begin{aligned}
\Pr(\text{death}|A_i, A^*) &\approx \sum_{h=1}^H \frac{N_{ih}}{N_i} \sum_{c=1}^{n_h} \frac{\hat{N}_{ihc}}{\hat{N}_{ih}} \Pr(\text{death}|\text{strata } h, \text{cluster } c, A_i, A^*) \\
\Rightarrow {}_{60}q_0^{(j)} &\approx \sum_{h=1}^H \frac{N_{ih}}{N_i} \sum_{c=1}^{n_h} \frac{\hat{N}_{ihc}}{\hat{N}_{ih}} {}_{60}q_{0hc}^{(j)} \\
&= \sum_{h=1}^H \frac{N_{ih}}{N_i} \sum_{c=1}^{n_h} \frac{\hat{N}_{ihc}}{\hat{N}_{ih}} \text{expit}(\beta_h^{(j)} + \varepsilon_i^{(j)} + \varepsilon_c^{(j)})
\end{aligned}$$

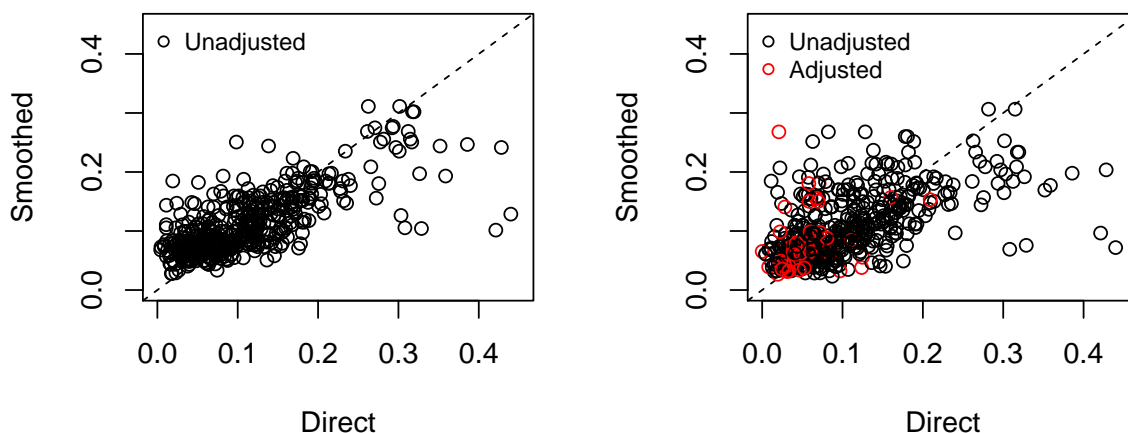


Figure A.1: Comparison of direct and smooth estimates when we throw out observed zeroes (Left) or adjust them (Right).

In Figure A.1 we see the results of fitting our space-time smoothing model in two settings: dropping the areas and time periods with zero estimates of U5MR or its variance or adjusting those areas and time periods prior to smoothing. The black dots represent survey estimates from areas and time periods that did not need to be adjusted. The red dots represent survey estimates from areas and time periods we need to predict due to small sample sizes.

On the left, we remove all observed zeroes from our data, thus there are no red dots in the plot. On the right, we see the results from our adjustment method. These plots show that the adjusted estimates and their smoothed counterparts do not lie outside the cloud of unadjusted estimates and their smoothed counterparts, indicating the adjustment provides reasonable values.

## A.2 *Bias in time*

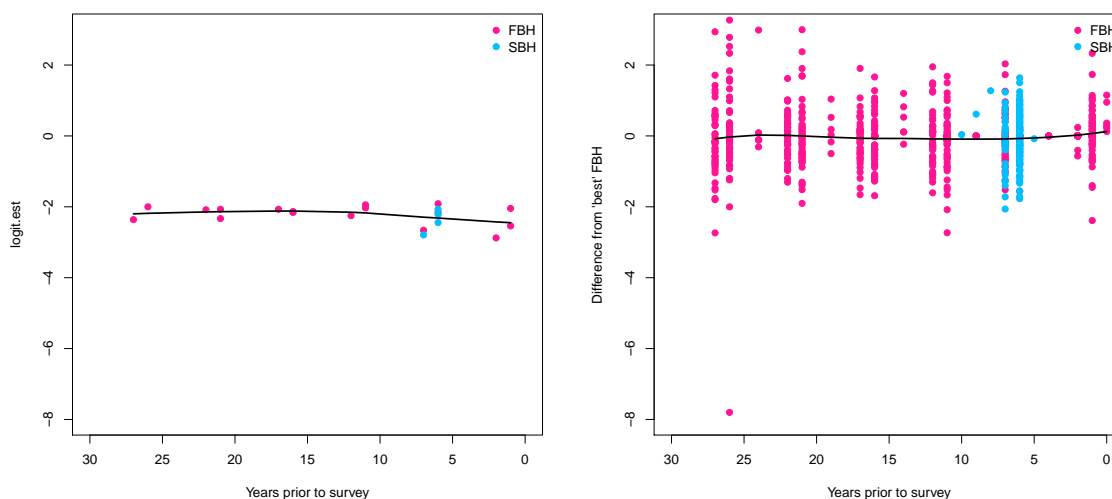


Figure A.2: Bias on the logit scale for Kenya for each estimate plotted against the years prior to data collection for which the estimate applies. Left: national estimates. Right: Admin-2 level estimates. Black line is a loess smooth applied to the biases in either panel.

## A.3 *Hyperparameters and Model parameters*

Parameter	Median	Proportion	
Kenya	ICAR	0.109	38.02
	IID Area	0.003	1.22
	RW2	0.108	37.78
	IID Time	0.001	0.46
	Interaction	0.064	22.52
Malawi	ICAR	0.022	8.26
	IID Area	0.000	0.03
	RW2	0.228	86.44
	IID Time	0.000	0.16
	Interaction	0.013	5.10

Table A.1: Proportion of variation as determined by marginal median variances for different random components.

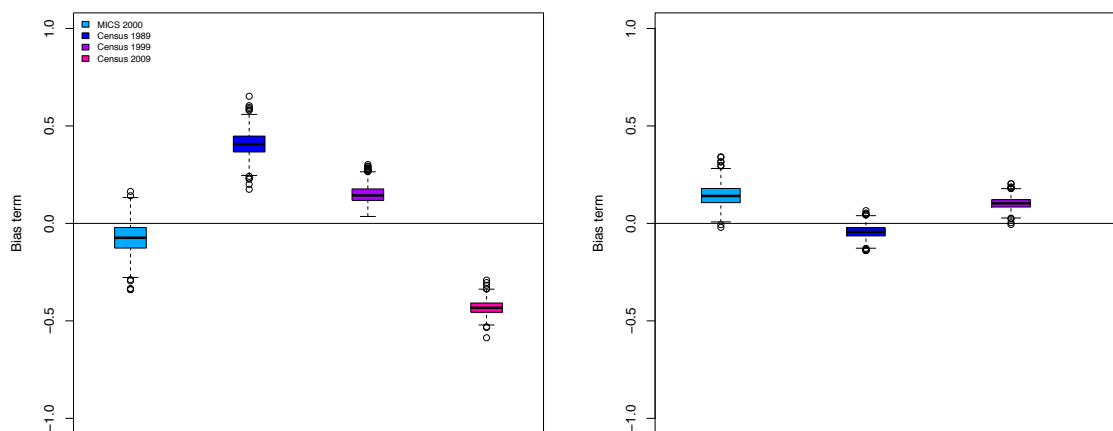


Figure A.3: Marginal posterior distribution of the bias term ( $\beta_s$ ) for each SBH data source in Kenya (Left) and Malawi (Right).

#### **A.4 National results**

Results of temporal smoothing model below for national direct and Brass estimates. Note: the Kenya MICS 2011 and 2013 provide estimates for only a small number of Admin-2 areas and, thus, are not included in the national analysis for Kenya. Figures A.4 and A.5 show period and yearly results with corresponding 95% intervals. In grey solid lines and dashed lines respectively are UN and IHME estimates for comparison. The national model below does not have any spatial components, but is otherwise analogous to the spatiotemporal model.

$$\begin{aligned} \text{logit}(\hat{q}_0^{ts}) | \eta_{ts} &\sim N(\eta_{ts}, \kappa_s^{-1} \times \hat{V}^{ts}), \\ \eta_{ts} &= \mu + \sum_s \beta_s \mathbf{1}(s \text{ is SBH}) + \alpha_t + \gamma_t \end{aligned} \tag{A.1}$$

#### **A.5 Period Results by Area: Kenya**

#### **A.6 Yearly Results by Area: Kenya**

#### **A.7 Period Results by Area: Malawi**

#### **A.8 Yearly Results by Area: Malawi**

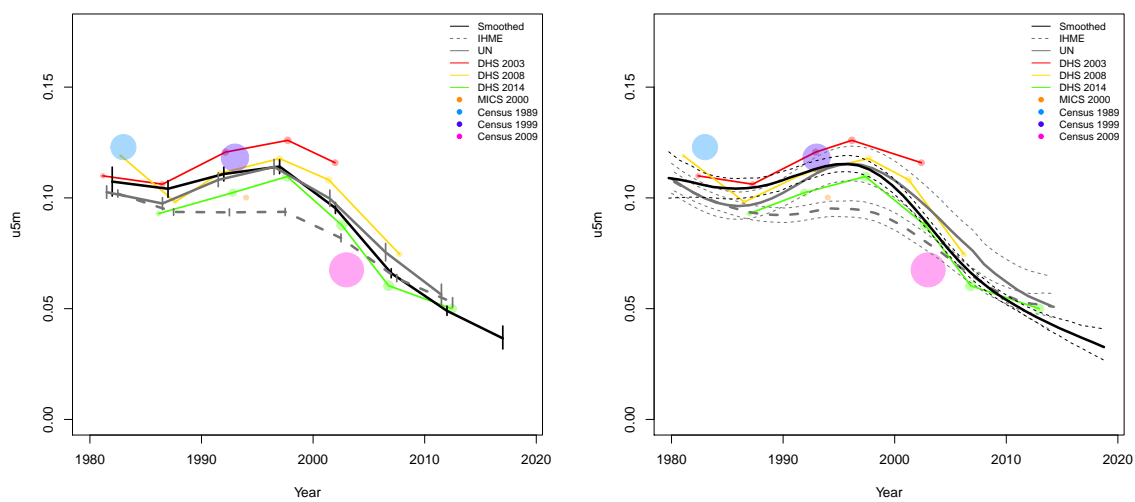


Figure A.4: National results for Kenya. Left: Five-year period smoothed results in black with larger data points for estimates with larger precision in color. Right: Yearly smoothed results in black.

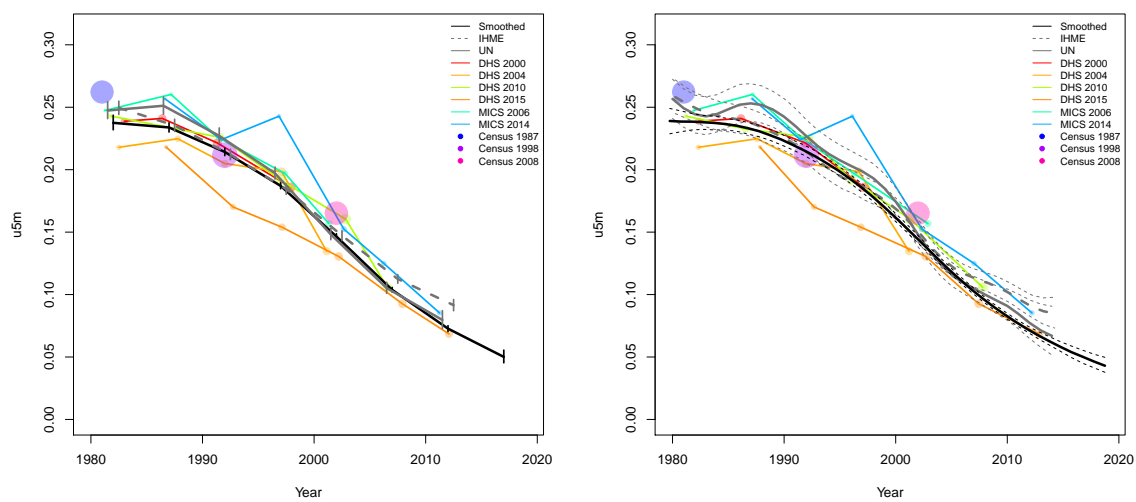


Figure A.5: National results for Malawi. Left: Five-year period smoothed results in black with larger data points for estimates with larger precision in color. Right: Yearly smoothed results in black.

Appendix B  
APPENDIX TO CHAPTER 5

*B.1 Benin*

Age	Survey	Clusters			Deaths			Agemonths		
		Urban	Rural	Total	Urban	Rural	Total	Urban	Rural	Total
0	1996	92	108	200	75	217	292	1789	5063	6852
	2001	118	129	247	125	390	515	3435	8528	11963
	2017	243	297	540	526	1018	1544	16857	26923	43780
1-11	1996	92	108	200	77	314	391	18250	51204	69454
	2001	118	129	247	143	492	635	35135	85796	120931
	2017	243	297	540	378	909	1287	172132	271937	444069
12-23	1996	92	108	200	40	142	182	18717	51818	70535
	2001	118	129	247	73	209	282	35865	86490	122355
	2017	243	297	540	203	444	647	172619	270749	443368
24-35	1996	92	108	200	26	82	108	18324	50193	68517
	2001	118	129	247	63	144	207	33996	81971	115967
	2017	243	297	540	172	340	512	159563	250872	410435
36-47	1996	92	108	200	22	99	121	17500	48150	65650
	2001	118	129	247	33	133	166	32458	76857	109315
	2017	243	297	540	108	270	378	147062	231979	379041
48-59	1996	92	108	200	11	49	60	16572	46106	62678
	2001	118	129	247	18	69	87	30866	72305	103171
	2017	243	297	540	60	116	176	135317	212964	348281

Table B.1: **Data summary for Benin.** Total numbers of clusters (Columns 3–5) with observations in each age group by survey in urban and rural areas and combined. Numbers of deaths (Columns 6–8) and number of agemonths (Columns 9–10) observed in each age group by survey in urban and rural areas and combined.

*B.1.1 Admin-1*

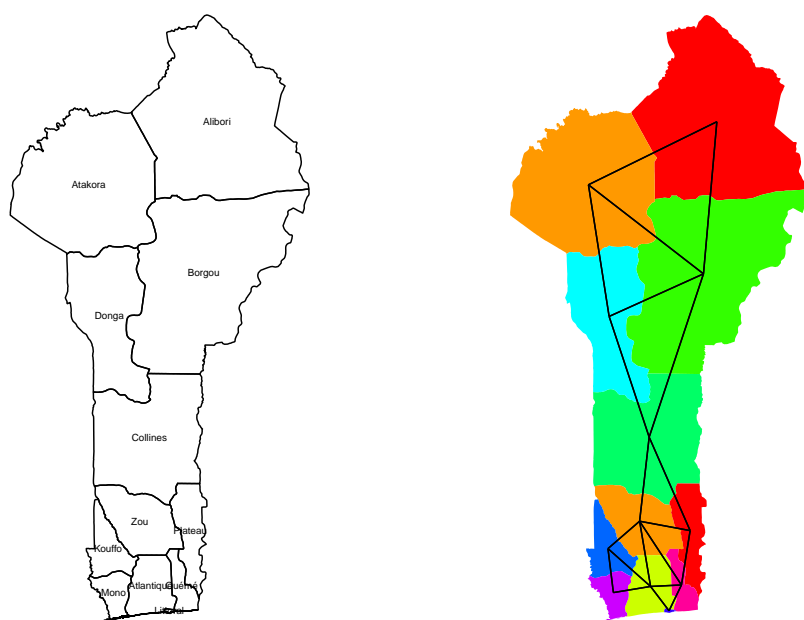


Figure B.1: **Left:** The names of the 12 Admin-1 areas of Benin . **Right:** The neighborhood structure of Admin-1 areas in Benin .

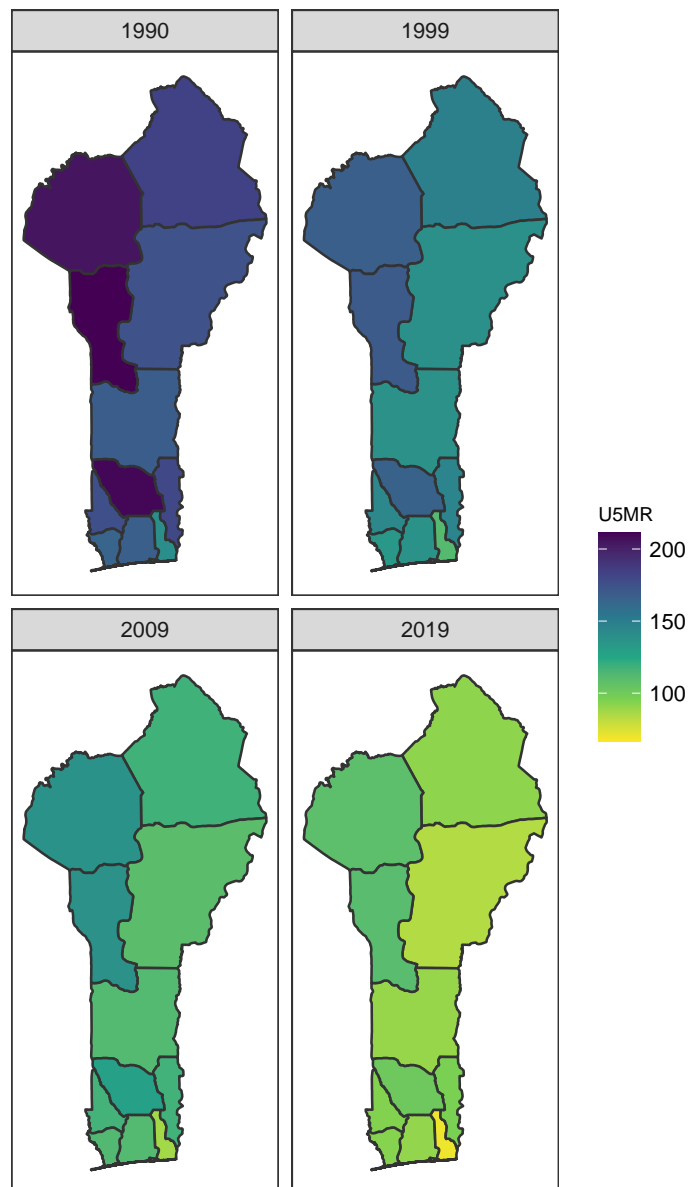


Figure B.2: Median U5MR estimates for years 1990, 1999, 2009, 2019 for Admin-1 areas in Benin .

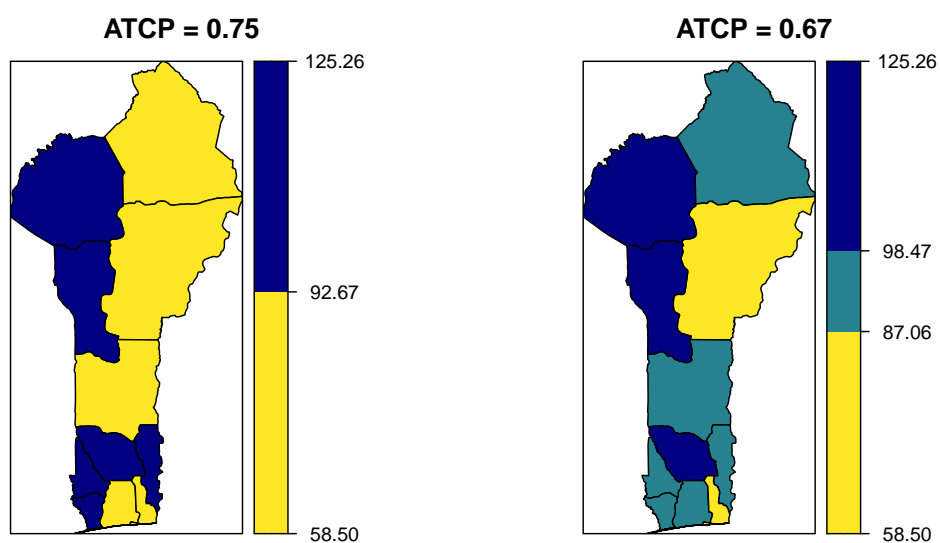
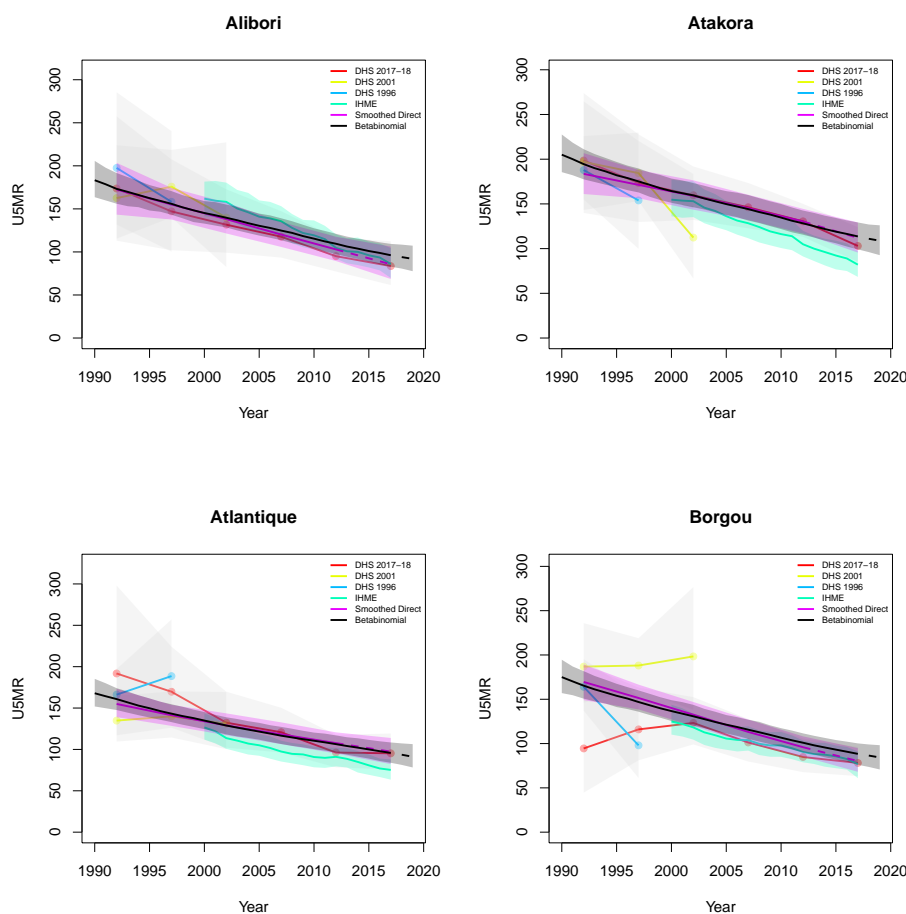


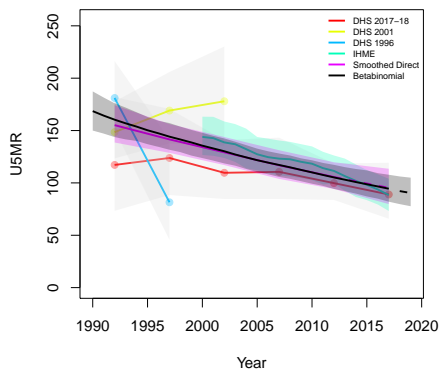
Figure B.3: Expression of uncertainty of U5MR (deaths per 1000 children) estimates for Admin-1 areas based on the average true classification probability (ATCP) in 2019 using  $K = 2, 3$  colors.

*Data and estimates over time by area*

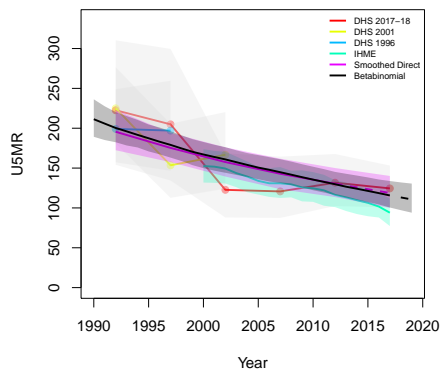
Colored lines with circular points and light grey uncertainty bands are 5-year survey-weighted estimates of U5MR for years 1990–1994 up to 2015–2019 depending on survey timing. For a survey that ends in the middle of a 5-year period, we plot the estimates at the mid-point of the years in that interval for which the survey provides data. Black lines and corresponding intervals represent posterior medians and 95% uncertainty intervals respectively for the betabinomial model. IHME’s estimates and corresponding intervals, where we can compare, are in aquamarine.



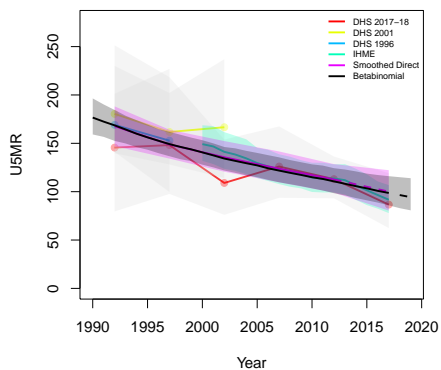
**Collines**



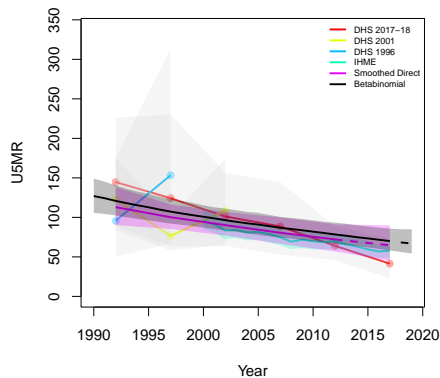
**Donga**



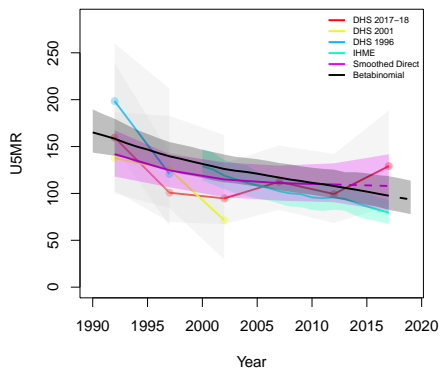
**Kouffo**



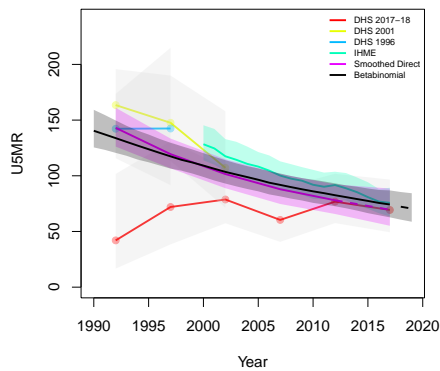
**Littoral**



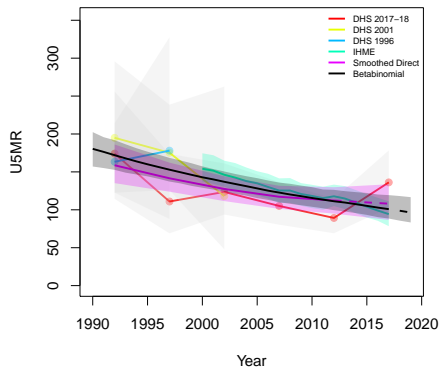
**Mono**



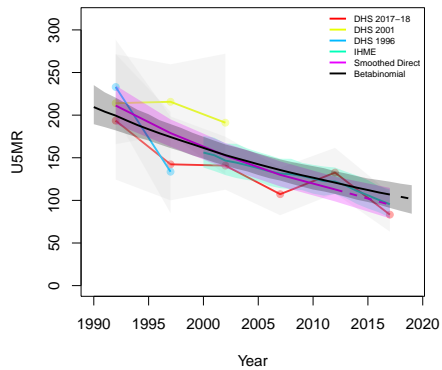
**Ouémé**



Plateau



Zou



*B.1.2 Admin-2*



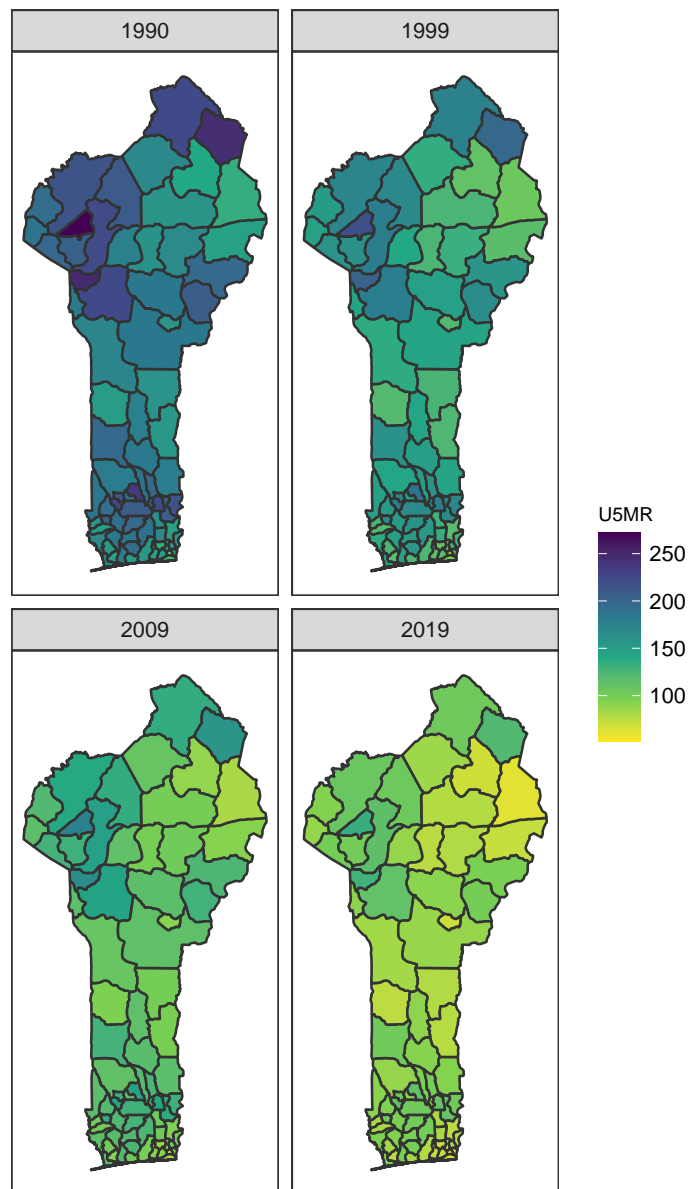


Figure B.5: Median U5MR estimates for years 1990, 1999, 2009, 2019 for Admin-2 areas in Benin .

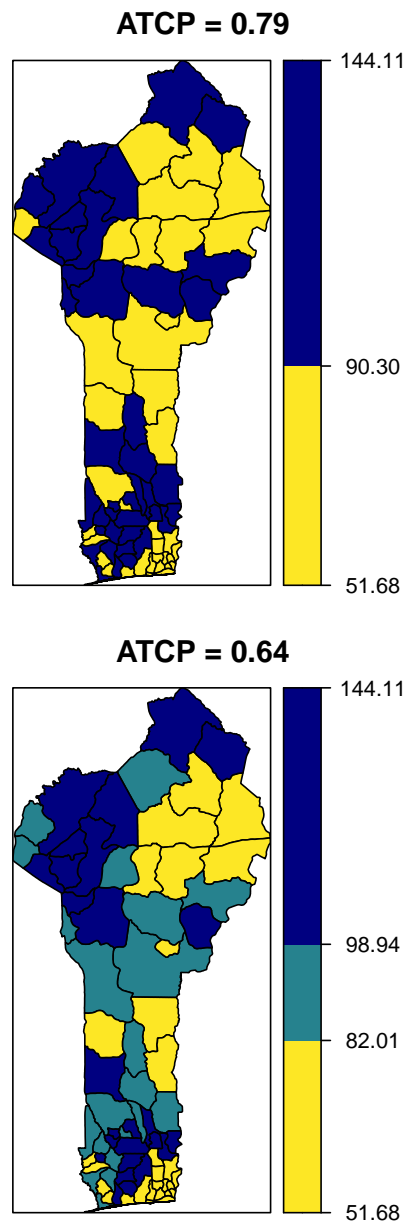
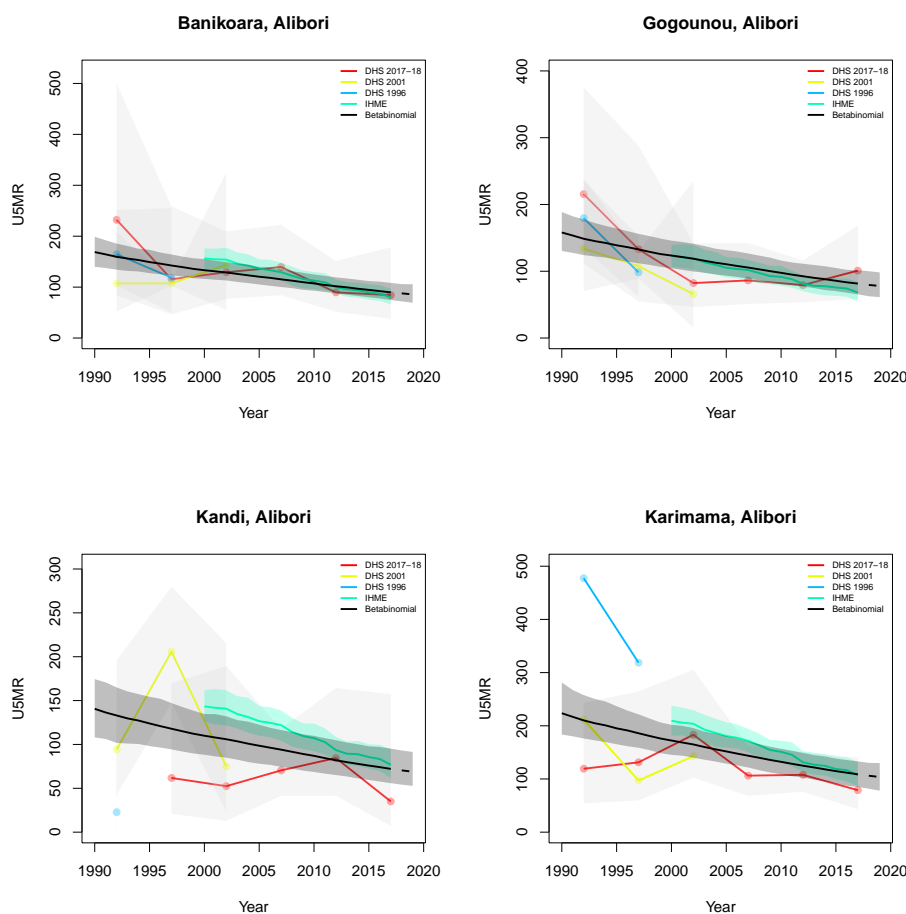


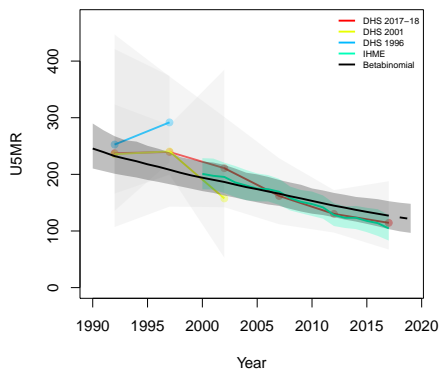
Figure B.6: Expression of uncertainty of U5MR (deaths per 1000 children) estimates for Admin-1 areas based on the average true classification probability (ATCP) in 2019 using  $K = 2, 3$  colors.

*Data and estimates over time by area*

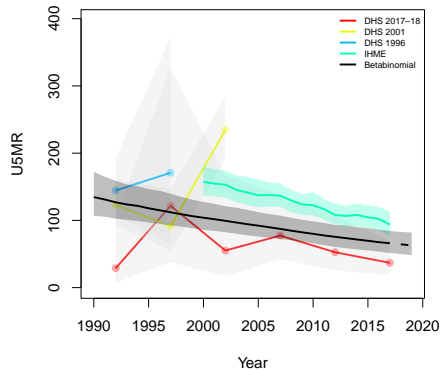
Colored lines with circular points and light grey uncertainty bands are 5-year survey-weighted estimates of U5MR for years 1990–1994 up to 2015–2019 depending on survey timing. For a survey that ends in the middle of a 5-year period, we plot the estimates at the mid-point of the years in that interval for which the survey provides data. Black lines and corresponding intervals represent posterior medians and 95% uncertainty intervals respectively for the betabinomial model. IHME’s estimates and corresponding intervals, where we can compare, are in aquamarine.



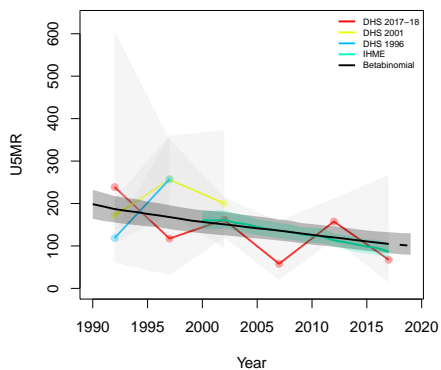
**Malanville, Alibori**



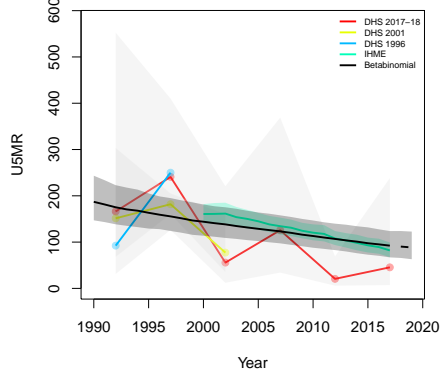
**Segbana, Alibori**



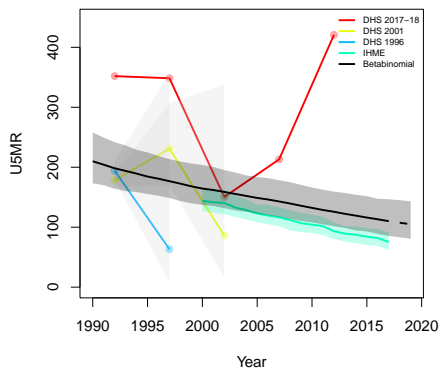
**Boukoubé, Atakora**



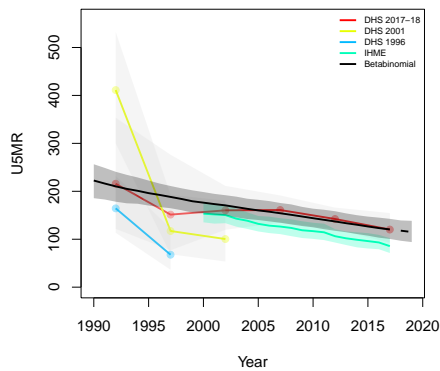
**Cobly, Atakora**



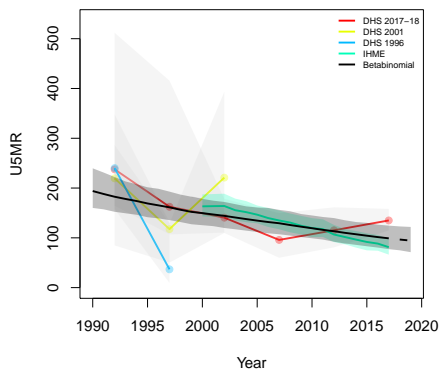
**Kérou, Atakora**



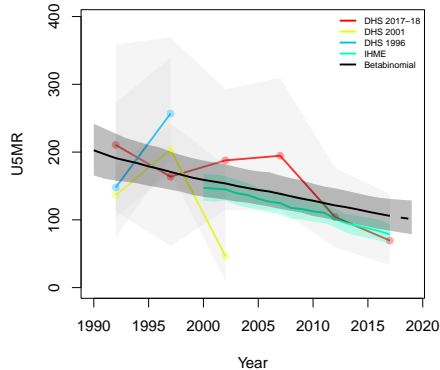
**Kouandé, Atakora**



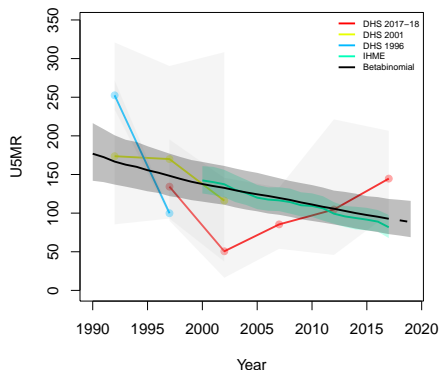
**Matéri, Atakora**



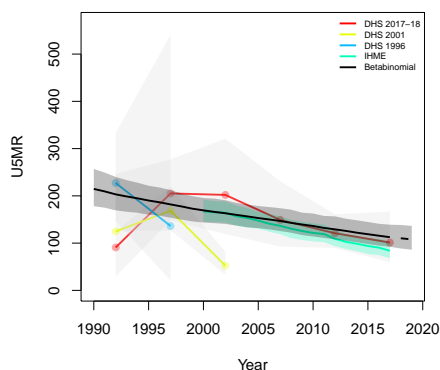
**Natitingou, Atakora**



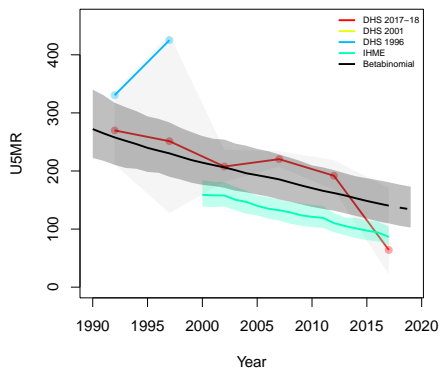
**Péunco, Atakora**



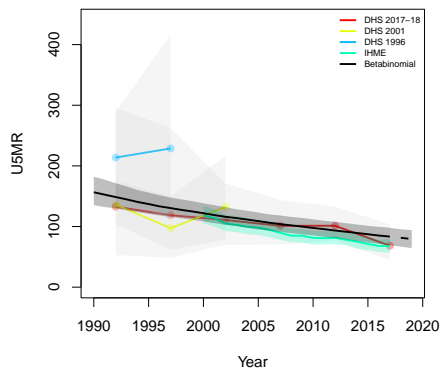
**Tanguiéta, Atakora**



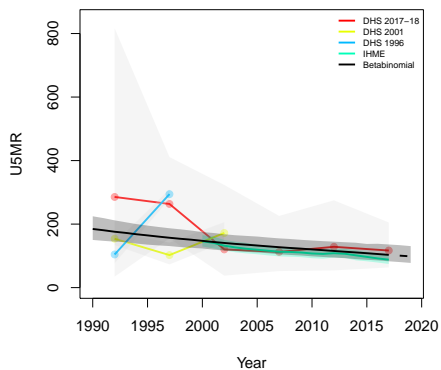
**Toucountouna, Atakora**



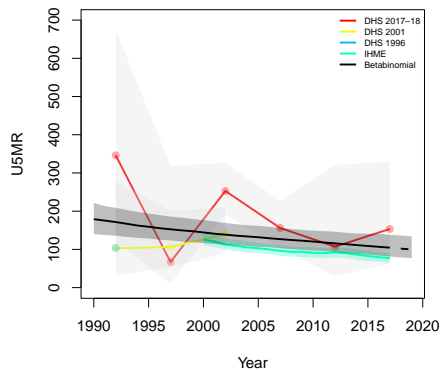
**Abomey-Calavi, Atlantique**



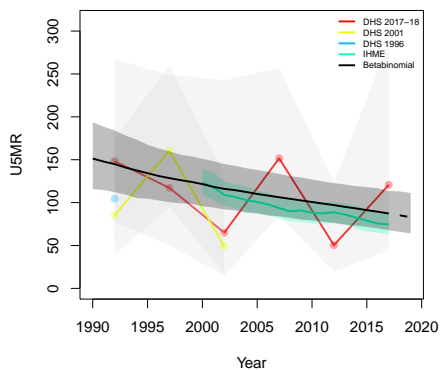
**Allada, Atlantique**



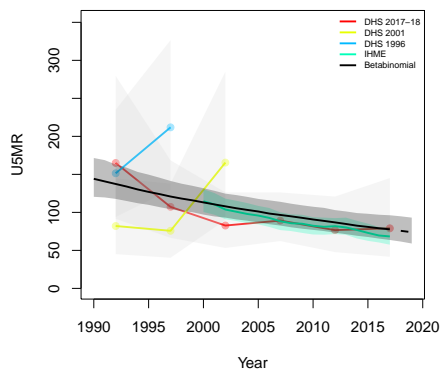
**Kpomassè, Atlantique**



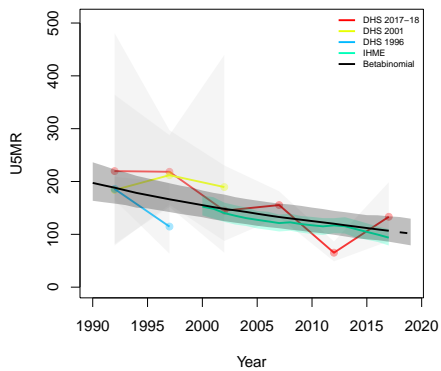
**Ouidah, Atlantique**



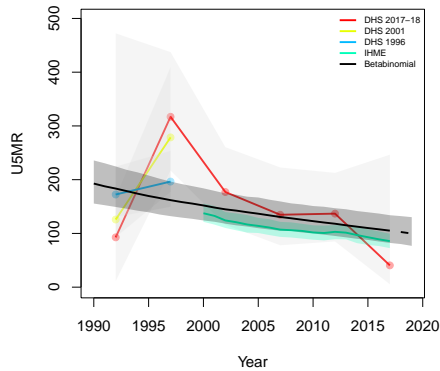
**Sô-Ava, Atlantique**



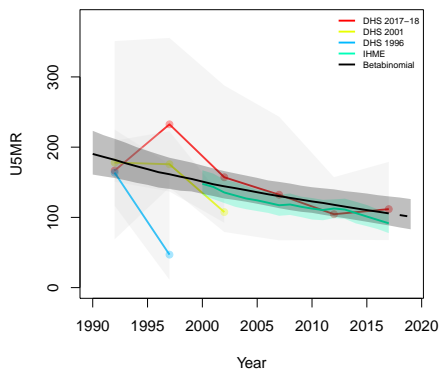
**Toffo, Atlantique**



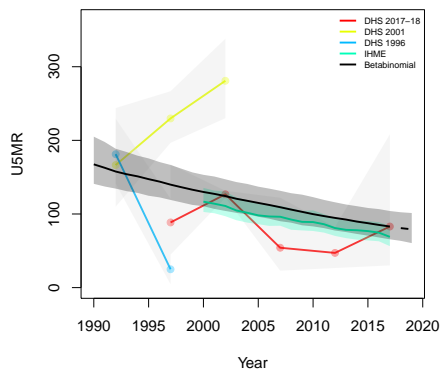
**Tori-Bossito, Atlantique**



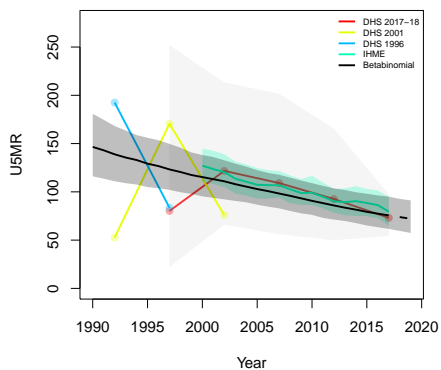
**Zè, Atlantique**



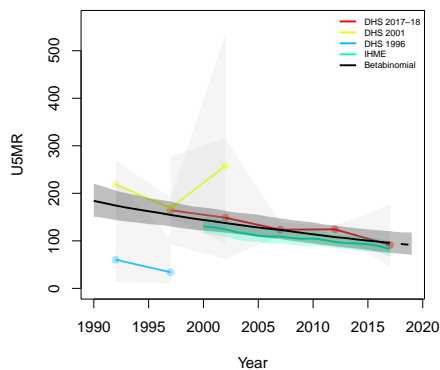
**Bembéréké, Borgou**



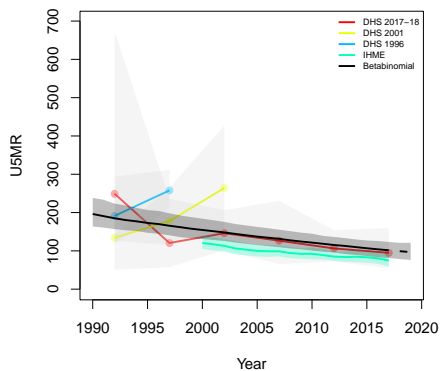
**Kalalé, Borgou**



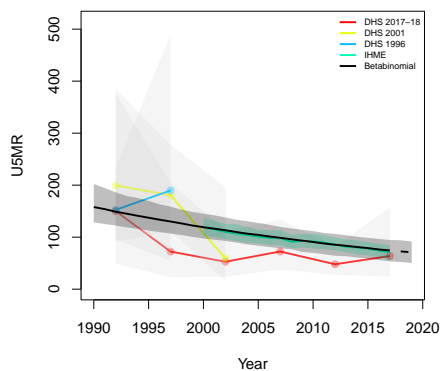
**N'Dali, Borgou**



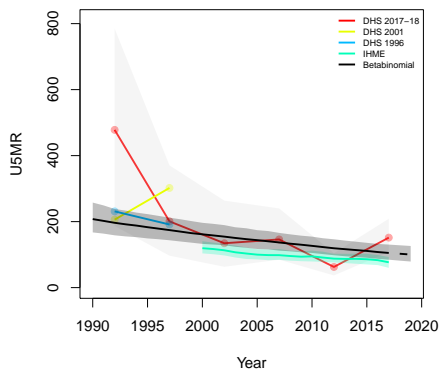
**Nikki, Borgou**



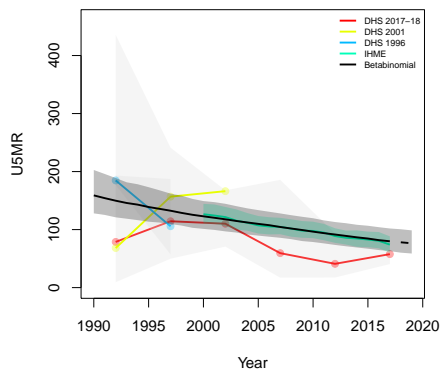
**Parakou, Borgou**



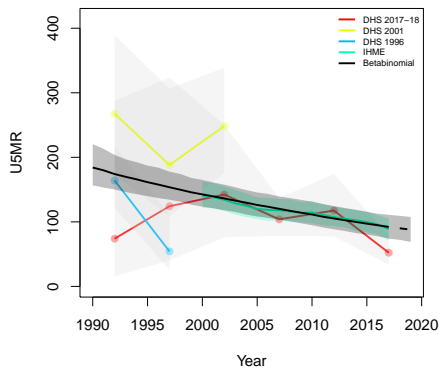
**Pèrèrè, Borgou**



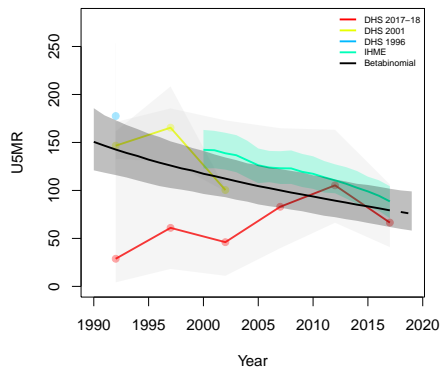
**Sinendé, Borgou**



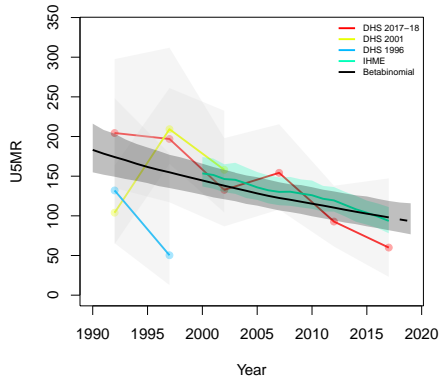
**Tchaourou, Borgou**



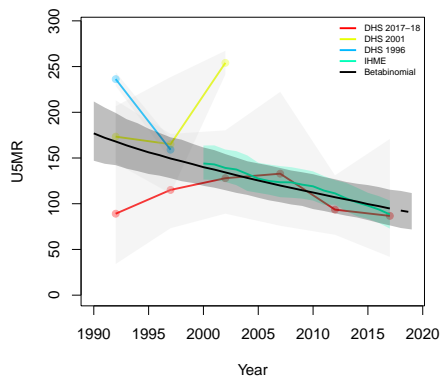
**Bantè, Collines**



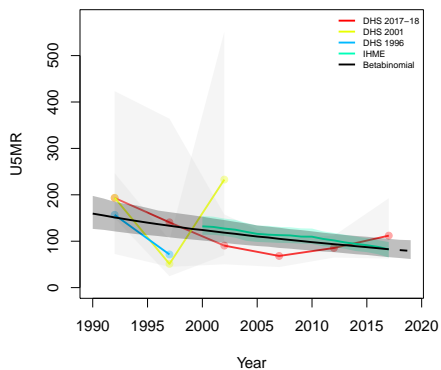
**Dassa-Zoumè, Collines**



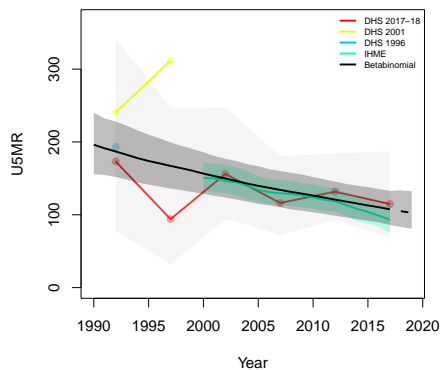
**Glazoué, Collines**



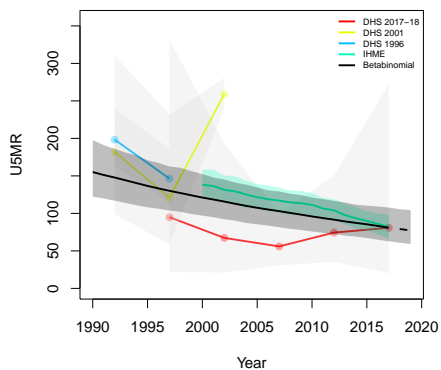
**Ouèssè, Collines**



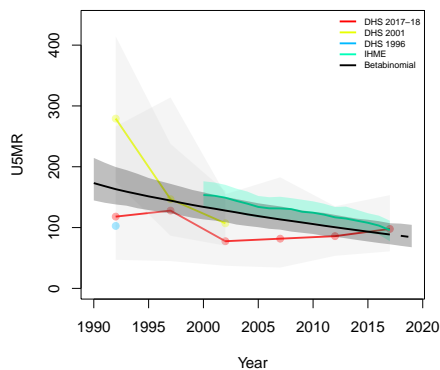
**Savalou, Collines**



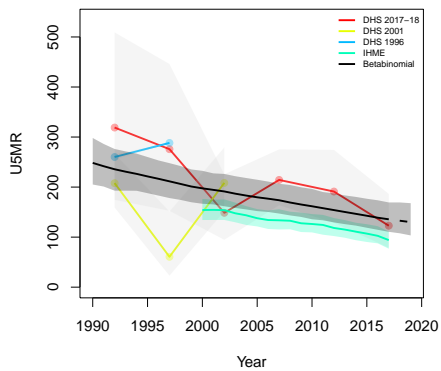
**Savè, Collines**



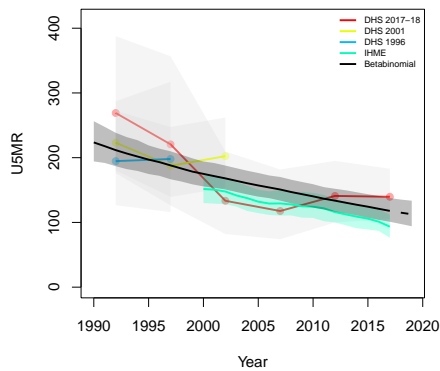
**Bassila, Donga**



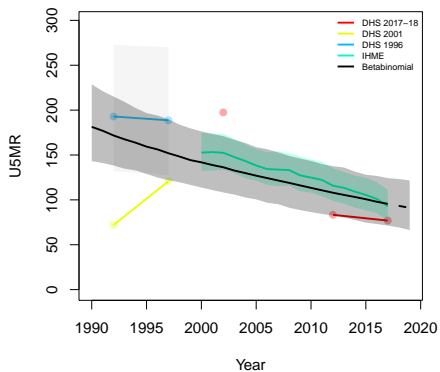
**Copargo, Donga**



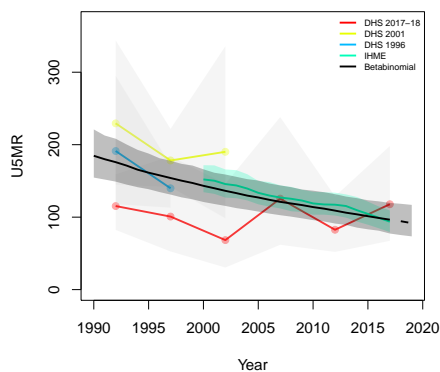
**Djouguou, Donga**



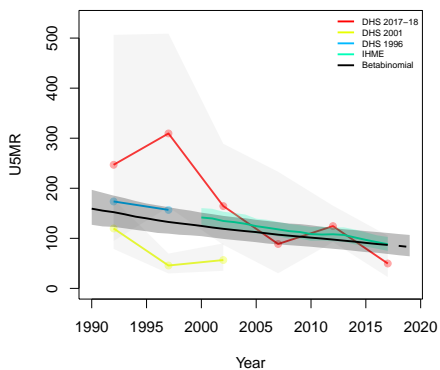
**Ouaké, Donga**



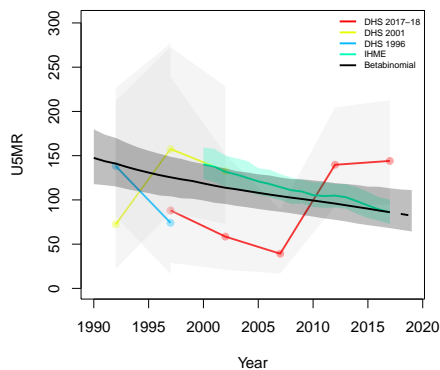
**Aplahoué, Kouffo**



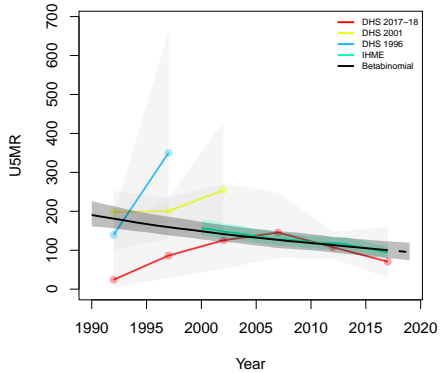
**Djakotomey, Kouffo**



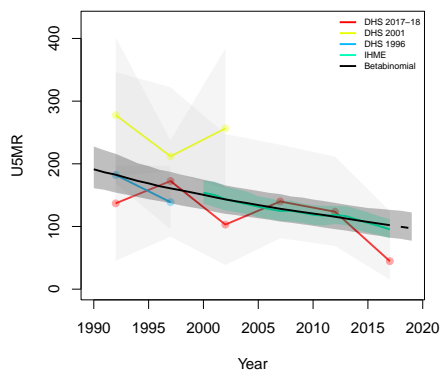
**Dogbo, Kouffo**

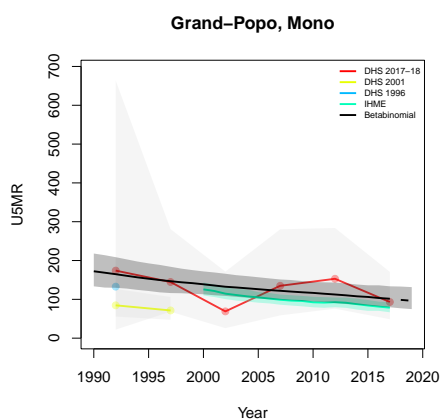
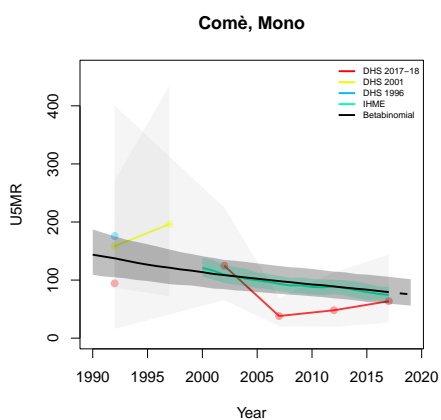
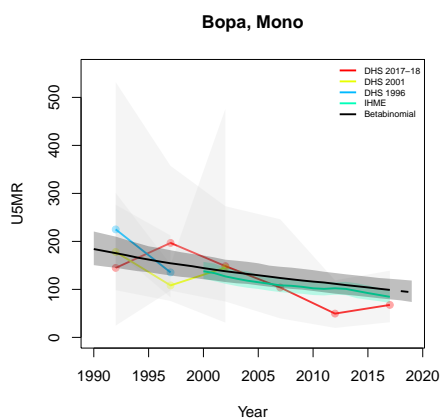
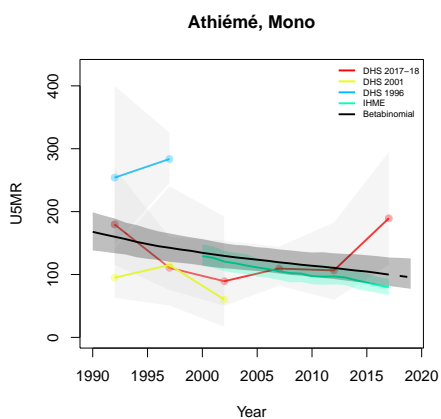
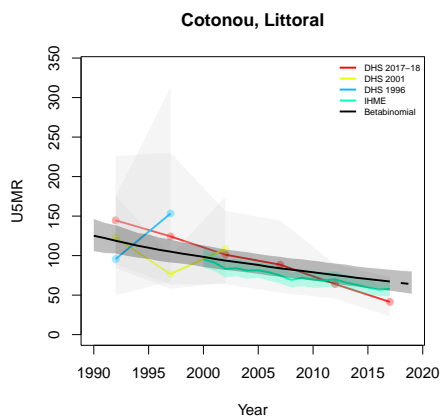
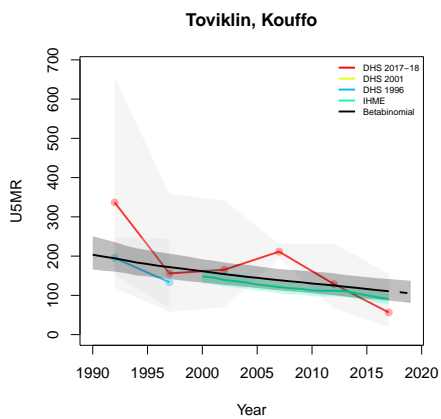


**Klouékanmè, Kouffo**

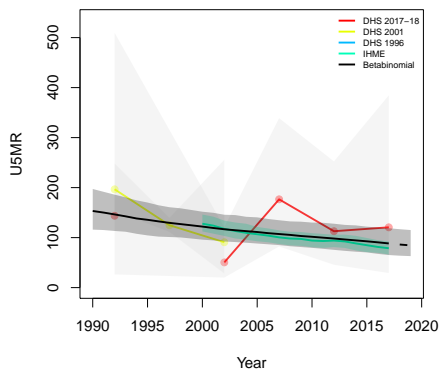


**Lalo, Kouffo**

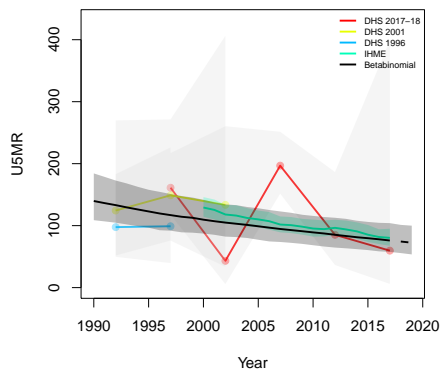




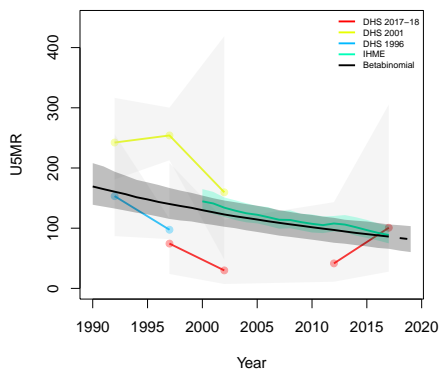
**Houéyogbé, Mono**



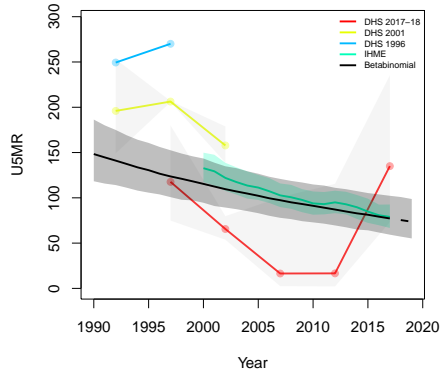
**Adjarra, Ouémé**



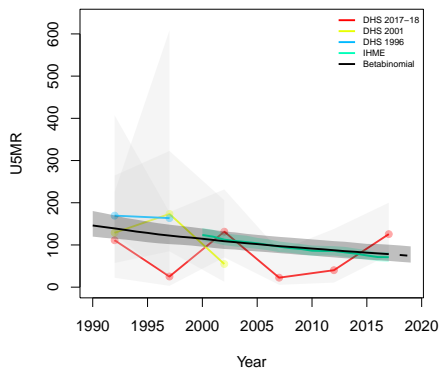
**Adjohoun, Ouémé**



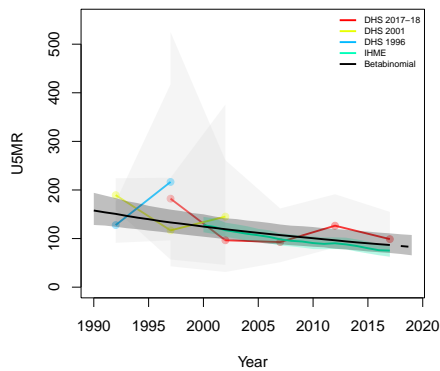
**Aguégués, Ouémé**



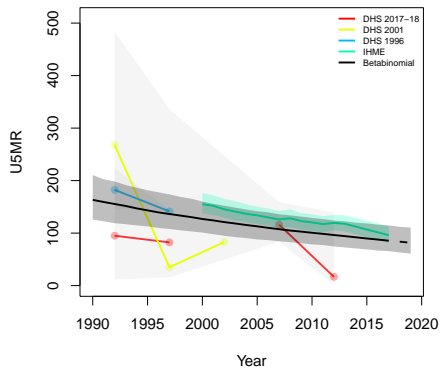
**Akpro-Missérétié, Ouémé**



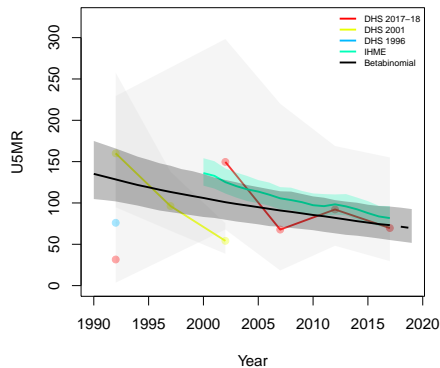
**Avrankou, Ouémé**



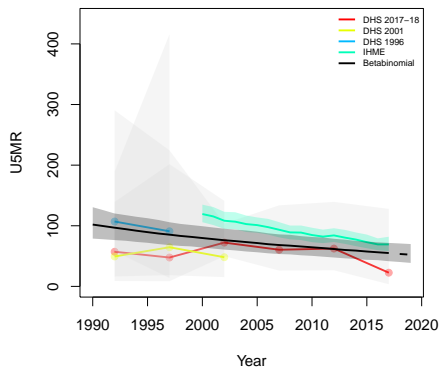
**Bonou, Ouémé**



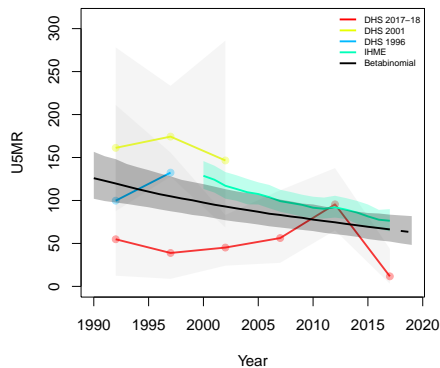
**Dangbo, Ouémé**



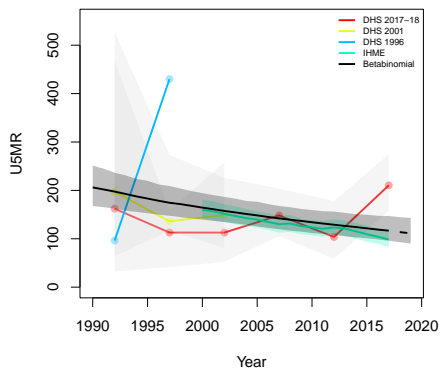
**Porto-Novo, Ouémé**



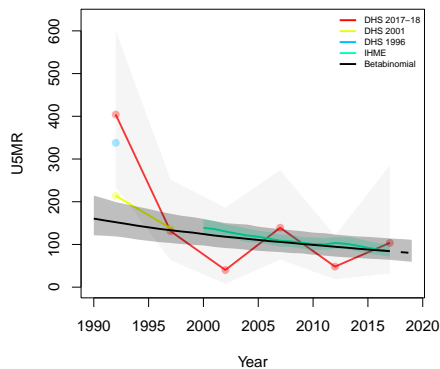
**Sèmè-Kpodji, Ouémé**



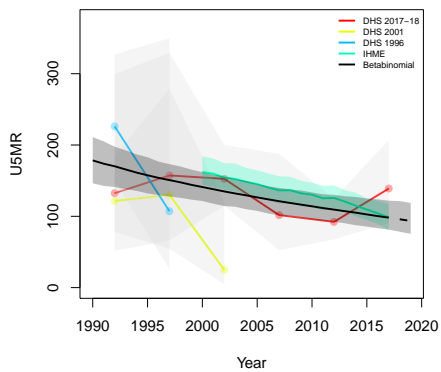
**Adja-Ouèrè, Plateau**



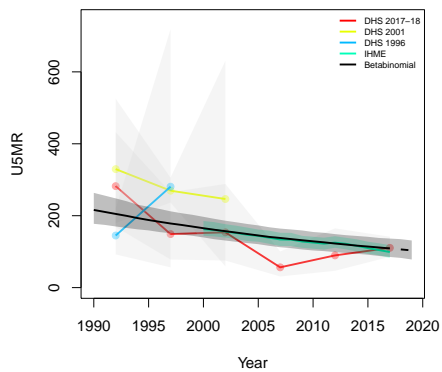
**Ifangni, Plateau**



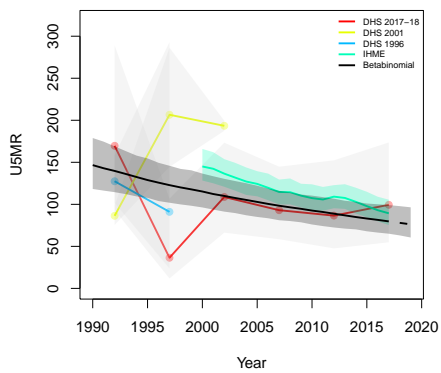
**Kétou, Plateau**



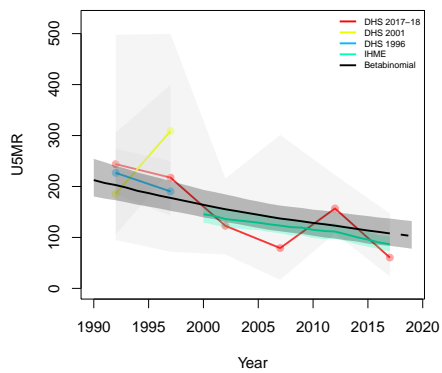
**Pobè, Plateau**



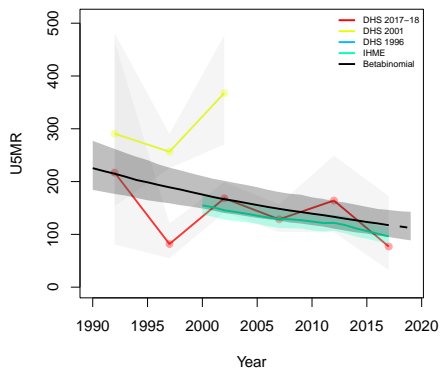
**Sakété, Plateau**



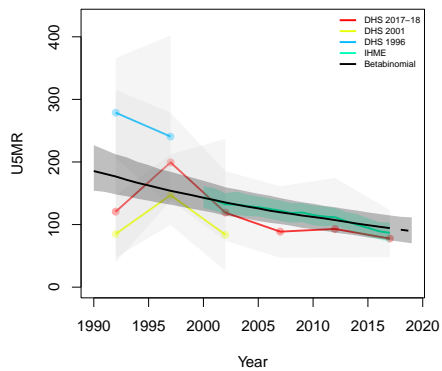
**Abomey, Zou**

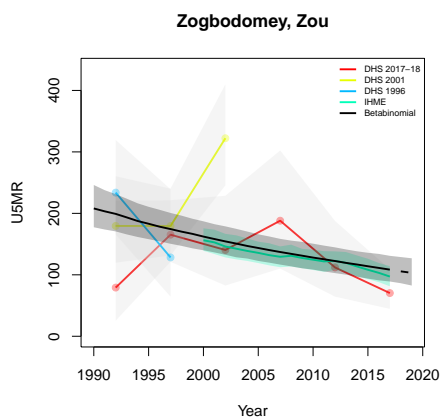
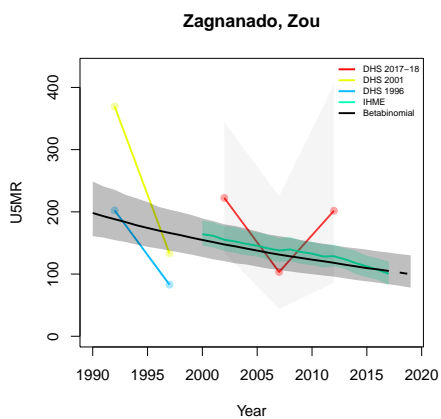
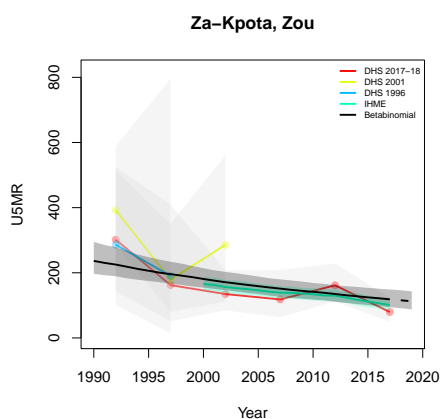
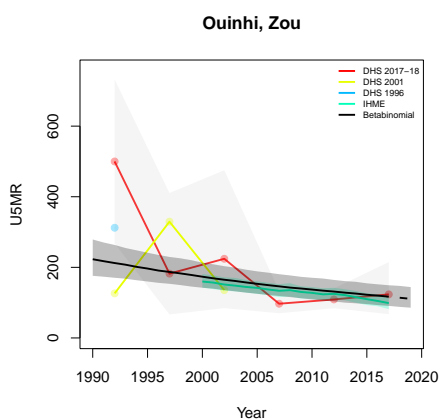
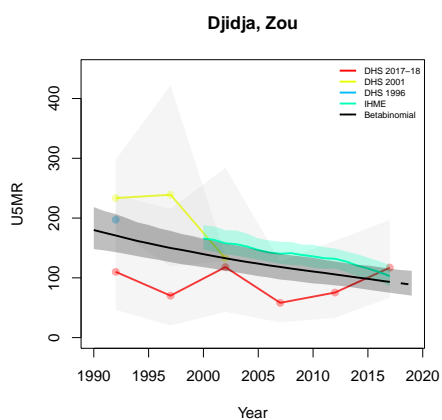
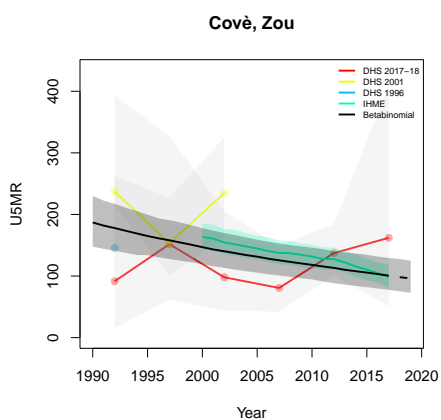


**Agbangnizoun, Zou**



**Bohicon, Zou**





**B.2 Burundi**

Age	Survey	Clusters			Deaths			Agemonths		
		Urban	Rural	Total	Urban	Rural	Total	Urban	Rural	Total
0	2010	75	301	376	93	745	838	3661	18693	22354
	2016	105	447	552	163	1107	1270	6900	37594	44494
1-11	2010	75	301	376	105	912	1017	37674	187209	224883
	2016	105	447	552	190	1347	1537	70639	383245	453884
12-23	2010	75	301	376	43	369	412	37584	185267	222851
	2016	105	447	552	79	565	644	70623	382268	452891
24-35	2010	75	301	376	31	270	301	34448	169483	203931
	2016	105	447	552	30	365	395	65530	353537	419067
36-47	2010	75	301	376	23	215	238	31511	154596	186107
	2016	105	447	552	34	316	350	60816	327220	388036
48-59	2010	75	301	376	20	124	144	28524	140504	169028
	2016	105	447	552	11	152	163	56268	300356	356624

Table B.2: **Data summary for Burundi.** Total numbers of clusters (Columns 3–5) with observations in each age group by survey in urban and rural areas and combined. Numbers of deaths (Columns 6–8) and number of agemonths (Columns 9–10) observed in each age group by survey in urban and rural areas and combined.

*B.2.1 Admin-1*

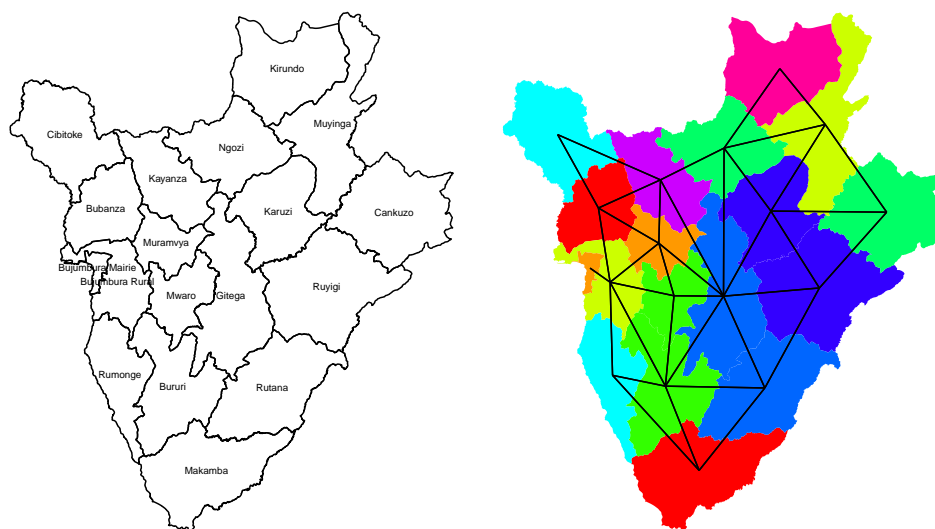


Figure B.7: **Left:** The names of the 18 Admin-1 areas of Burundi . **Right:** The neighborhood structure of Admin-1 areas in Burundi .

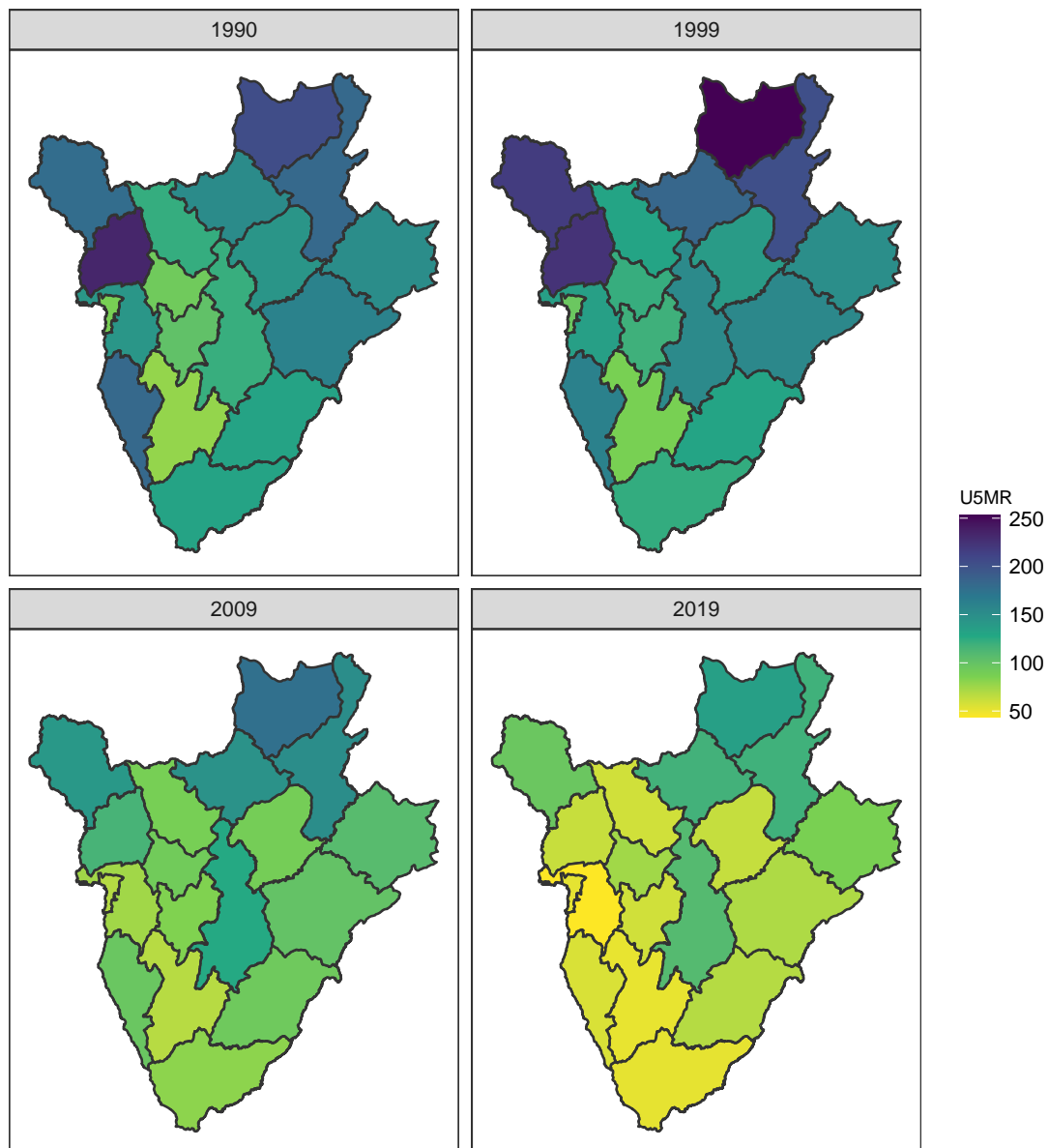


Figure B.8: Median U5MR estimates for years 1990, 1999, 2009, 2019 for Admin-1 areas in Burundi .

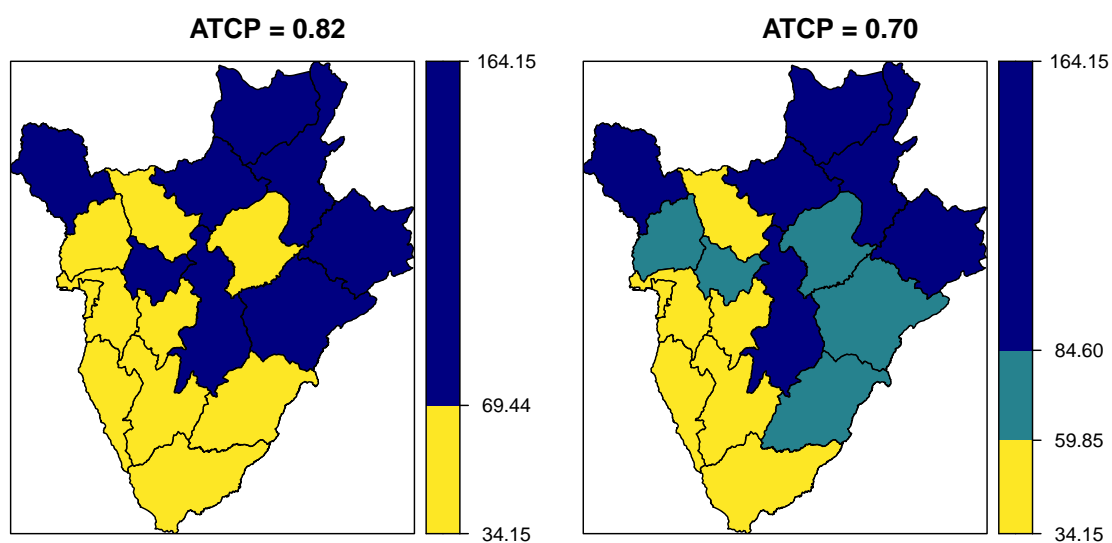
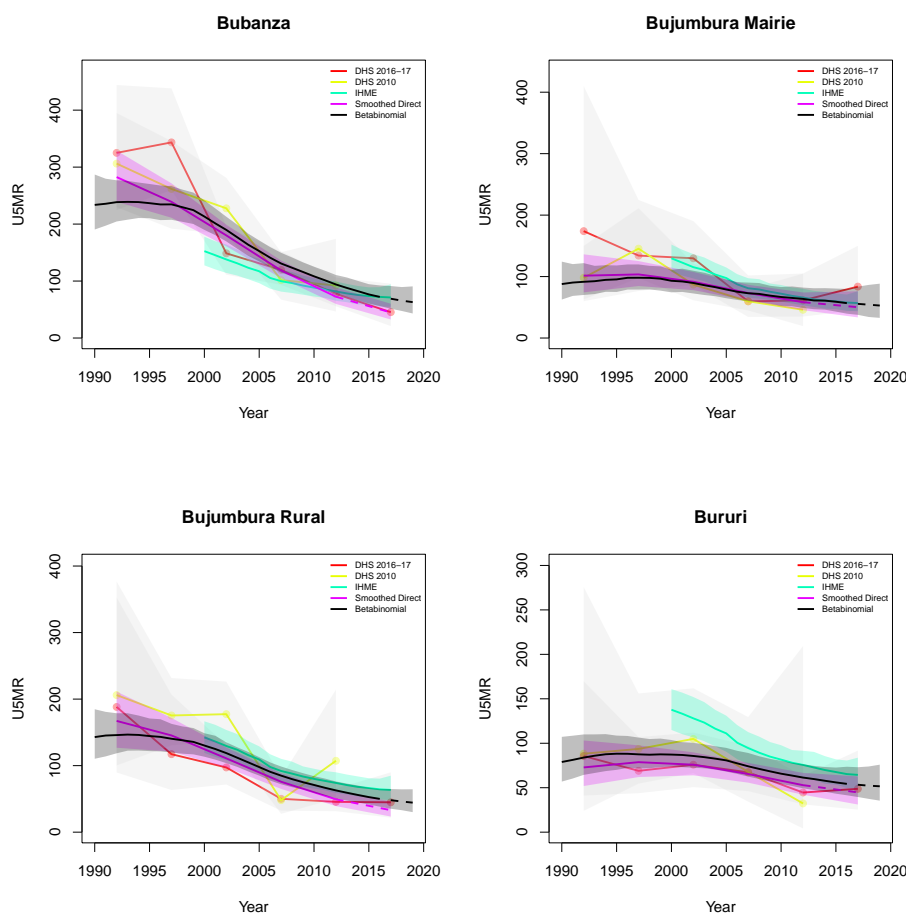
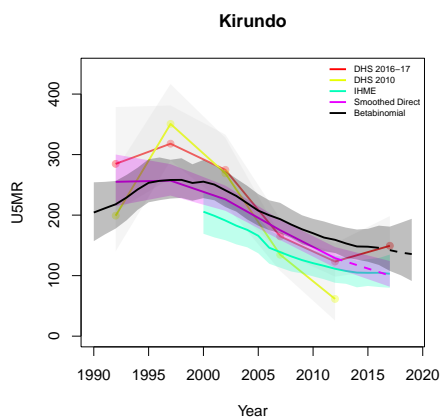
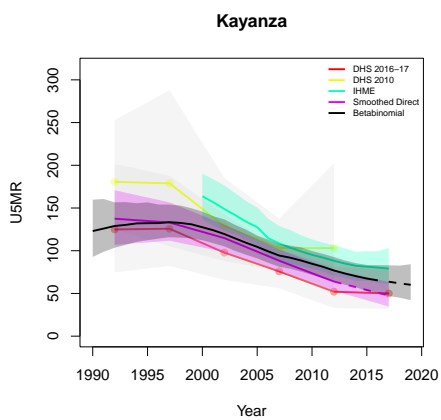
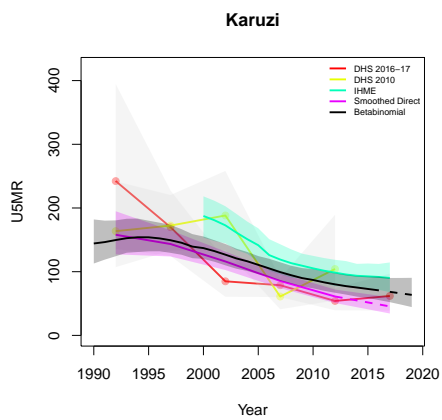
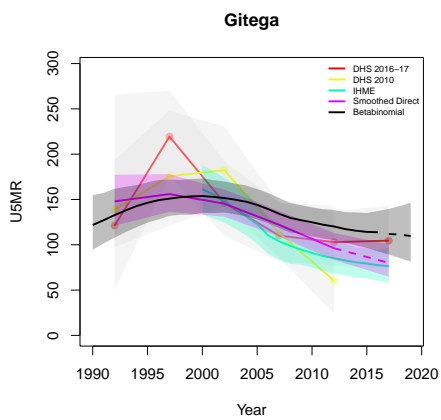
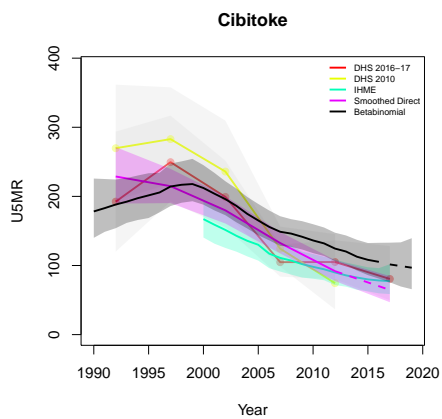
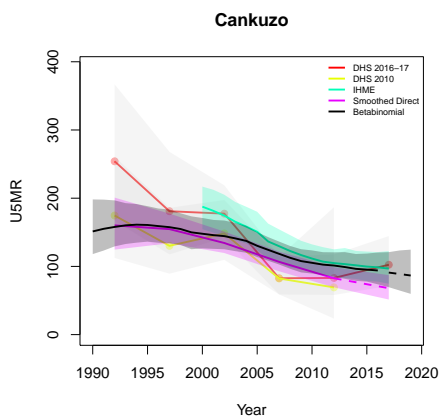


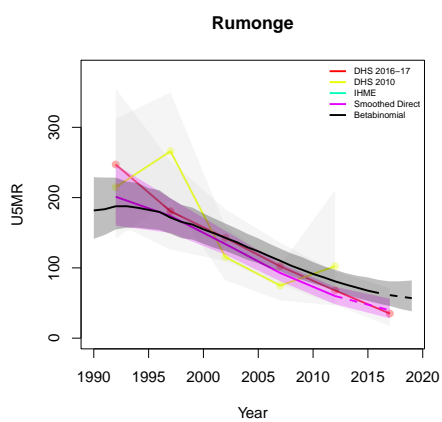
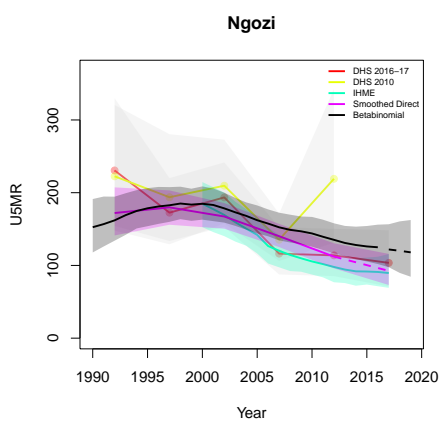
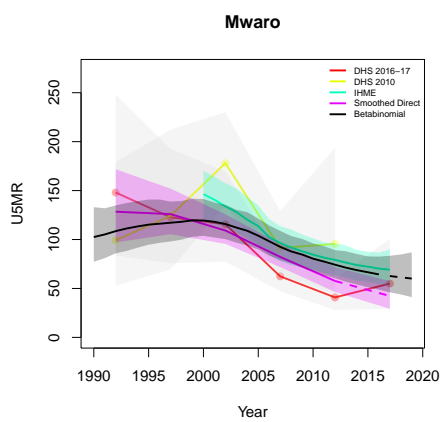
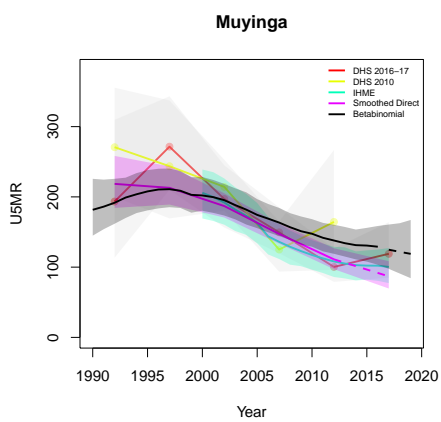
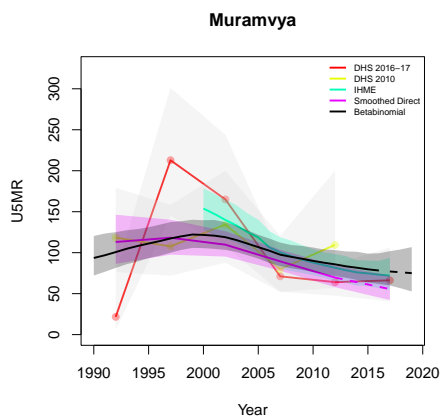
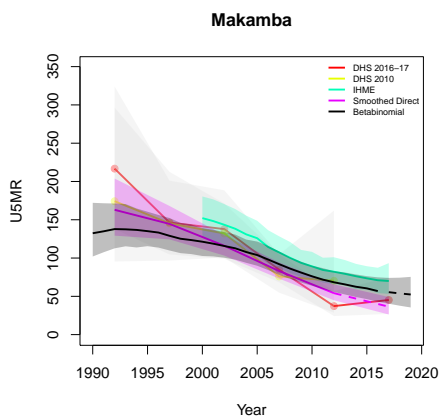
Figure B.9: Expression of uncertainty of U5MR (deaths per 1000 children) estimates for Admin-1 areas based on the average true classification probability (ATCP) in 2019 using  $K = 2, 3$  colors.

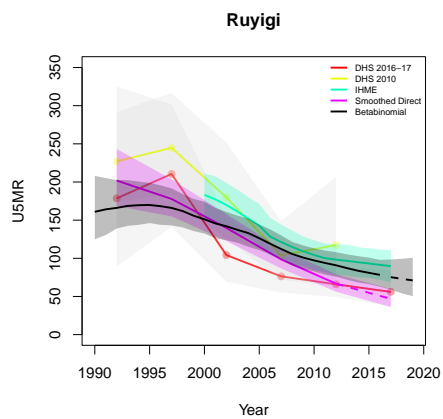
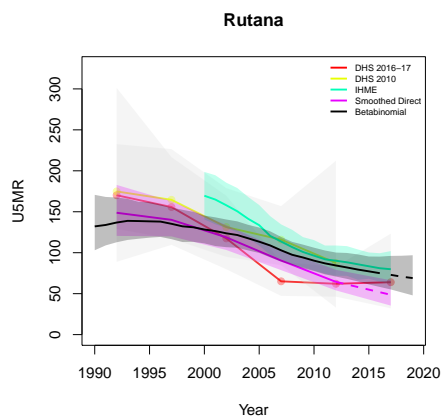
*Data and estimates over time by area*

Colored lines with circular points and light grey uncertainty bands are 5-year survey-weighted estimates of U5MR for years 1990–1994 up to 2015–2019 depending on survey timing. For a survey that ends in the middle of a 5-year period, we plot the estimates at the mid-point of the years in that interval for which the survey provides data. Black lines and corresponding intervals represent posterior medians and 95% uncertainty intervals respectively for the betabinomial model. IHME’s estimates and corresponding intervals, where we can compare, are in aquamarine.









*B.2.2 Admin-2*

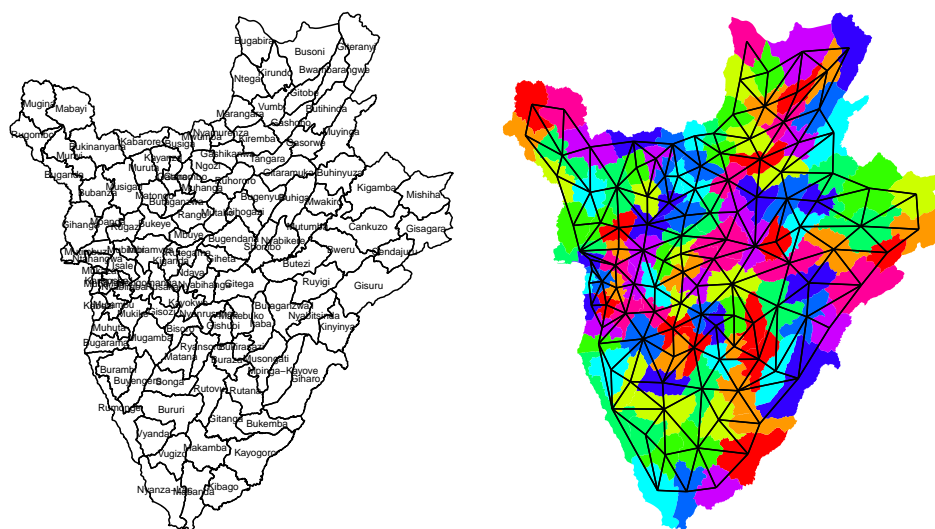


Figure B.10: **Left:** The names of the 119 Admin-2 areas of Burundi . **Right:** The neighborhood structure of Admin-2 areas in Burundi .

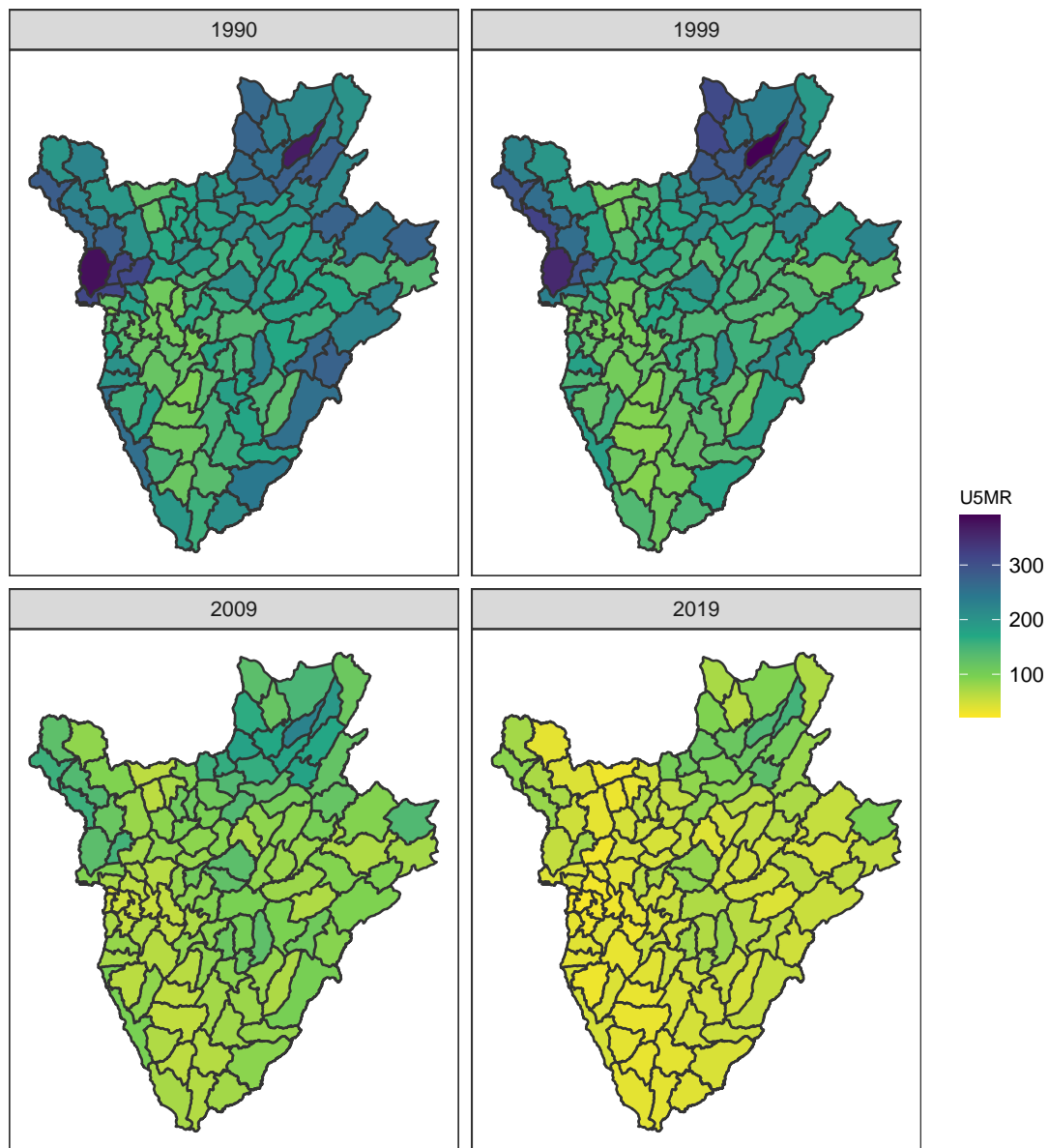


Figure B.11: Median U5MR estimates for years 1990, 1999, 2009, 2019 for Admin-2 areas in Burundi .

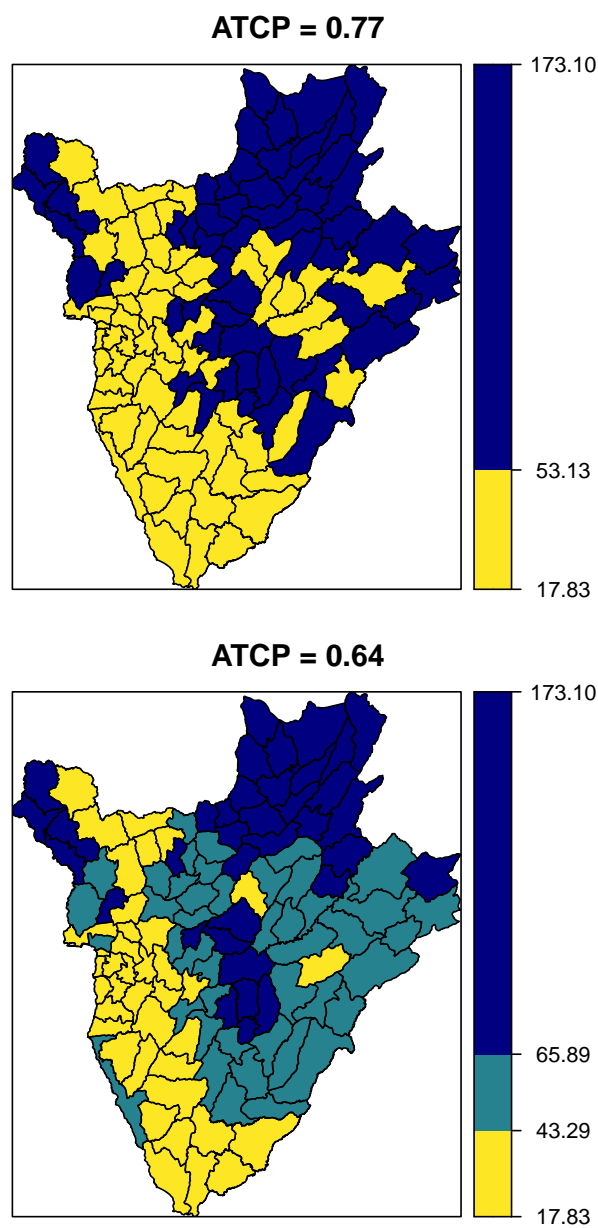
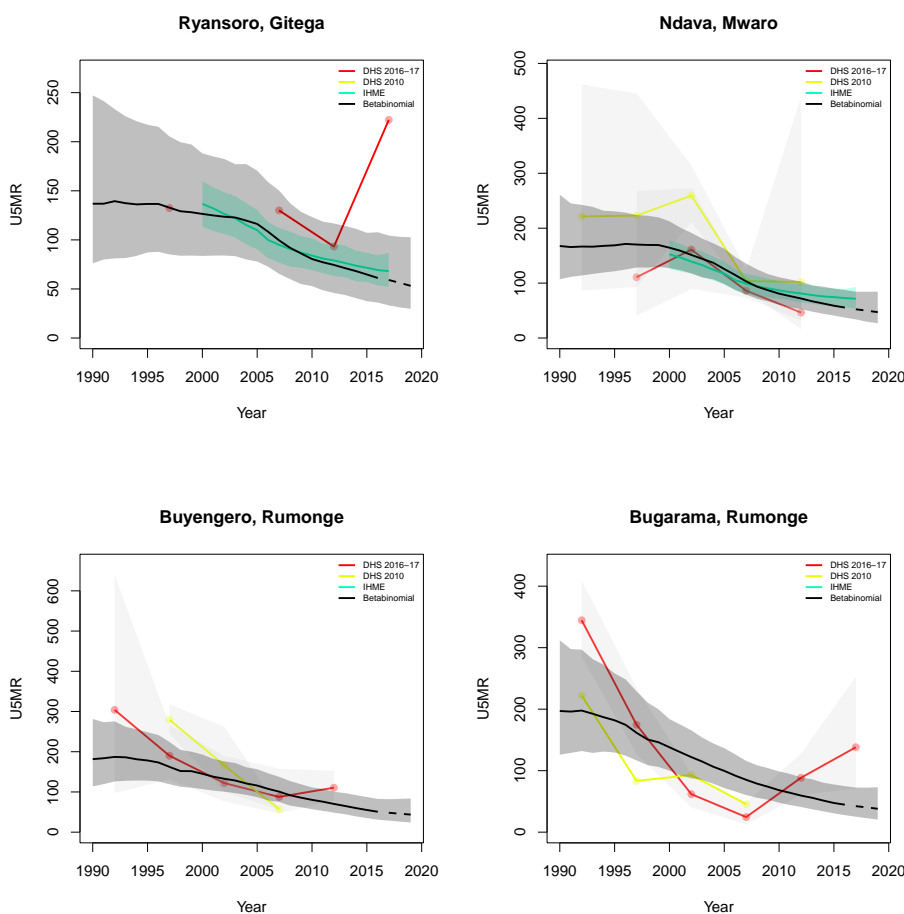


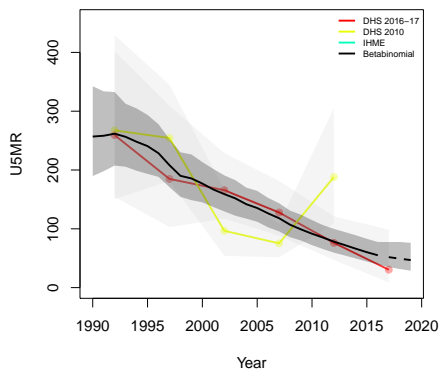
Figure B.12: Expression of uncertainty of U5MR (deaths per 1000 children) estimates for Admin-1 areas based on the average true classification probability (ATCP) in 2019 using  $K = 2, 3$  colors.

*Data and estimates over time by area*

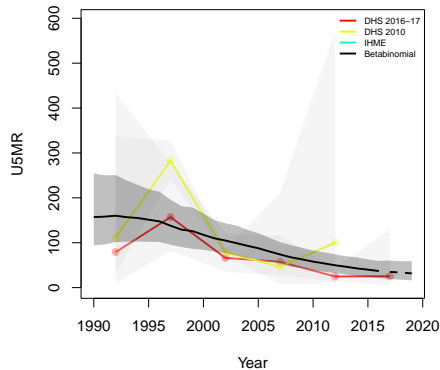
Colored lines with circular points and light grey uncertainty bands are 5-year survey-weighted estimates of U5MR for years 1990–1994 up to 2015–2019 depending on survey timing. For a survey that ends in the middle of a 5-year period, we plot the estimates at the mid-point of the years in that interval for which the survey provides data. Black lines and corresponding intervals represent posterior medians and 95% uncertainty intervals respectively for the betabinomial model. IHME’s estimates and corresponding intervals, where we can compare, are in aquamarine.



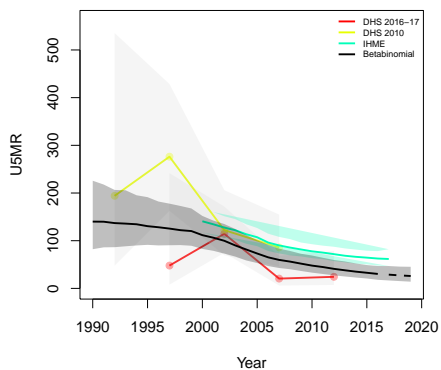
**Rumonge, Rumonge**



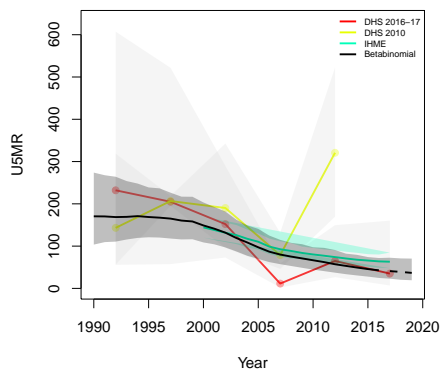
**Burambi, Rumonge**



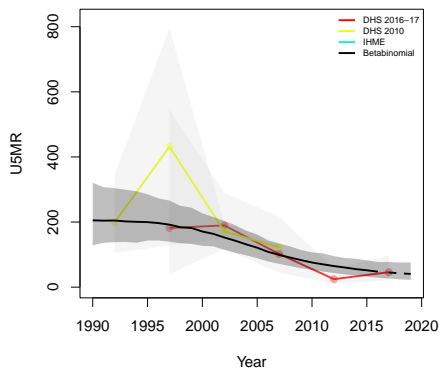
**Kanyosha1, Bujumbura Rural**



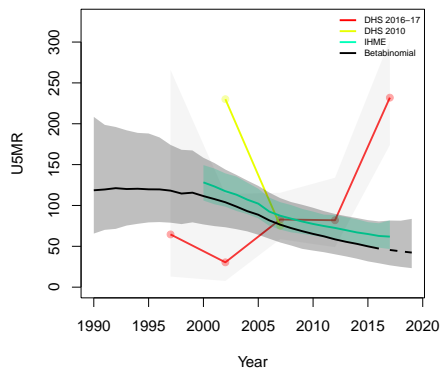
**Kabezi, Bujumbura Rural**



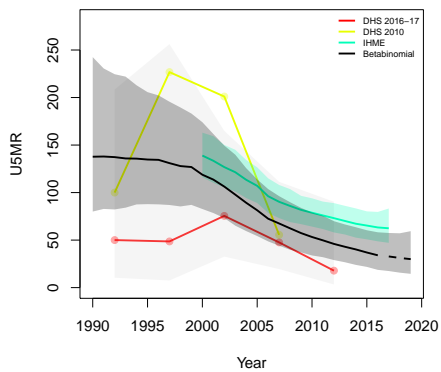
**Muhuta, Rumonge**



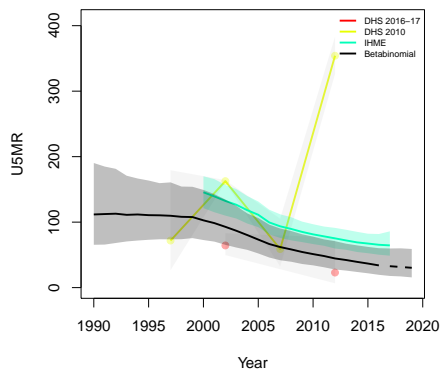
**Mukike, Bujumbura Rural**



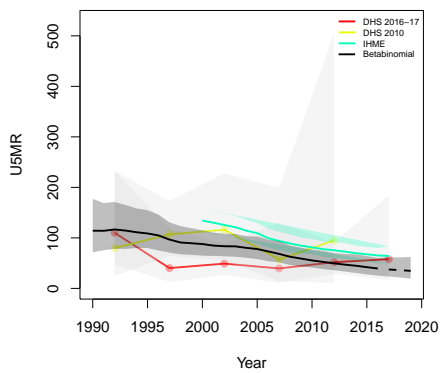
**Mutambu, Bujumbura Rural**



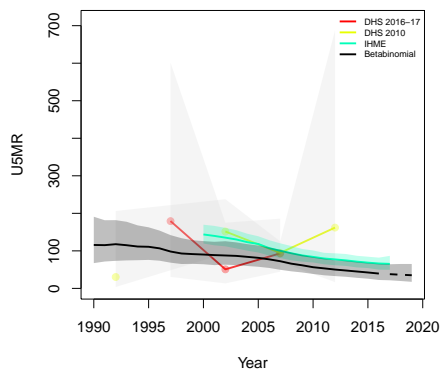
**Nyabiraba, Bujumbura Rural**



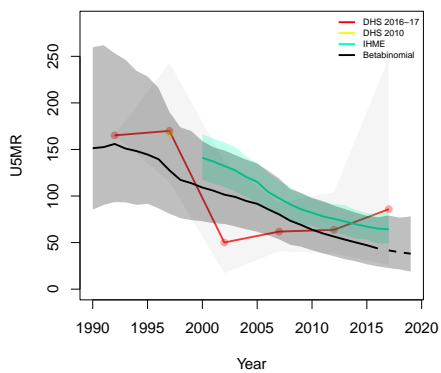
**Bururi, Bururi**



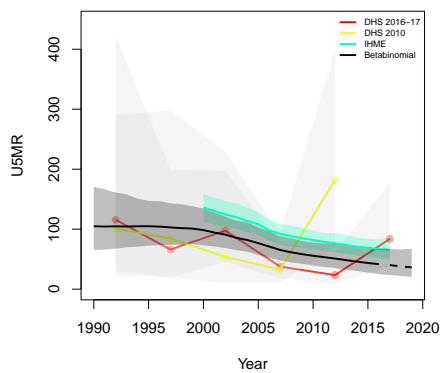
**Vugizo, Makamba**

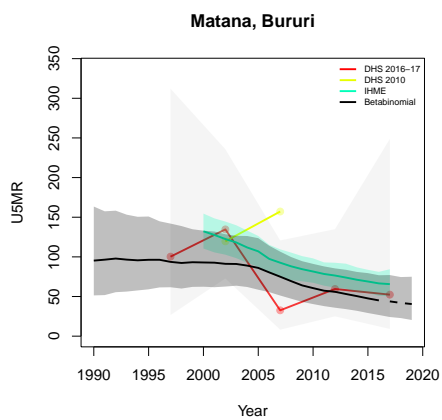
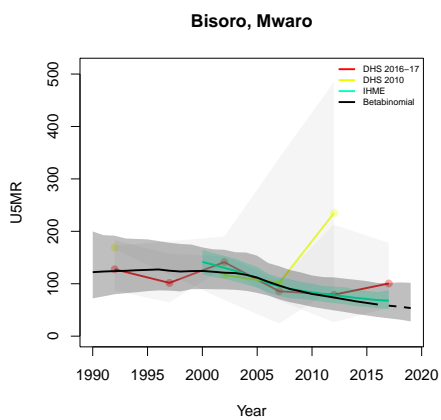
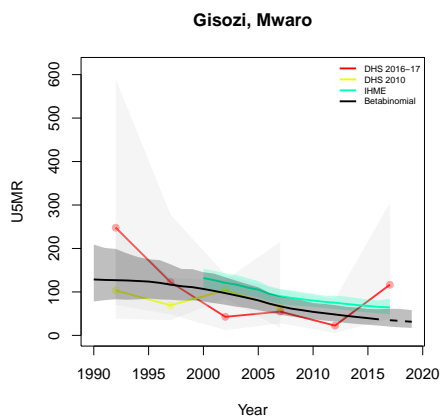
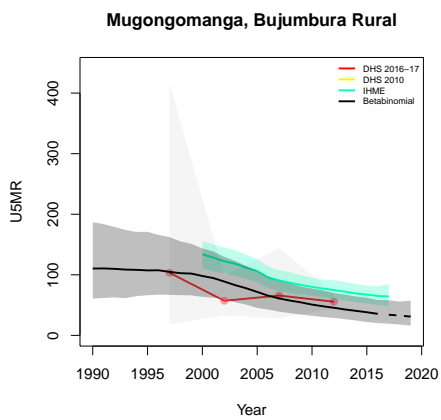
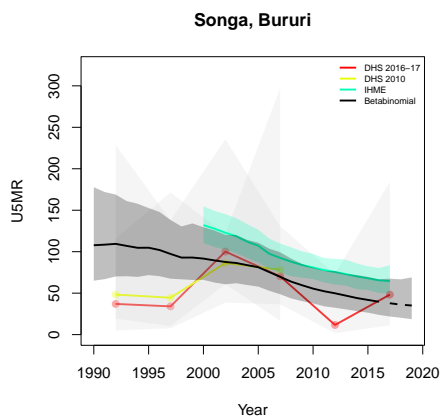
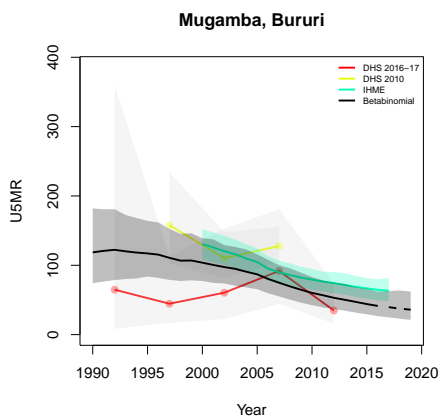


**Vyanda, Bururi**

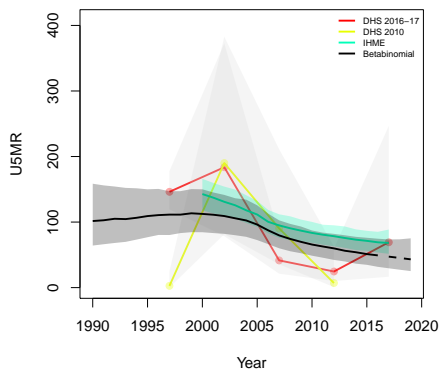


**Rusaka, Mwaro**

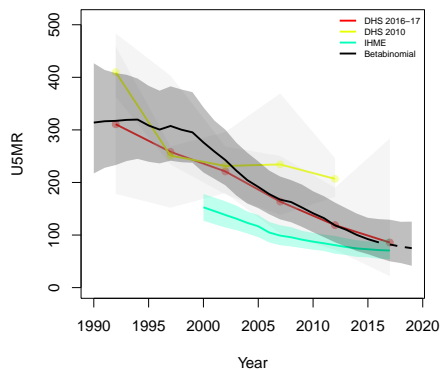




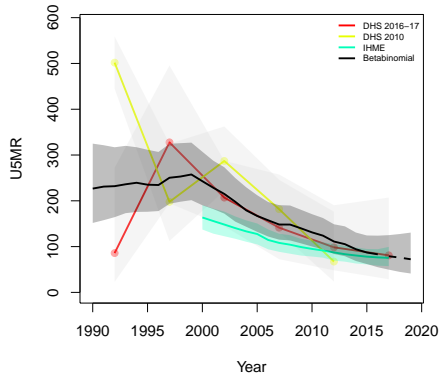
**Kayokwe, Mwaro**



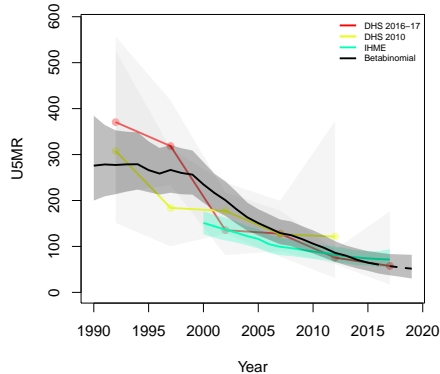
**Mpanda, Bubanza**



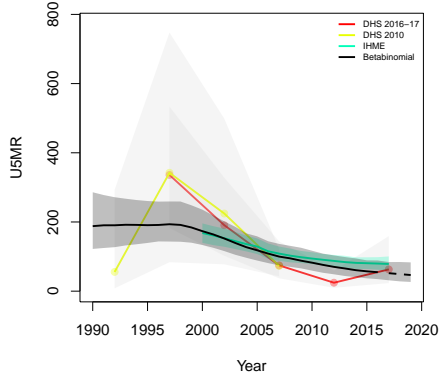
**Murwi, Cibitoke**



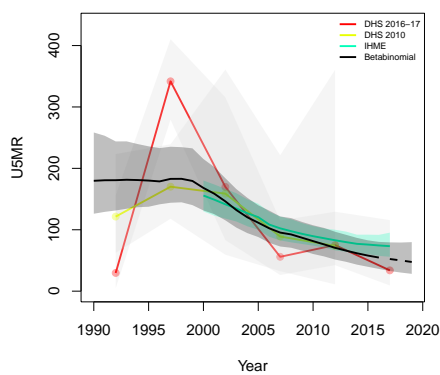
**Bubanza, Bubanza**



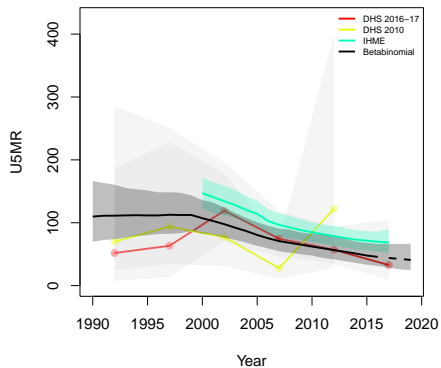
**Rango, Kayanza**



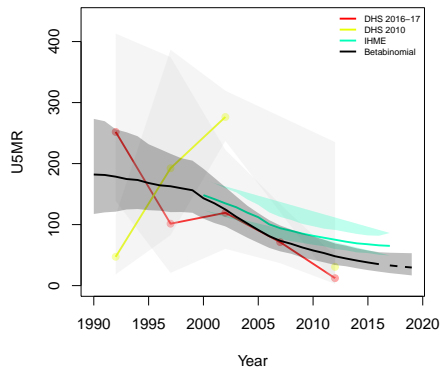
**Bukeye, Muramvya**



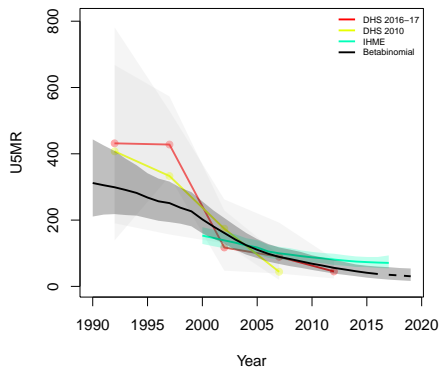
**Muramvya, Muramvya**



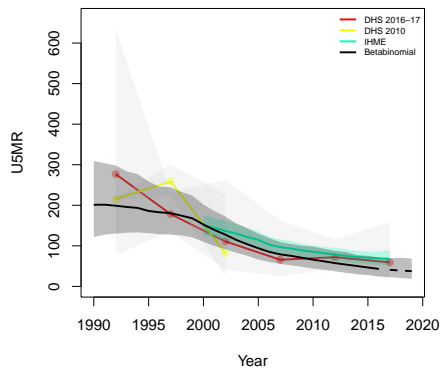
**Isale, Bujumbura Rural**



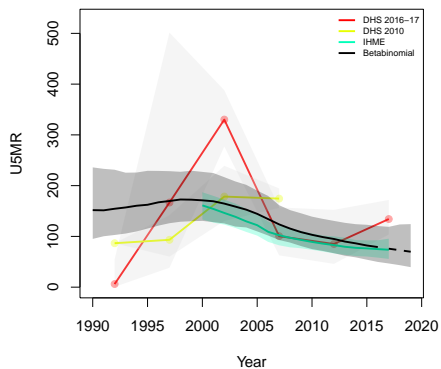
**Rugazi, Bubanza**



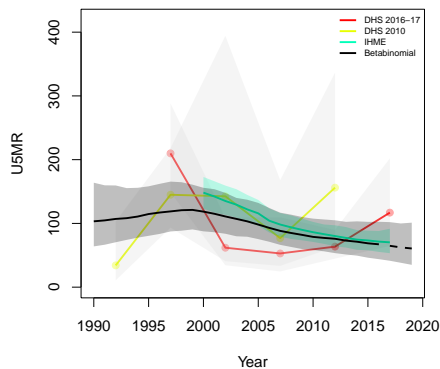
**Mubimbi, Bujumbura Rural**

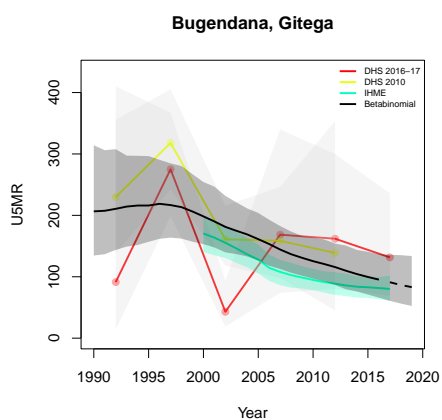
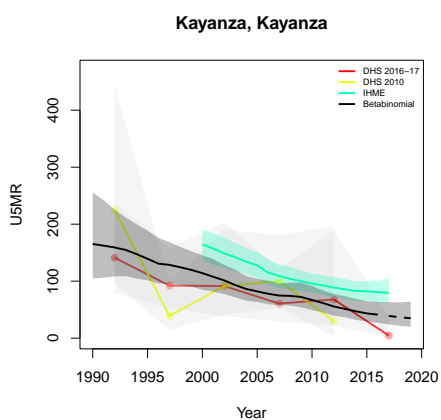
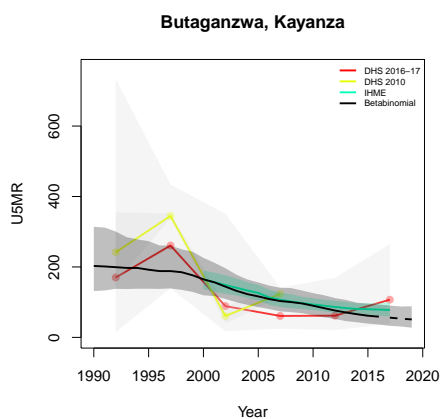
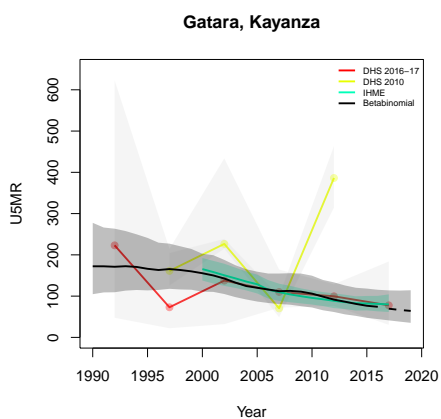
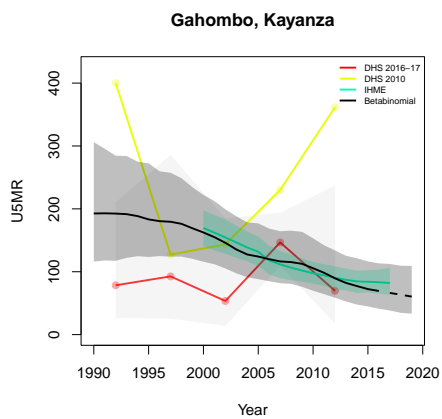
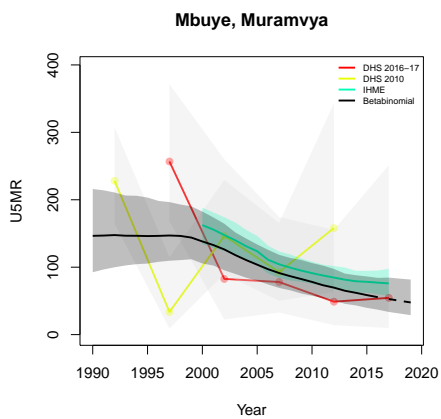


**Rutegama, Muramvya**

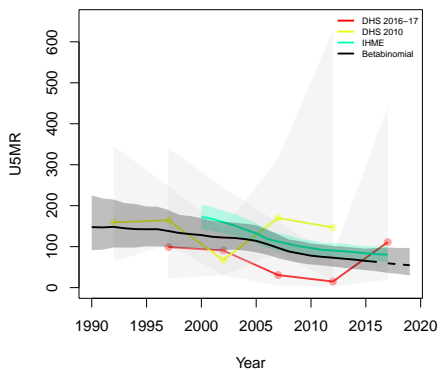


**Kiganda, Muramvya**

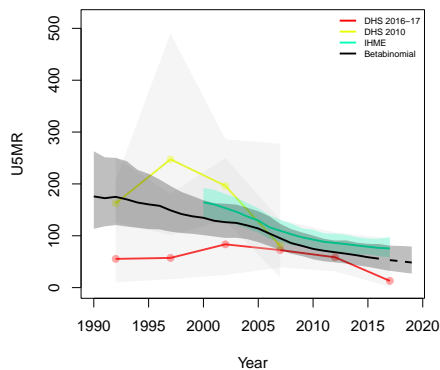




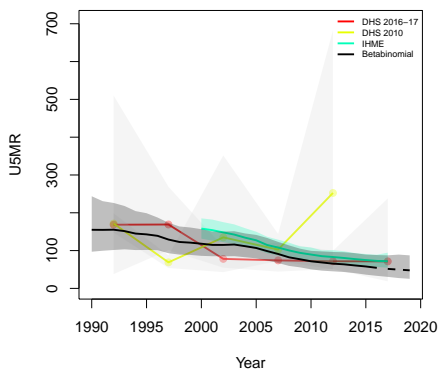
**Musongati, Rutana**



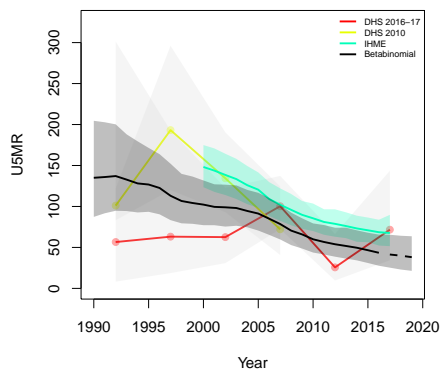
**Rutana, Rutana**



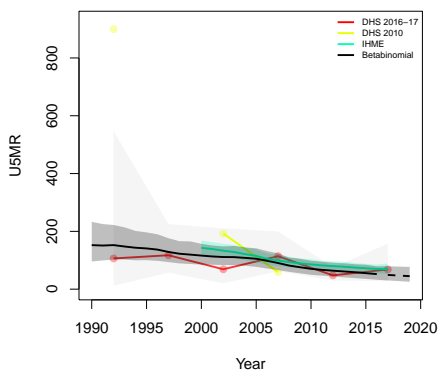
**Gitanga, Rutana**



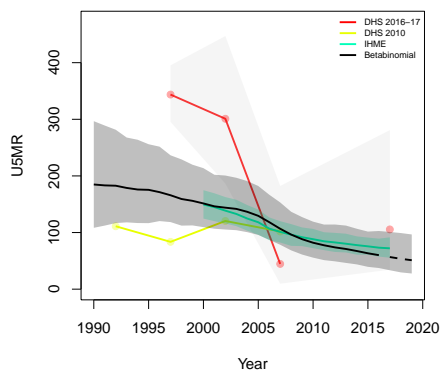
**Makamba, Makamba**



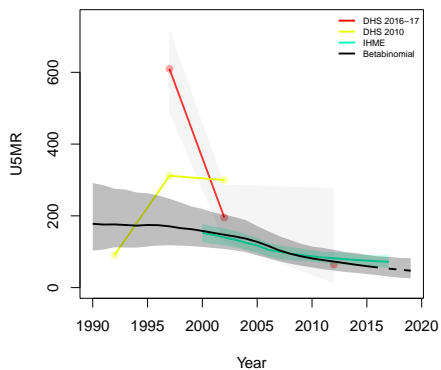
**Rutovu, Bururi**



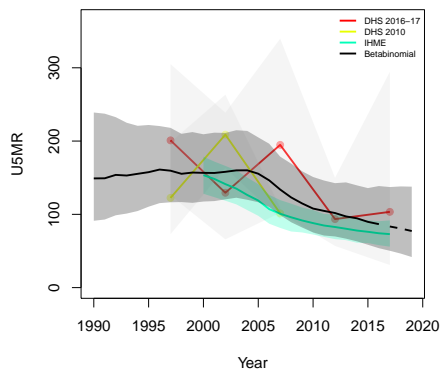
**Buraza, Gitega**



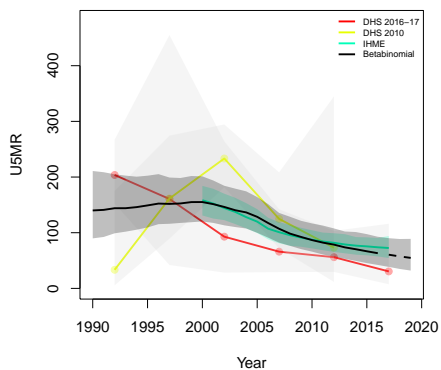
**Nyanrusange, Gitega**



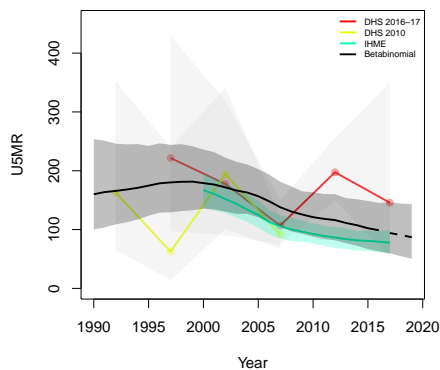
**Gishubi, Gitega**



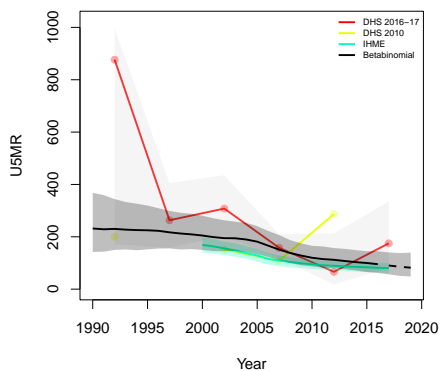
**Nyabihanga, Mwaro**



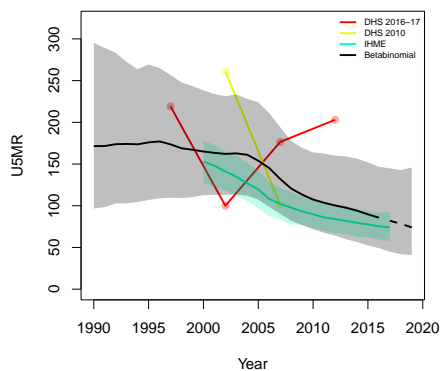
**Giheta, Gitega**



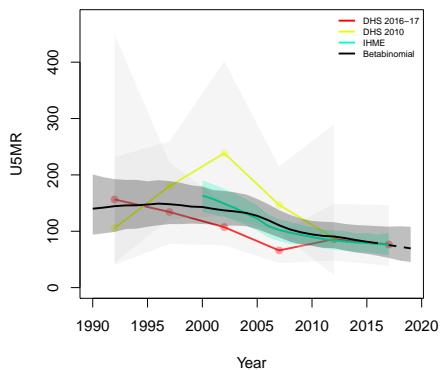
**Itaba, Gitega**



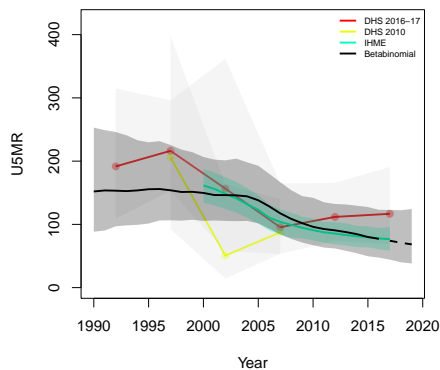
**Bukirasazi, Gitega**



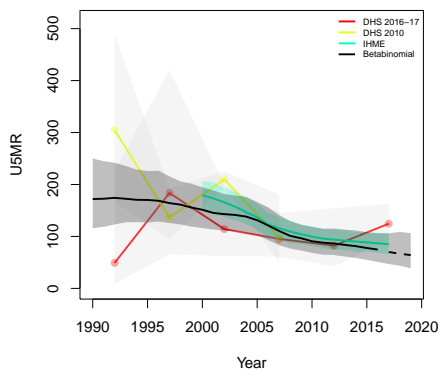
**Gitega, Gitega**



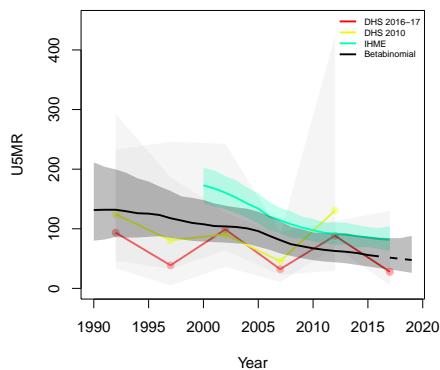
**Makebuko, Gitega**



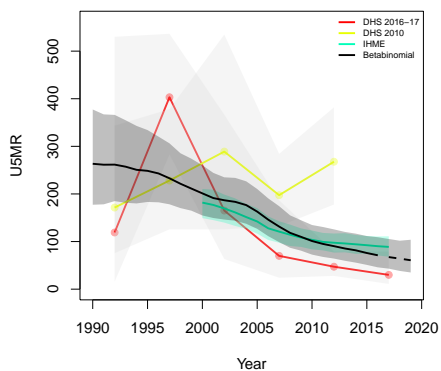
**Butaganzwa1, Ruyigi**



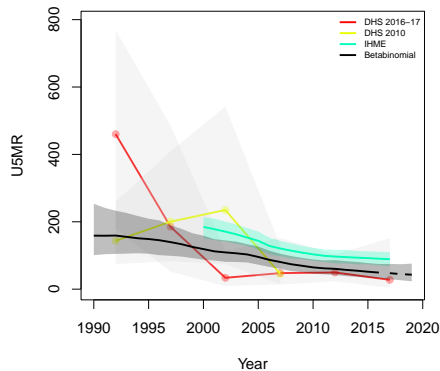
**Mpinga-Kayove, Rutana**



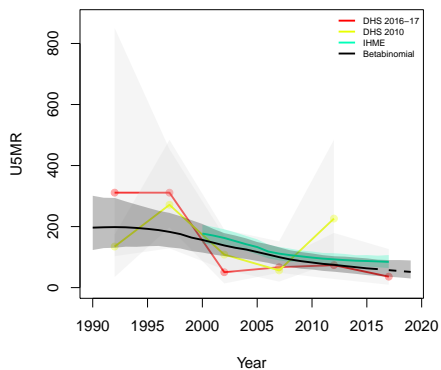
**Nyabitsinda, Ruyigi**



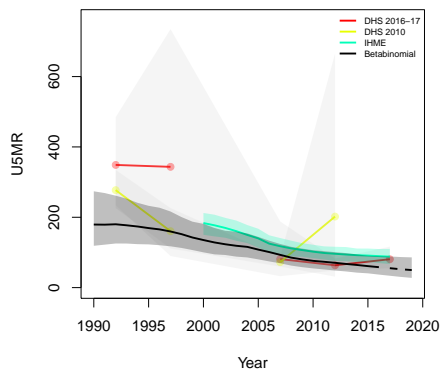
**Ruyigi, Ruyigi**



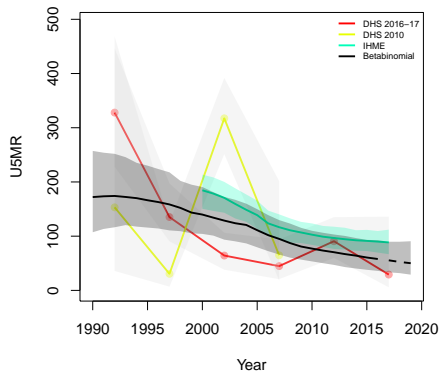
**Shombo, Karuzi**



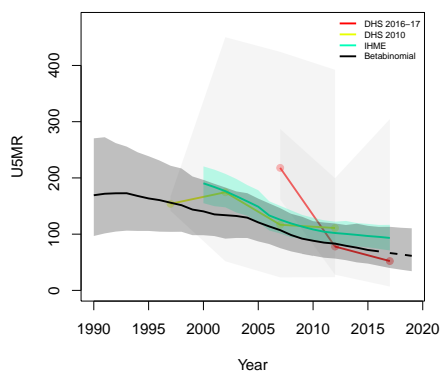
**Butezi, Ruyigi**



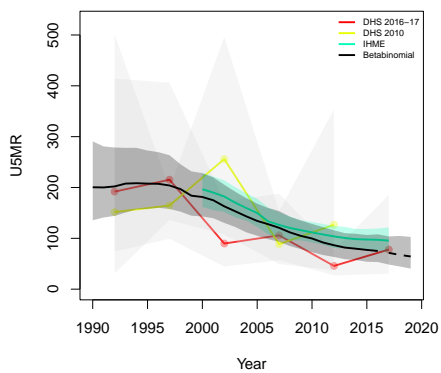
**Nyabikere, Karuzi**



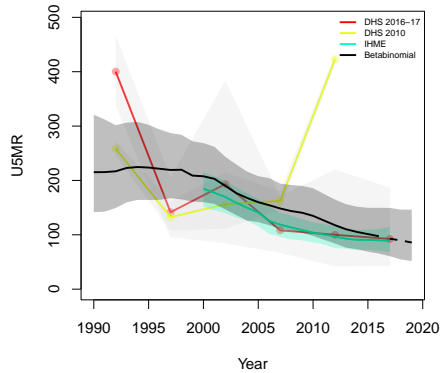
**Bweru, Ruyigi**



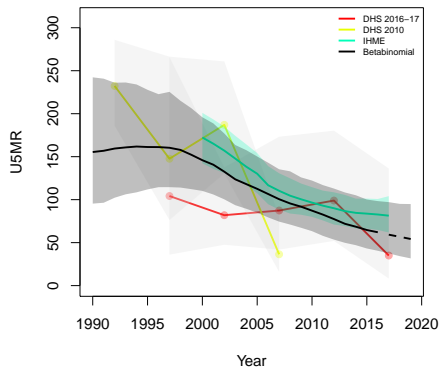
**Gitaramuka, Karuzi**



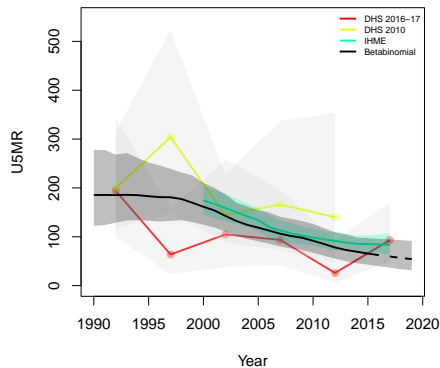
**Gashikanwa, Ngozi**



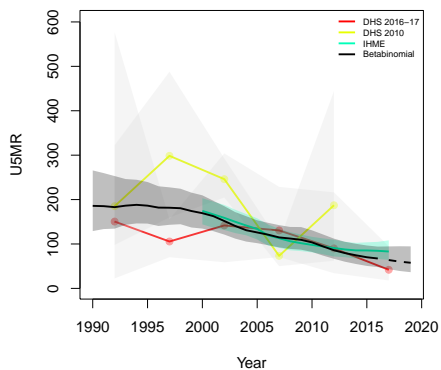
**Mutaho, Gitega**



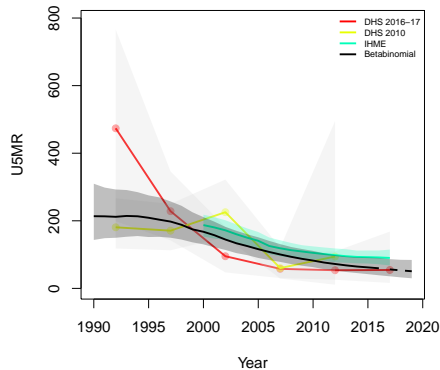
**Muhanga, Kayanza**



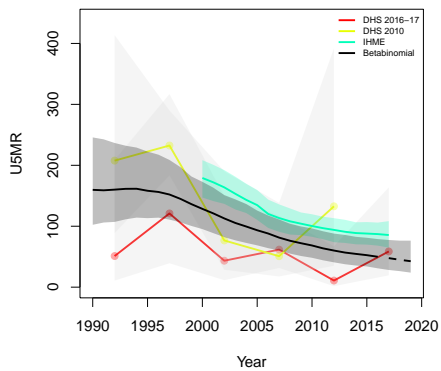
**Ngozi, Ngozi**



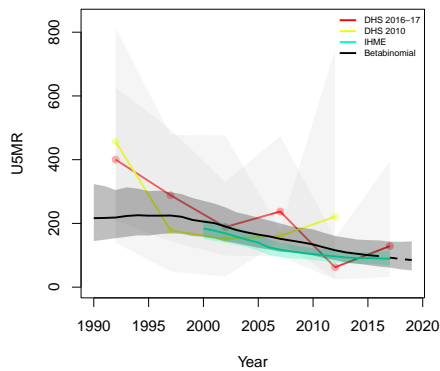
**Bugenyuzi, Karuzi**



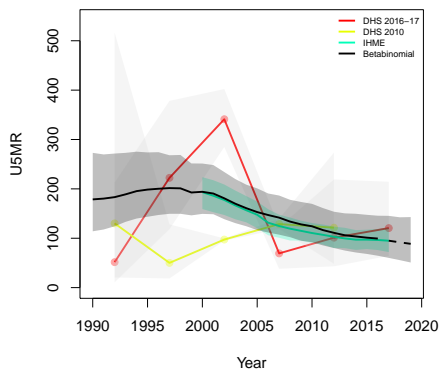
**Gihogazi, Karuzi**



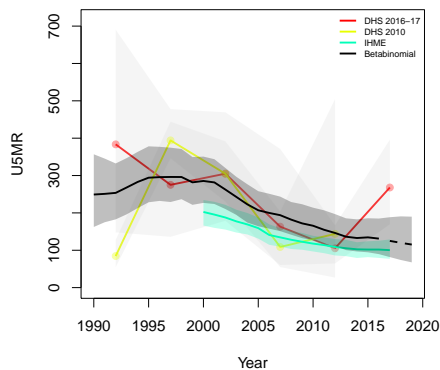
**Ruhororo, Ngozi**



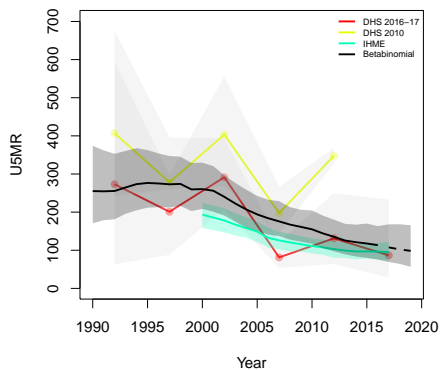
**Tangara, Ngozi**



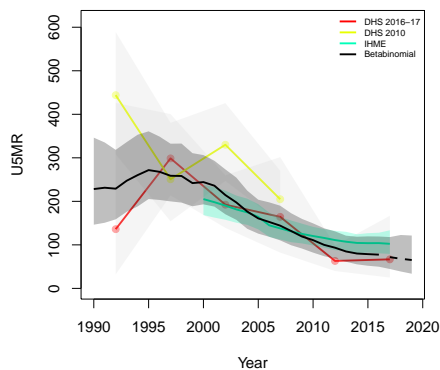
**Vumbi, Kirundo**



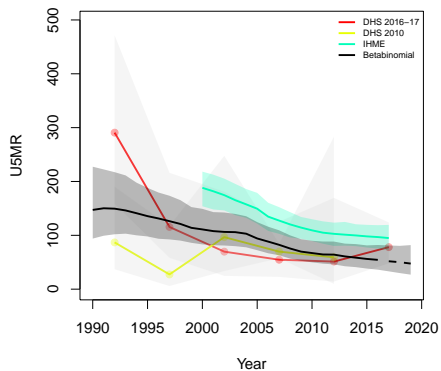
**Kiremba, Ngozi**



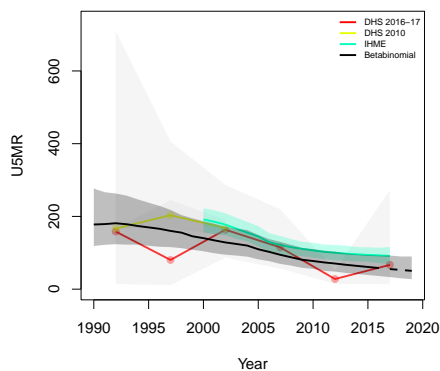
**Kirundo, Kirundo**



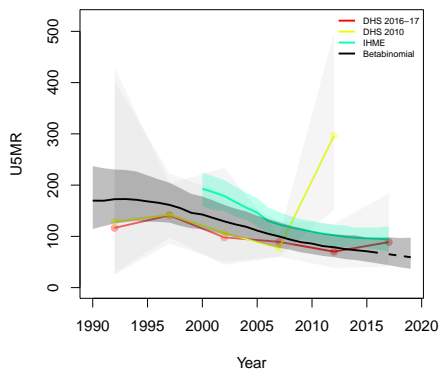
**Cankuzo, Cankuzo**



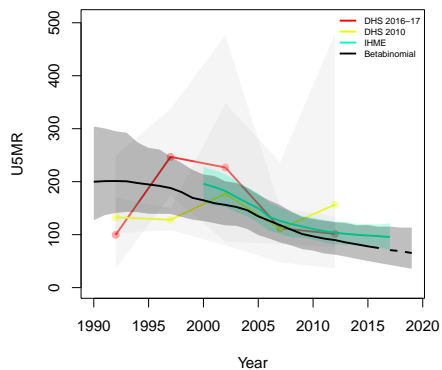
**Mutumba, Karuzi**



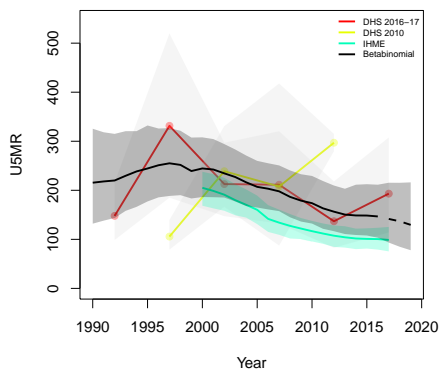
**Buhiga, Karuzi**



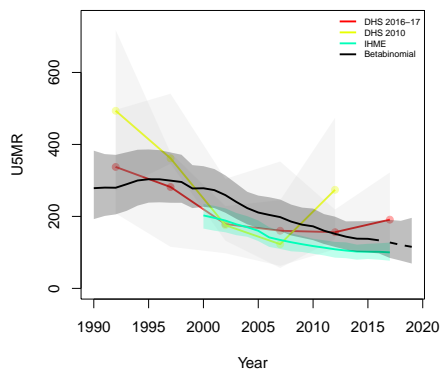
**Mwakiro, Muyinga**



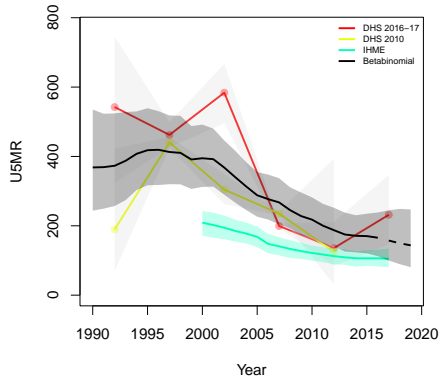
**Gasorwe, Muyinga**



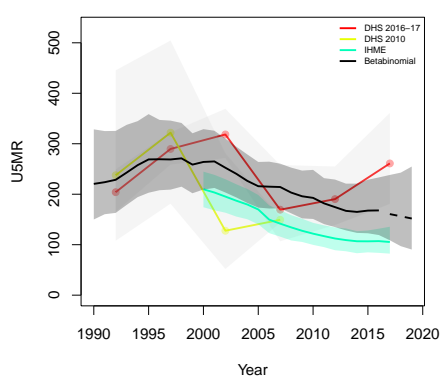
**Gashoho, Muyinga**



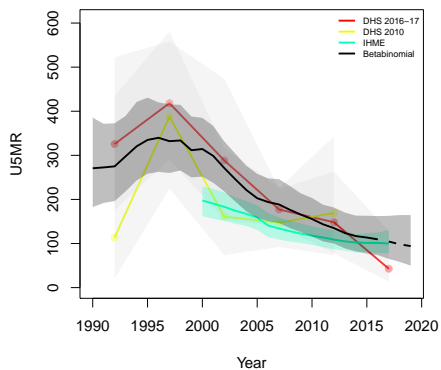
**Gitobe, Kirundo**



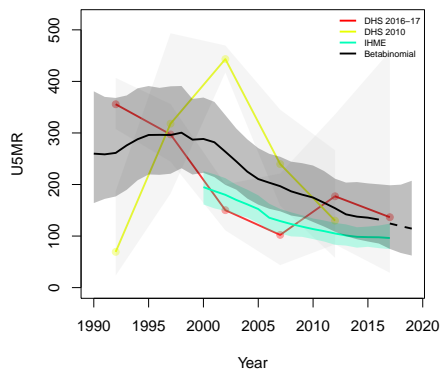
**Bwambarangwe, Kirundo**



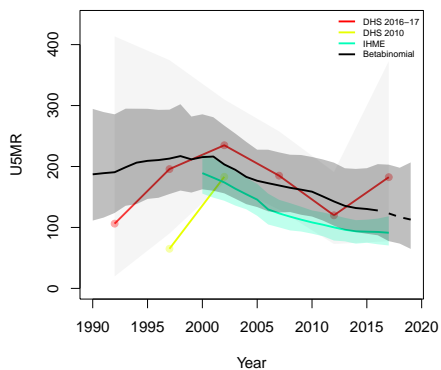
**Ntega, Kirundo**



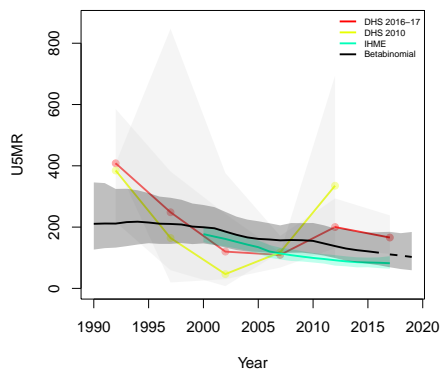
**Marangara, Ngozi**



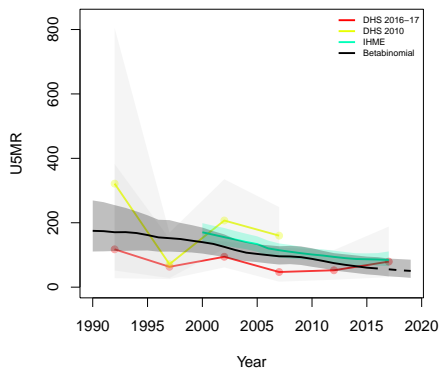
**Nyamurenza, Ngozi**



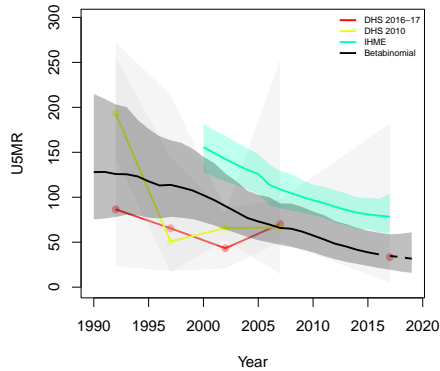
**Mwumba, Ngozi**



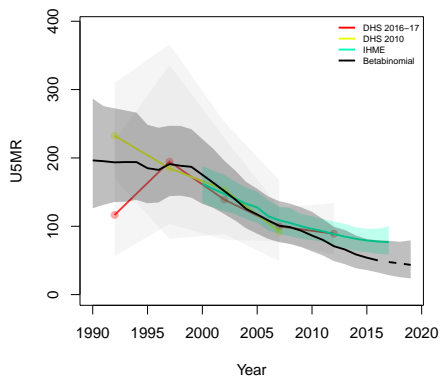
**Busiga, Ngozi**



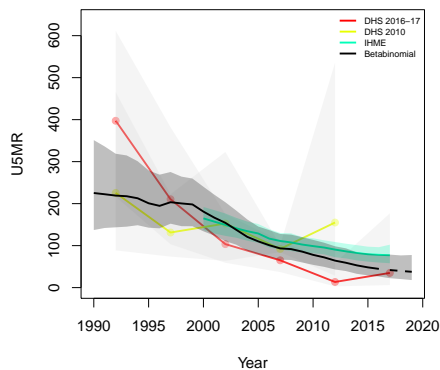
**Kabarore, Kayanza**



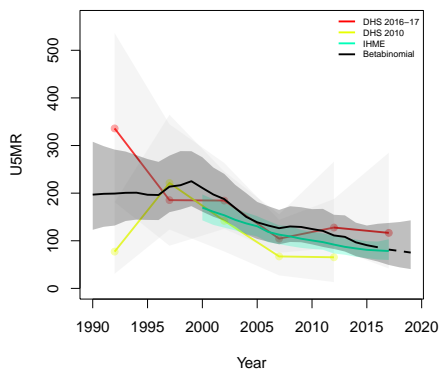
**Bukinanyana, Cibitoke**



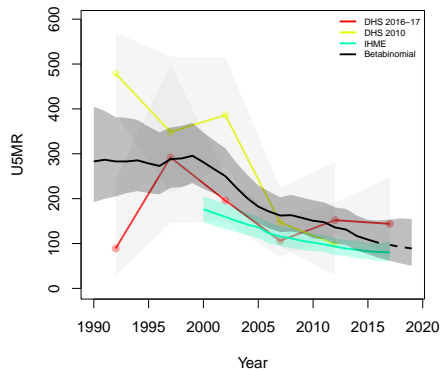
**Mabayi, Cibitoke**



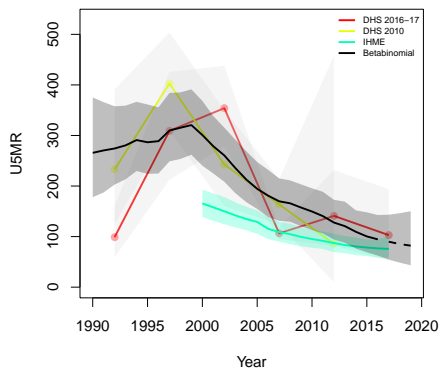
**Mugina, Cibitoke**



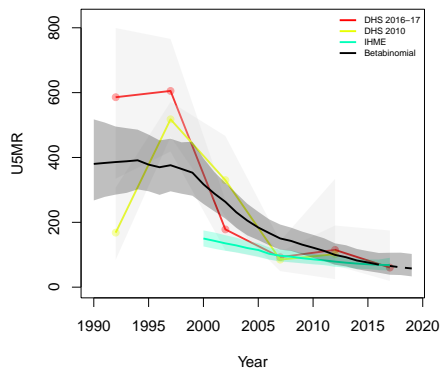
**Rugombo, Cibitoke**



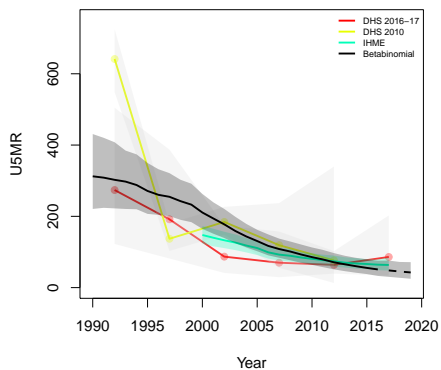
**Buganda, Cibitoke**



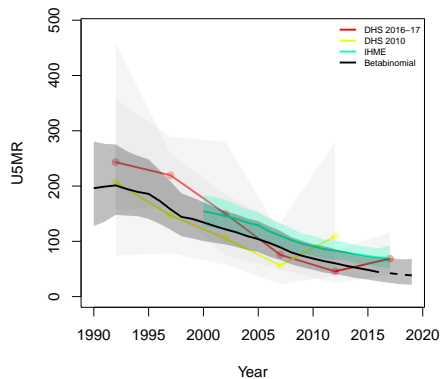
**Gihanga, Bubanza**



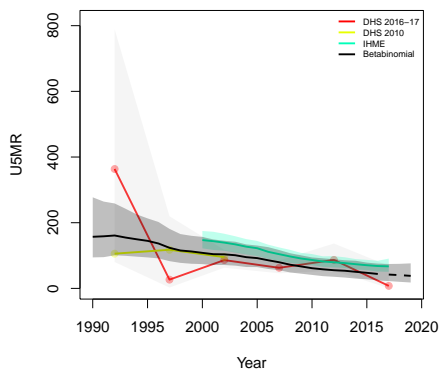
**Mutimbuzi, Bujumbura Rural**



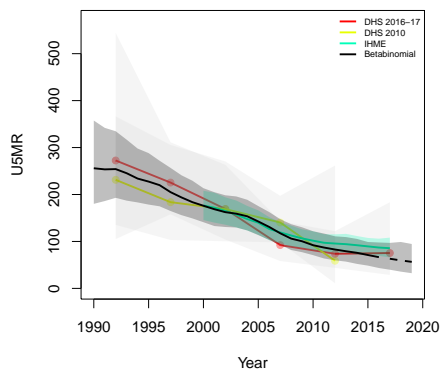
**Nyanza-Lac, Makamba**



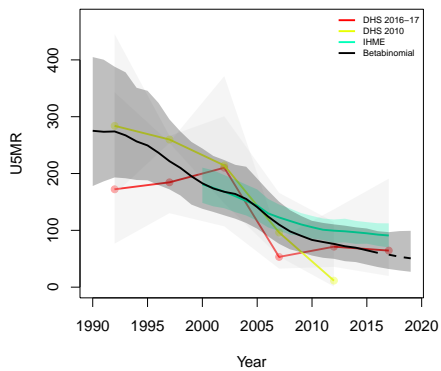
**Mabanda, Makamba**



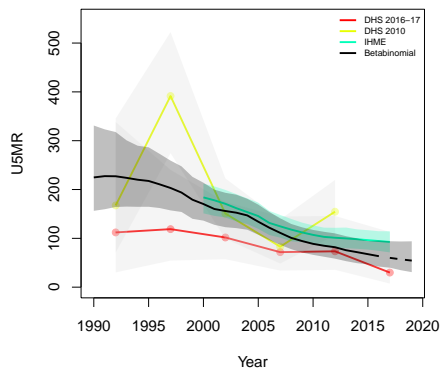
**Giharo, Rutana**



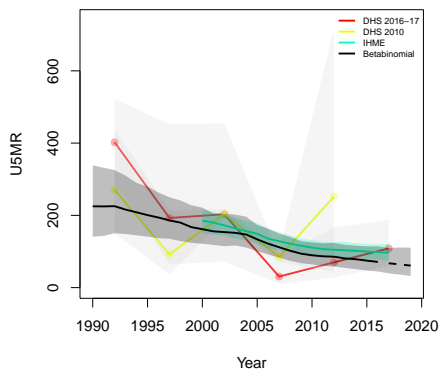
**Kinyinya, Ruyigi**



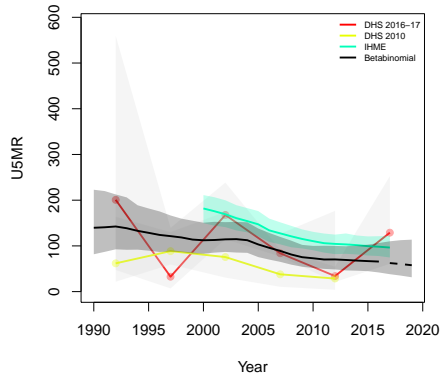
**Gisuru, Ruyigi**



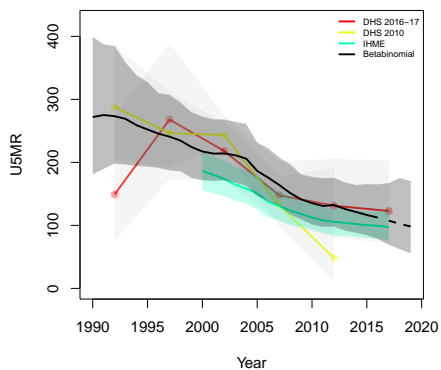
**Cendajuru, Cankuzo**



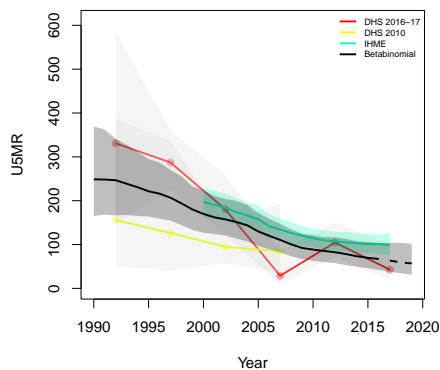
**Gisagara, Cankuzo**



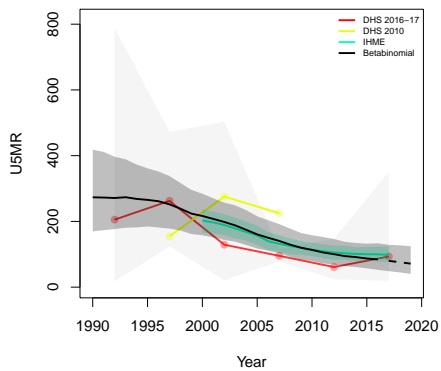
**Mishiha, Cankuzo**



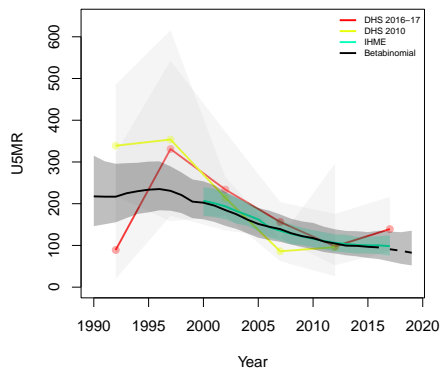
**Kigamba, Cankuzo**



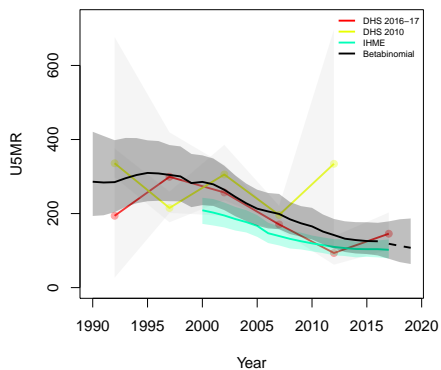
**Buhinyuza, Muyinga**



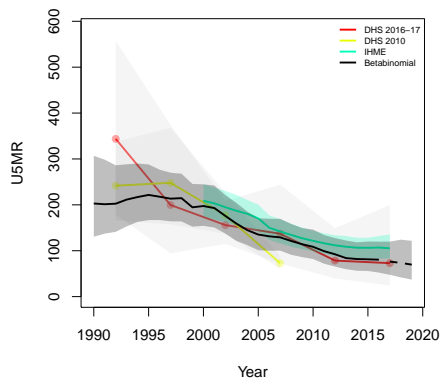
**Muyinga, Muyinga**



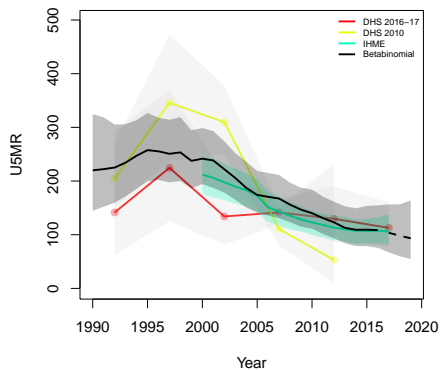
**Butihinda, Muyinga**



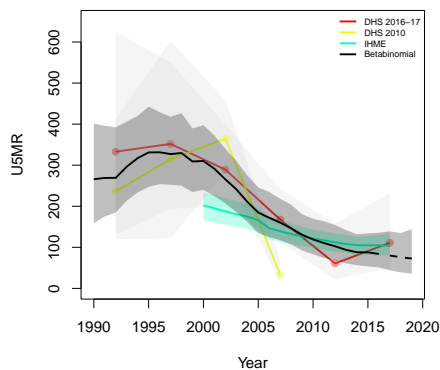
**Giteranyi, Muyinga**



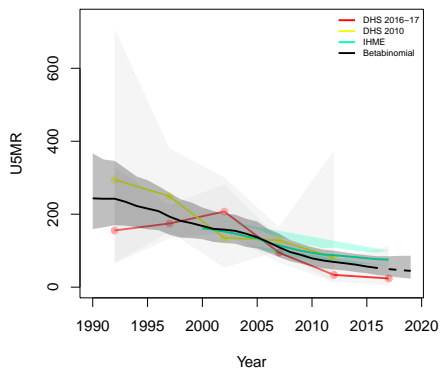
**Busoni, Kirundo**



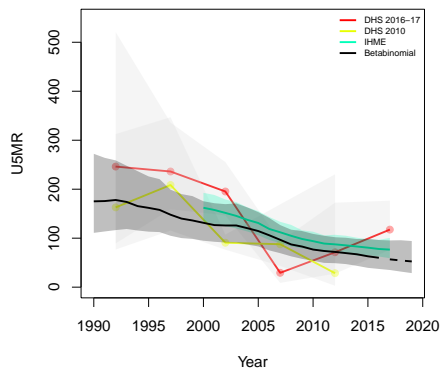
**Bugabira, Kirundo**



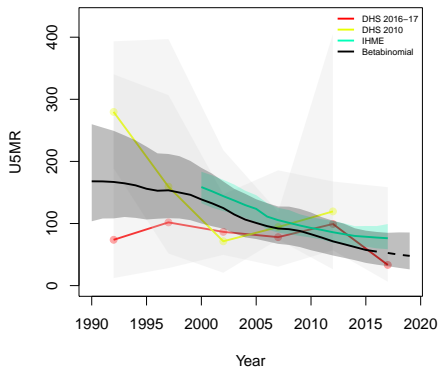
**Kayogoro, Makamba**



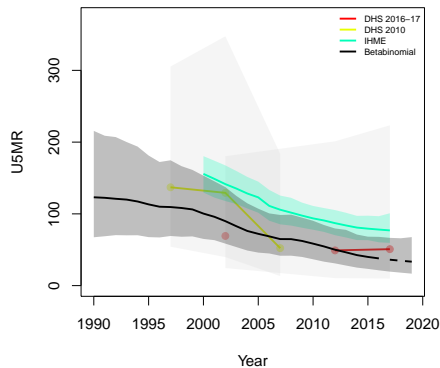
**Bukemba, Rutana**



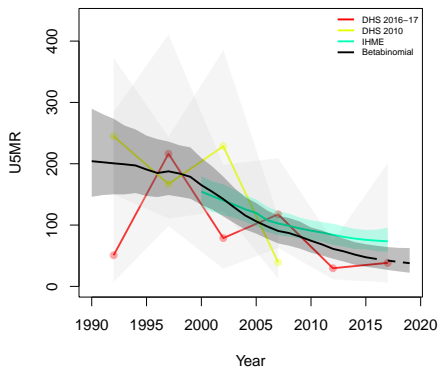
**Matongo, Kayanza**



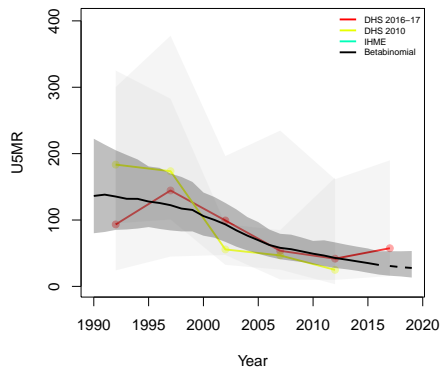
**Muruta, Kayanza**



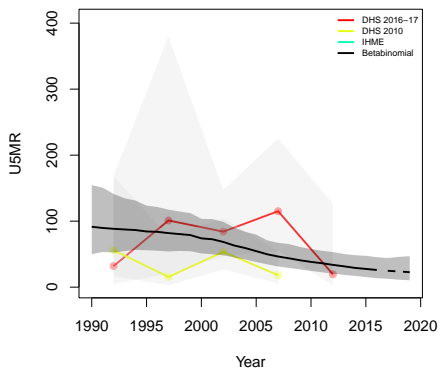
**Musigati, Bubanza**



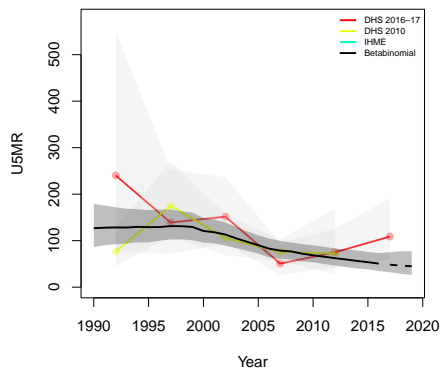
**Muha, Bujumbura Mairie**

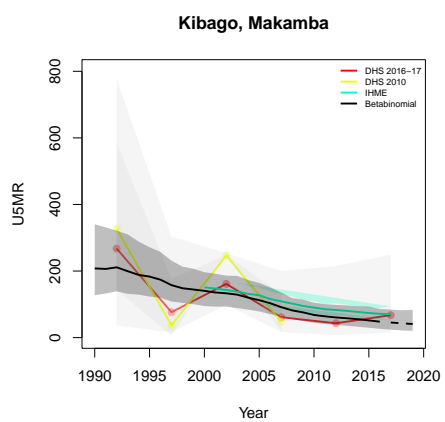


**Mukaza, Bujumbura Mairie**



**Ntangwa, Bujumbura Mairie**





### ***B.3 Ethiopia***

Age	Survey	Clusters			Deaths			Agemonths		
		Urban	Rural	Total	Urban	Rural	Total	Urban	Rural	Total
0	2000	138	397	535	149	1079	1228	3477	18653	22130
	2005	144	384	528	127	1061	1188	4276	24550	28826
	2011	163	408	571	197	1695	1892	5942	33099	39041
	2016	202	420	622	207	1391	1598	7325	31236	38561
1-11	2000	138	397	535	176	1008	1184	35736	186414	222150
	2005	144	384	528	121	1028	1149	44774	247208	291982
	2011	163	408	571	160	1333	1493	61231	328813	390044
	2016	202	420	622	178	978	1156	75735	314797	390532
12-23	2000	138	397	535	84	485	569	37078	188655	225733
	2005	144	384	528	62	464	526	47098	250710	297808
	2011	163	408	571	95	573	668	62784	332955	395739
	2016	202	420	622	72	351	423	76641	319444	396085
24-35	2000	138	397	535	45	394	439	36392	181161	217553
	2005	144	384	528	34	364	398	46280	239329	285609
	2011	163	408	571	45	514	559	59781	314920	374701
	2016	202	420	622	49	320	369	72065	300895	372960
36-47	2000	138	397	535	24	342	366	35621	171355	206976
	2005	144	384	528	26	370	396	45840	226104	271944
	2011	163	408	571	48	417	465	56485	292012	348497
	2016	202	420	622	41	237	278	68071	280502	348573
48-59	2000	138	397	535	23	224	247	34838	160598	195436
	2005	144	384	528	21	211	232	44466	209422	253888
	2011	163	408	571	24	221	245	53324	270377	323701
	2016	202	420	622	19	138	157	63639	260425	324064

Table B.3: **Data summary for Ethiopia.** Total numbers of clusters (Columns 3–5) with observations in each age group by survey in urban and rural areas and combined. Numbers of deaths (Columns 6–8) and number of agemonths (Columns 9–10) observed in each age group by survey in urban and rural areas and combined.

*B.3.1 Admin-1*

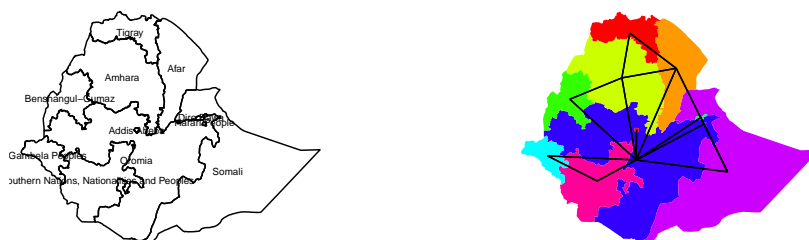


Figure B.13: **Left:** The names of the 11 Admin-1 areas of Ethiopia . **Right:** The neighborhood structure of Admin-1 areas in Ethiopia .

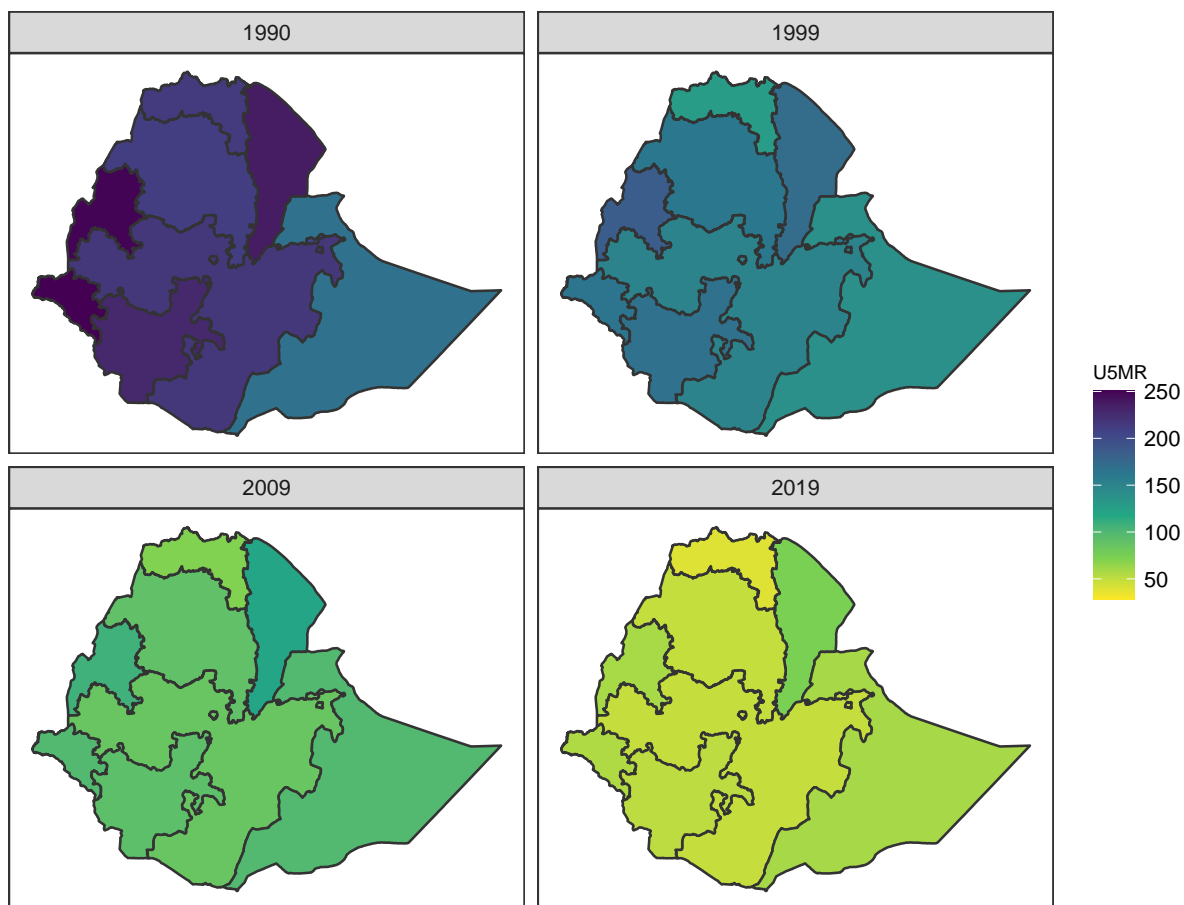


Figure B.14: Median U5MR estimates for years 1990, 1999, 2009, 2019 for Admin-1 areas in Ethiopia .

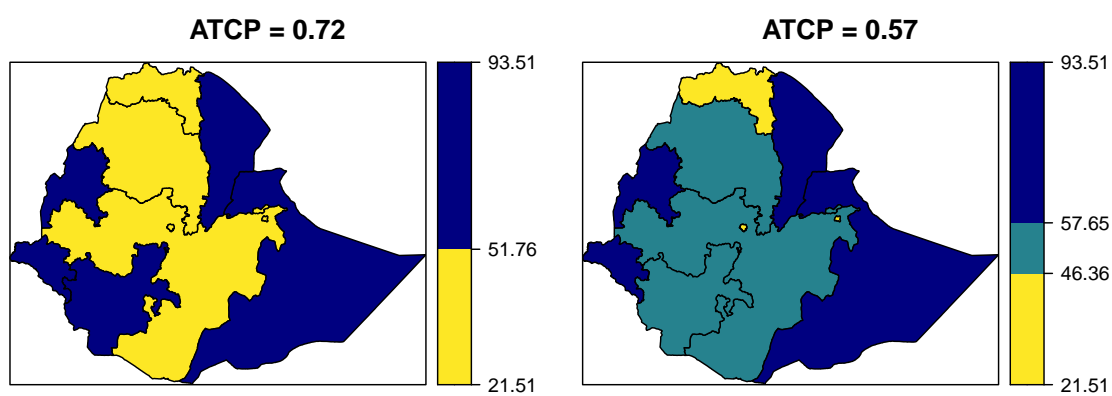
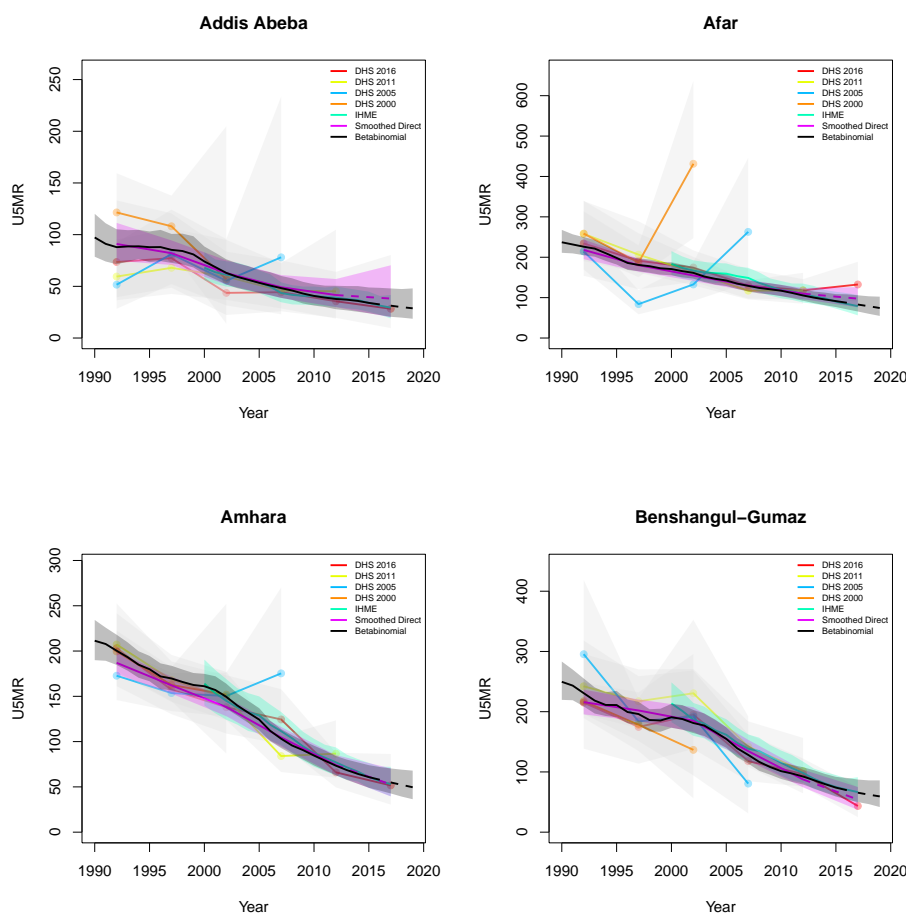


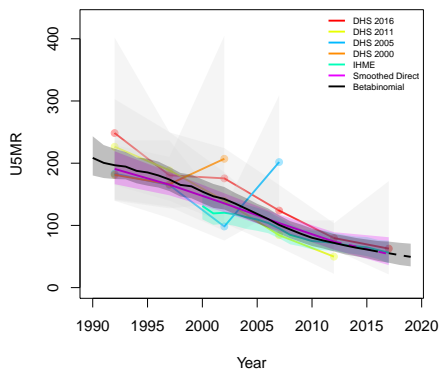
Figure B.15: Expression of uncertainty of U5MR (deaths per 1000 children) estimates for Admin-1 areas based on the average true classification probability (ATCP) in 2019 using  $K = 2, 3$  colors.

*Data and estimates over time by area*

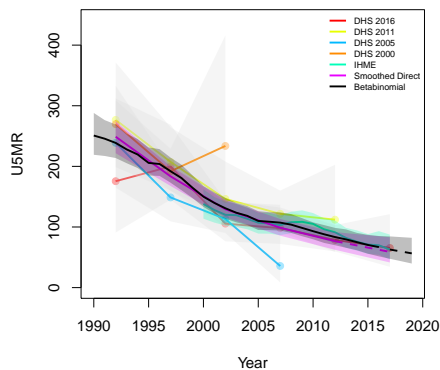
Colored lines with circular points and light grey uncertainty bands are 5-year survey-weighted estimates of U5MR for years 1990–1994 up to 2015–2019 depending on survey timing. For a survey that ends in the middle of a 5-year period, we plot the estimates at the mid-point of the years in that interval for which the survey provides data. Black lines and corresponding intervals represent posterior medians and 95% uncertainty intervals respectively for the betabinomial model. IHME’s estimates and corresponding intervals, where we can compare, are in aquamarine.



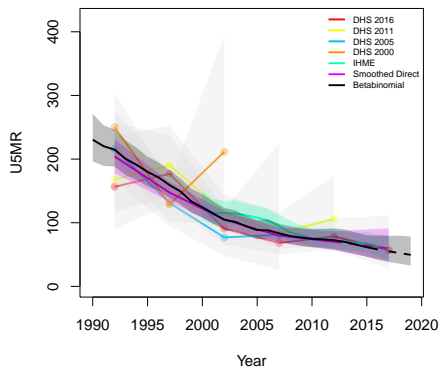
**Dire Dawa**



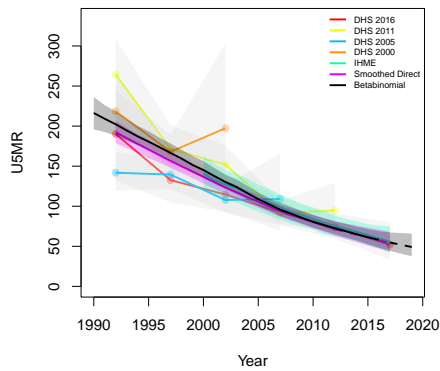
**Gambela Peoples**



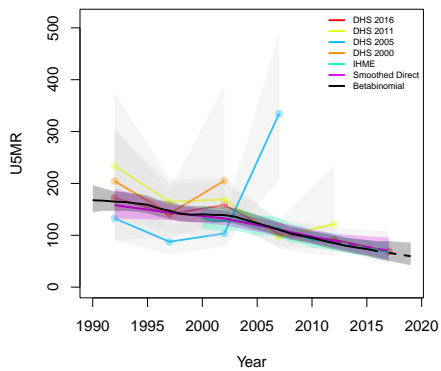
**Harari People**



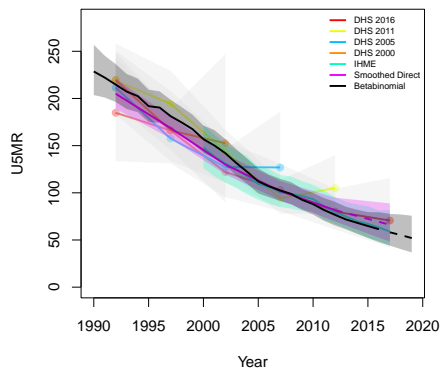
**Oromia**

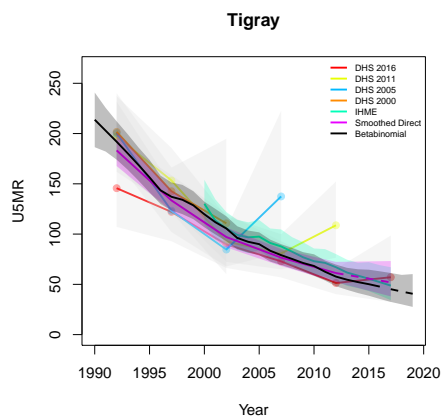


**Somali**



**Southern Nations, Nationalities and Peoples**





*B.3.2 Admin-2*

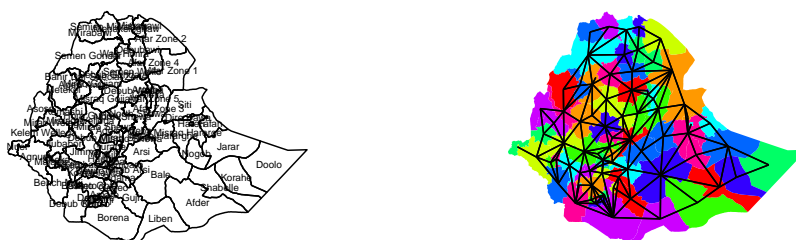


Figure B.16: **Left:** The names of the 79 Admin-2 areas of Ethiopia . **Right:** The neighborhood structure of Admin-2 areas in Ethiopia .

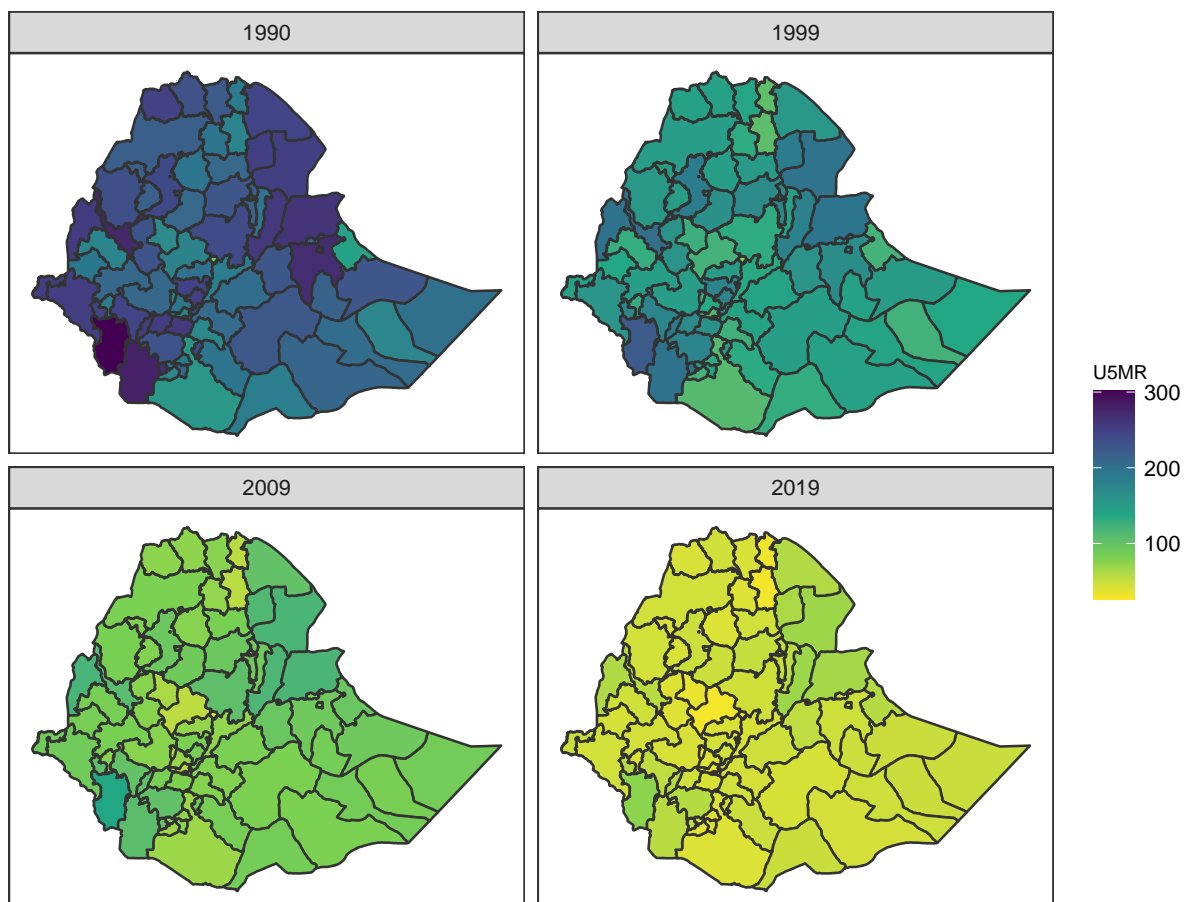


Figure B.17: Median U5MR estimates for years 1990, 1999, 2009, 2019 for Admin-2 areas in Ethiopia .

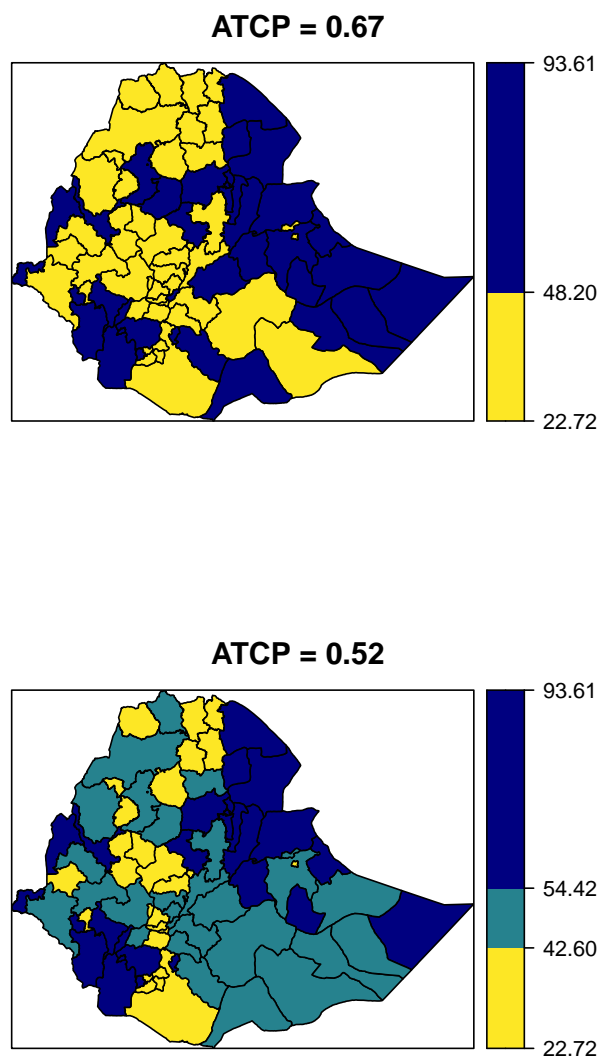
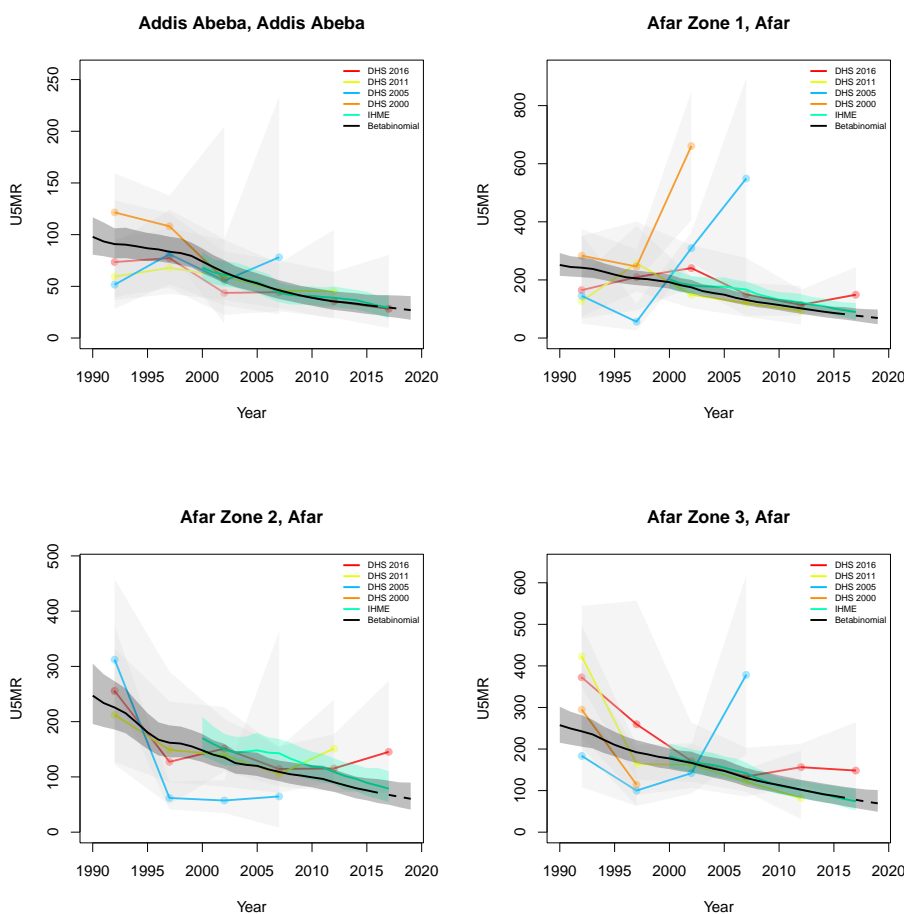


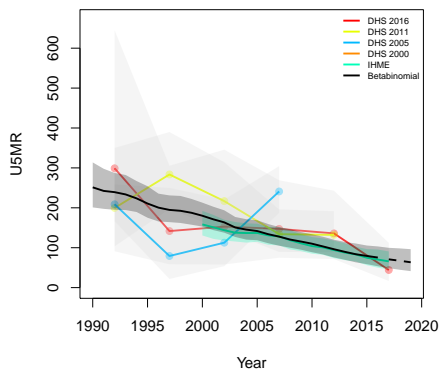
Figure B.18: Expression of uncertainty of U5MR (deaths per 1000 children) estimates for Admin-1 areas based on the average true classification probability (ATCP) in 2019 using  $K = 2, 3$  colors.

*Data and estimates over time by area*

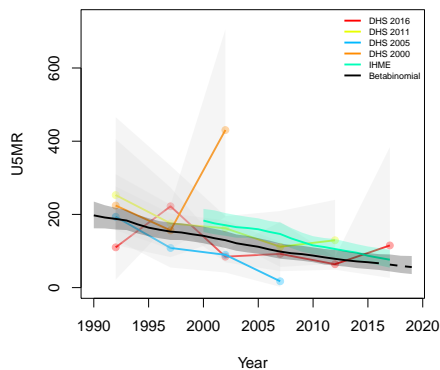
Colored lines with circular points and light grey uncertainty bands are 5-year survey-weighted estimates of U5MR for years 1990–1994 up to 2015–2019 depending on survey timing. For a survey that ends in the middle of a 5-year period, we plot the estimates at the mid-point of the years in that interval for which the survey provides data. Black lines and corresponding intervals represent posterior medians and 95% uncertainty intervals respectively for the betabinomial model. IHME’s estimates and corresponding intervals, where we can compare, are in aquamarine.



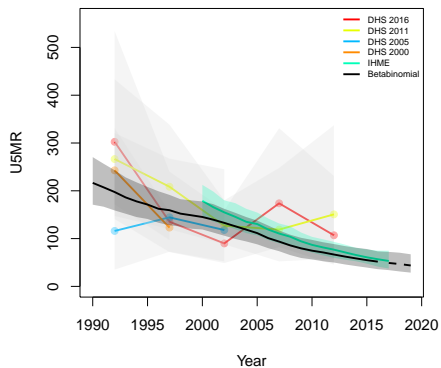
**Afar Zone 4, Afar**



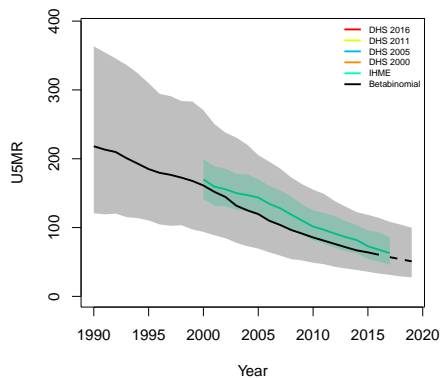
**Afar Zone 5, Afar**



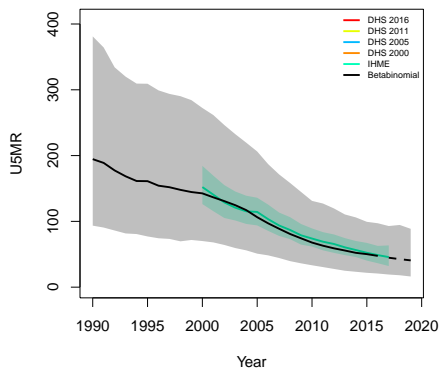
**Agew Awi, Amhara**



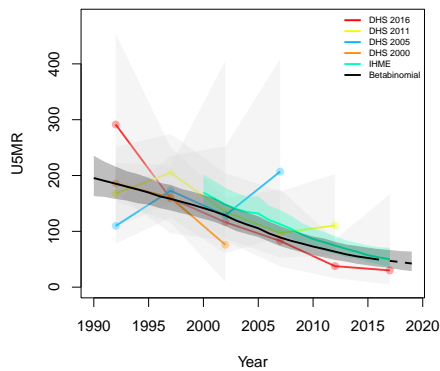
**Argoba, Amhara**



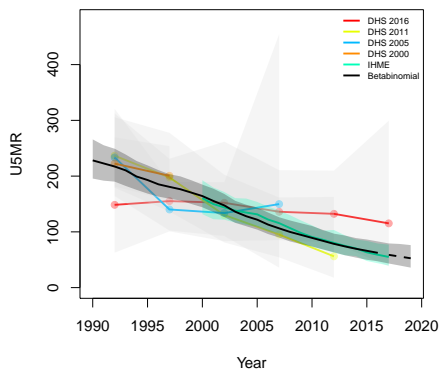
**Bahir Dar Special Zone, Amhara**



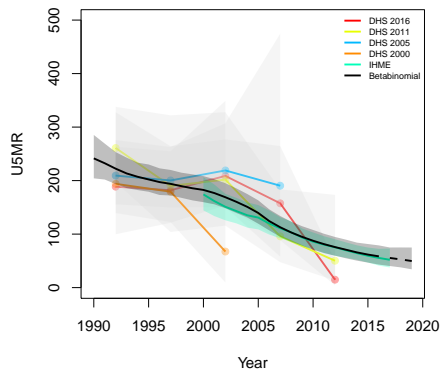
**Debub Gondar, Amhara**



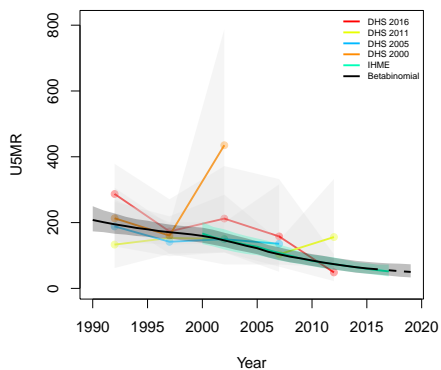
**Debub Wollo, Amhara**



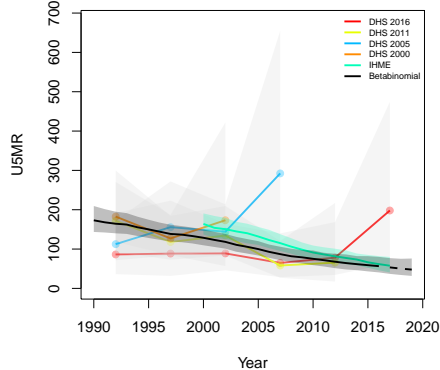
**Mirab Gojjam, Amhara**



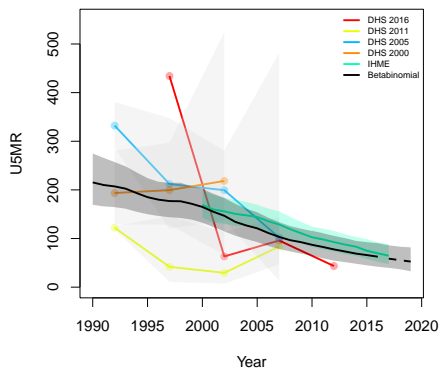
**Misraq Gojjam, Amhara**



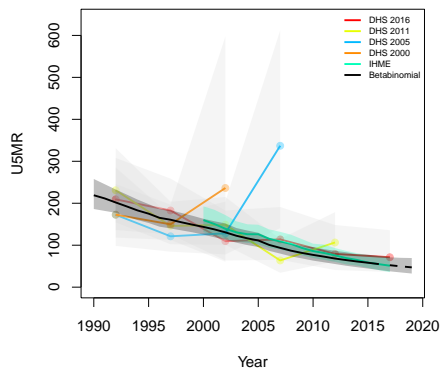
**North Shewa, Amhara**



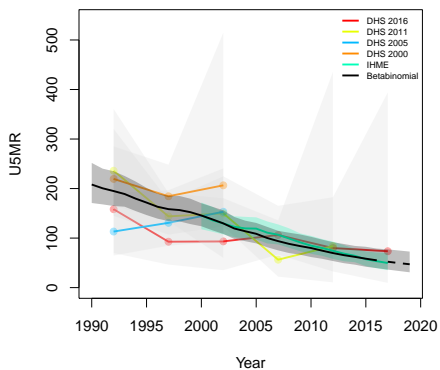
**Oromia, Amhara**



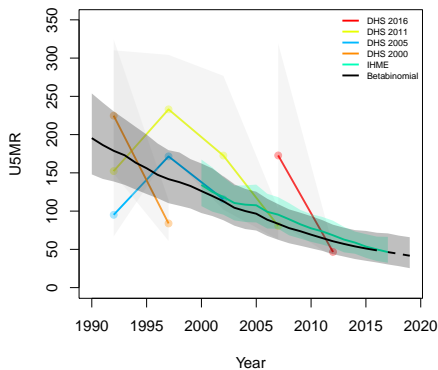
**Semen Gondar, Amhara**



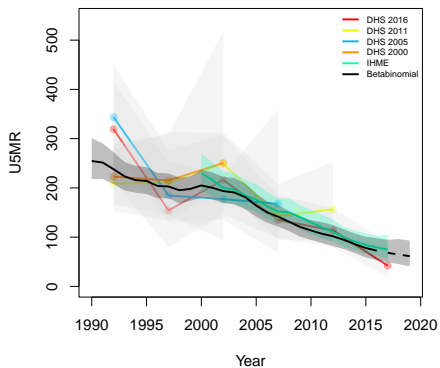
**Semen Wello, Amhara**



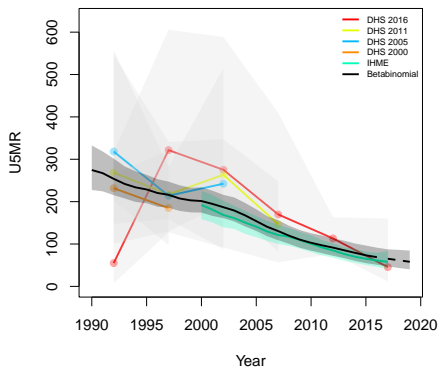
**Wag Himra, Amhara**



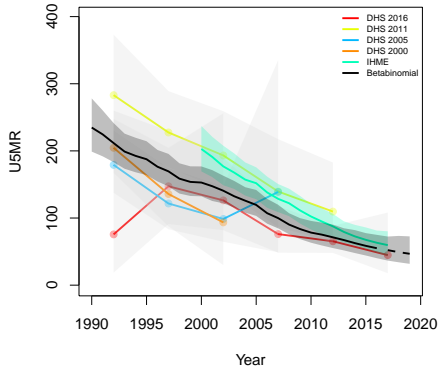
**Asosa, Benshangul-Gumaz**



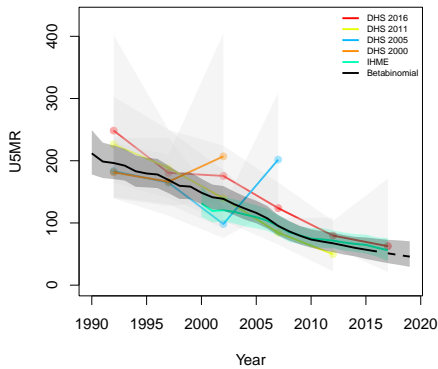
**Kemashi, Benshangul-Gumaz**



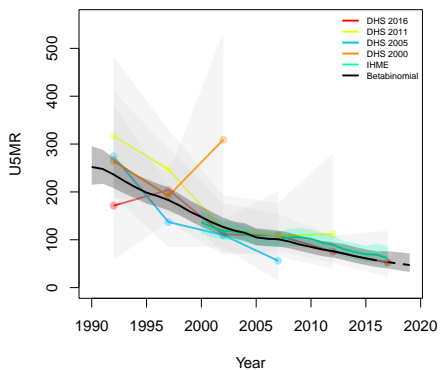
**Metekel, Benshangul-Gumaz**



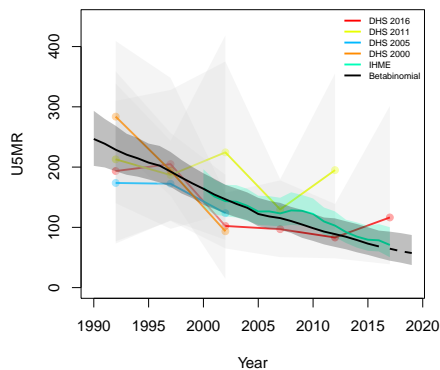
**Dire Dawa, Dire Dawa**



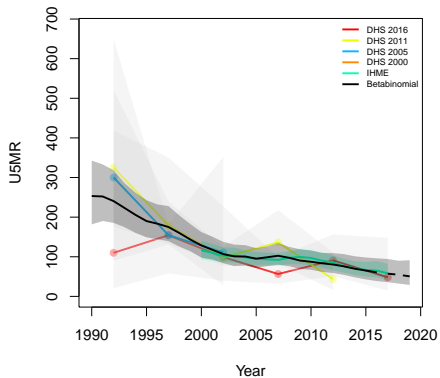
**Agnuak, Gambela Peoples**



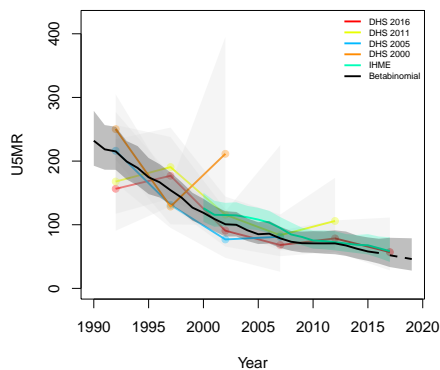
**Majang, Gambela Peoples**



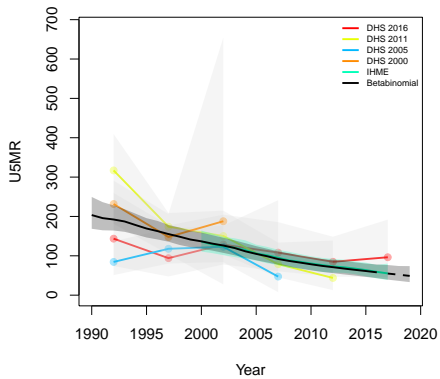
**Nuer, Gambela Peoples**



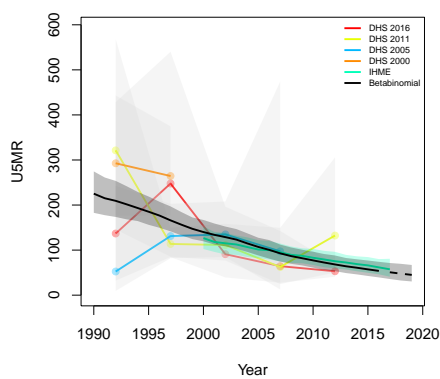
**Hareri, Harari People**



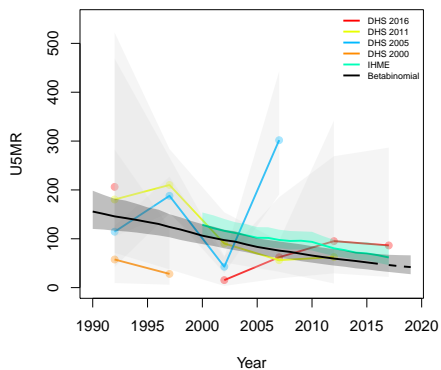
**Arsi, Oromia**



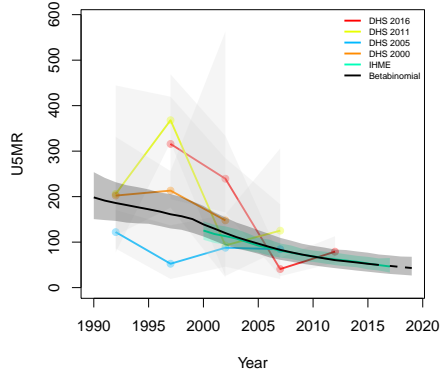
**Bale, Oromia**



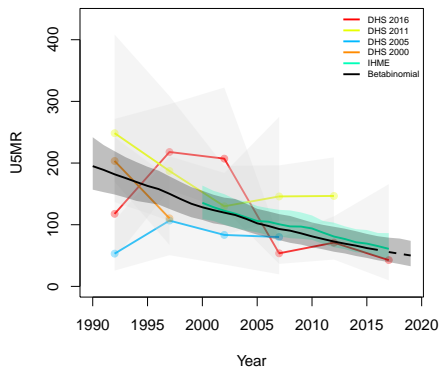
**Borena, Oromia**



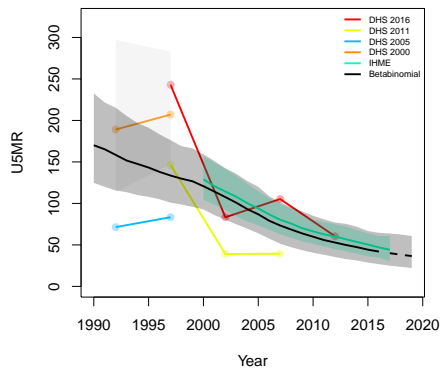
**Debub Mirab Shewa, Oromia**



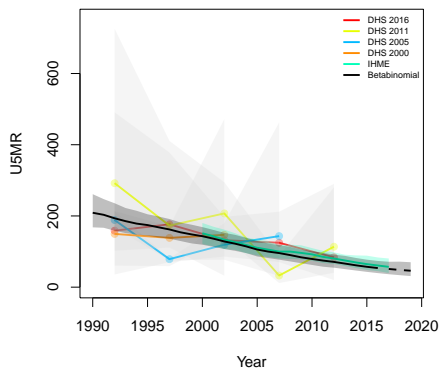
**Guji, Oromia**



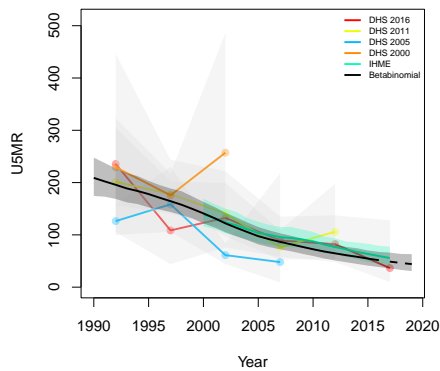
**Horo Guduru, Oromia**



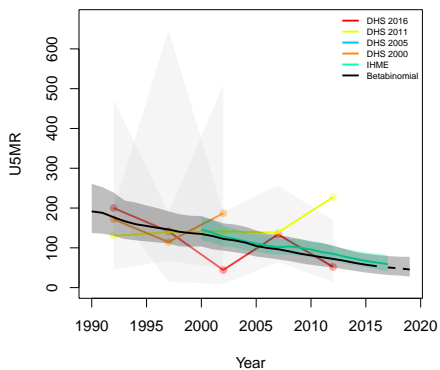
**Ilubabor, Oromia**



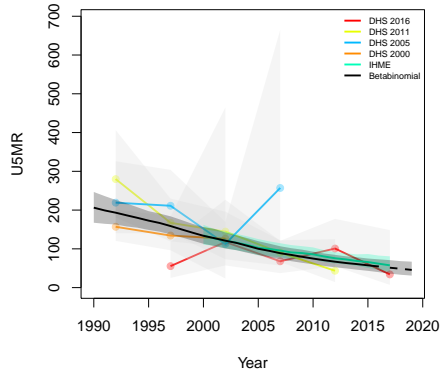
**Jimma, Oromia**



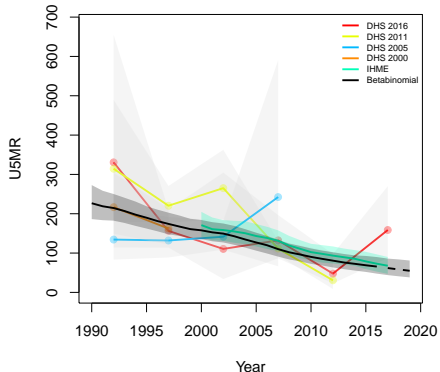
**Kelem Wellega, Oromia**



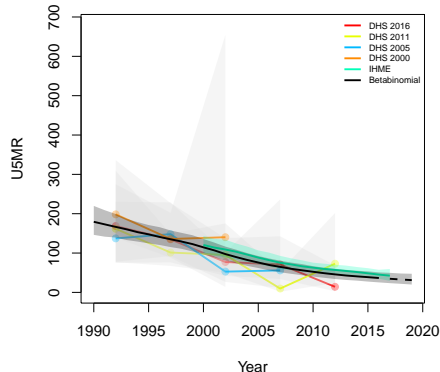
**Mirab Arsi, Oromia**



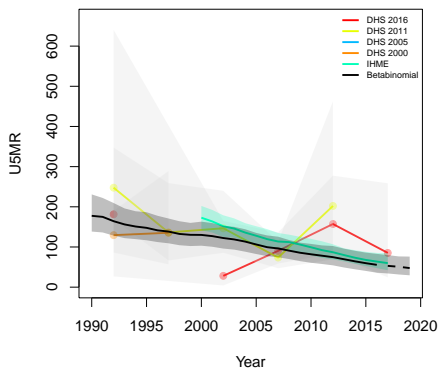
**Mirab Hararghe, Oromia**



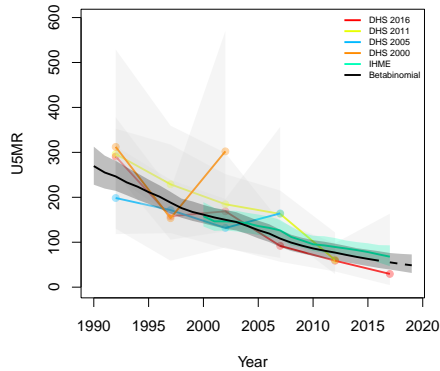
**Mirab Shewa, Oromia**



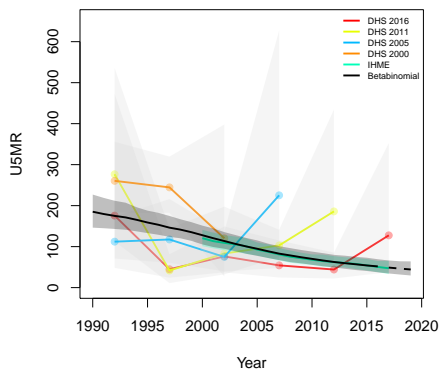
**Mirab Welega, Oromia**



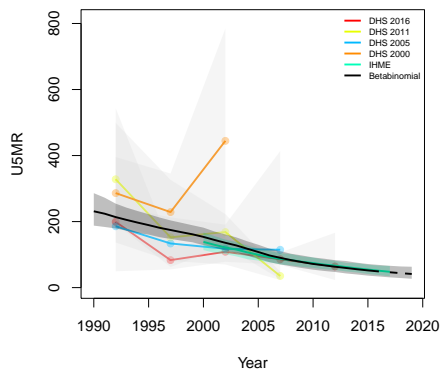
**Misraq Harerge, Oromia**



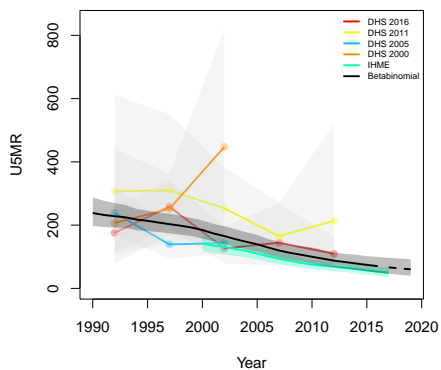
**Misraq Shewa, Oromia**



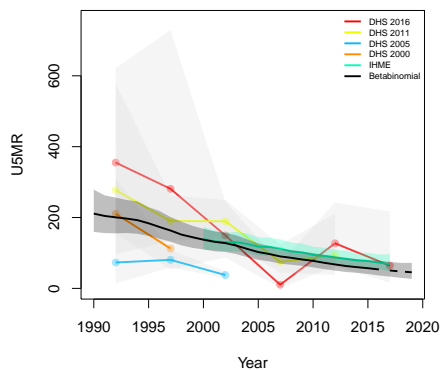
**Misraq Wellega, Oromia**



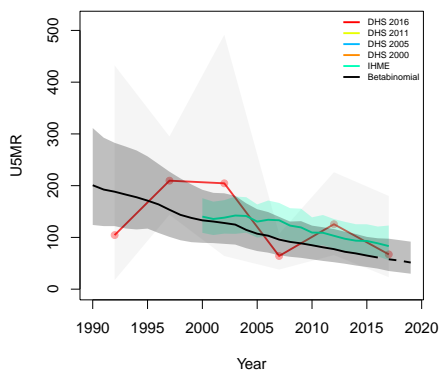
**North Shewa, Oromia**



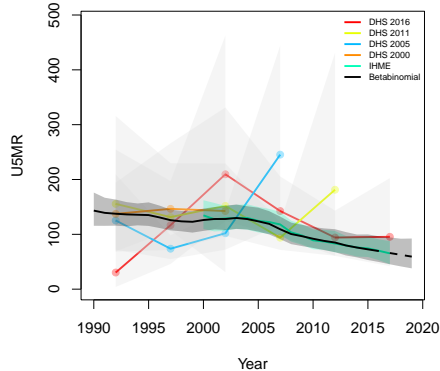
**Afder, Somali**

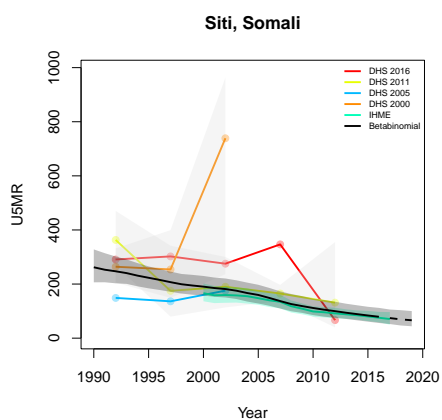
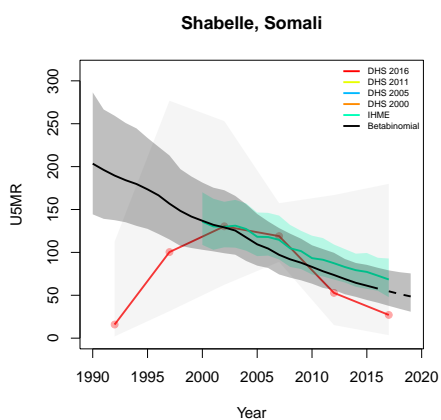
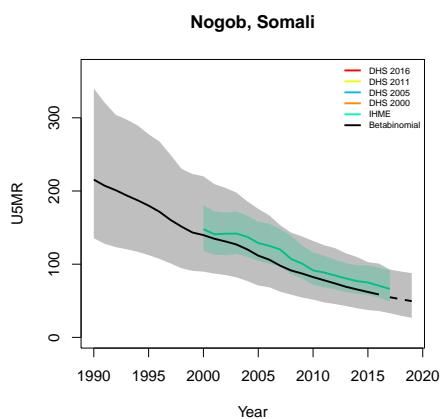
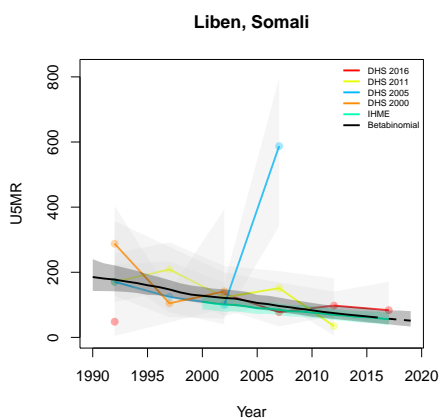
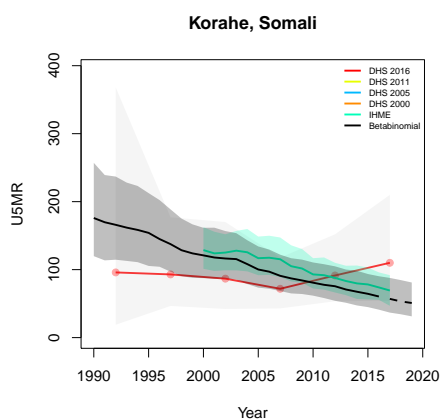
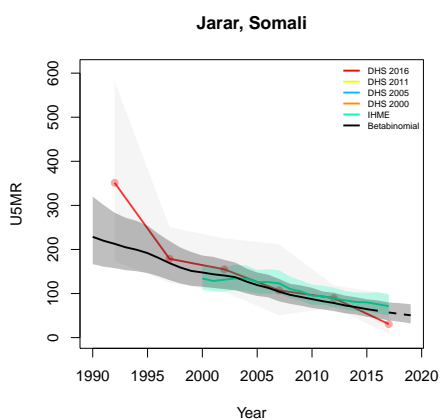


**Doolo, Somali**

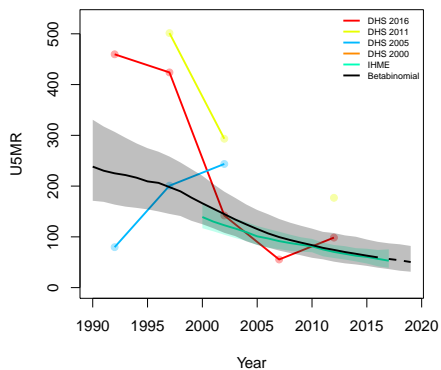


**Fafan, Somali**

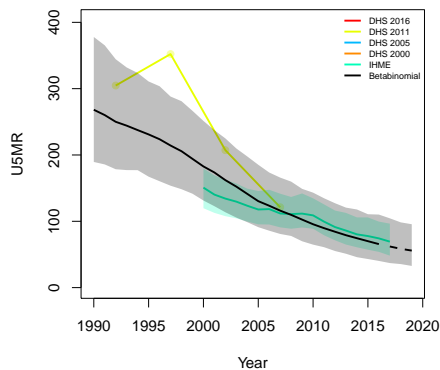




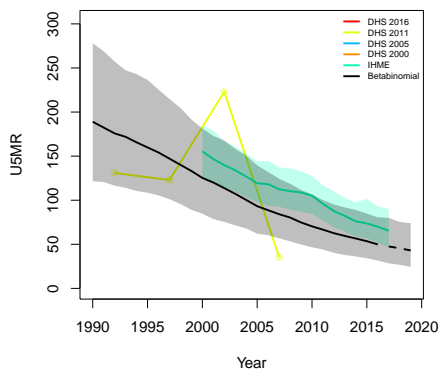
**Alaba, Southern Nations, Nationalities and Peoples**



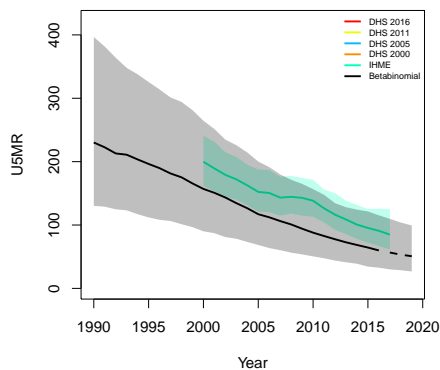
**Alle, Southern Nations, Nationalities and Peoples**



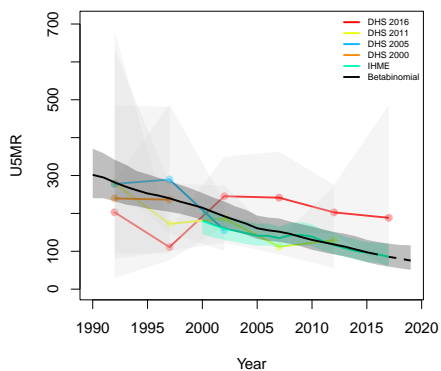
**Amaro, Southern Nations, Nationalities and Peoples**



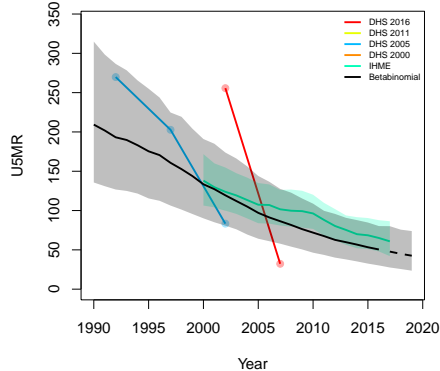
**Basketo, Southern Nations, Nationalities and Peoples**



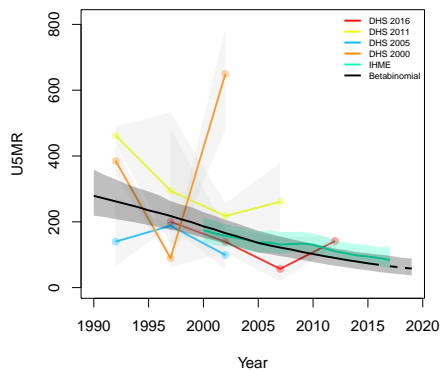
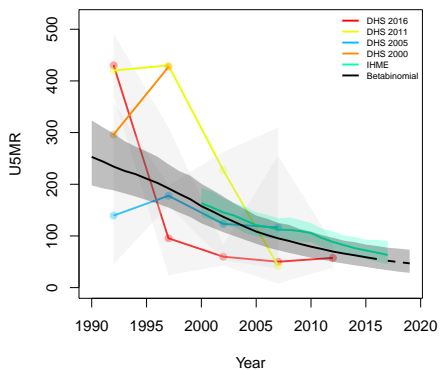
**Bench Maji, Southern Nations, Nationalities and Peoples**



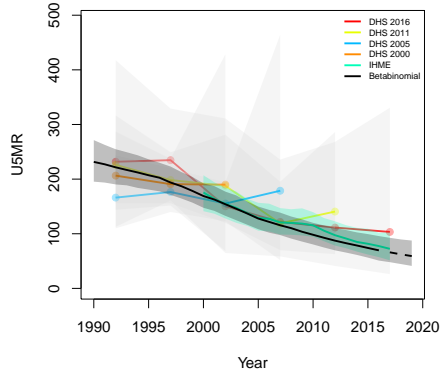
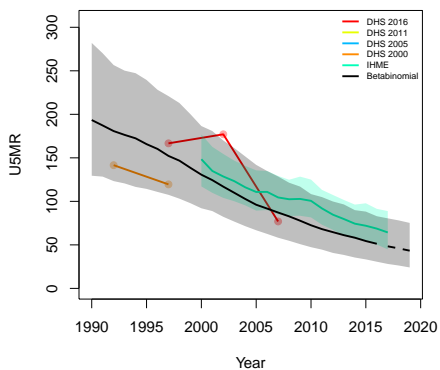
**Burji, Southern Nations, Nationalities and Peoples**



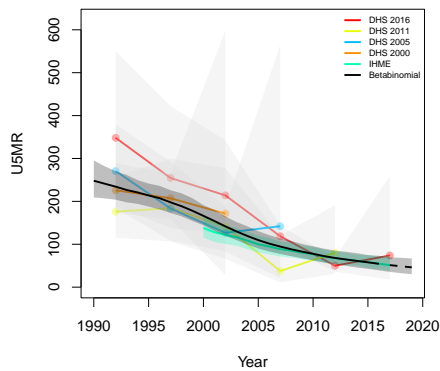
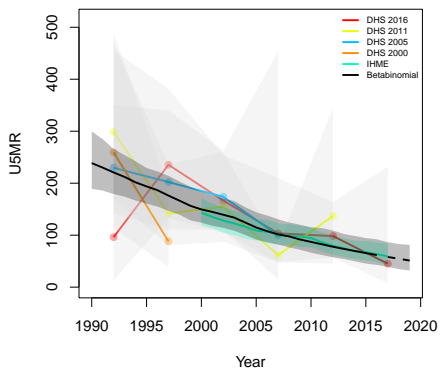
**Dawro, Southern Nations, Nationalities and People: Dehub Omo, Southern Nations, Nationalities and People**



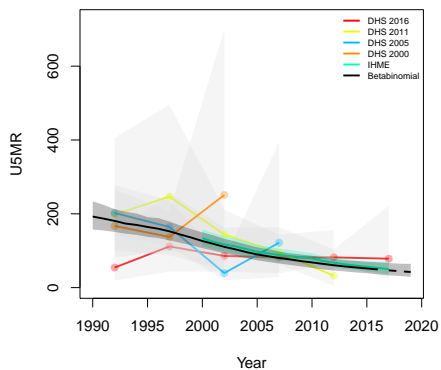
**Derashe, Southern Nations, Nationalities and People: Gamo Gofa, Southern Nations, Nationalities and People**



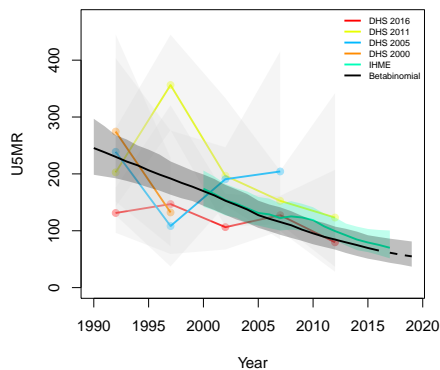
**Gedeo, Southern Nations, Nationalities and People: Gurage, Southern Nations, Nationalities and People**



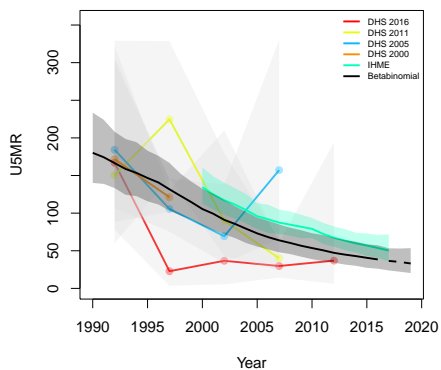
**Hadiya, Southern Nations, Nationalities and People**



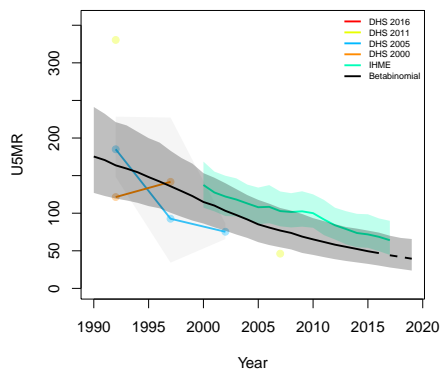
**Keffa, Southern Nations, Nationalities and Peoples**



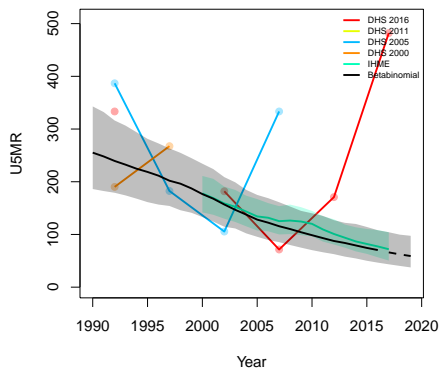
**mbata Tembaro, Southern Nations, Nationalities and P**



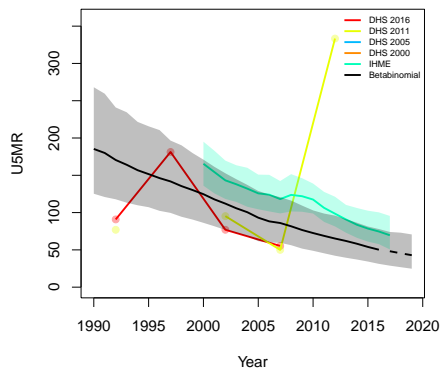
**Konso, Southern Nations, Nationalities and People**



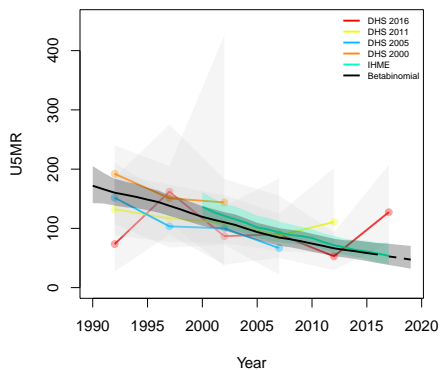
**Konta, Southern Nations, Nationalities and People:**



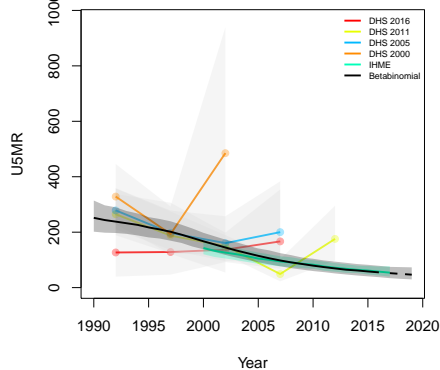
**Sheka, Southern Nations, Nationalities and People:**



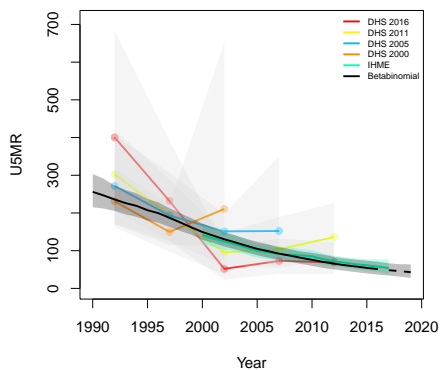
**Sidama, Southern Nations, Nationalities and People**



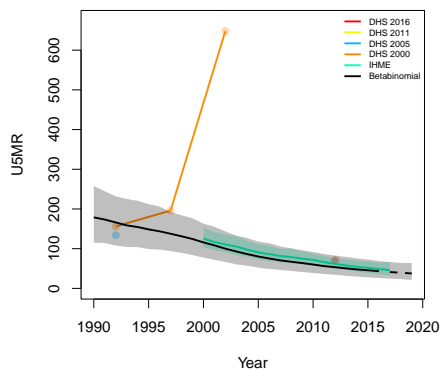
**Silti, Southern Nations, Nationalities and Peoples**



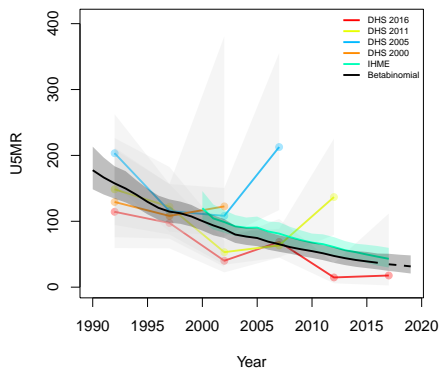
**Wolayita, Southern Nations, Nationalities and People**



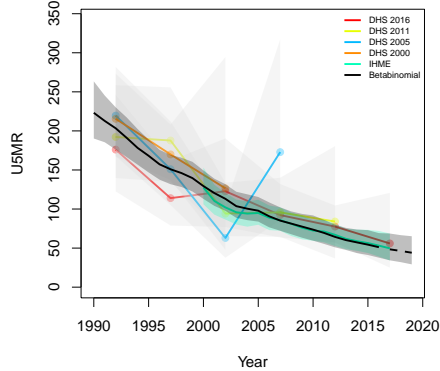
**Yem, Southern Nations, Nationalities and Peoples**

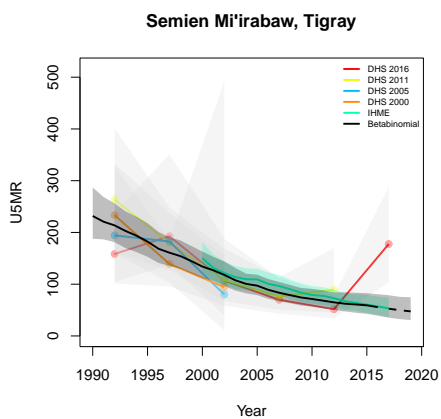
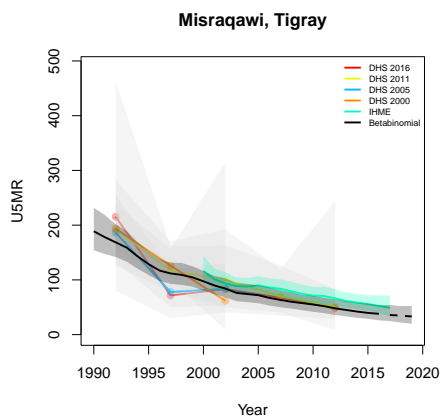
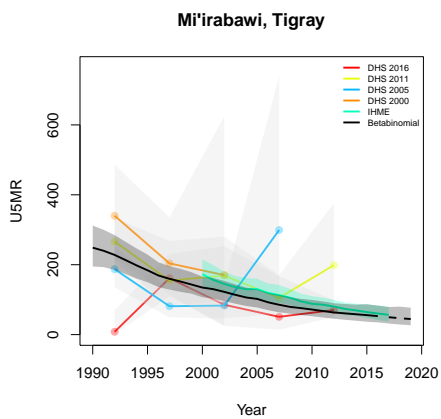


**Debubawi, Tigray**



**Mehakelegnaw, Tigray**





**B.4 Ghana**

Age	Survey	Clusters			Deaths			Agemonths		
		Urban	Rural	Total	Urban	Rural	Total	Urban	Rural	Total
0	1993	149	250	399	29	85	114	758	1991	2749
	1998	138	262	400	29	161	190	1307	4646	5953
	2003	174	236	410	105	308	413	2758	7100	9858
	2008	179	225	404	115	234	349	3085	6398	9483
	2014	216	207	423	288	394	682	8578	12778	21356
1-11	1993	149	250	399	15	55	70	8275	21347	29622
	1998	138	262	400	27	160	187	13909	47728	61637
	2003	174	236	410	48	192	240	28893	72458	101351
	2008	179	225	404	60	177	237	31663	65452	97115
	2014	216	207	423	151	295	446	88403	131603	220006
12-23	1993	149	250	399	9	36	45	8910	23290	32200
	1998	138	262	400	13	90	103	14937	49711	64648
	2003	174	236	410	35	123	158	30159	74373	104532
	2008	179	225	404	35	97	132	32553	66623	99176
	2014	216	207	423	73	148	221	90729	134476	225205
24-35	1993	149	250	399	6	29	35	8565	22783	31348
	1998	138	262	400	17	79	96	14702	48124	62826
	2003	174	236	410	30	104	134	29345	71207	100552
	2008	179	225	404	21	77	98	31186	63931	95117
	2014	216	207	423	50	148	198	85765	126756	212521
36-47	1993	149	250	399	11	22	33	8615	22711	31326
	1998	138	262	400	8	62	70	14729	46297	61026
	2003	174	236	410	21	74	95	28400	67284	95684
	2008	179	225	404	15	57	72	29794	60883	90677
	2014	216	207	423	44	75	119	80492	118630	199122
48-59	1993	149	250	399	4	15	19	8258	21338	29596
	1998	138	262	400	6	29	35	14774	44407	59181
	2003	174	236	410	11	42	53	27678	63397	91075
	2008	179	225	404	11	30	41	28440	57468	85908

*B.4.1 Admin-1*

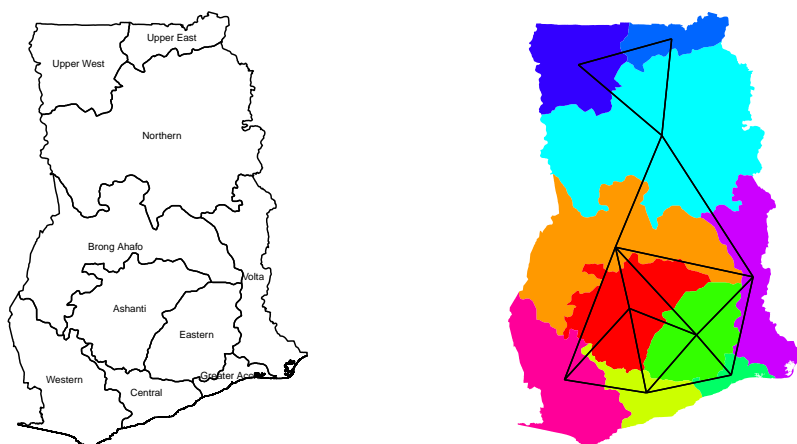


Figure B.19: **Left:** The names of the 10 Admin-1 areas of Ghana . **Right:** The neighborhood structure of Admin-1 areas in Ghana .

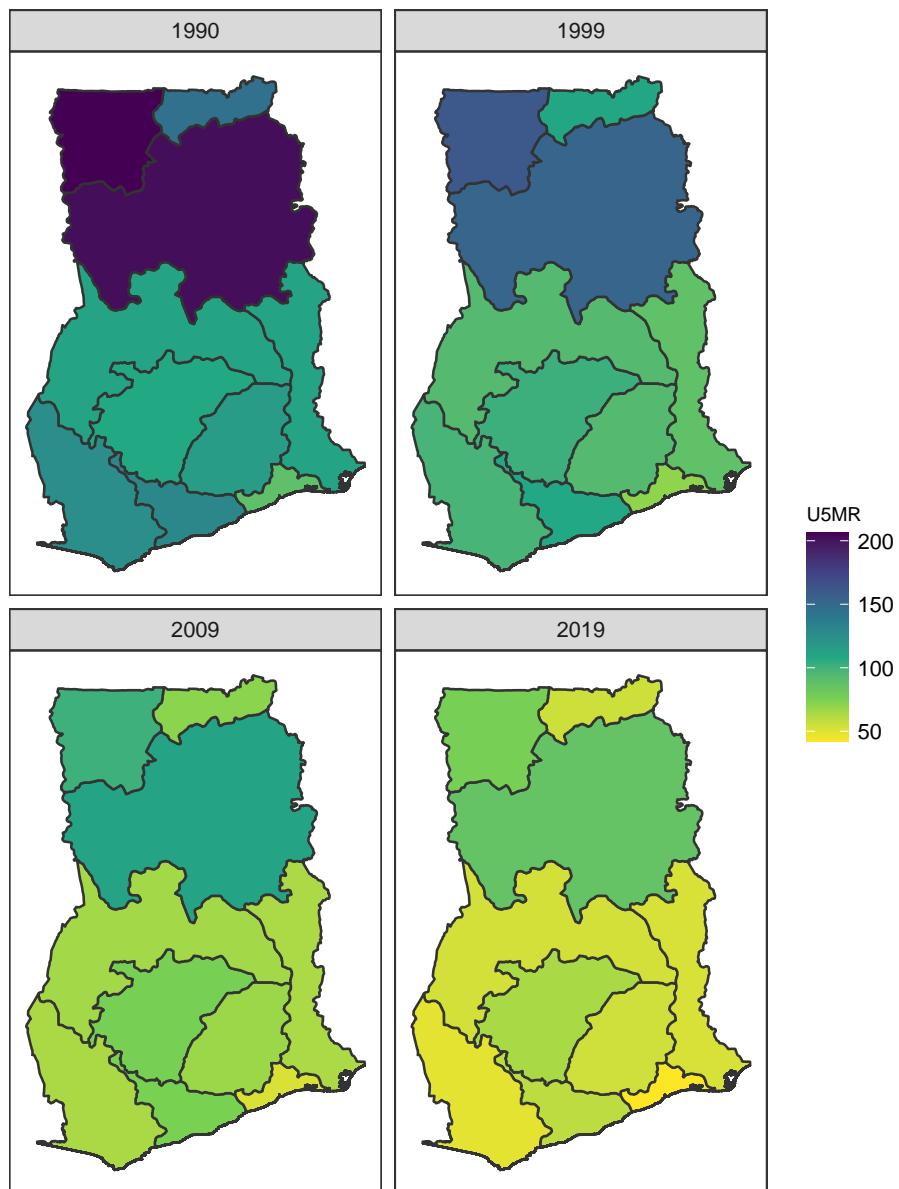


Figure B.20: Median U5MR estimates for years 1990, 1999, 2009, 2019 for Admin-1 areas in Ghana .

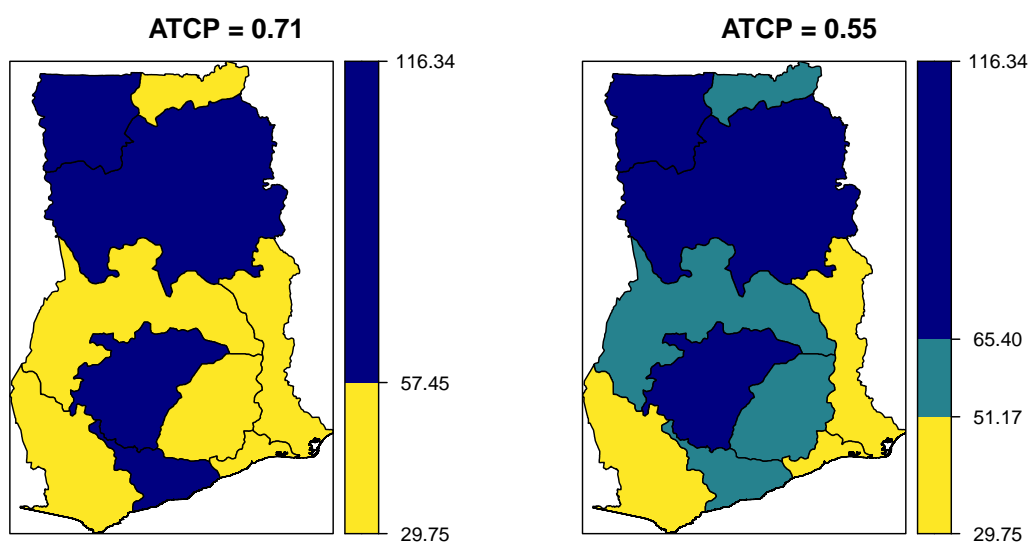
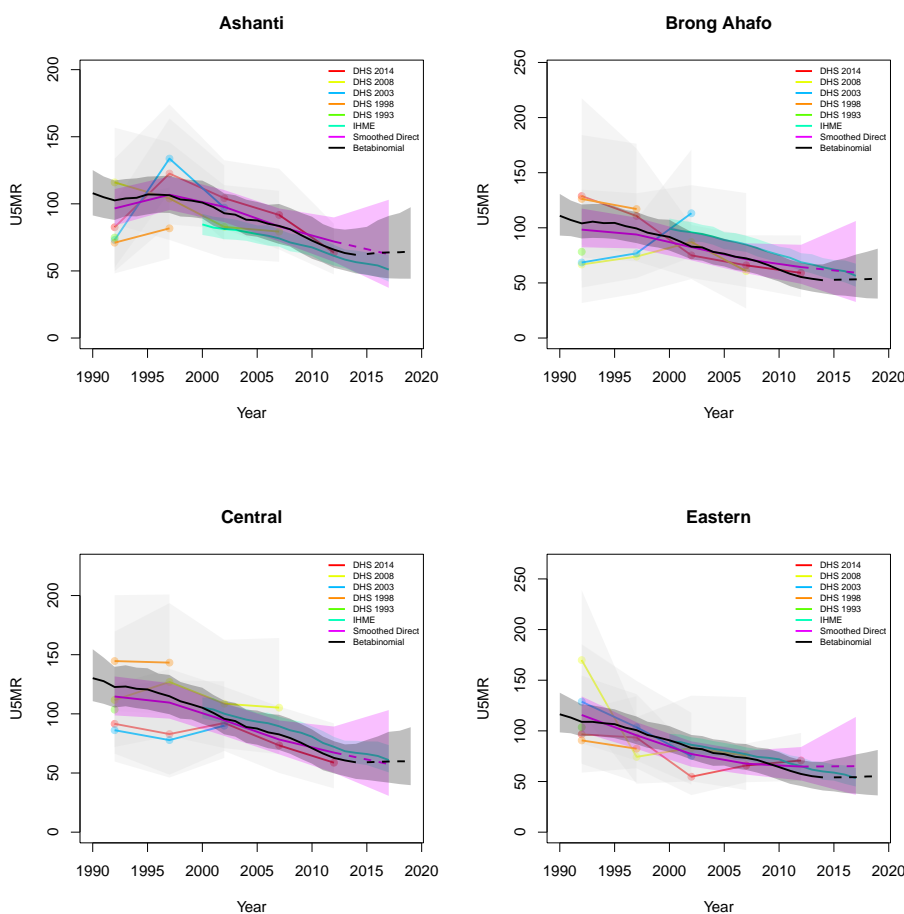
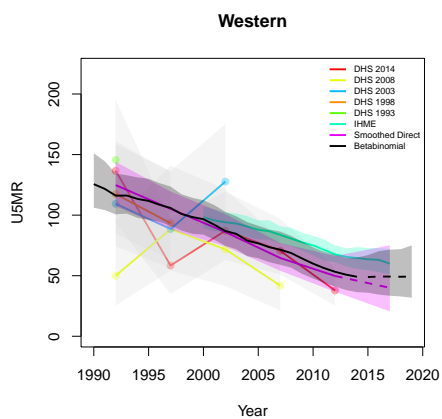
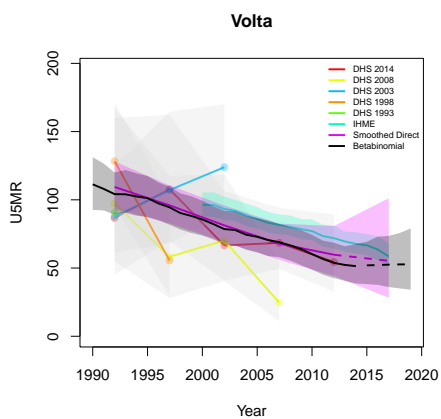
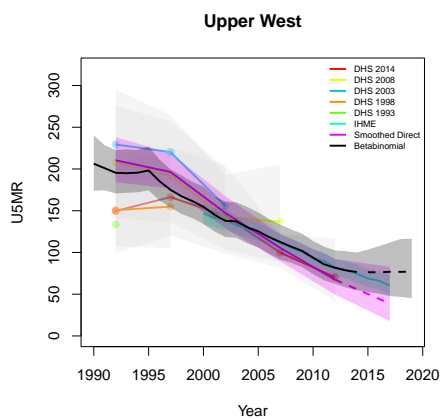
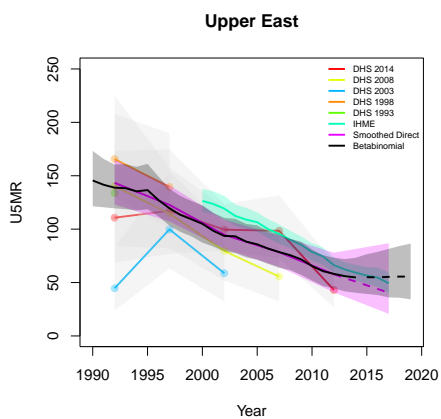
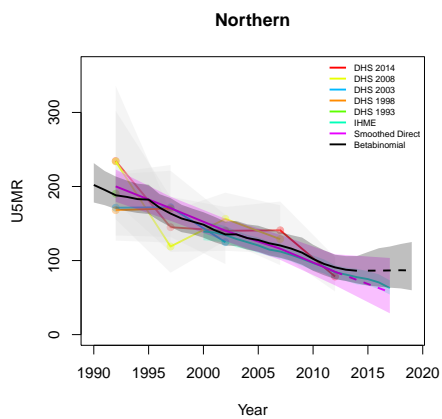
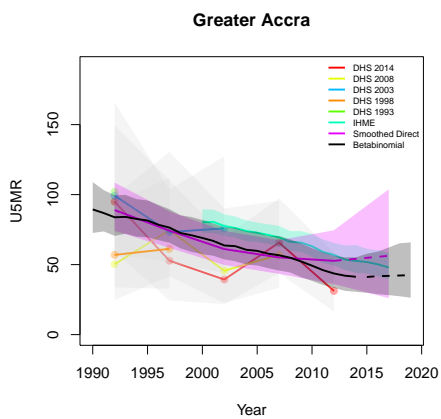


Figure B.21: Expression of uncertainty of U5MR (deaths per 1000 children) estimates for Admin-1 areas based on the average true classification probability (ATCP) in 2019 using  $K = 2, 3$  colors.

*Data and estimates over time by area*

Colored lines with circular points and light grey uncertainty bands are 5-year survey-weighted estimates of U5MR for years 1990–1994 up to 2015–2019 depending on survey timing. For a survey that ends in the middle of a 5-year period, we plot the estimates at the mid-point of the years in that interval for which the survey provides data. Black lines and corresponding intervals represent posterior medians and 95% uncertainty intervals respectively for the betabinomial model. IHME’s estimates and corresponding intervals, where we can compare, are in aquamarine.





*B.4.2 Admin-2*



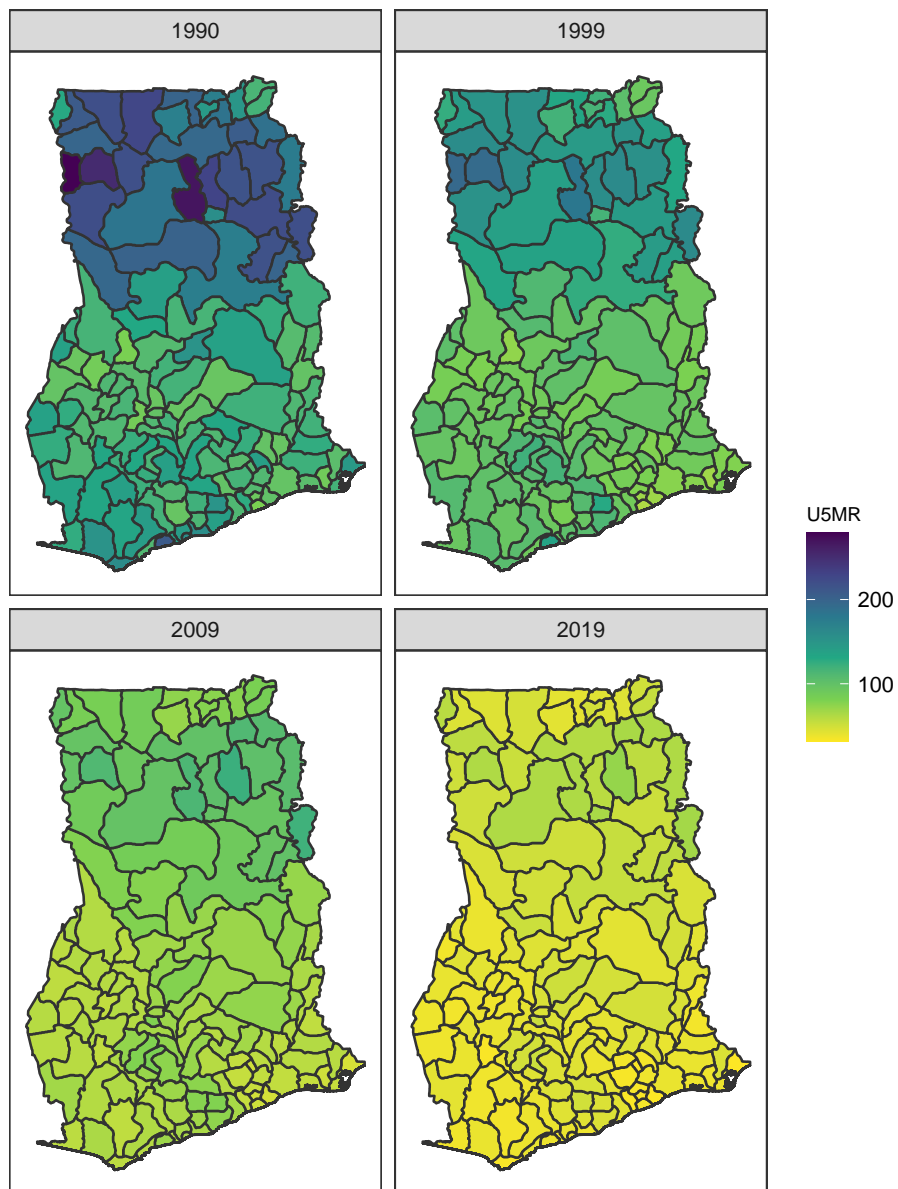


Figure B.23: Median U5MR estimates for years 1990, 1999, 2009, 2019 for Admin-2 areas in Ghana .

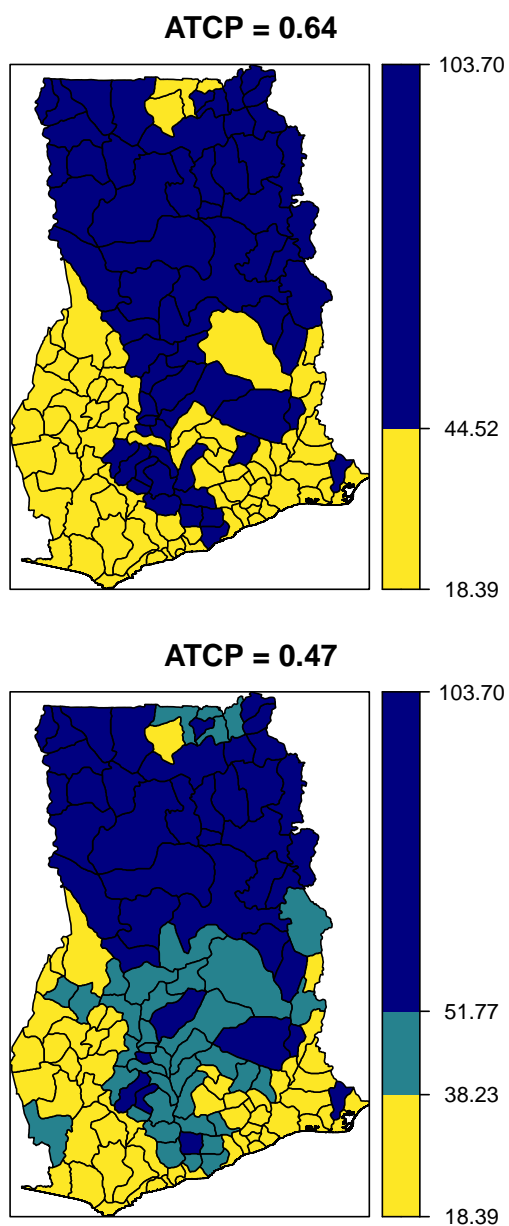
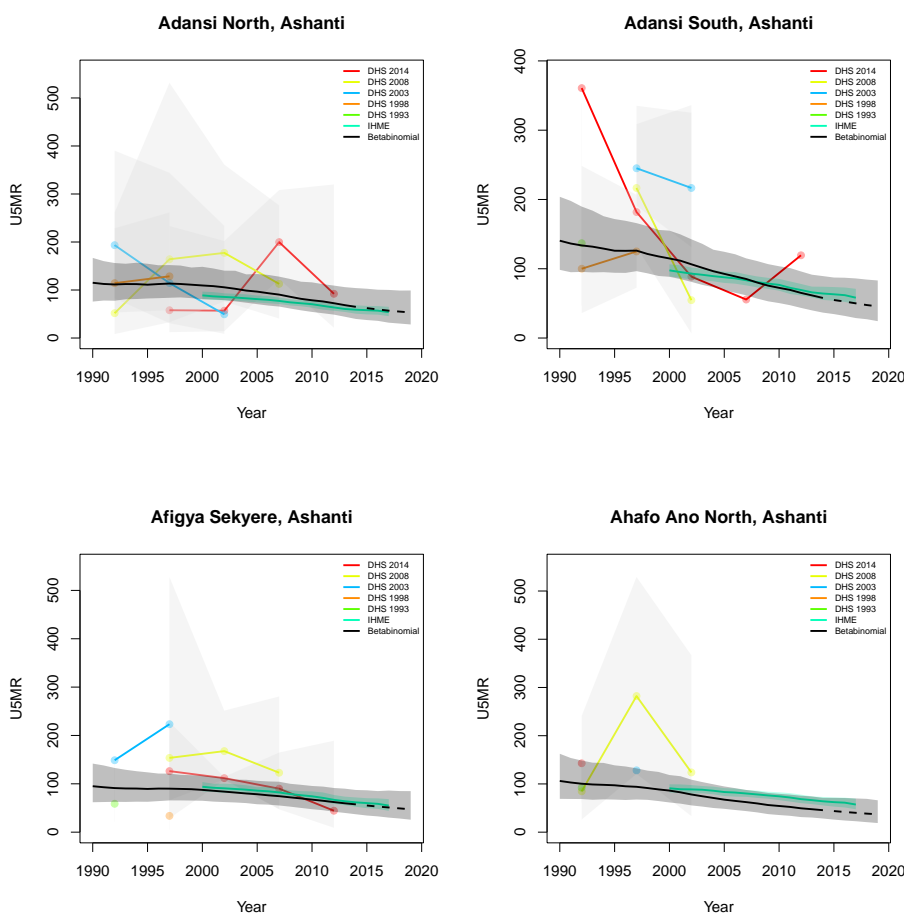


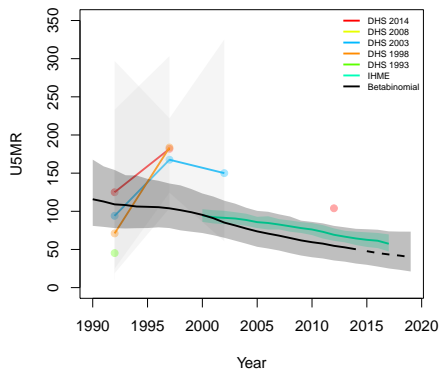
Figure B.24: Expression of uncertainty of U5MR (deaths per 1000 children) estimates for Admin-1 areas based on the average true classification probability (ATCP) in 2019 using  $K = 2, 3$  colors.

*Data and estimates over time by area*

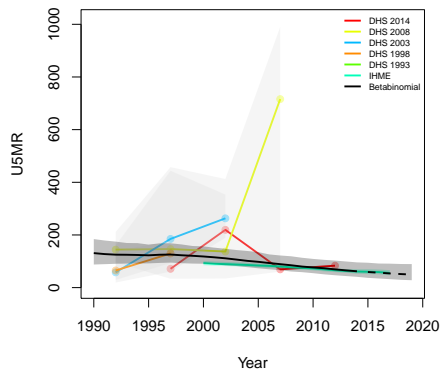
Colored lines with circular points and light grey uncertainty bands are 5-year survey-weighted estimates of U5MR for years 1990–1994 up to 2015–2019 depending on survey timing. For a survey that ends in the middle of a 5-year period, we plot the estimates at the mid-point of the years in that interval for which the survey provides data. Black lines and corresponding intervals represent posterior medians and 95% uncertainty intervals respectively for the betabinomial model. IHME’s estimates and corresponding intervals, where we can compare, are in aquamarine.



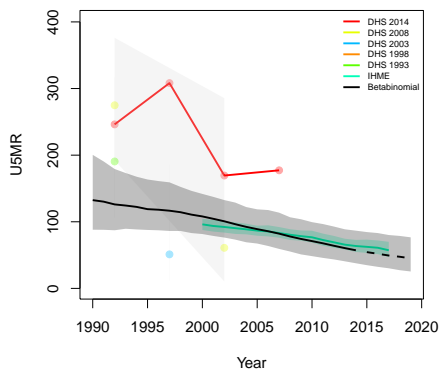
**Ahafo Ano South, Ashanti**



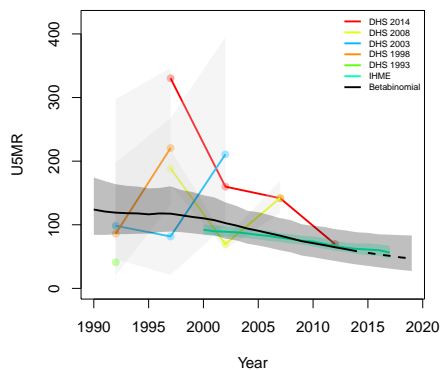
**Amansie Central, Ashanti**



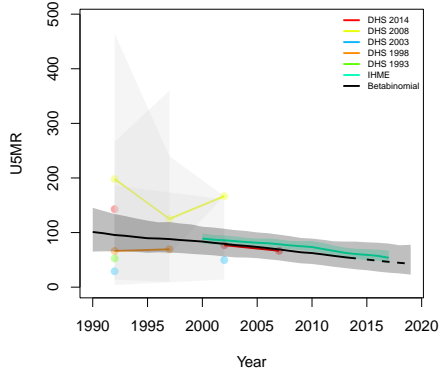
**Amansie East, Ashanti**



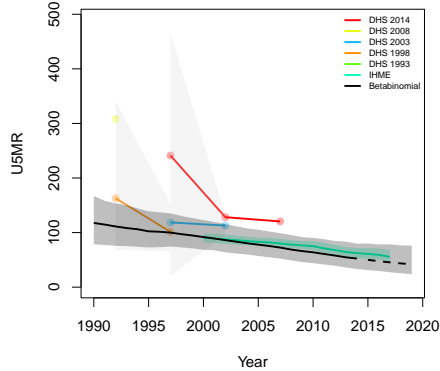
**Amansie West, Ashanti**



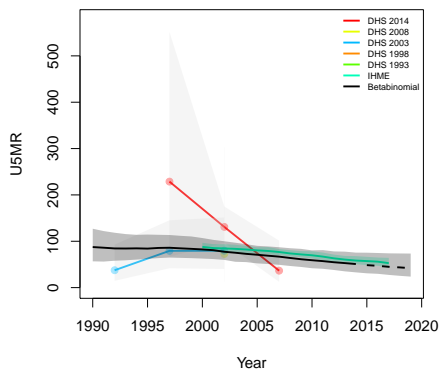
**Asante Akim North, Ashanti**



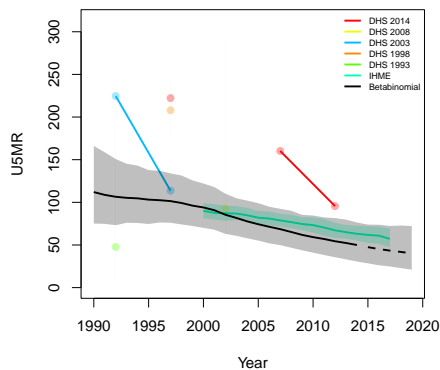
**Asante Akim South, Ashanti**



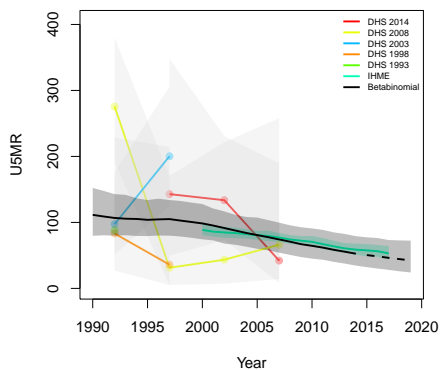
**Atwima, Ashanti**



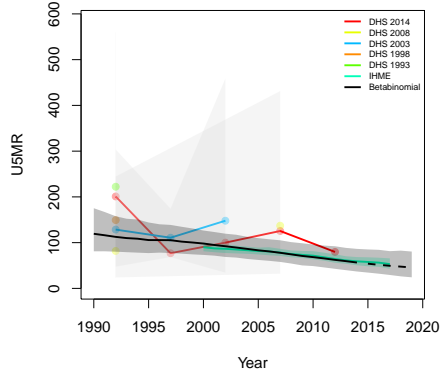
**Atwima Mponua, Ashanti**



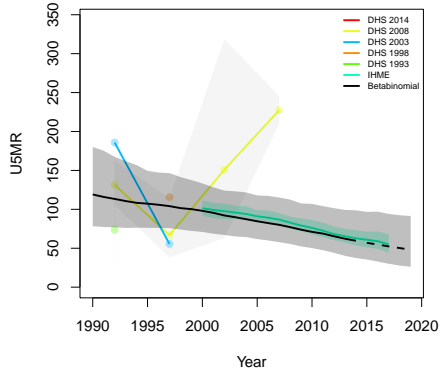
**Bosomtwe-Kwanwoma, Ashanti**



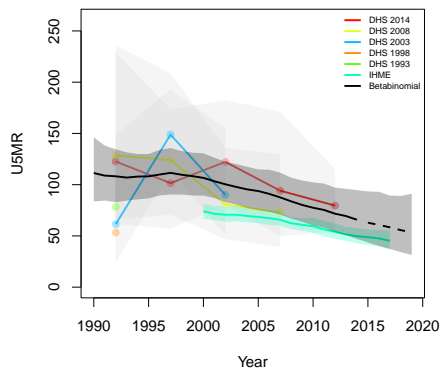
**Ejisu-Juabeng, Ashanti**



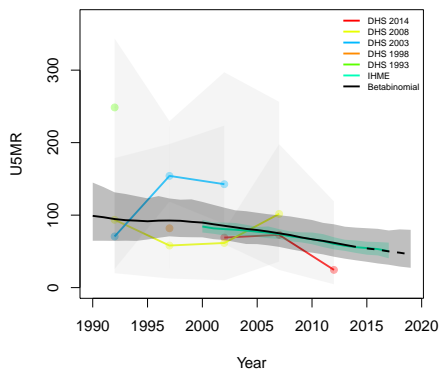
**Ejura Sekyedumase, Ashanti**



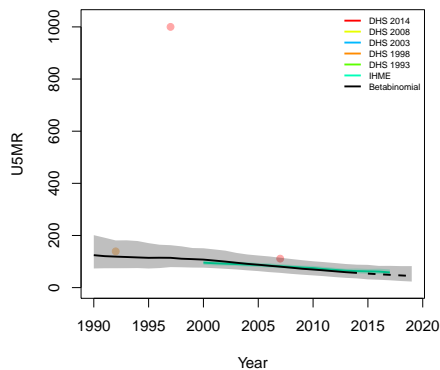
**Kumasi, Ashanti**



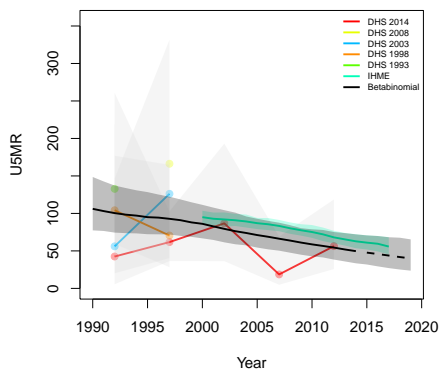
**Kwabre, Ashanti**



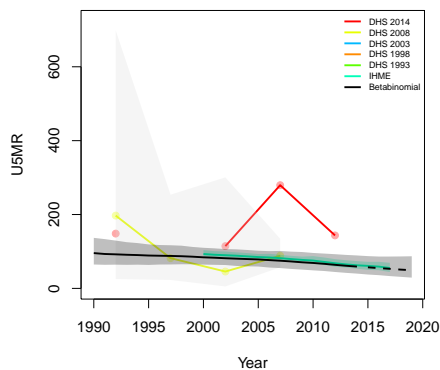
**Obuasi Municipal, Ashanti**



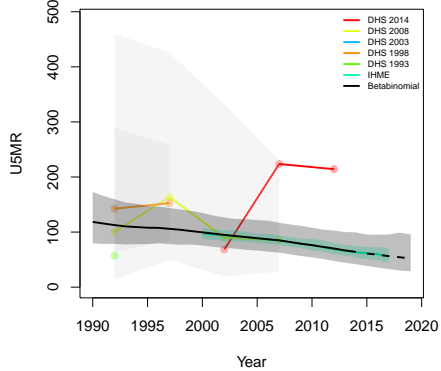
**Offinso, Ashanti**



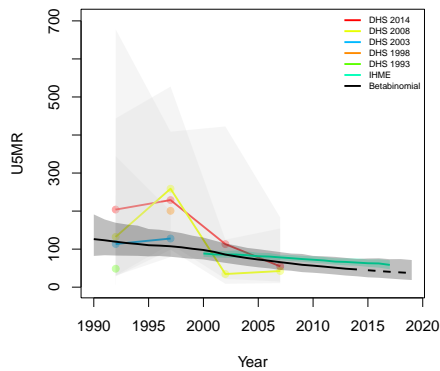
**Sekyere East, Ashanti**



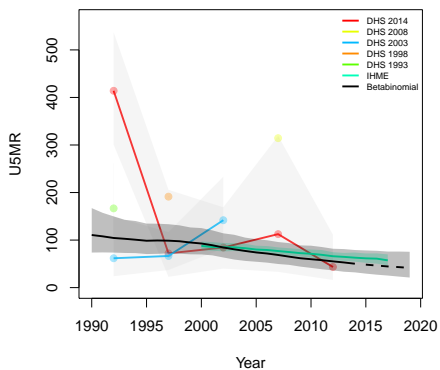
**Sekyere West, Ashanti**



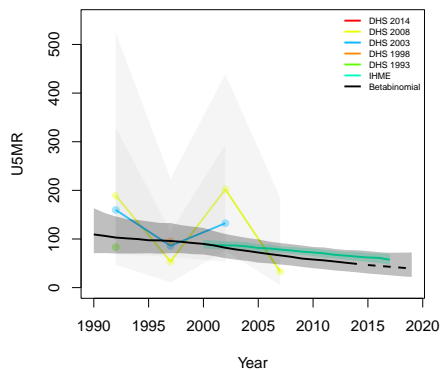
**Asunafo North, Brong Ahafo**



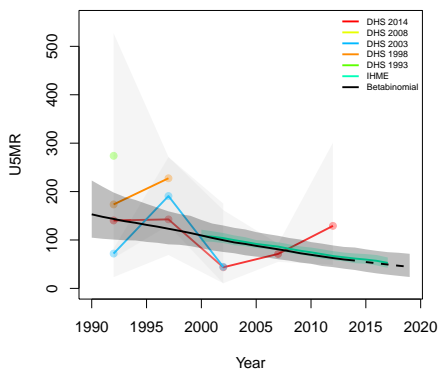
**Asunafo South, Brong Ahafo**



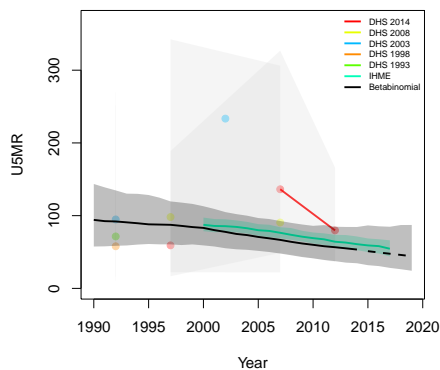
**Asutifi, Brong Ahafo**



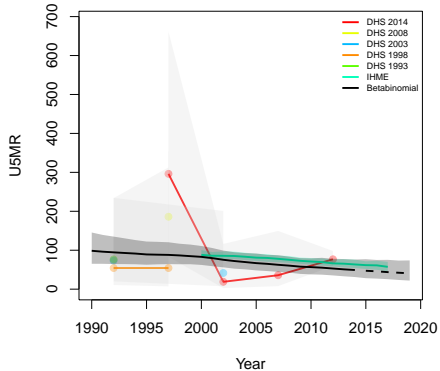
**Atebubu-Amantin, Brong Ahafo**



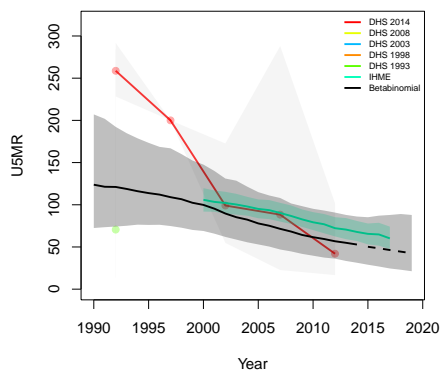
**Berekum, Brong Ahafo**



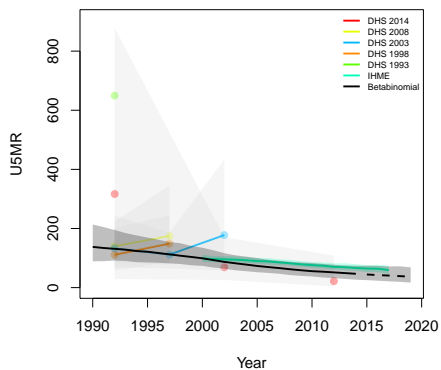
**Dormaa, Brong Ahafo**



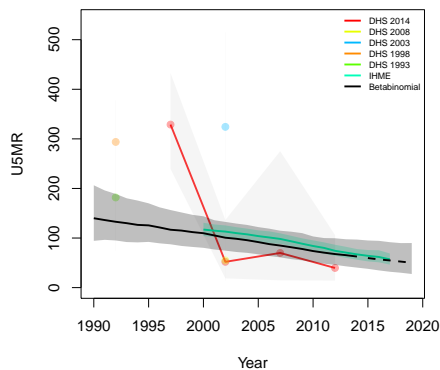
**Jaman North, Brong Ahafo**



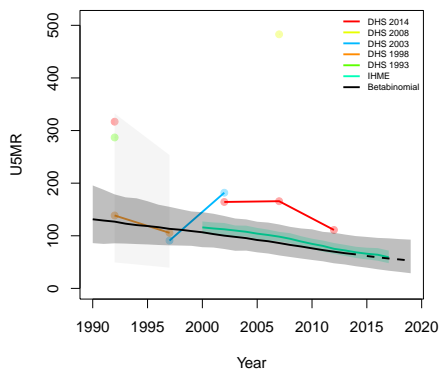
Jaman South, Brong Ahafo



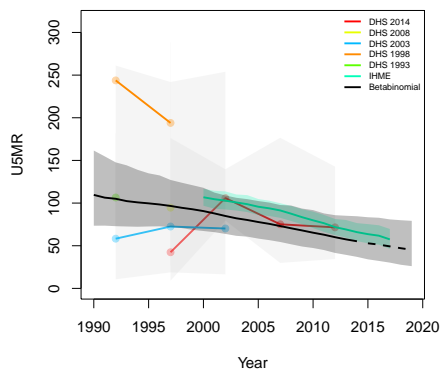
Kintampo North, Brong Ahafo



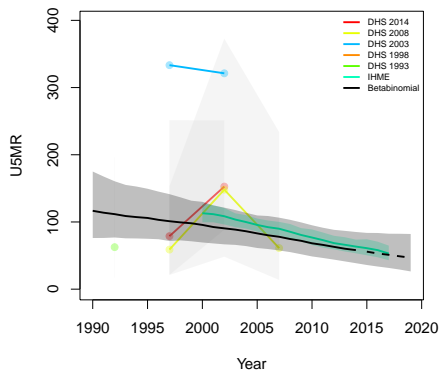
Kintampo South, Brong Ahafo



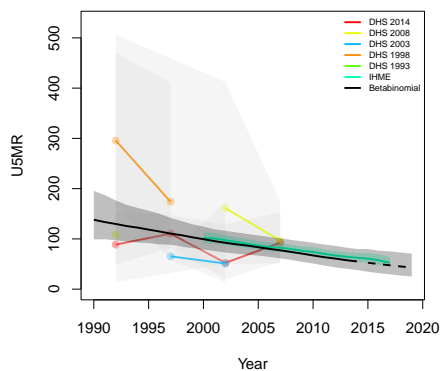
Nkoranza, Brong Ahafo



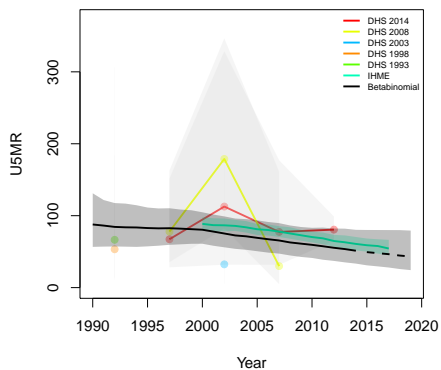
Pru, Brong Ahafo



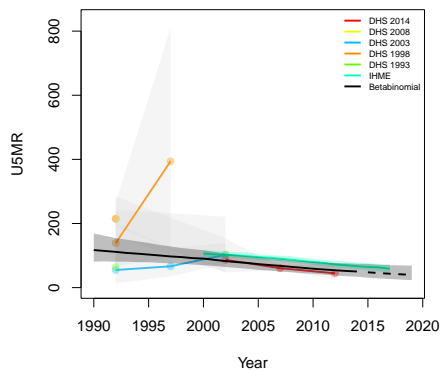
Sene, Brong Ahafo



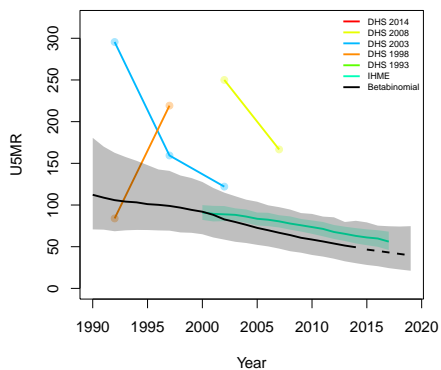
Sunyani, Brong Ahafo



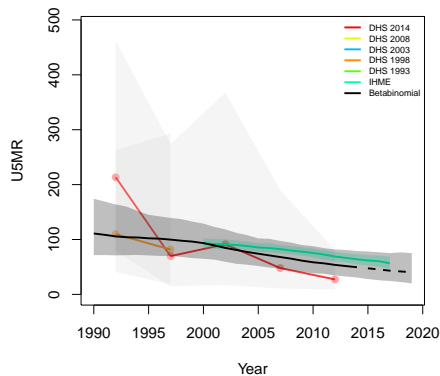
Tain, Brong Ahafo



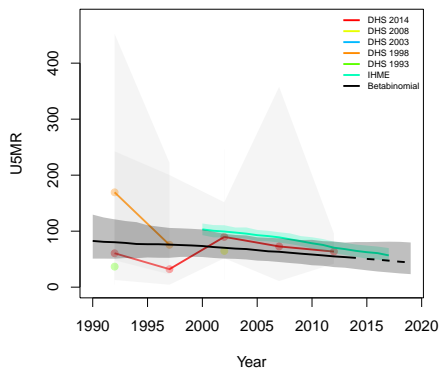
Tano North, Brong Ahafo



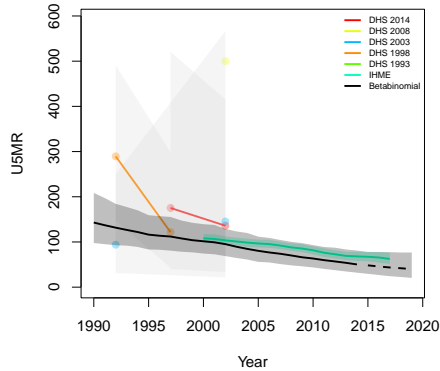
Tano South, Brong Ahafo



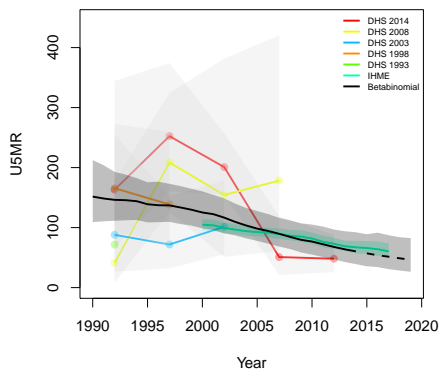
Techiman, Brong Ahafo



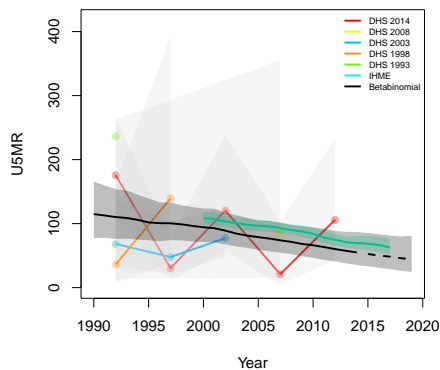
Abura-Asebu-Kwamankese, Central



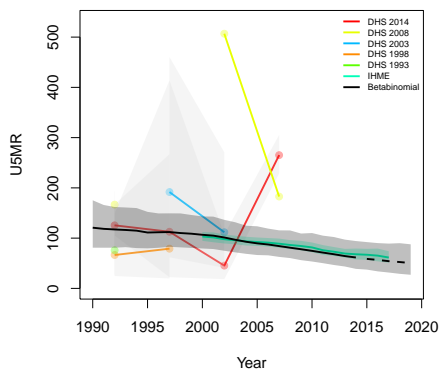
**Agona, Central**



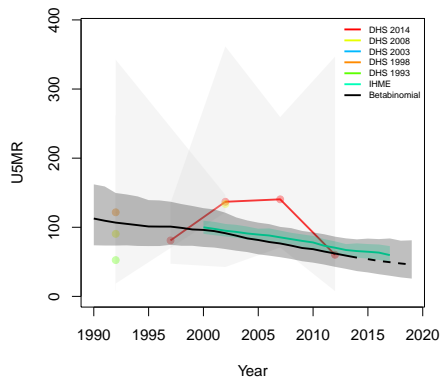
**Ajumako-Enyan-Esiam, Central**



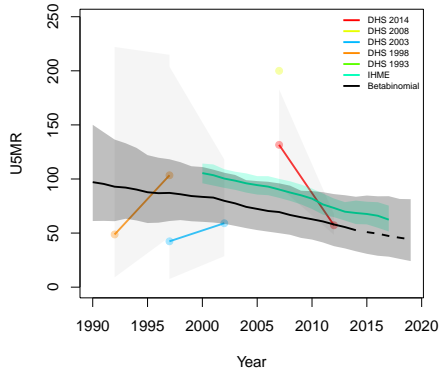
**Asikuma Odoben Brakwa, Central**



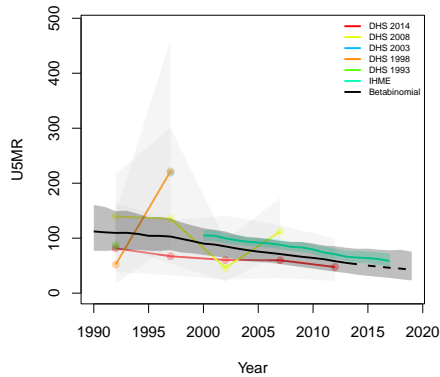
**Assin North, Central**



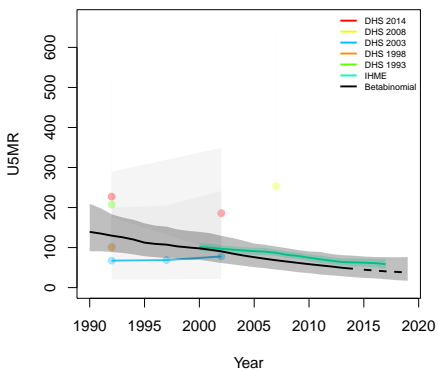
**Assin South, Central**



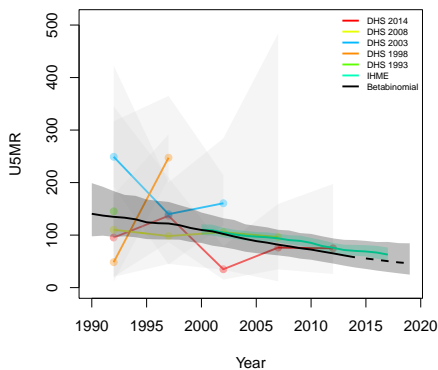
**Awutu Efutu Senya, Central**



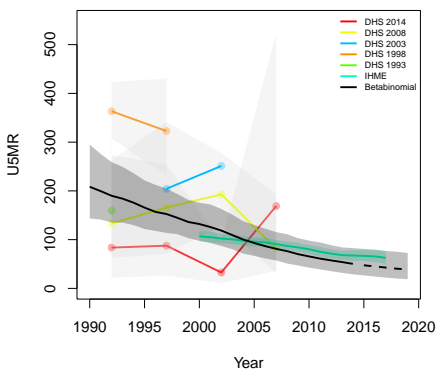
**Cape Coast, Central**



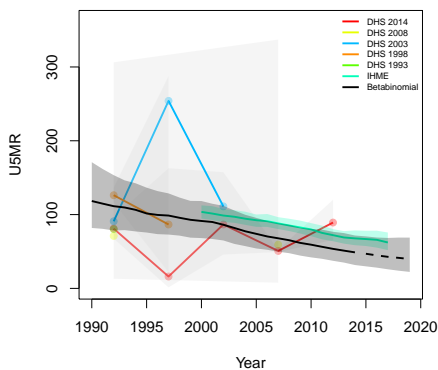
**Gomoa, Central**



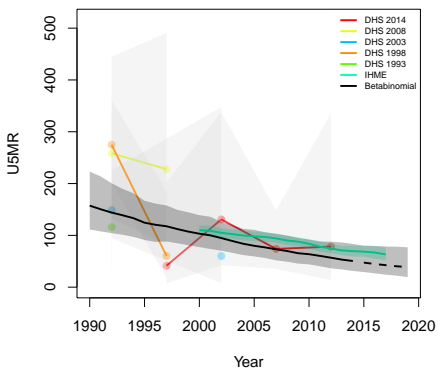
**Komenda–Edina–Eguafo–Abirem, Central**



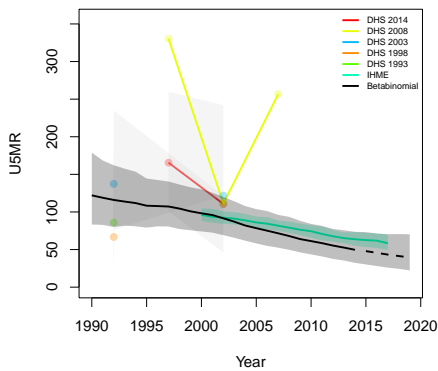
**Lower Denkyira, Central**



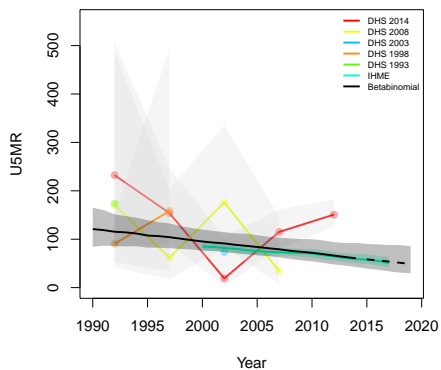
**Mfantiman, Central**



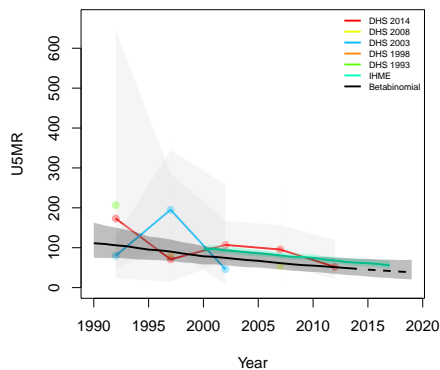
**Upper Denkyira, Central**



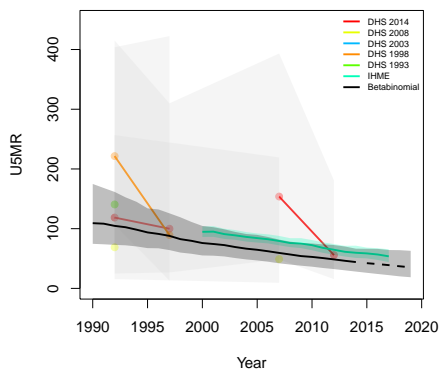
**Afram Plains, Eastern**



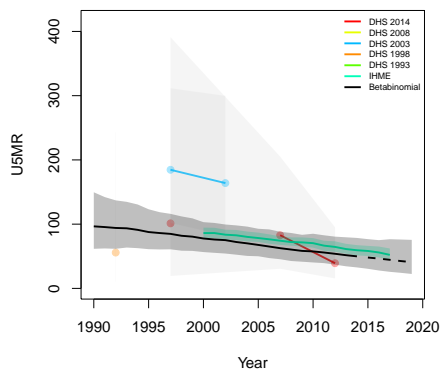
**Akwapim North, Eastern**



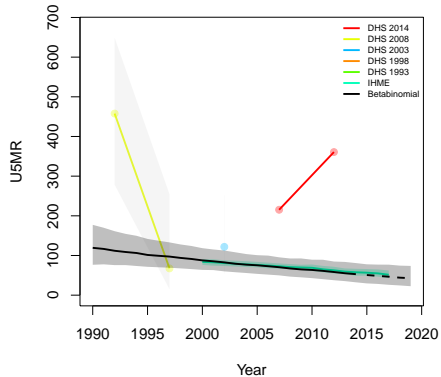
**Akwapim South, Eastern**



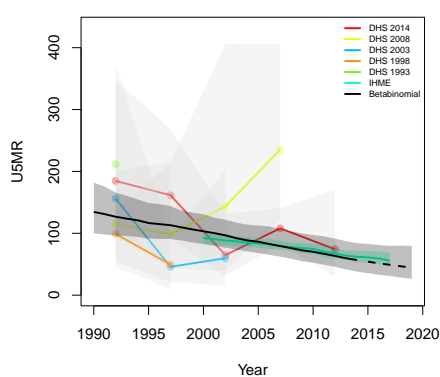
**Asuogyaman, Eastern**



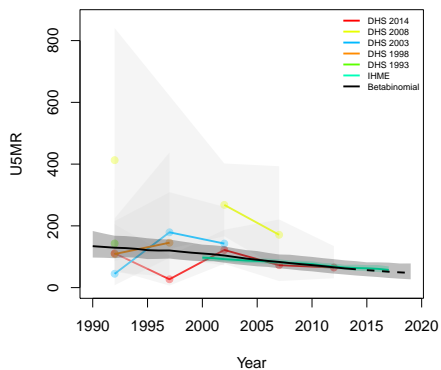
**Atiwa, Eastern**



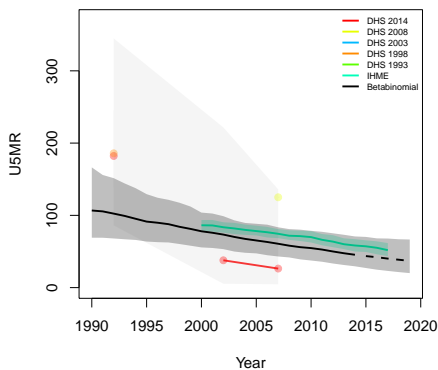
**Birim North, Eastern**



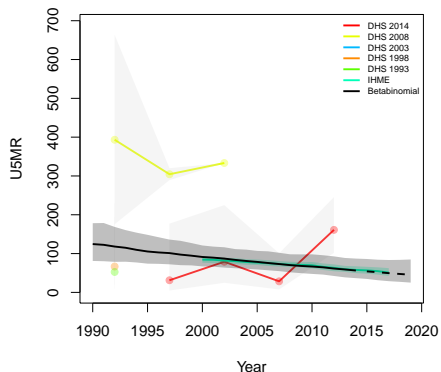
**Birim South, Eastern**



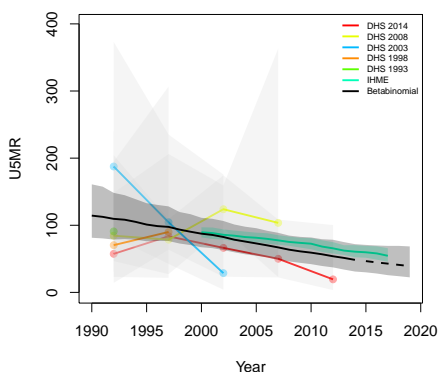
**East Akim, Eastern**



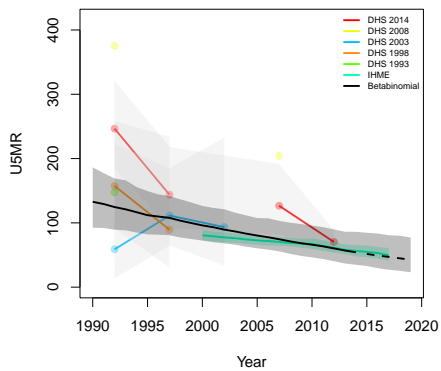
**Fanteakwa, Eastern**



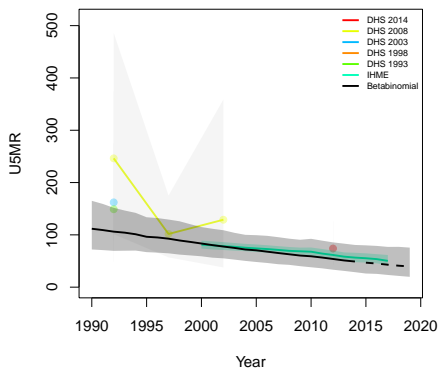
**Kwabibirem, Eastern**



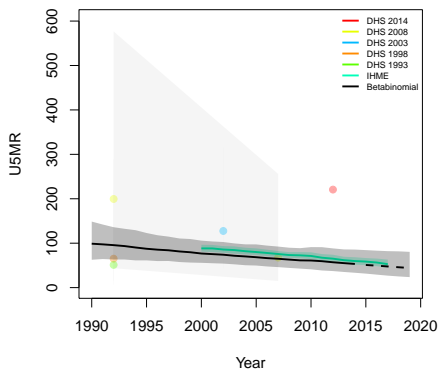
**Kwahu South, Eastern**



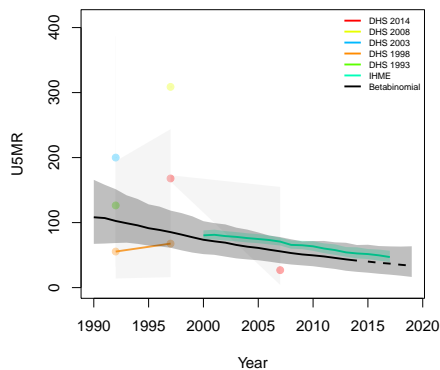
**Kwahu West, Eastern**



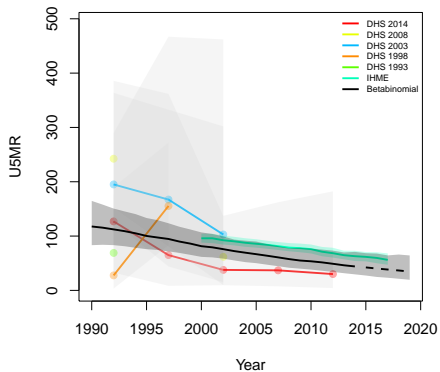
**Manya Krobo, Eastern**



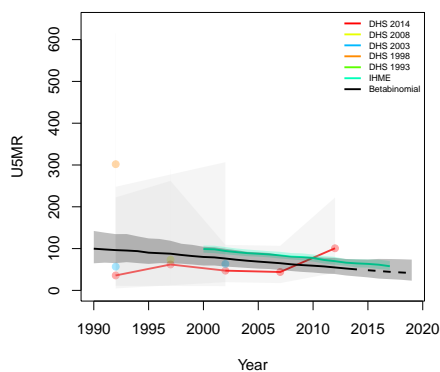
**New Juaben, Eastern**



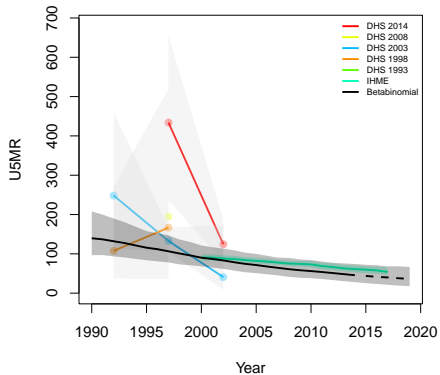
**Suhum Kraboa Coalta, Eastern**



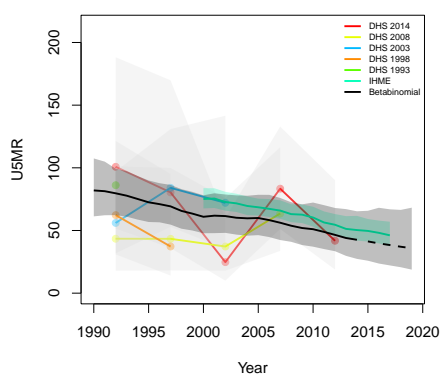
**West Akim, Eastern**



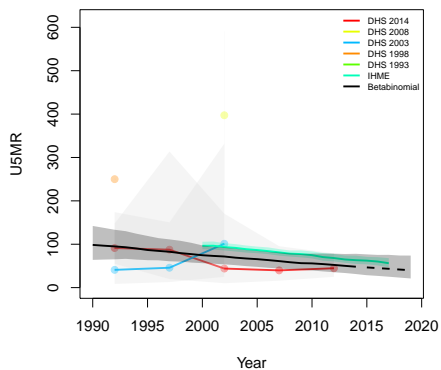
**Yilo Krobo, Eastern**



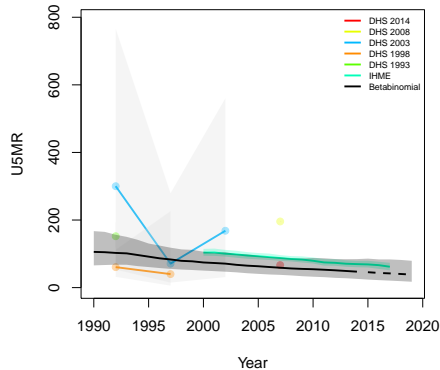
**Accra, Greater Accra**



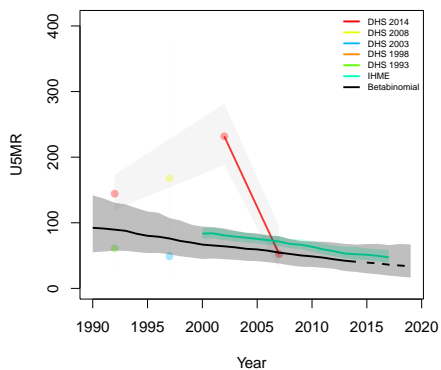
Dangbe East, Greater Accra



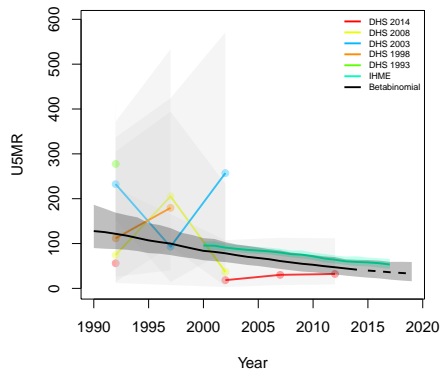
Dangbe West, Greater Accra



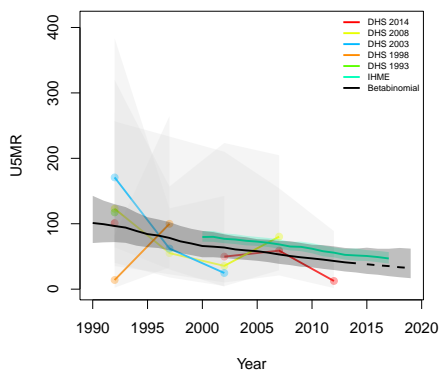
Ga East, Greater Accra



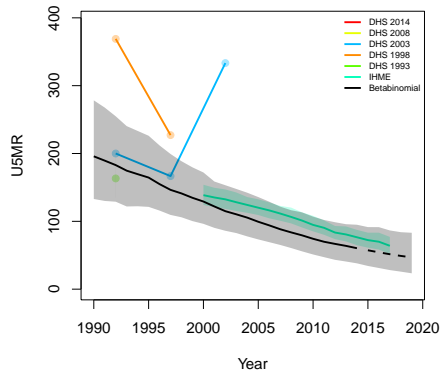
Ga West, Greater Accra



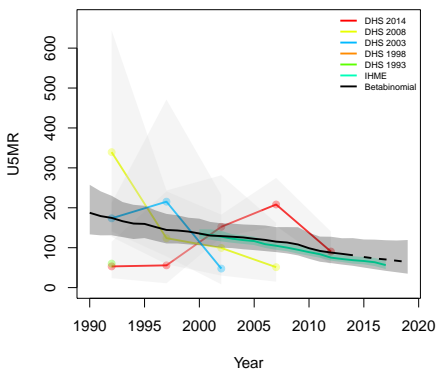
Tema, Greater Accra



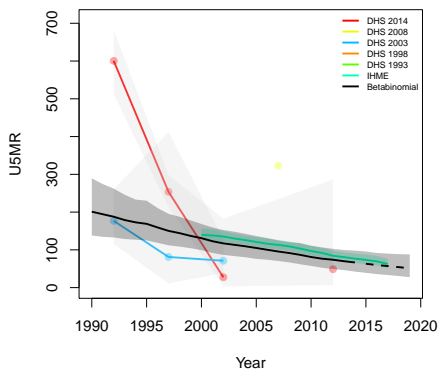
Bole, Northern



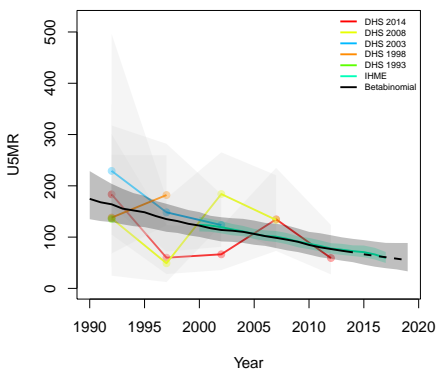
**Bunkpurugu Yunyoo, Northern**



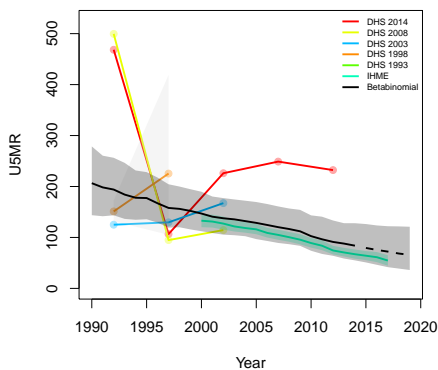
**Central Gonja, Northern**



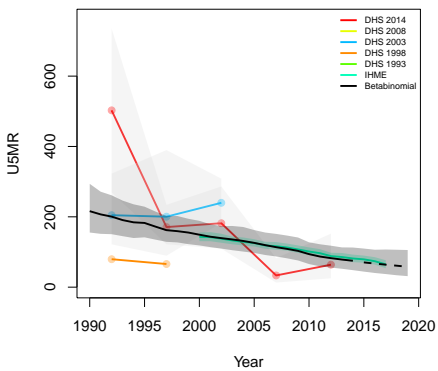
**East Gonja, Northern**



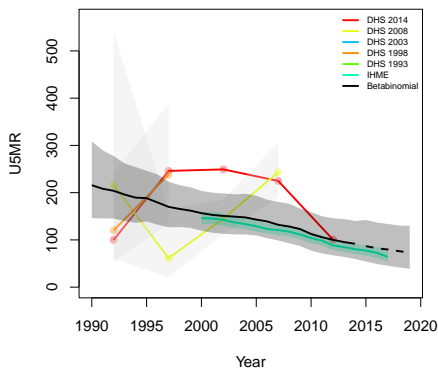
**East Mamprusi, Northern**



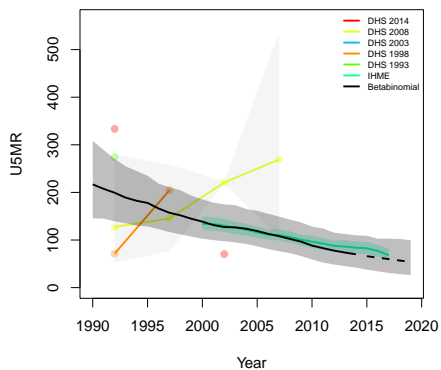
**Gushiegu, Northern**



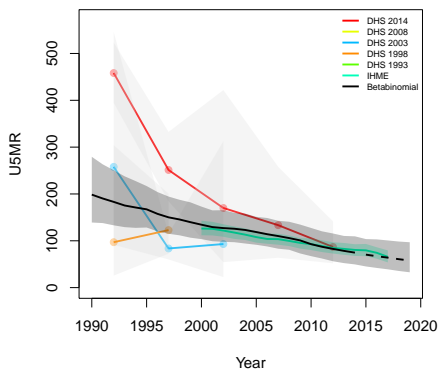
**Karaga, Northern**



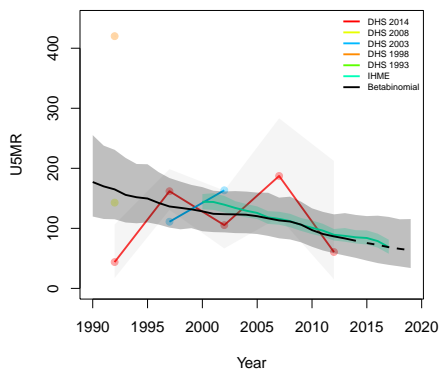
**Nanumba North, Northern**



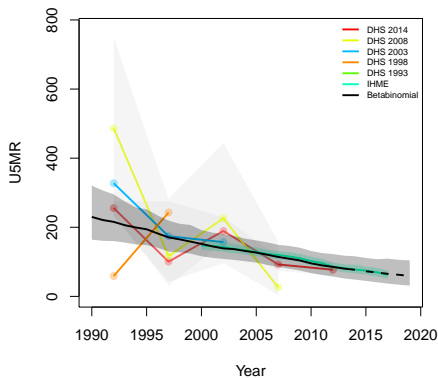
**Nanumba South, Northern**



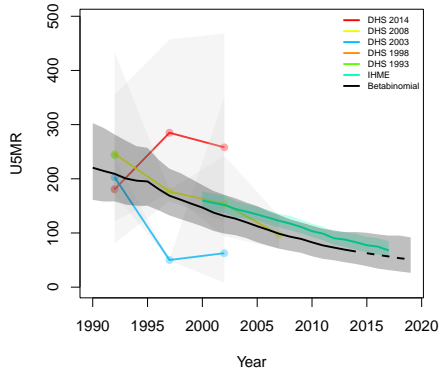
**Saboba Chereponi, Northern**



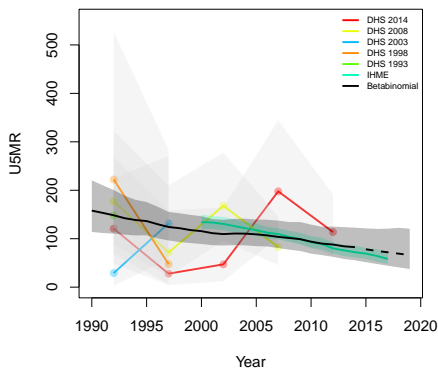
**Savelugu Nanton, Northern**



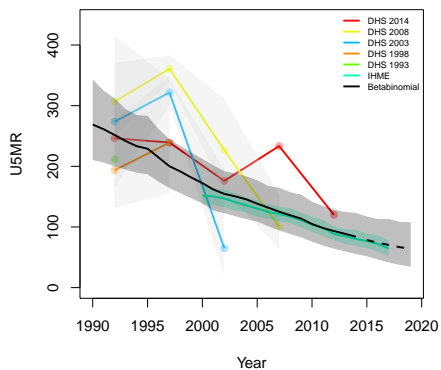
**Sawa-Tuna-Kalba, Northern**



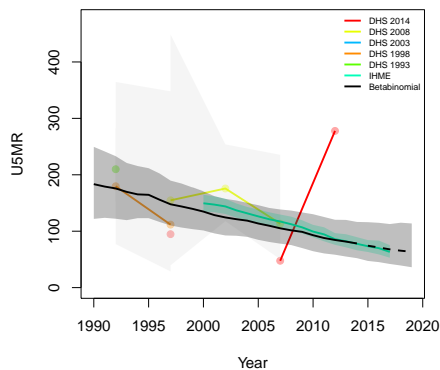
**Tamale, Northern**



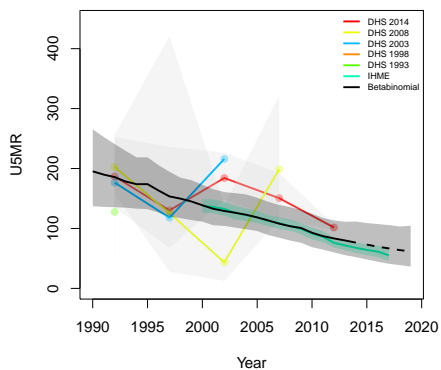
**Tolon–Kumbungu, Northern**



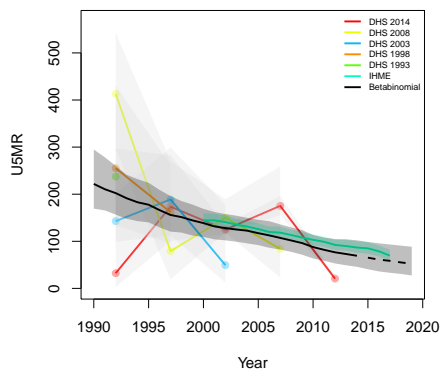
**West Gonja, Northern**



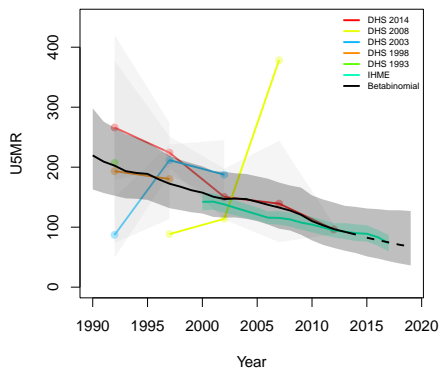
**West Mamprusi, Northern**



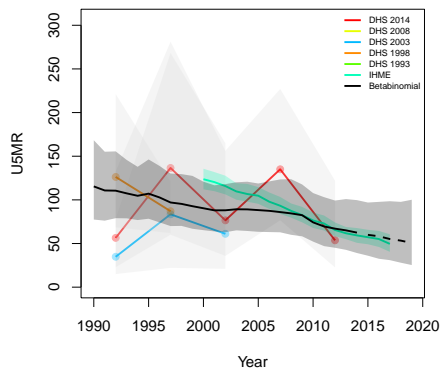
**Yendi, Northern**



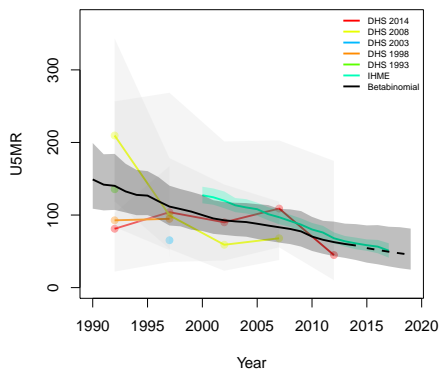
**Zabzugu Tatale, Northern**



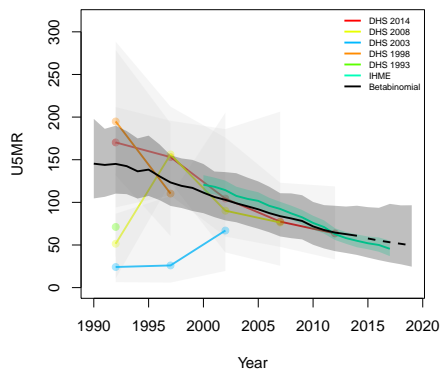
**Bawku Municipal, Upper East**



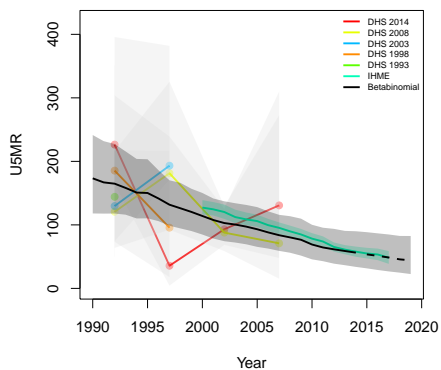
**Bawku West, Upper East**



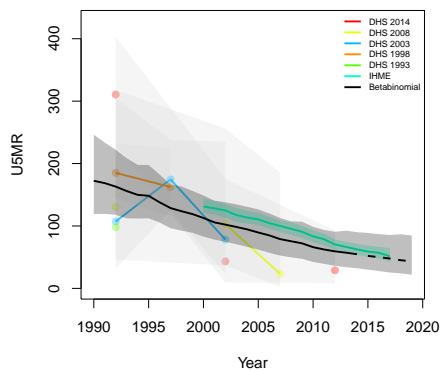
**Bolgatanga, Upper East**



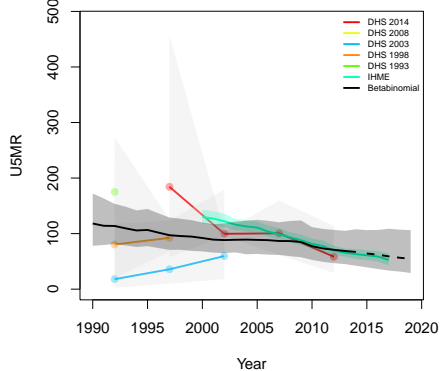
**Bongo, Upper East**



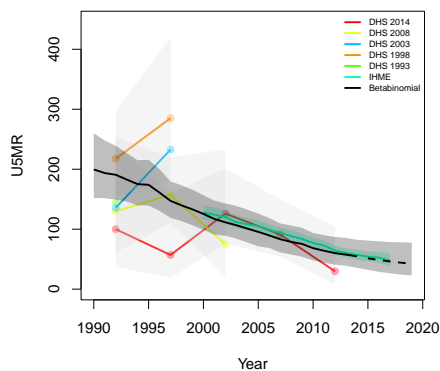
**Builsa, Upper East**



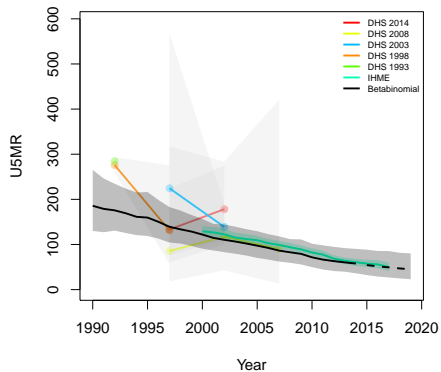
**Garu Tempene, Upper East**



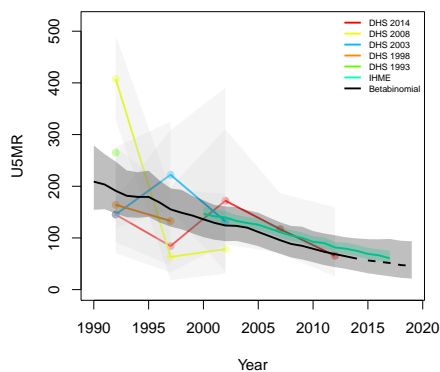
**Kassena Nankana, Upper East**



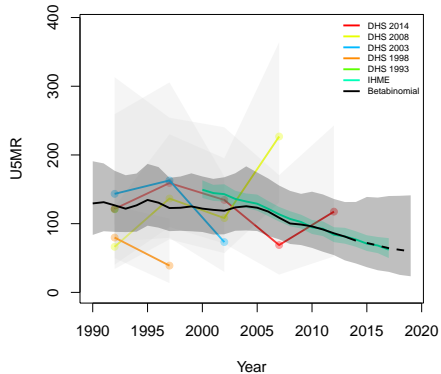
Talensi Nabdam, Upper East



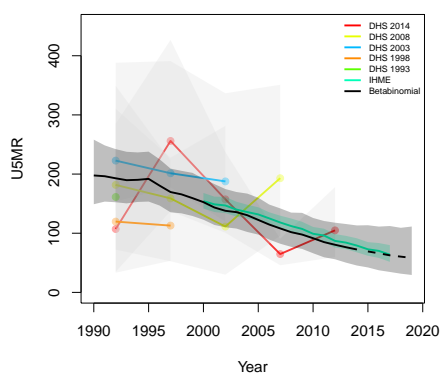
Jirapa Lambussie, Upper West



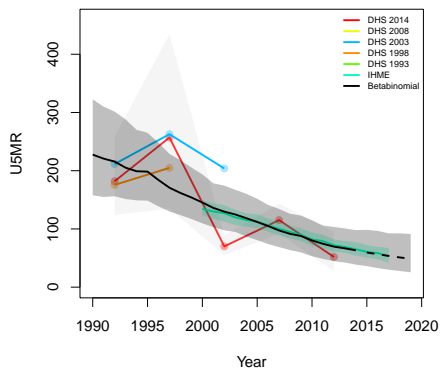
Lawra, Upper West



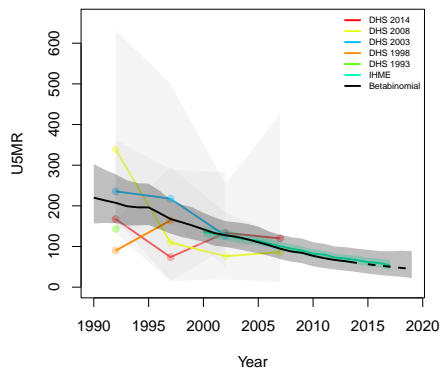
Nadowli, Upper West



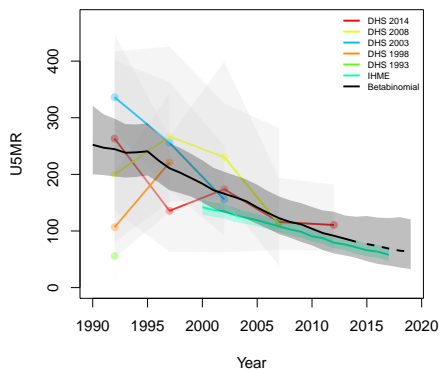
Sissala East, Upper West



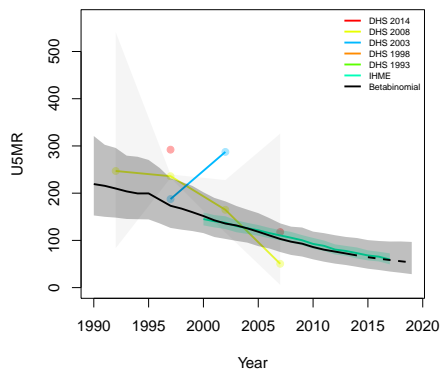
Sissala West, Upper West



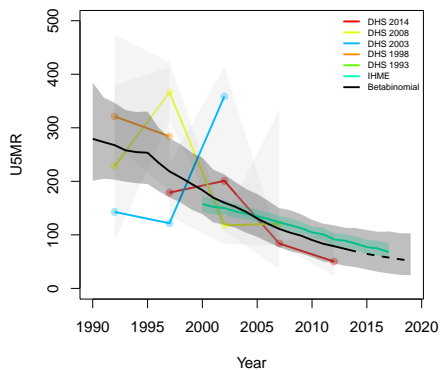
**Wa, Upper West**



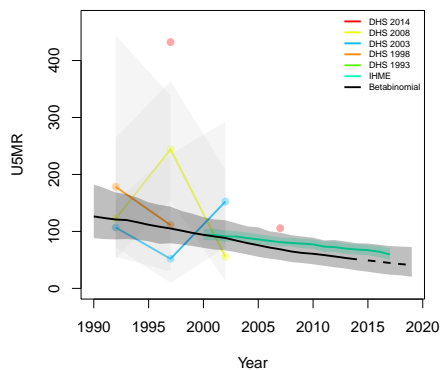
**Wa East, Upper West**



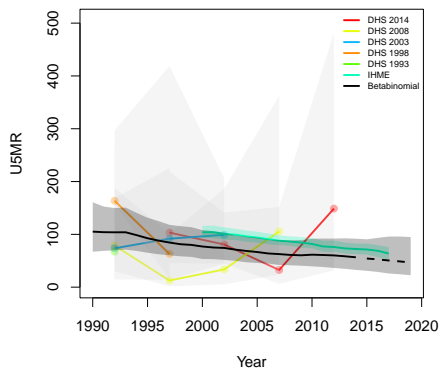
**Wa West, Upper West**



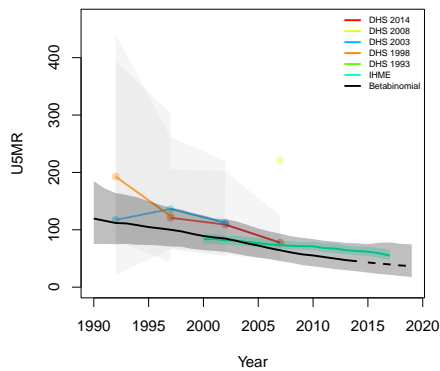
**Adaklu Anyigbe, Volta**

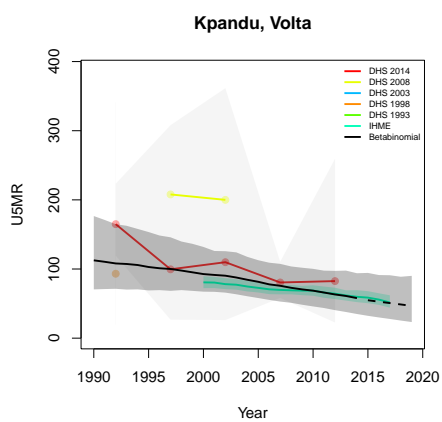
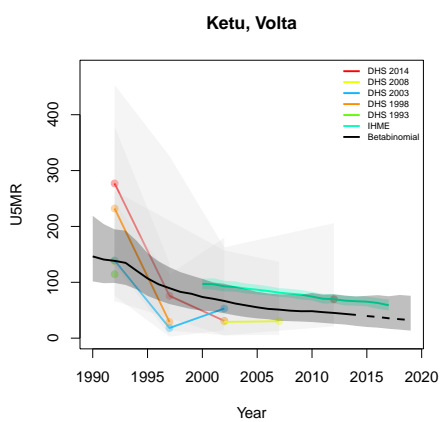
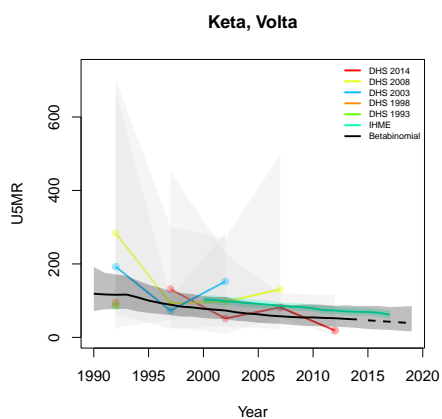
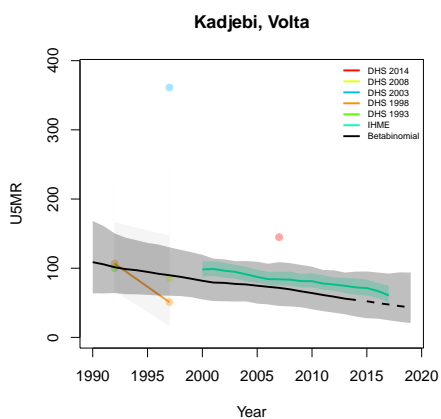
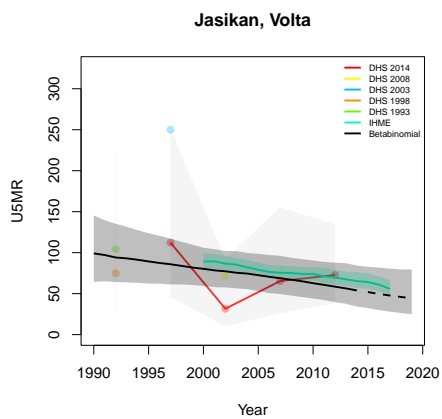
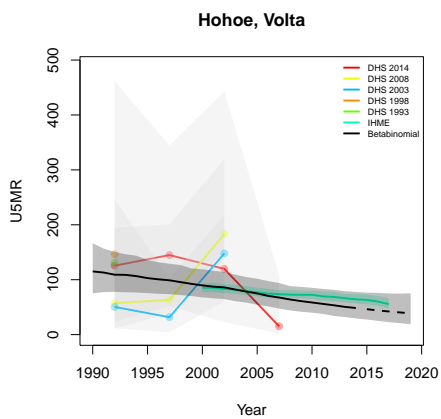


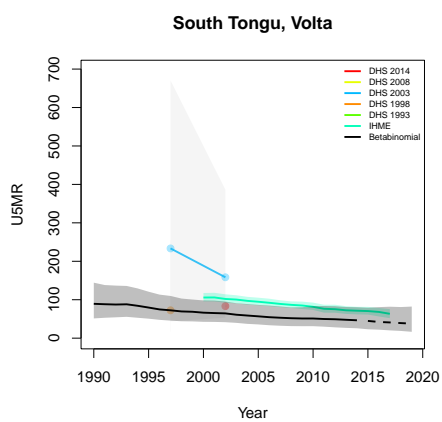
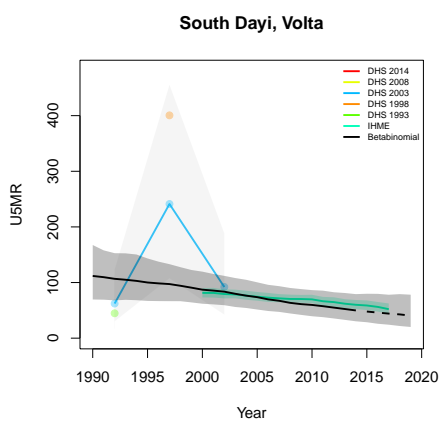
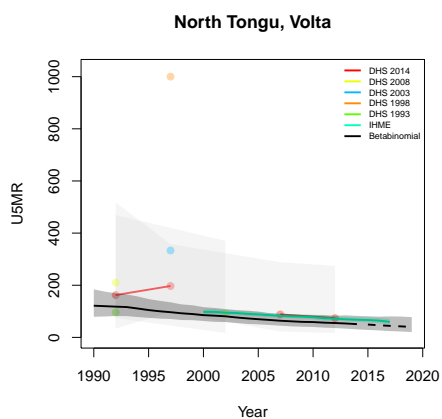
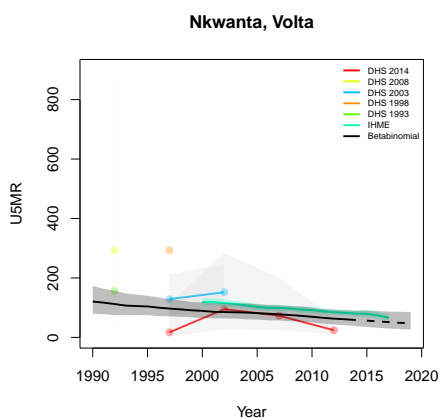
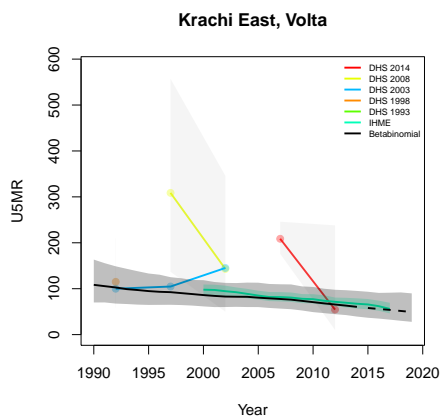
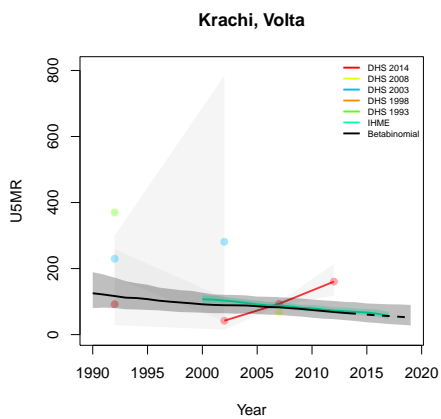
**Akatsi, Volta**



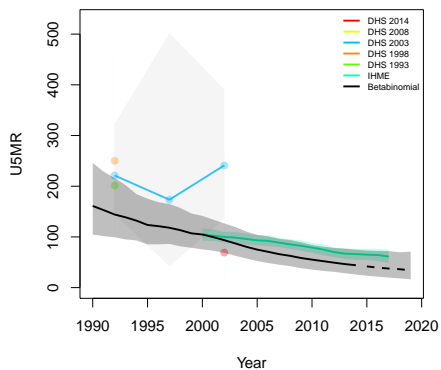
**Ho, Volta**



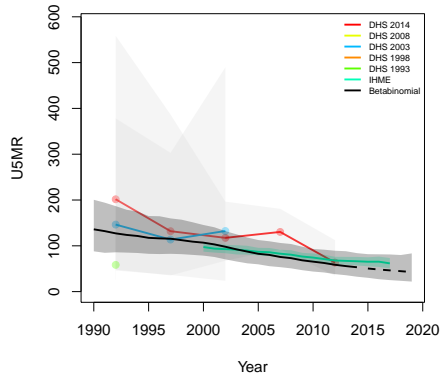




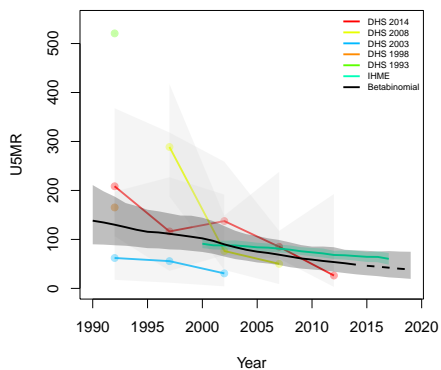
**Ahanta West, Western**



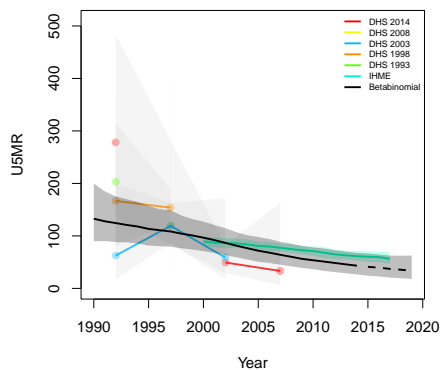
**Aowin-Suaman, Western**



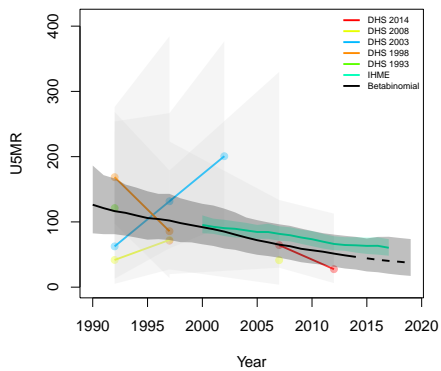
**Bia, Western**



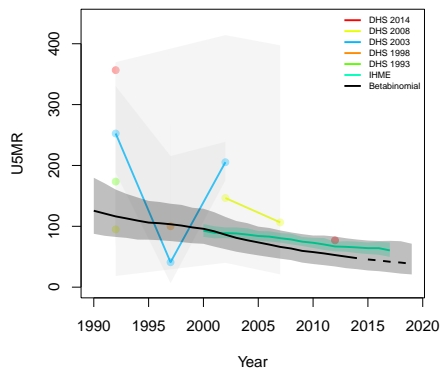
**Bibiani Anhwiaso Bekwai, Western**



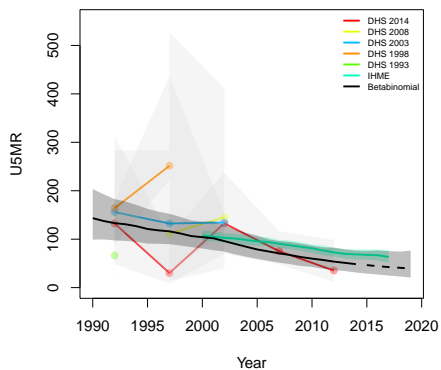
**Jomoro, Western**



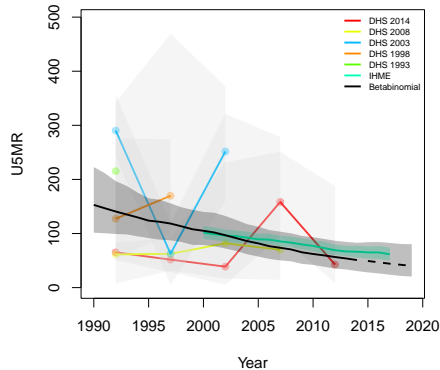
**Juabeso, Western**



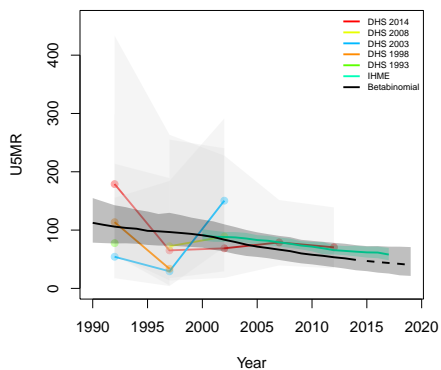
**Mpohor Wassa East, Western**



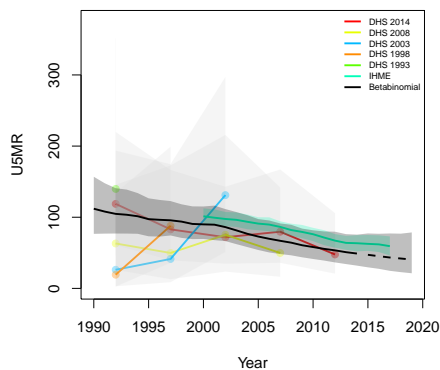
**Nzema East, Western**



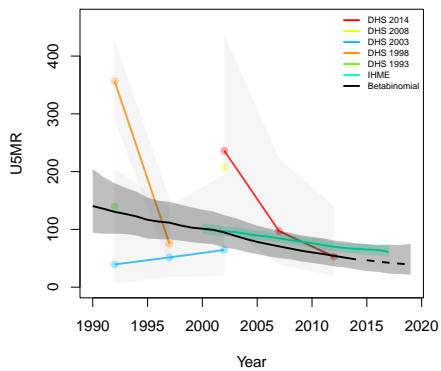
**Sefwi Wiawso, Western**



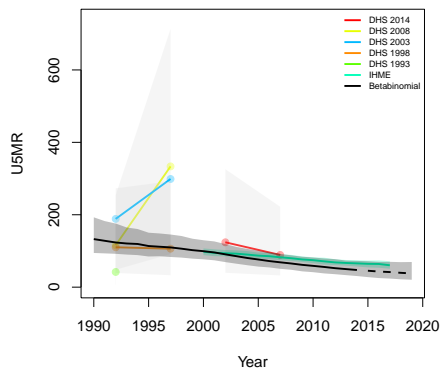
**Shama Ahanta East, Western**

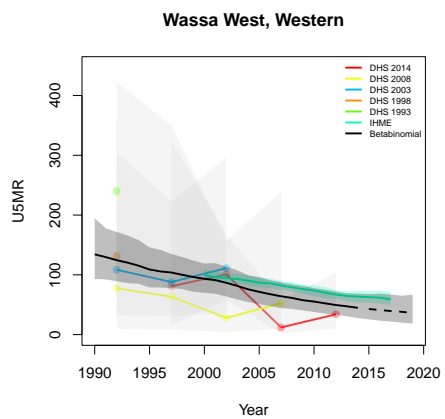


**Wasa Amenfi East, Western**



**Wasa Amenfi West, Western**





## B.5 Kenya

Age	Survey	Clusters			Deaths			Agements		
		Urban	Rural	Total	Urban	Rural	Total	Urban	Rural	Total
0	2003	129	270	399	116	367	483	3577	10613	14190
	2008	132	265	397	116	447	563	3962	14560	18522
	2014	615	969	1584	607	1230	1837	23951	54730	78681
1-11	2003	129	270	399	132	415	547	36671	109181	145852
	2008	132	265	397	100	432	532	40609	149845	190454
	2014	615	969	1584	463	1209	1672	248866	571415	820281
12-23	2003	129	270	399	53	158	211	37546	111622	149168
	2008	132	265	397	39	192	231	40875	153811	194686
	2014	615	969	1584	202	467	669	254503	587820	842323
24-35	2003	129	270	399	27	102	129	35947	107305	143252
	2008	132	265	397	19	126	145	38443	146835	185278
	2014	615	969	1584	92	267	359	240762	558679	799441
36-47	2003	129	270	399	23	60	83	34042	103009	137051
	2008	132	265	397	6	74	80	35895	139151	175046
	2014	615	969	1584	81	214	295	225217	528231	753448
48-59	2003	129	270	399	8	35	43	32279	98373	130652
	2008	132	265	397	5	43	48	33669	132099	165768
	2014	615	969	1584	48	122	170	210286	497579	707865

Table B.5: **Data summary for Kenya.** Total numbers of clusters (Columns 3–5) with observations in each age group by survey in urban and rural areas and combined. Numbers of deaths (Columns 6–8) and number of agements (Columns 9–10) observed in each age group by survey in urban and rural areas and combined.

*B.5.1 Admin-1*

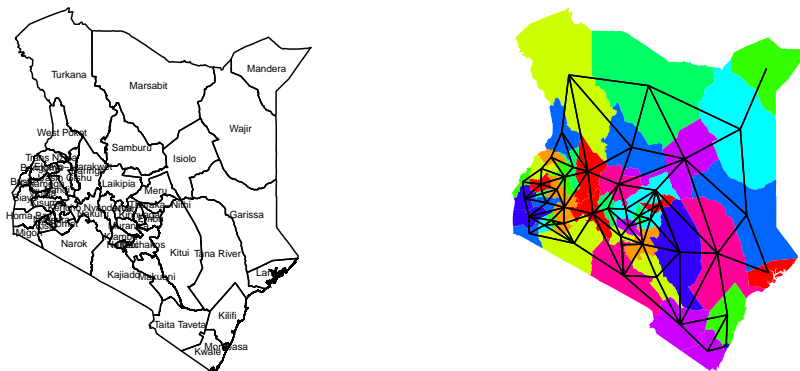


Figure B.25: **Left:** The names of the 47 Admin-1 areas of Kenya . **Right:** The neighborhood structure of Admin-1 areas in Kenya .

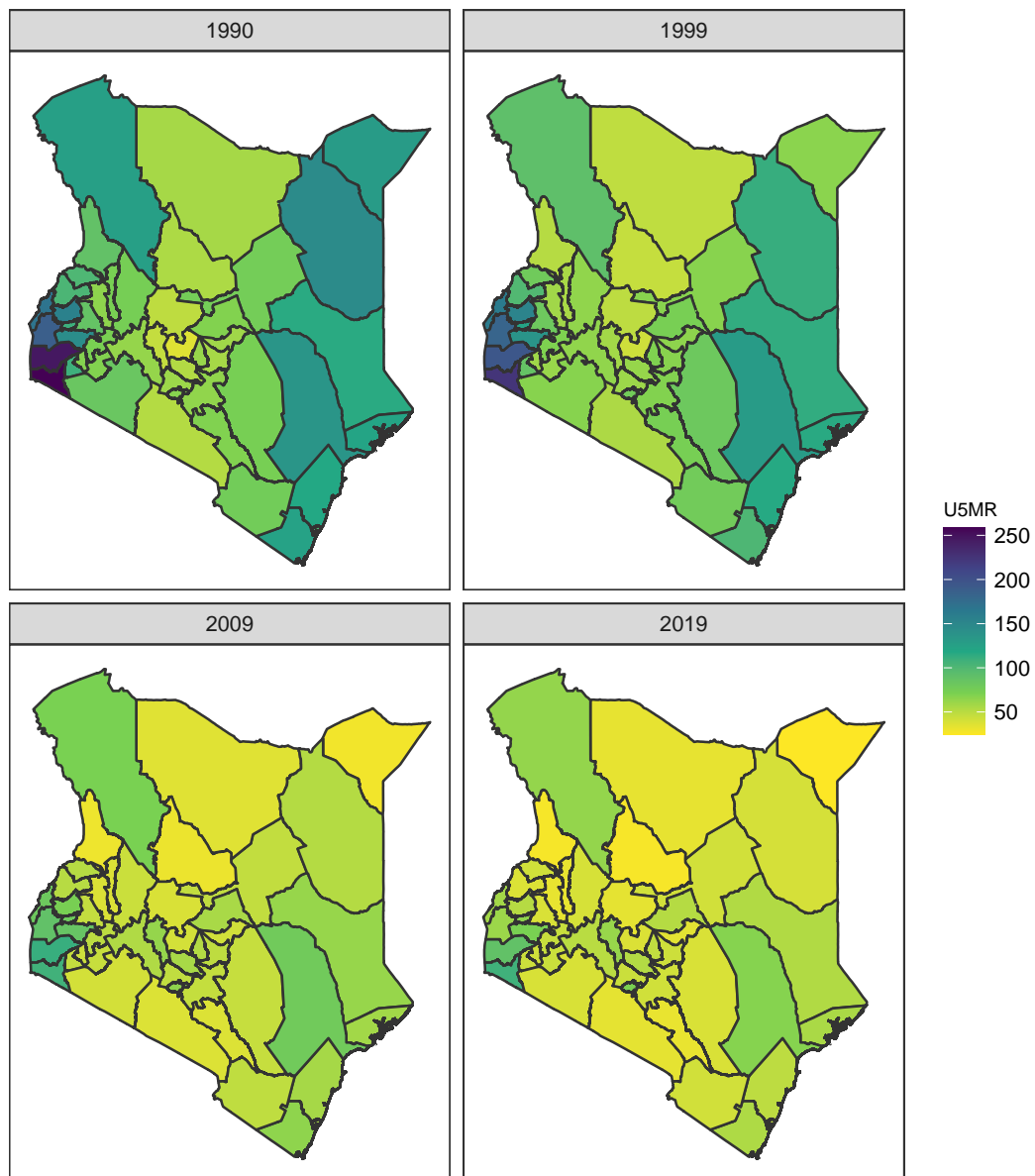


Figure B.26: Median U5MR estimates for years 1990, 1999, 2009, 2019 for Admin-1 areas in Kenya .

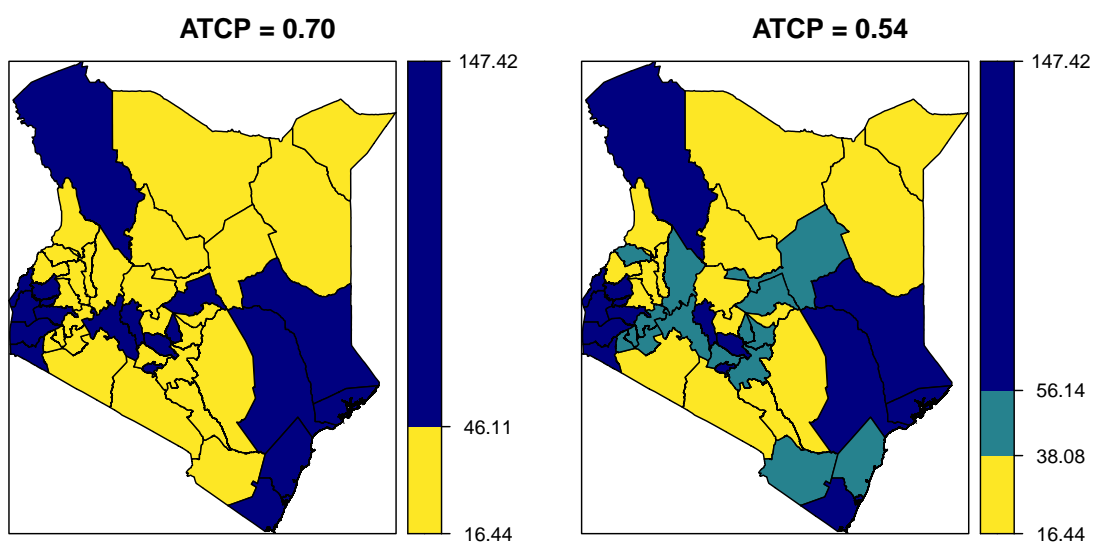
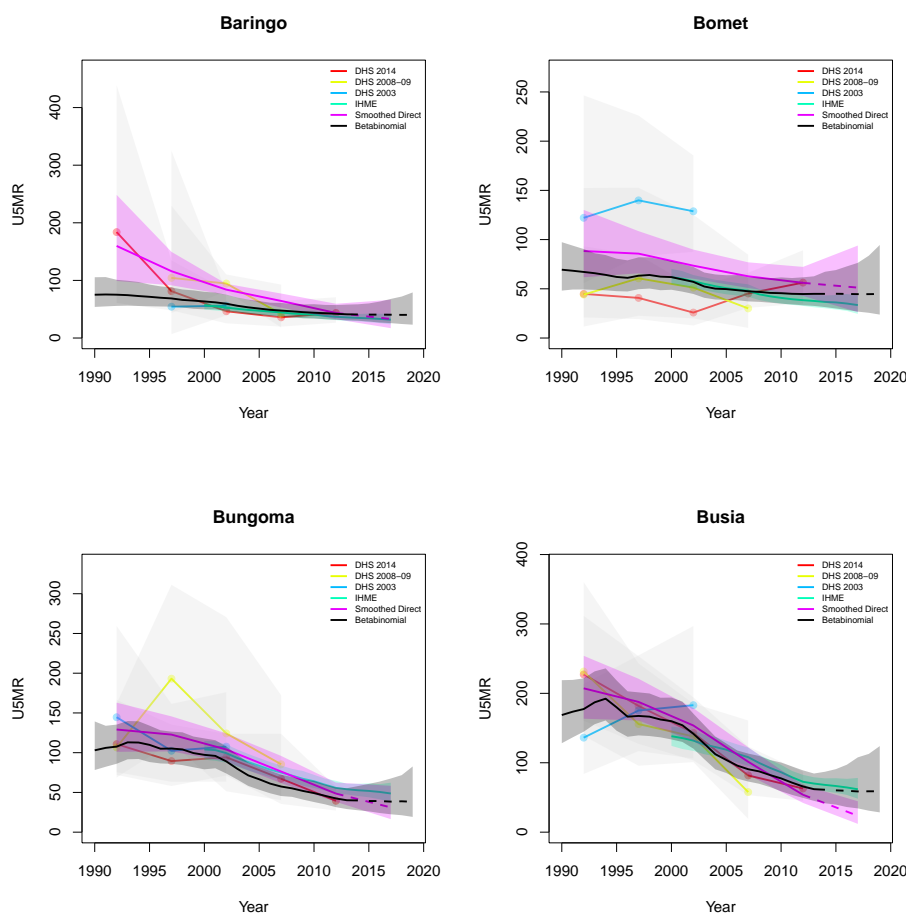


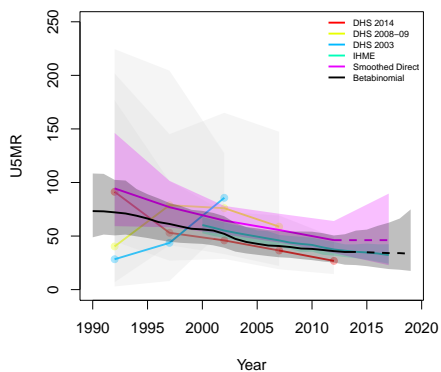
Figure B.27: Expression of uncertainty of U5MR (deaths per 1000 children) estimates for Admin-1 areas based on the average true classification probability (ATCP) in 2019 using  $K = 2, 3$  colors.

*Data and estimates over time by area*

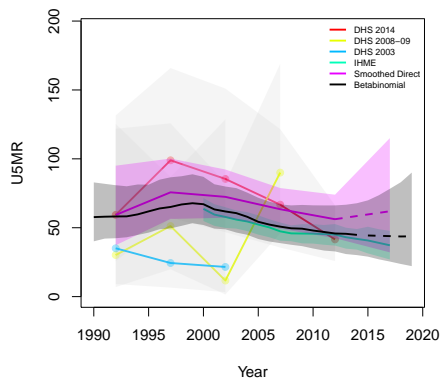
Colored lines with circular points and light grey uncertainty bands are 5-year survey-weighted estimates of U5MR for years 1990–1994 up to 2015–2019 depending on survey timing. For a survey that ends in the middle of a 5-year period, we plot the estimates at the mid-point of the years in that interval for which the survey provides data. Black lines and corresponding intervals represent posterior medians and 95% uncertainty intervals respectively for the betabinomial model. IHME’s estimates and corresponding intervals, where we can compare, are in aquamarine.



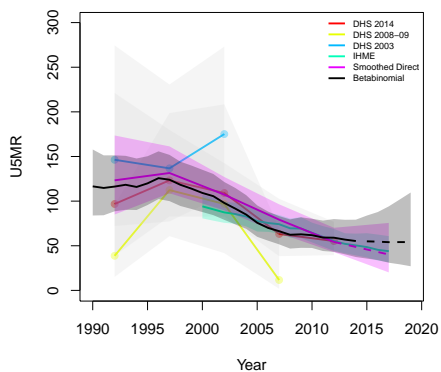
**Elgeyo-Marakwet**



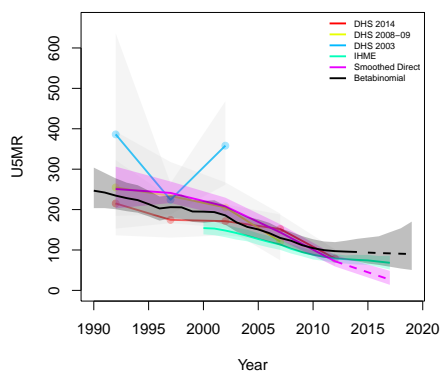
**Embu**



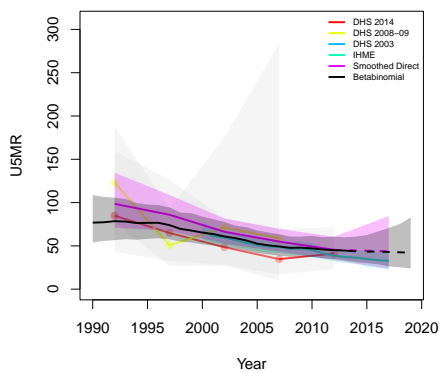
**Garissa**



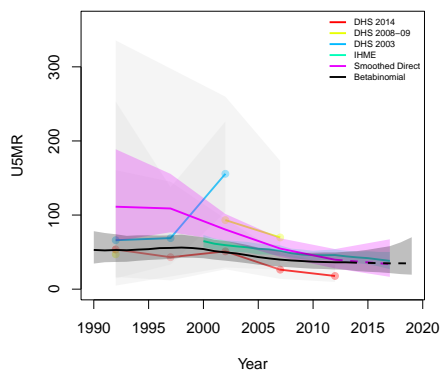
**Homa Bay**

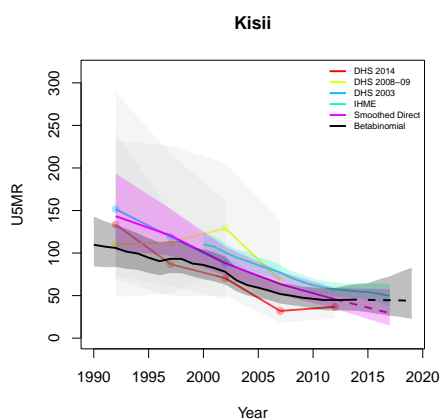
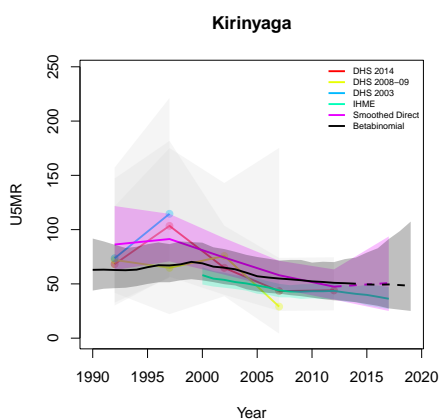
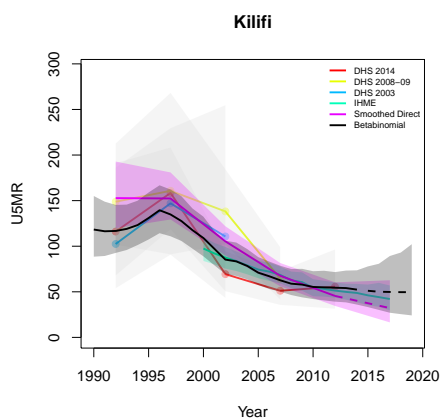
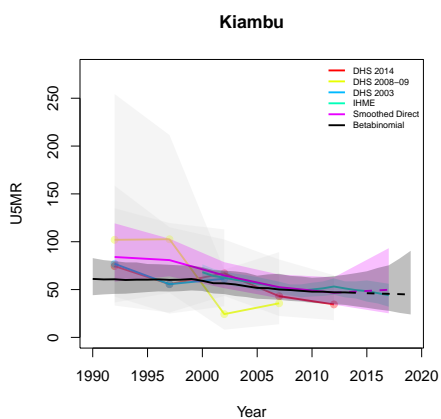
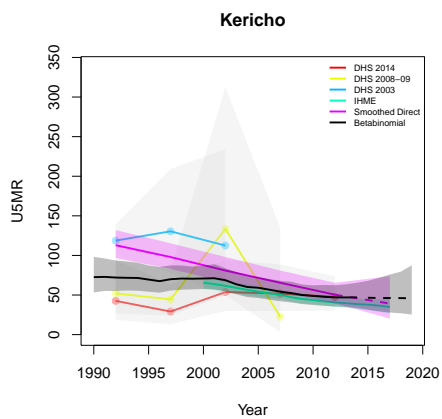
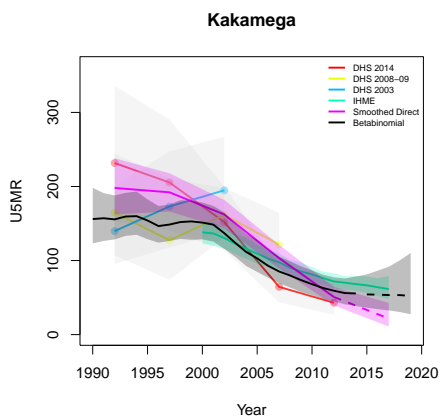


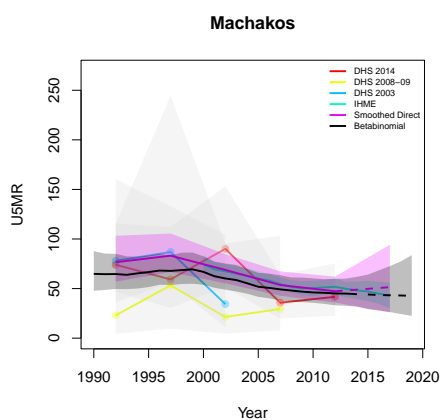
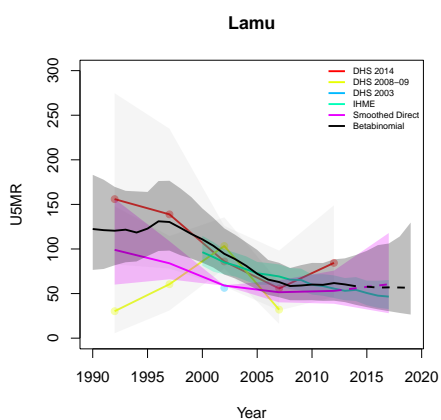
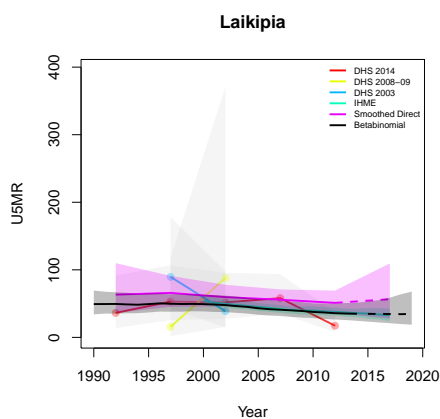
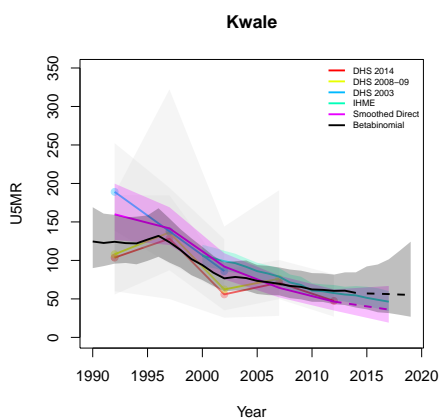
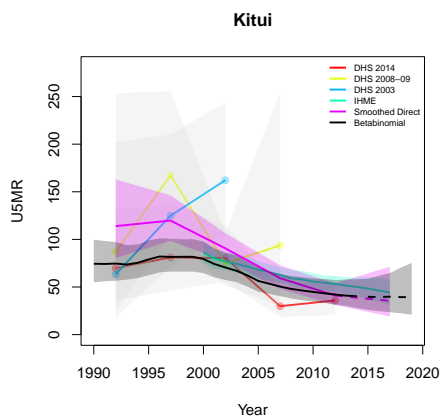
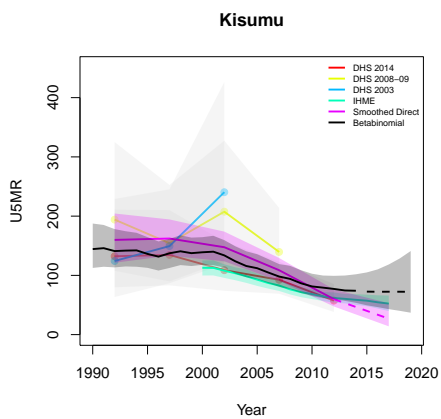
**Isiolo**



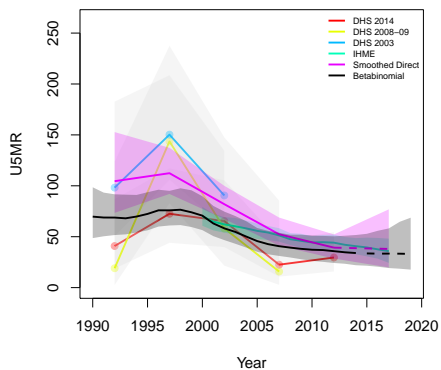
**Kajiado**



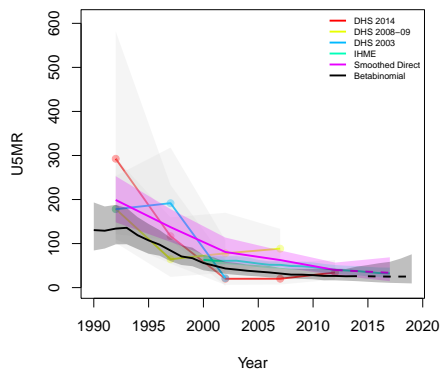




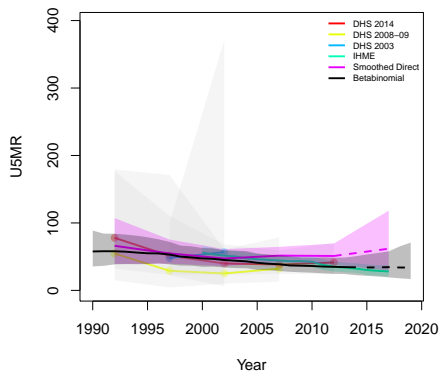
**Makueni**



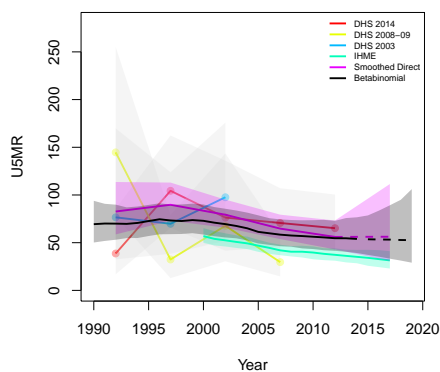
**Mandera**



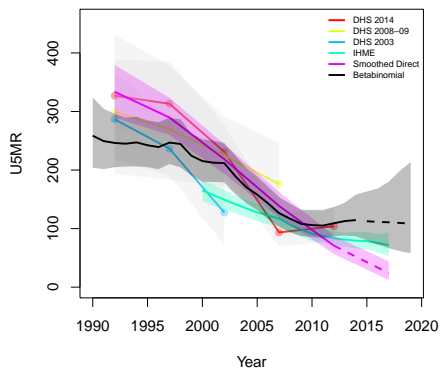
**Marsabit**



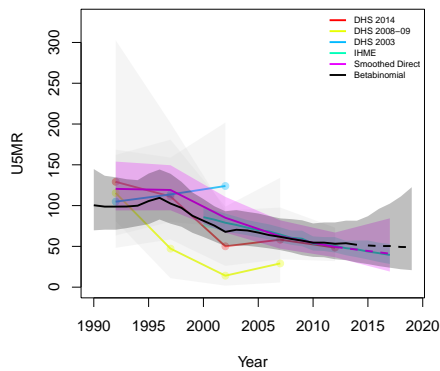
**Meru**

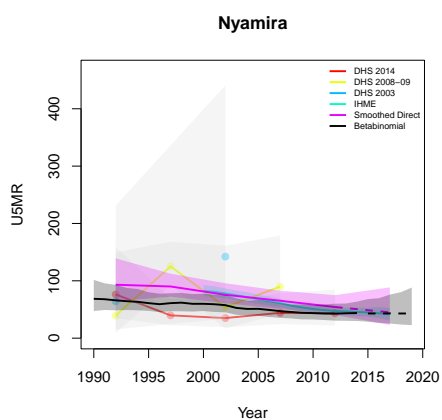
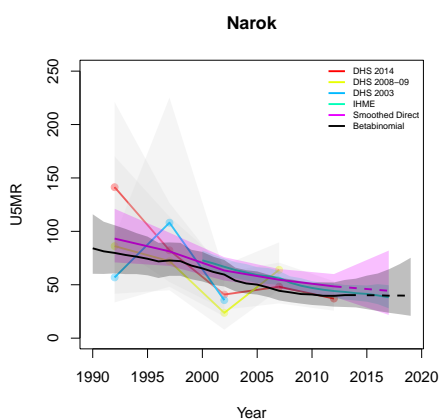
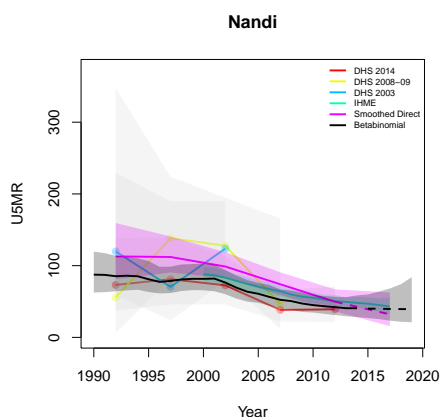
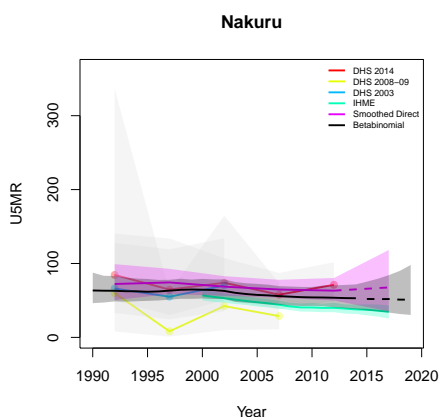
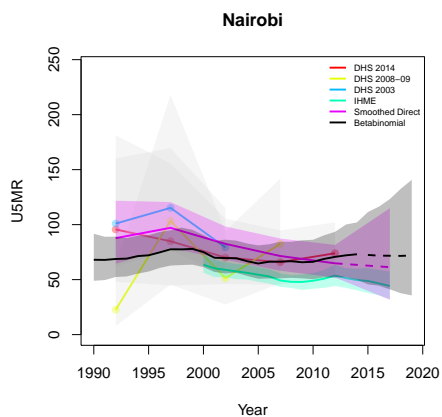
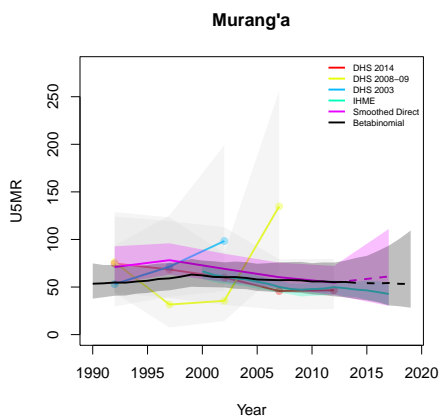


**Migori**

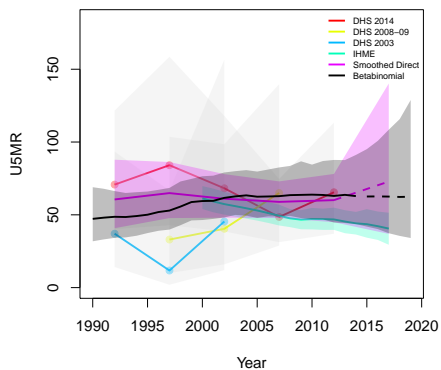


**Mombasa**

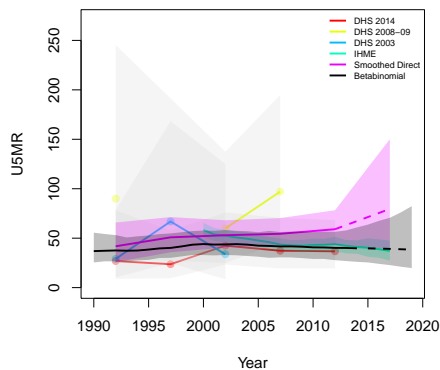




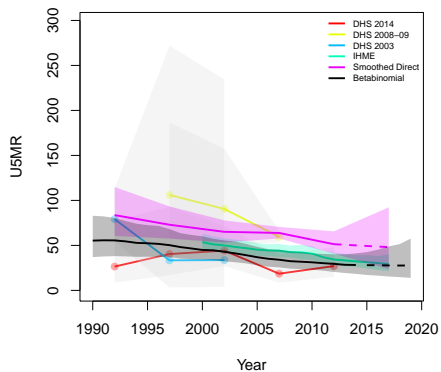
**Nyandarua**



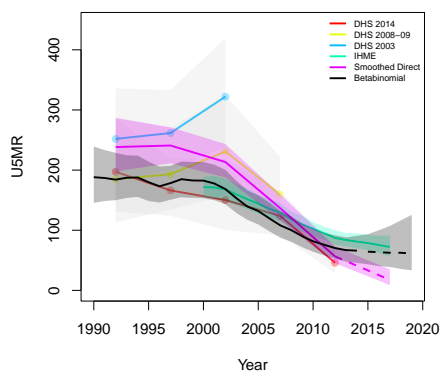
**Nyeri**



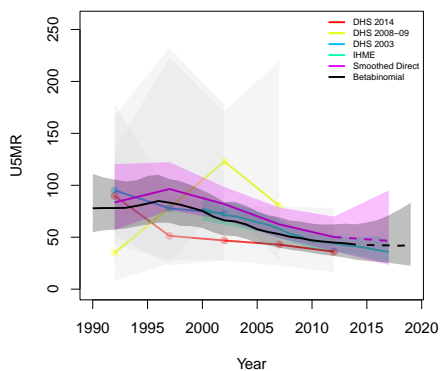
**Samburu**



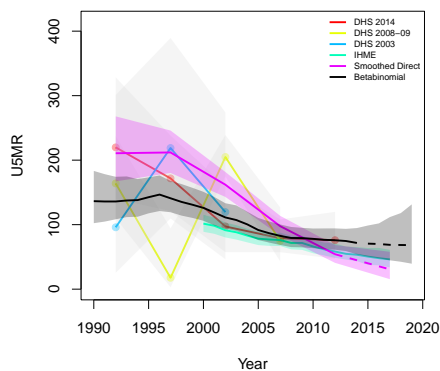
**Siaya**



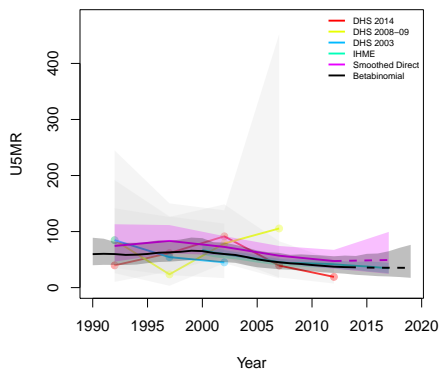
**Taita Taveta**



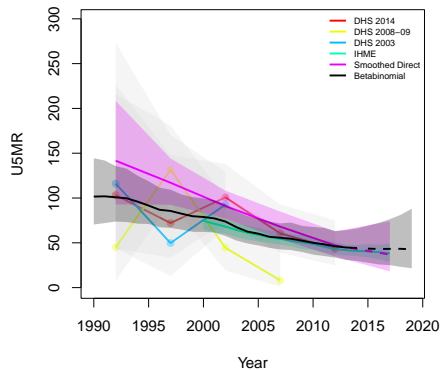
**Tana River**



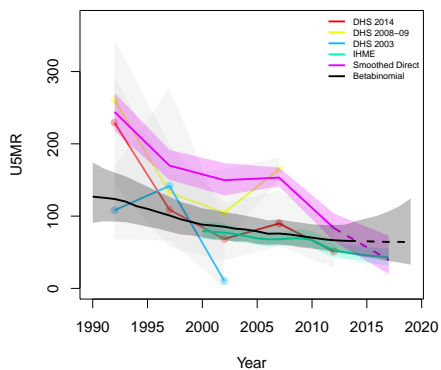
**Tharaka-Nithi**



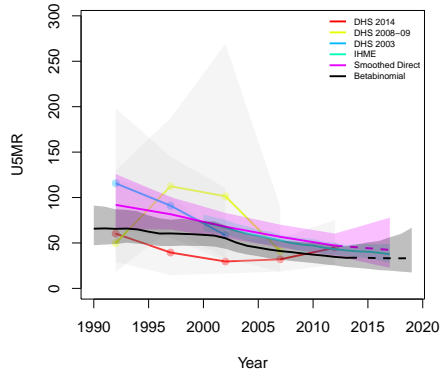
**Trans Nzoia**



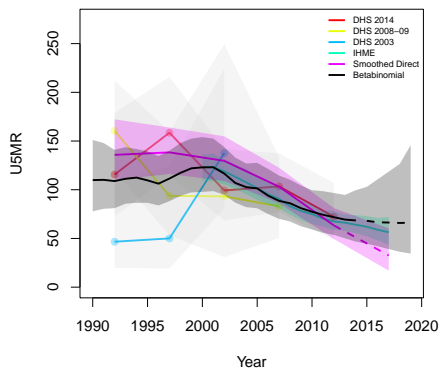
**Turkana**



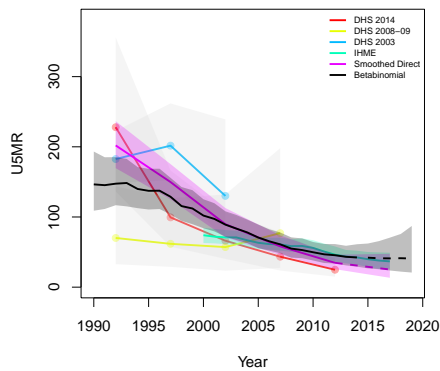
**Uasin Gishu**

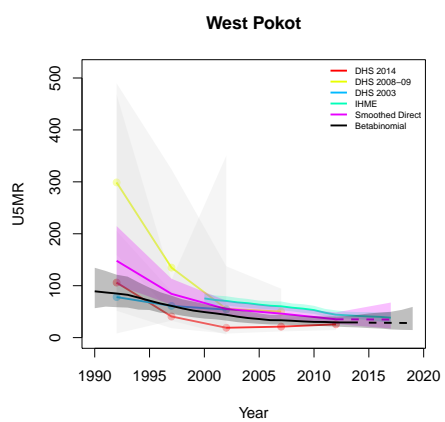


**Vihiga**



**Wajir**





**B.6 Lesotho**

Age	Survey	Clusters			Deaths			Agemonths		
		Urban	Rural	Total	Urban	Rural	Total	Urban	Rural	Total
0	2004	104	277	381	59	355	414	1881	7334	9215
	2009	94	301	395	83	370	453	2265	9600	11865
	2014	118	281	399	74	300	374	2830	8206	11036
1-11	2004	104	277	381	68	261	329	19603	73613	93216
	2009	94	301	395	64	340	404	23151	97008	120159
	2014	118	281	399	108	243	351	29130	83229	112359
12-23	2004	104	277	381	13	57	70	20568	75801	96369
	2009	94	301	395	15	97	112	24115	98581	122696
	2014	118	281	399	28	84	112	29506	83929	113435
24-35	2004	104	277	381	6	29	35	20317	73550	93867
	2009	94	301	395	10	33	43	23295	93573	116868
	2014	118	281	399	13	33	46	27729	78866	106595
36-47	2004	104	277	381	4	20	24	19929	70772	90701
	2009	94	301	395	5	29	34	22419	89272	111691
	2014	118	281	399	5	22	27	26360	74621	100981
48-59	2004	104	277	381	7	17	24	19512	68677	88189
	2009	94	301	395	1	7	8	21489	85128	106617
	2014	118	281	399	2	12	14	24783	70735	95518

Table B.6: **Data summary for Lesotho.** Total numbers of clusters (Columns 3–5) with observations in each age group by survey in urban and rural areas and combined. Numbers of deaths (Columns 6–8) and number of agemonths (Columns 9–10) observed in each age group by survey in urban and rural areas and combined.

*B.6.1 Admin-1*



Figure B.28: **Left:** The names of the 10 Admin-1 areas of Lesotho . **Right:** The neighborhood structure of Admin-1 areas in Lesotho .

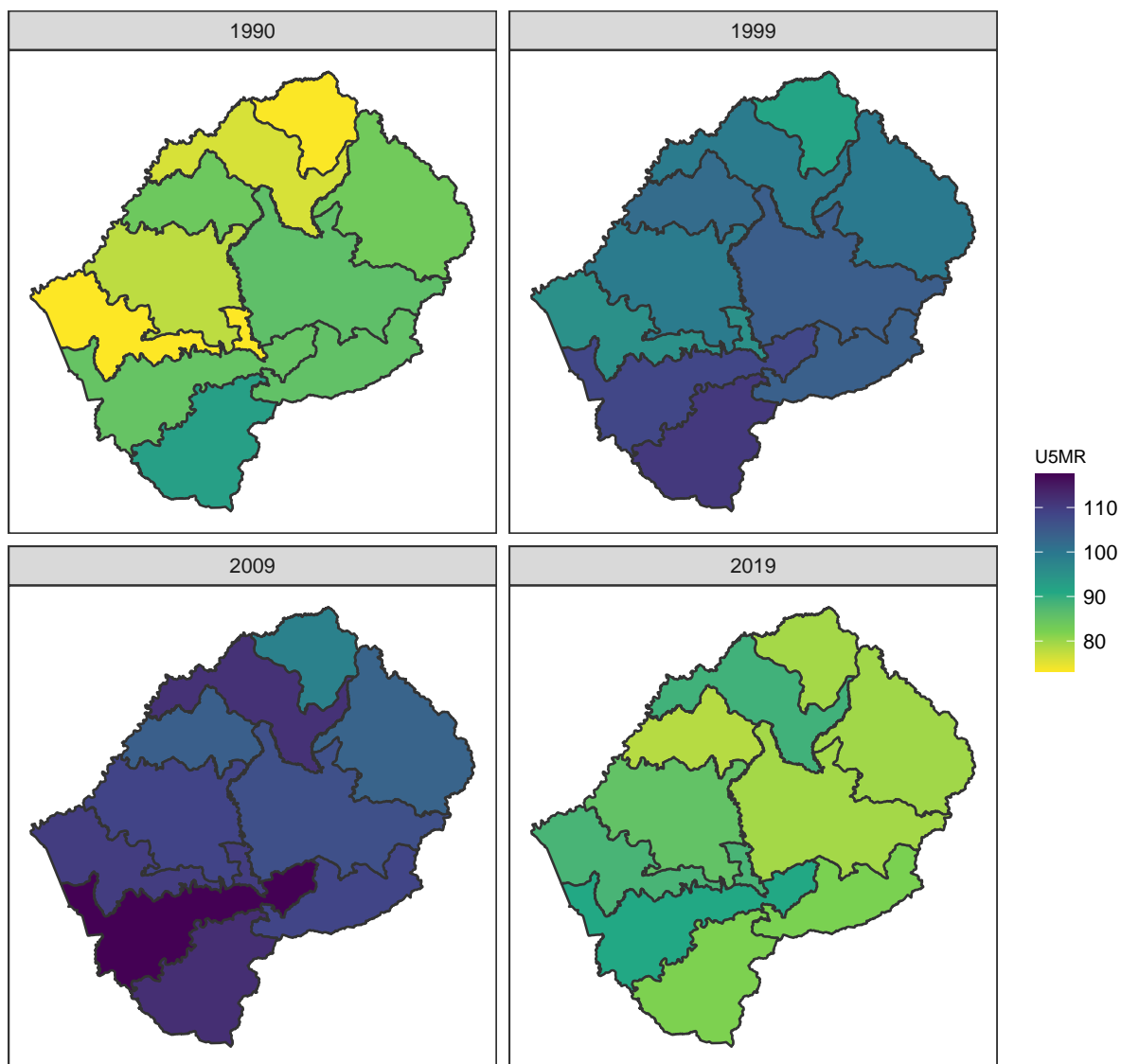


Figure B.29: Median U5MR estimates for years 1990, 1999, 2009, 2019 for Admin-1 areas in Lesotho .

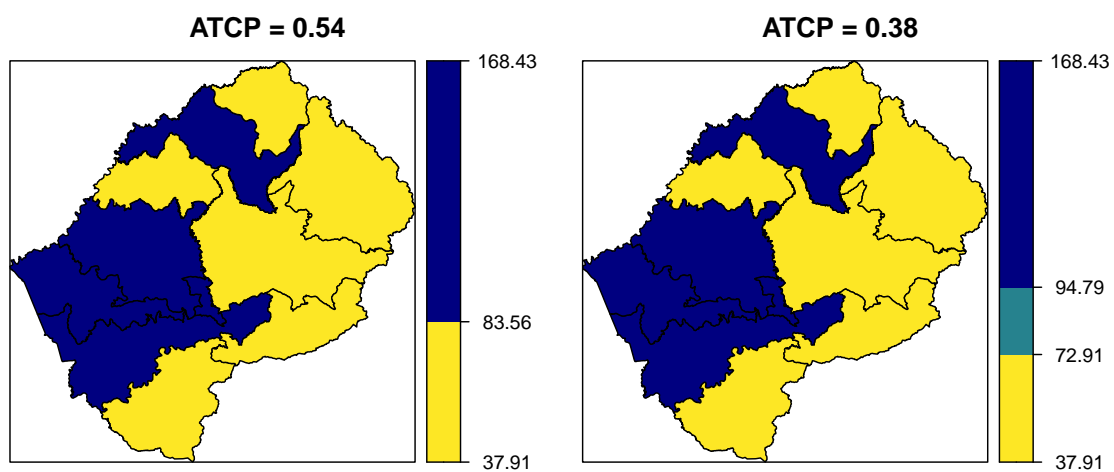
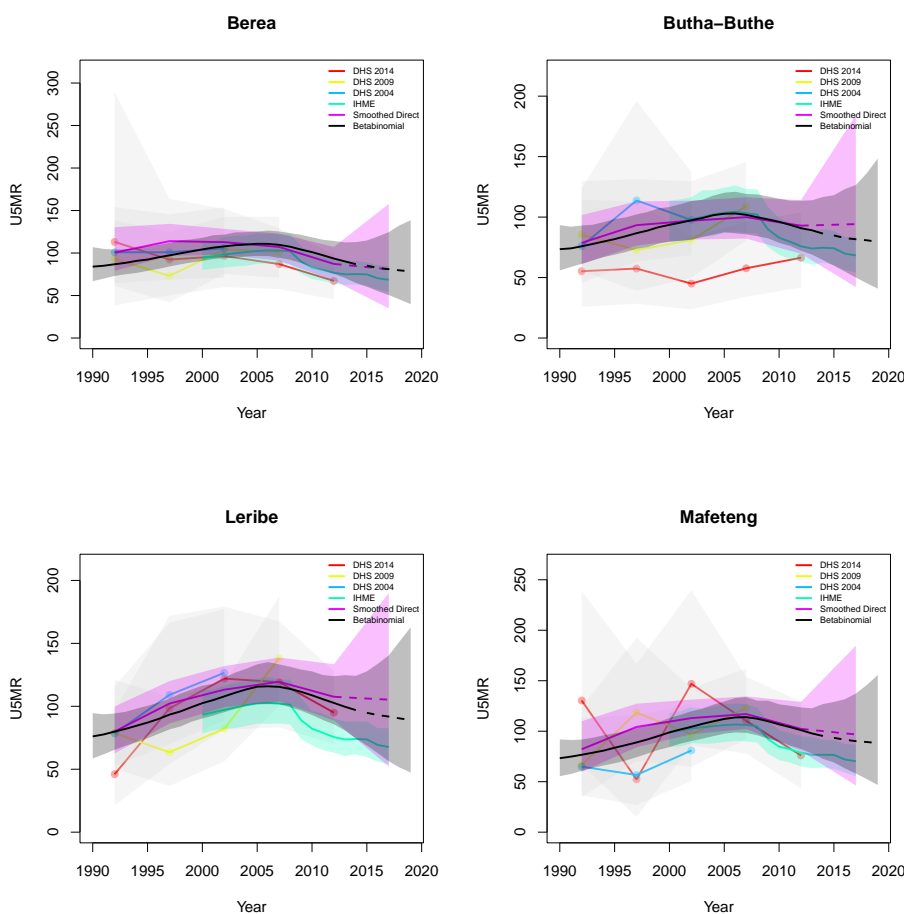
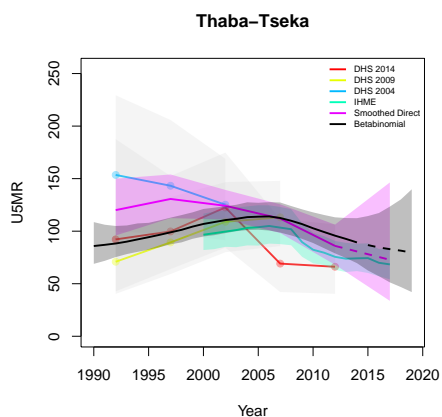
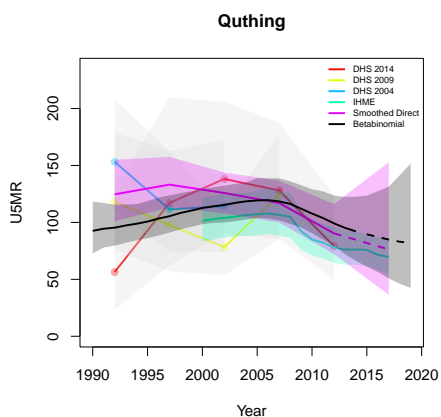
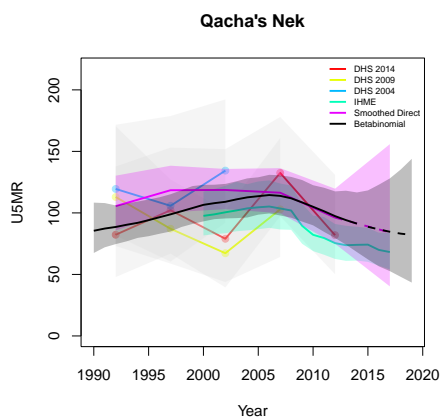
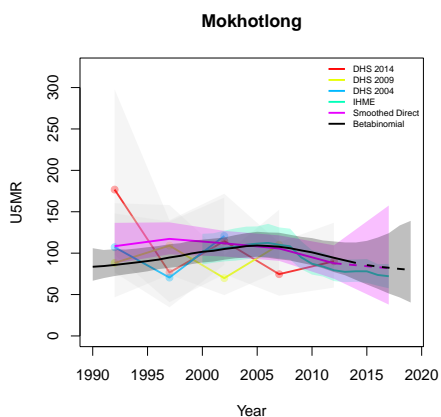
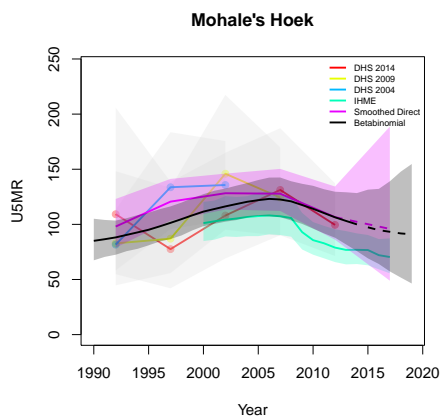
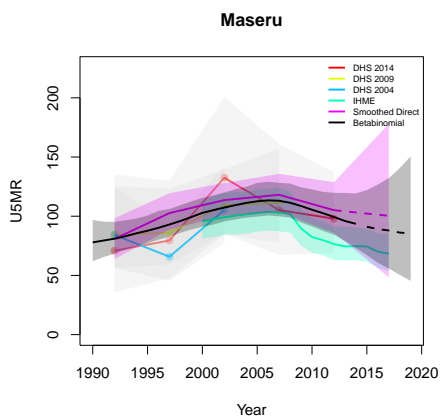


Figure B.30: Expression of uncertainty of U5MR (deaths per 1000 children) estimates for Admin-1 areas based on the average true classification probability (ATCP) in 2019 using  $K = 2, 3$  colors.

*Data and estimates over time by area*

Colored lines with circular points and light grey uncertainty bands are 5-year survey-weighted estimates of U5MR for years 1990–1994 up to 2015–2019 depending on survey timing. For a survey that ends in the middle of a 5-year period, we plot the estimates at the mid-point of the years in that interval for which the survey provides data. Black lines and corresponding intervals represent posterior medians and 95% uncertainty intervals respectively for the betabinomial model. IHME’s estimates and corresponding intervals, where we can compare, are in aquamarine.





**B.7 Liberia**

Age	Survey	Clusters			Deaths			Agemonths		
		Urban	Rural	Total	Urban	Rural	Total	Urban	Rural	Total
0	2007	110	181	291	226	441	667	6059	10560	16619
	2013	119	203	322	409	765	1174	8914	18864	27778
1-11	2007	110	181	291	361	734	1095	61132	104511	165643
	2013	119	203	322	461	1156	1617	89262	188388	277650
12-23	2007	110	181	291	158	236	394	60184	102854	163038
	2013	119	203	322	216	500	716	88630	185275	273905
24-35	2007	110	181	291	77	162	239	56884	97045	153929
	2013	119	203	322	115	276	391	83327	174038	257365
36-47	2007	110	181	291	62	75	137	53355	90578	143933
	2013	119	203	322	90	177	267	78039	162510	240549
48-59	2007	110	181	291	20	55	75	50407	84723	135130
	2013	119	203	322	43	96	139	73087	151833	224920

Table B.7: **Data summary for Liberia.** Total numbers of clusters (Columns 3–5) with observations in each age group by survey in urban and rural areas and combined. Numbers of deaths (Columns 6–8) and number of agemonths (Columns 9–10) observed in each age group by survey in urban and rural areas and combined.

**B.7.1 Admin-1**

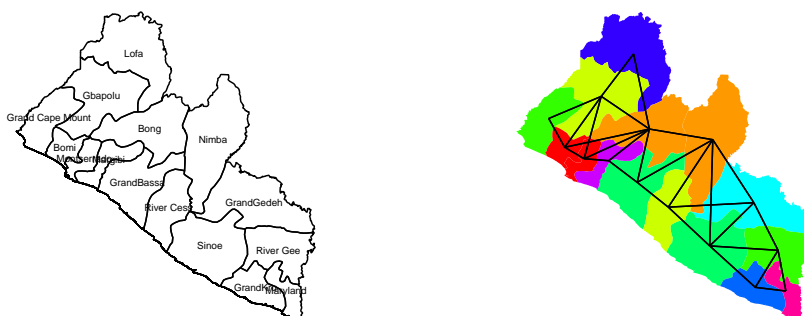


Figure B.31: **Left:** The names of the 15 Admin-1 areas of Liberia . **Right:** The neighborhood structure of Admin-1 areas in Liberia .

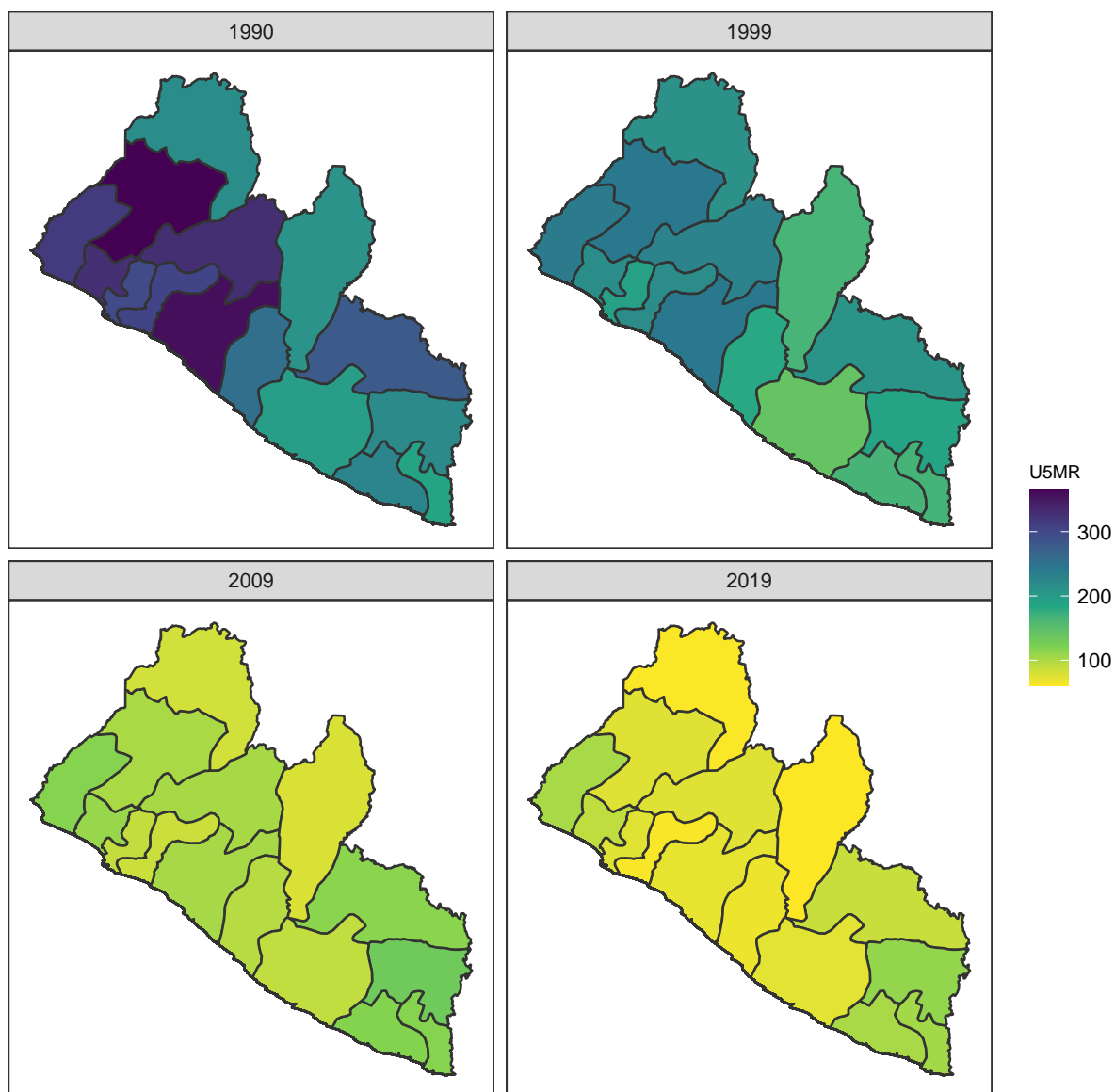


Figure B.32: Median U5MR estimates for years 1990, 1999, 2009, 2019 for Admin-1 areas in Liberia .

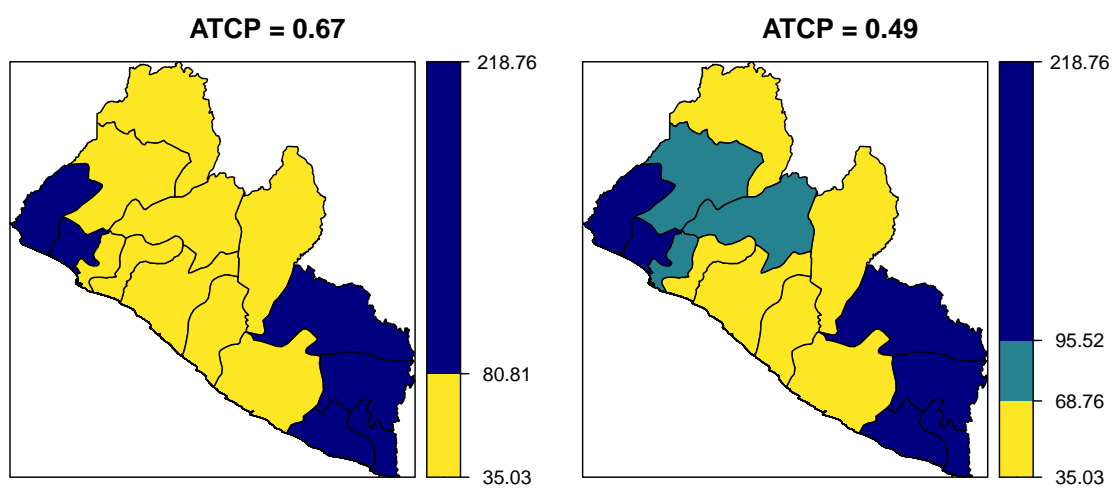
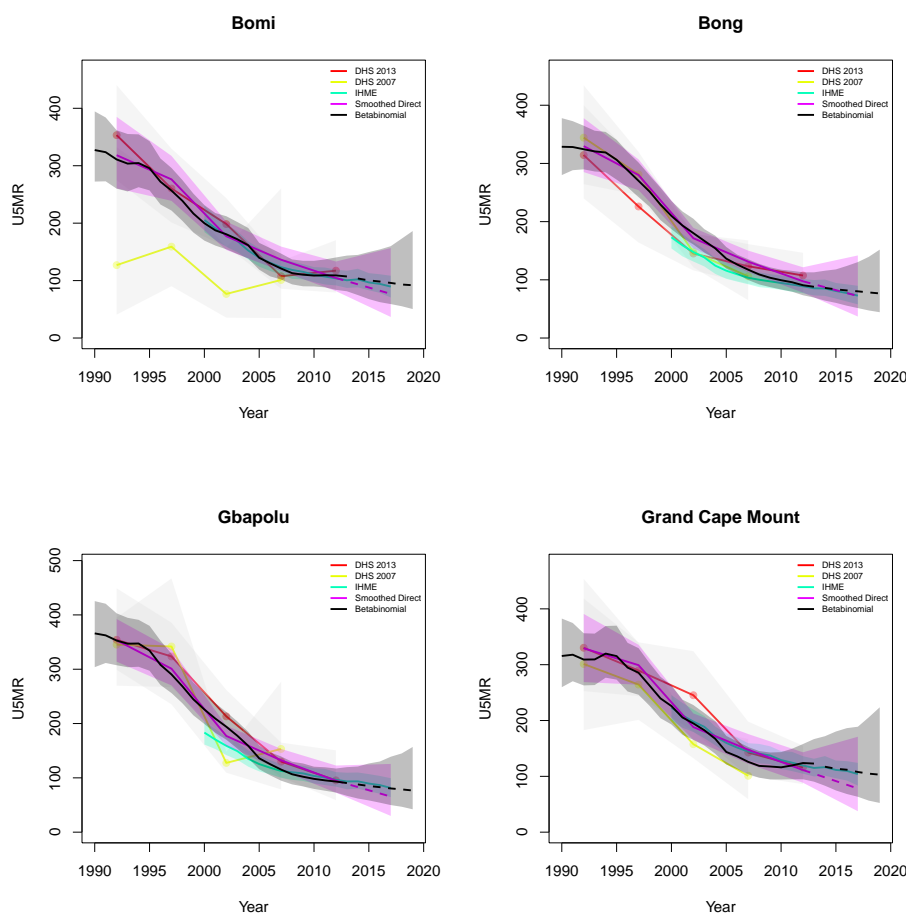
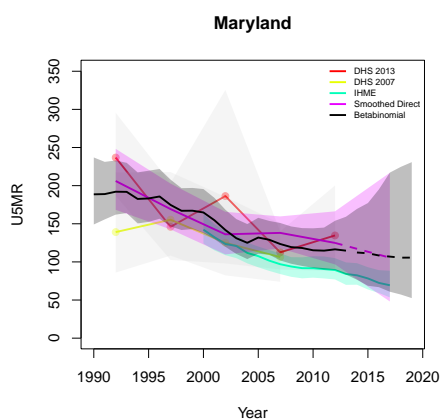
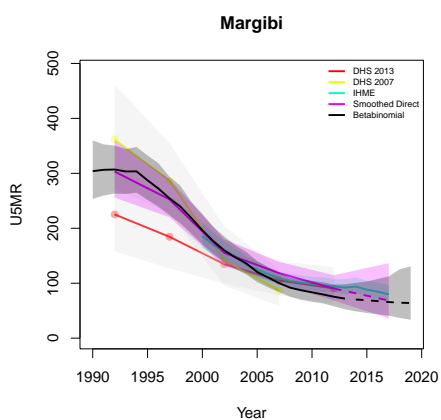
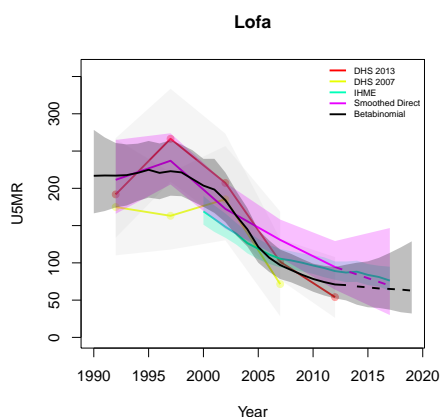
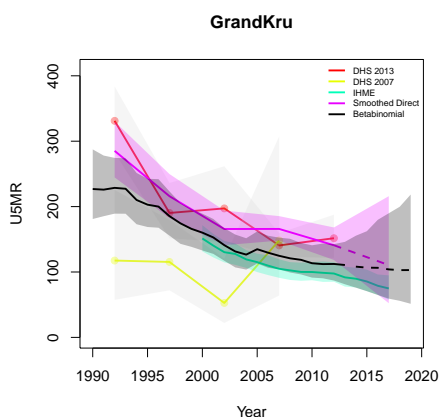
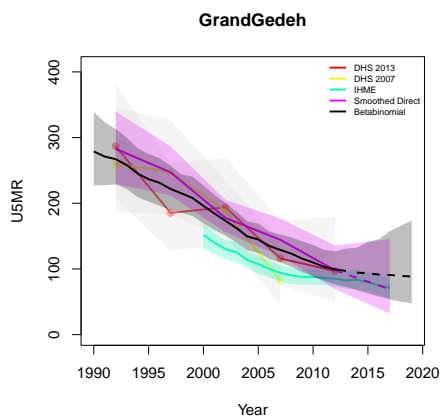
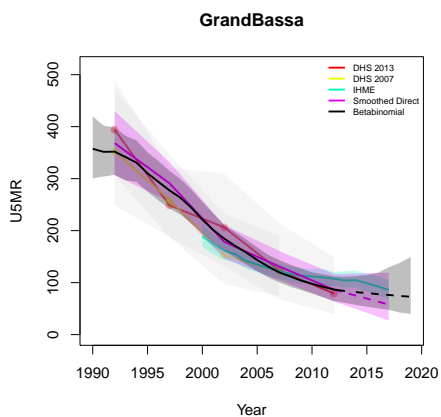


Figure B.33: Expression of uncertainty of U5MR (deaths per 1000 children) estimates for Admin-1 areas based on the average true classification probability (ATCP) in 2019 using  $K = 2, 3$  colors.

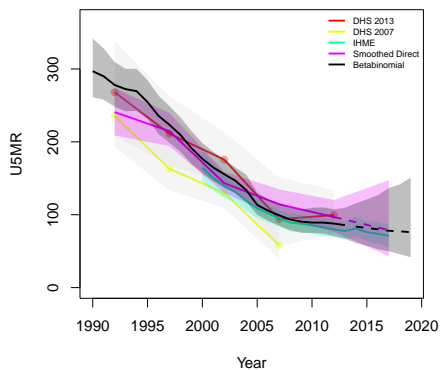
*Data and estimates over time by area*

Colored lines with circular points and light grey uncertainty bands are 5-year survey-weighted estimates of U5MR for years 1990–1994 up to 2015–2019 depending on survey timing. For a survey that ends in the middle of a 5-year period, we plot the estimates at the mid-point of the years in that interval for which the survey provides data. Black lines and corresponding intervals represent posterior medians and 95% uncertainty intervals respectively for the betabinomial model. IHME’s estimates and corresponding intervals, where we can compare, are in aquamarine.

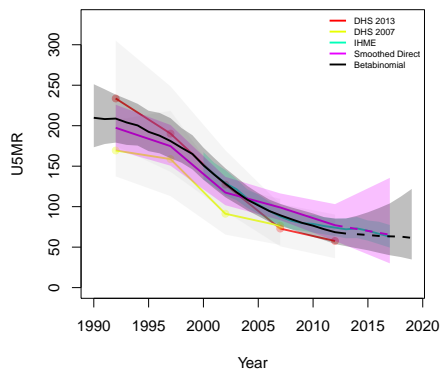




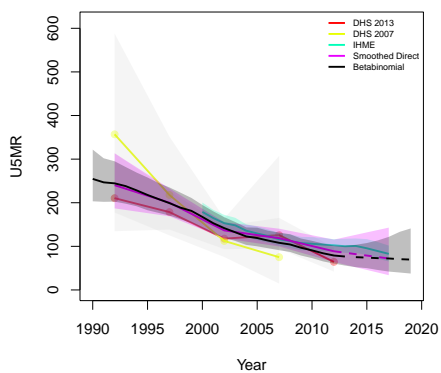
**Montserrat**



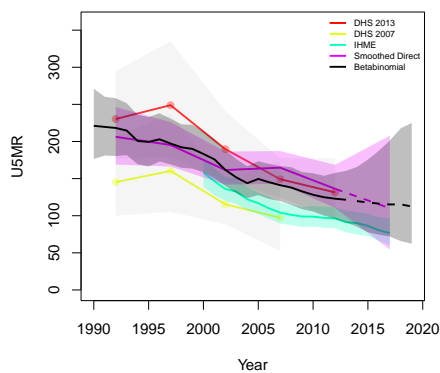
**Nimba**



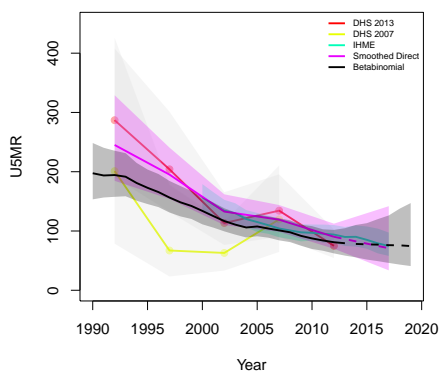
**River Cess**



**River Gee**



**Sinoe**



*B.7.2 Admin-2*



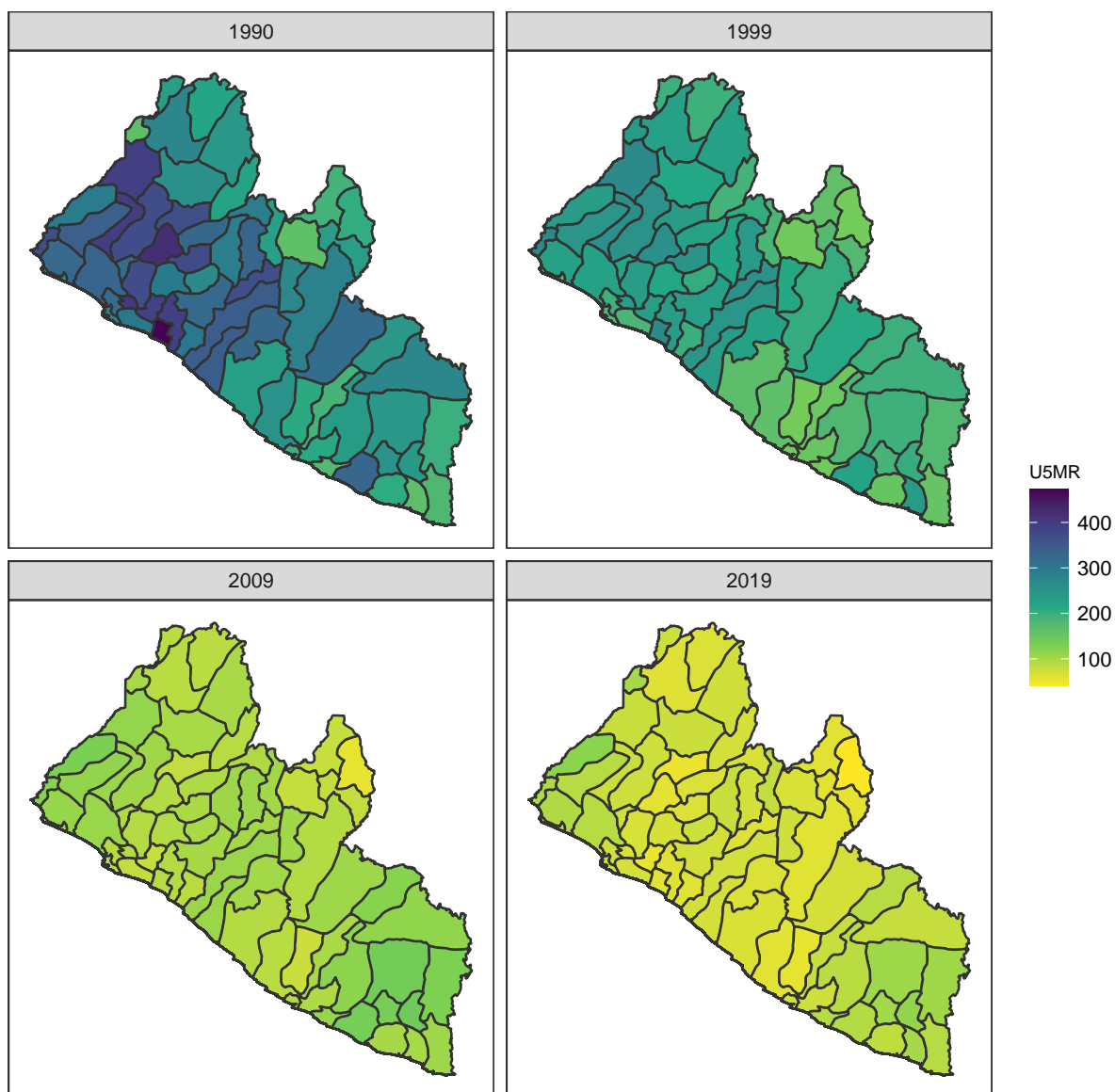


Figure B.35: Median U5MR estimates for years 1990, 1999, 2009, 2019 for Admin-2 areas in Liberia .

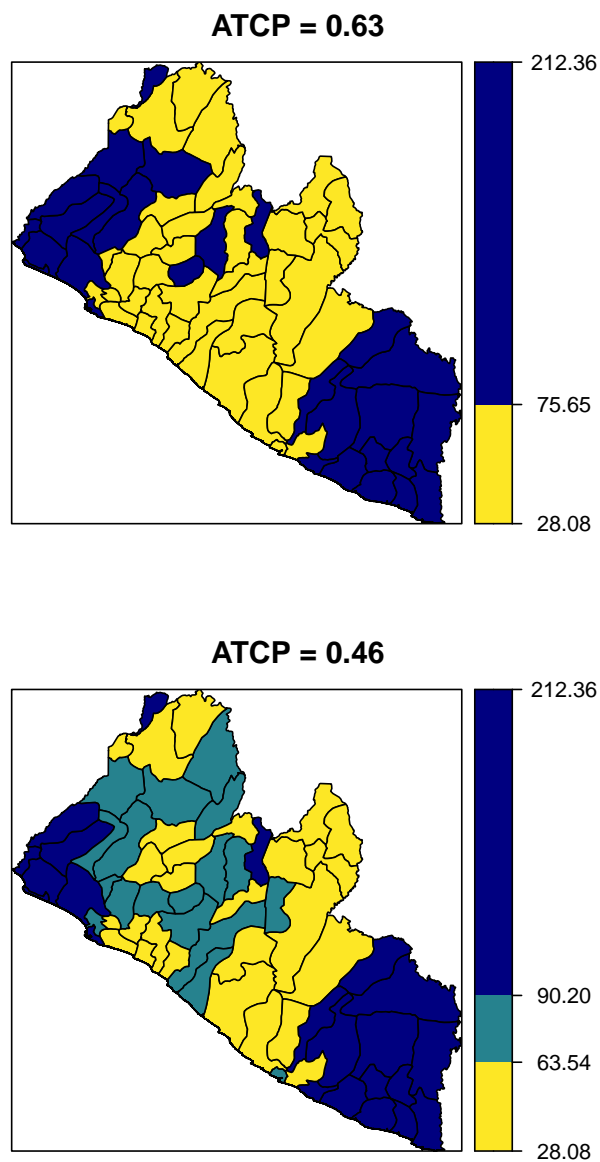
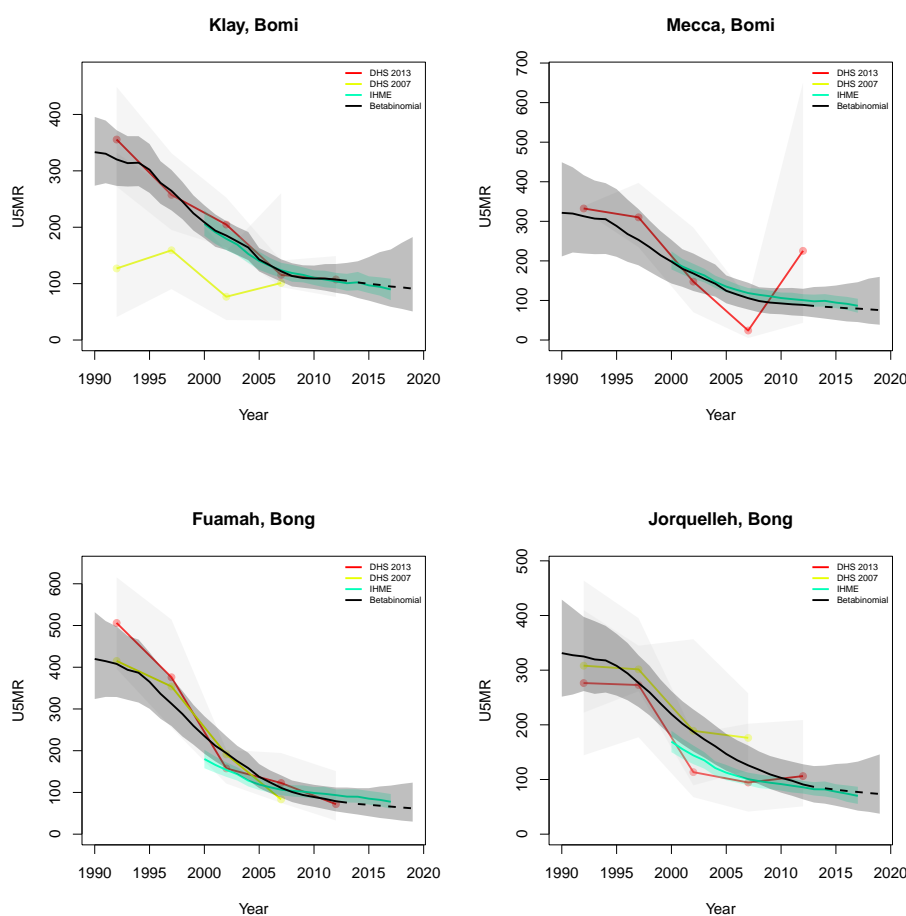


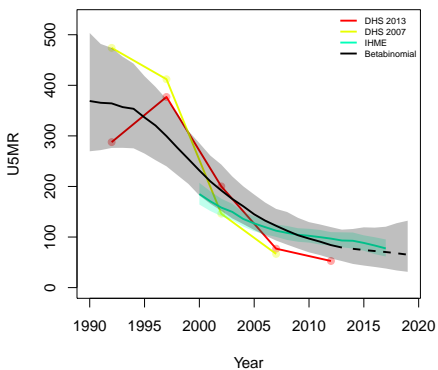
Figure B.36: Expression of uncertainty of U5MR (deaths per 1000 children) estimates for Admin-1 areas based on the average true classification probability (ATCP) in 2019 using  $K = 2, 3$  colors.

*Data and estimates over time by area*

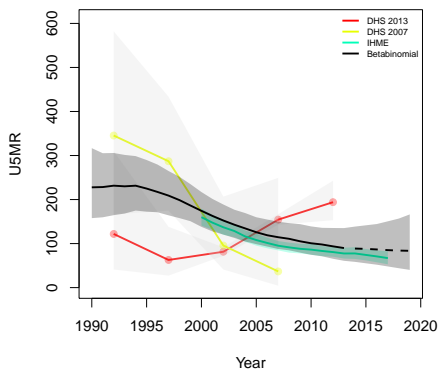
Colored lines with circular points and light grey uncertainty bands are 5-year survey-weighted estimates of U5MR for years 1990–1994 up to 2015–2019 depending on survey timing. For a survey that ends in the middle of a 5-year period, we plot the estimates at the mid-point of the years in that interval for which the survey provides data. Black lines and corresponding intervals represent posterior medians and 95% uncertainty intervals respectively for the betabinomial model. IHME’s estimates and corresponding intervals, where we can compare, are in aquamarine.



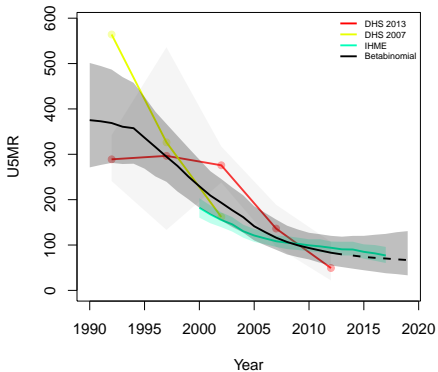
**Kokoyah, Bong**



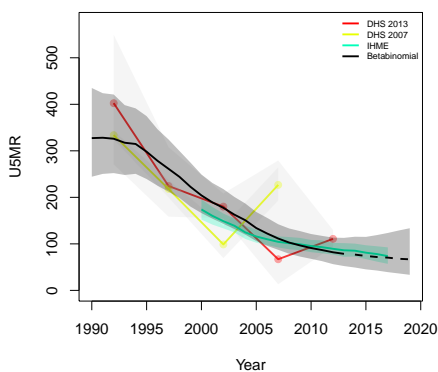
**Panta-Kpa, Bong**



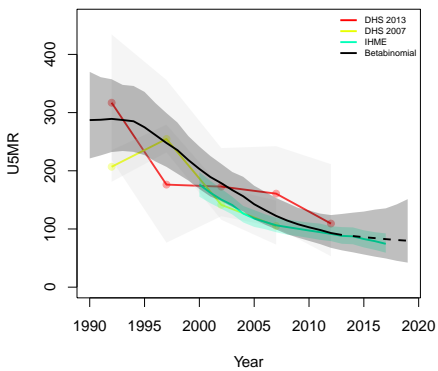
**Salala, Bong**



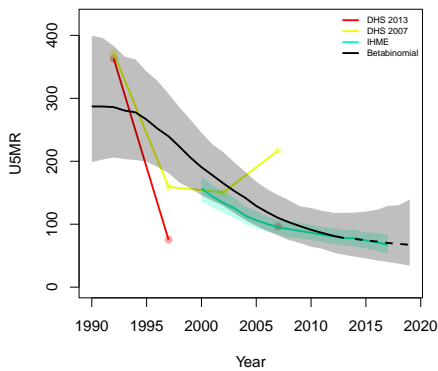
**Sanayea, Bong**



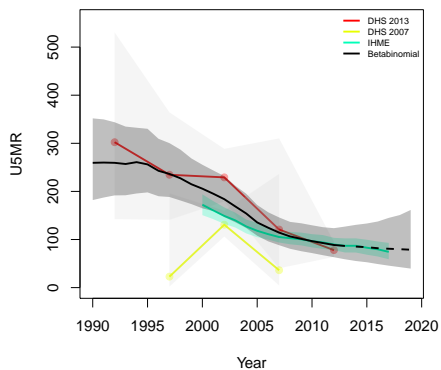
**Suakoko, Bong**



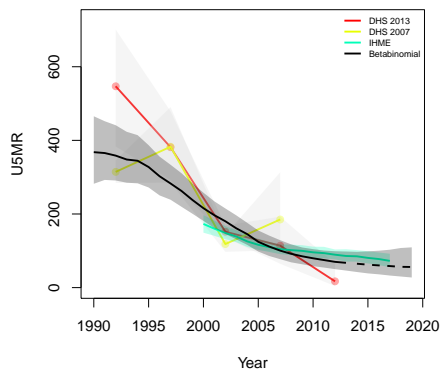
**Zota, Bong**



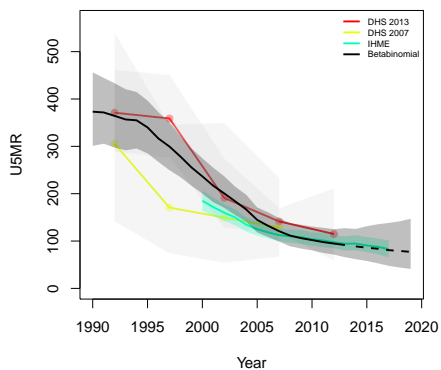
**Belleh, Gbapolu**



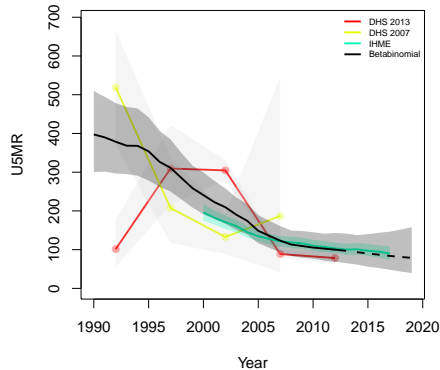
**Bokomu, Gbapolu**



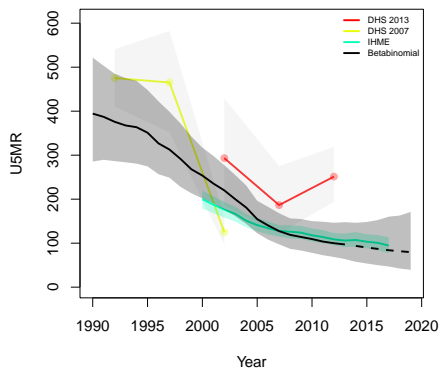
**Bopolu, Gbapolu**



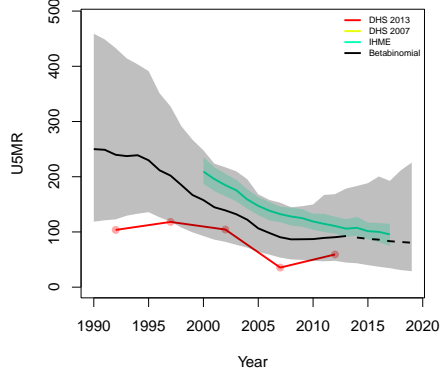
**Gbarma, Gbapolu**



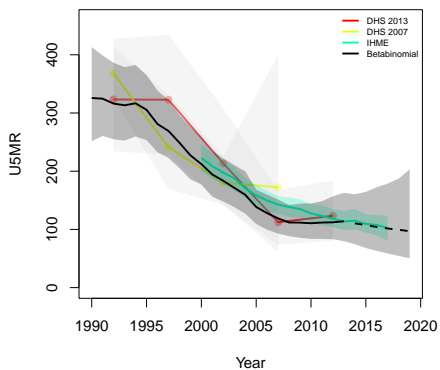
**Kongba, Gbapolu**



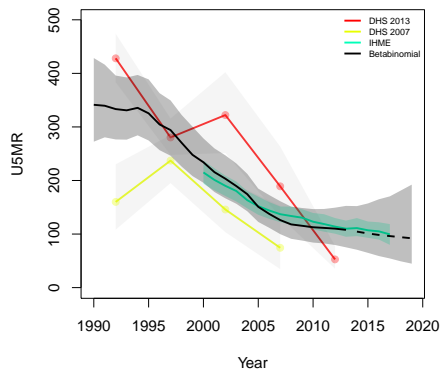
**Commwealth, Grand Cape Mount**



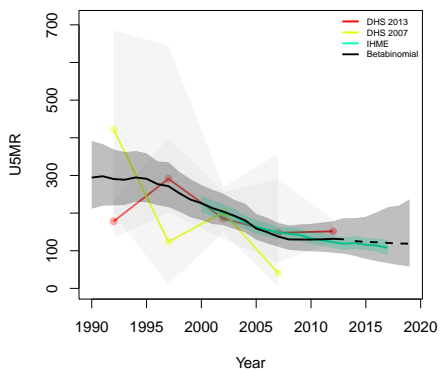
**Garwula, Grand Cape Mount**



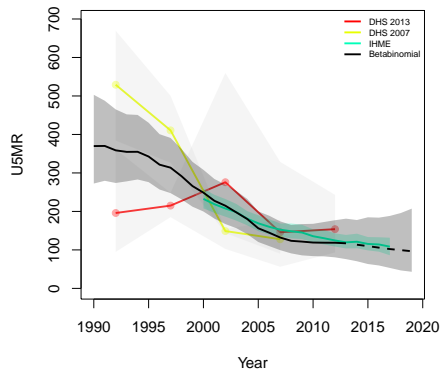
**Golakonneh, Grand Cape Mount**



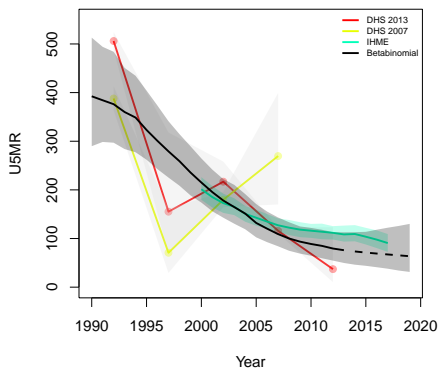
**Porkpa, Grand Cape Mount**



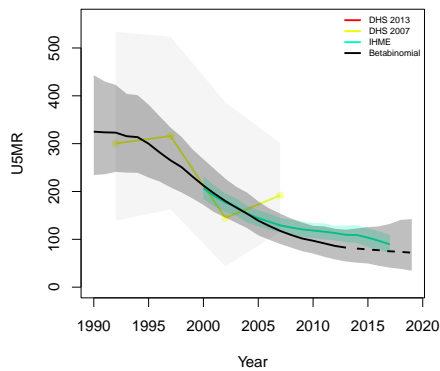
**Tewor, Grand Cape Mount**



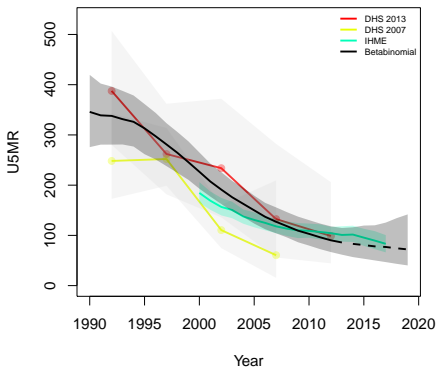
**District # 1, GrandBassa**



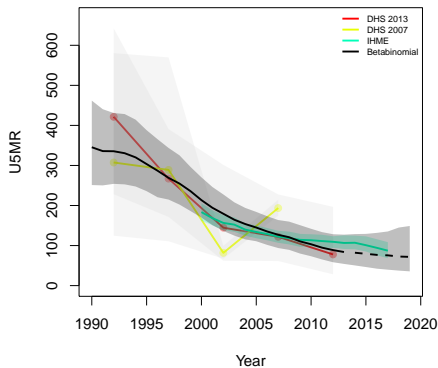
**District # 2, GrandBassa**



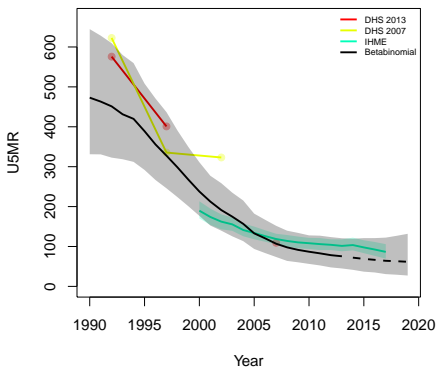
**District # 3, GrandBassa**



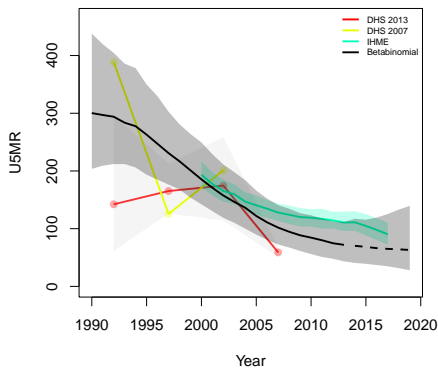
**District # 4, GrandBassa**



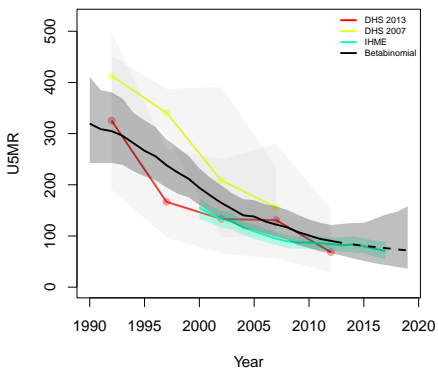
**Owensgrove, GrandBassa**



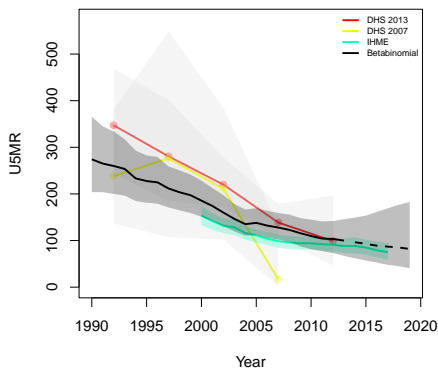
**Stjohnriver, GrandBassa**

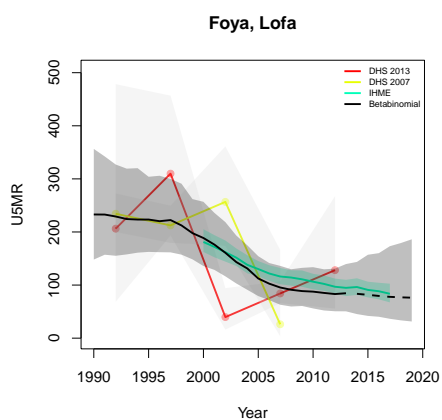
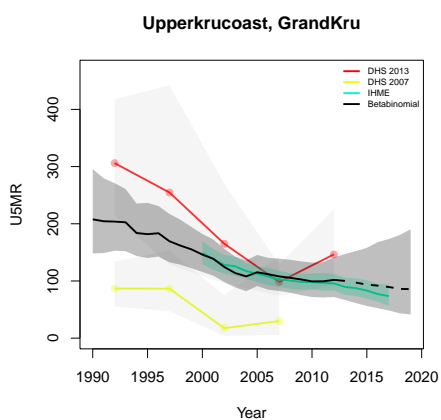
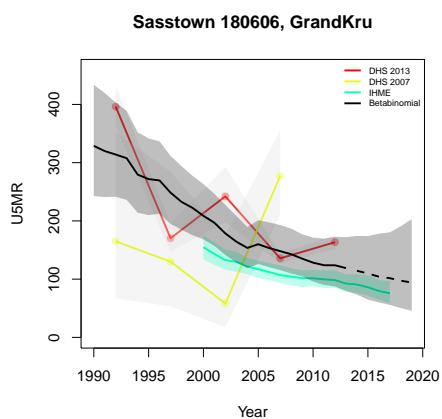
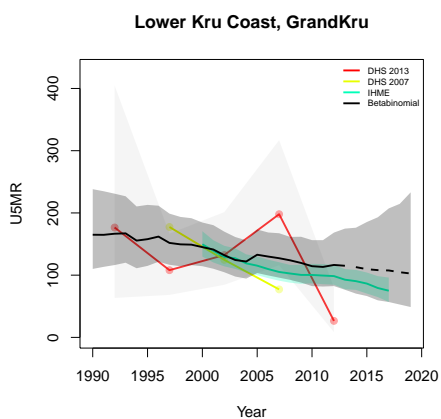
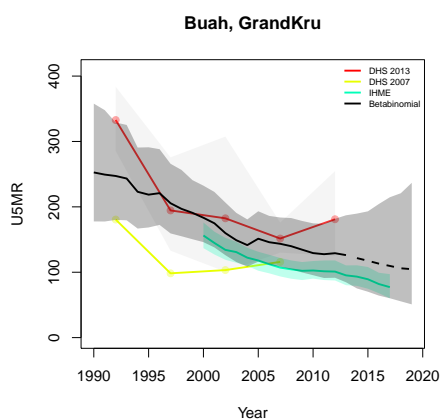
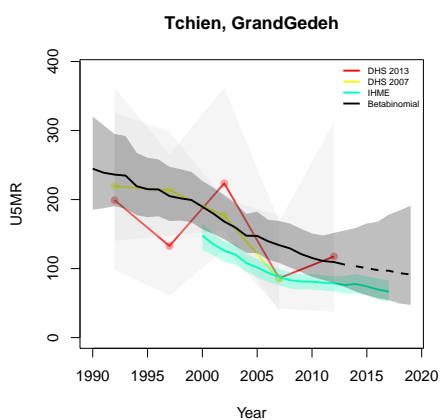


**Gbarzon, GrandGedeh**

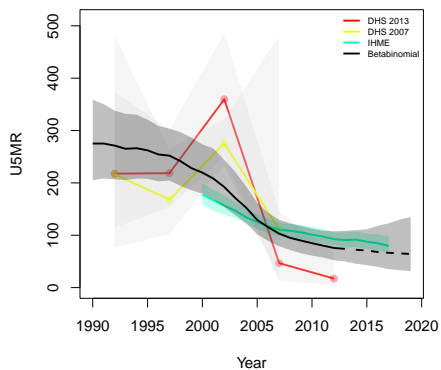


**Konobo, GrandGedeh**

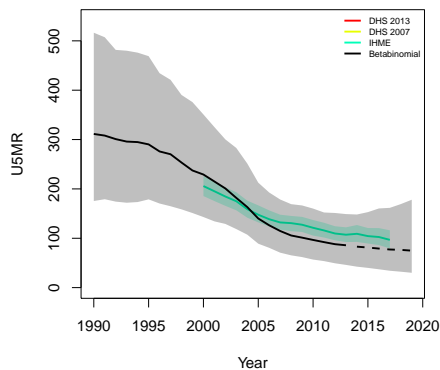




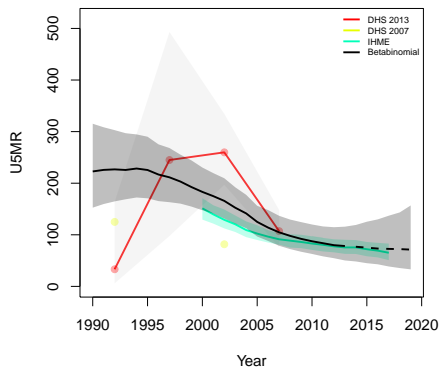
**Kolahun, Lofa**



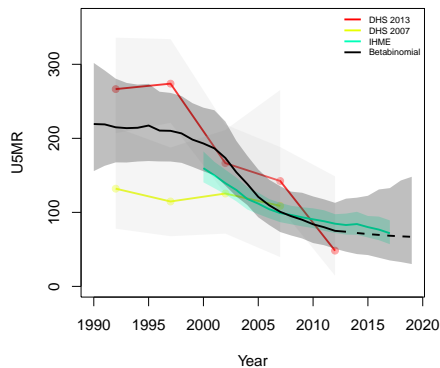
**Lower Kru Coast, Lofa**



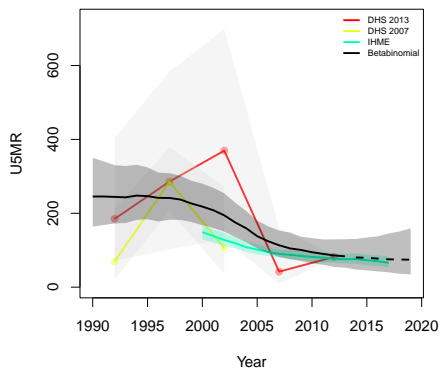
**Salayea, Lofa**



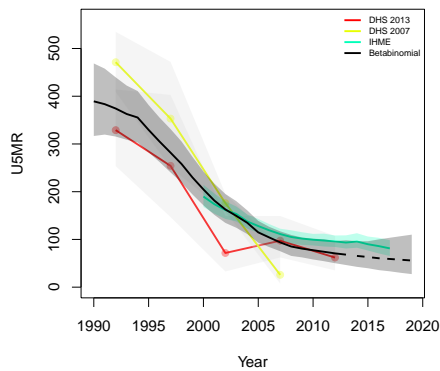
**Voinjama, Lofa**



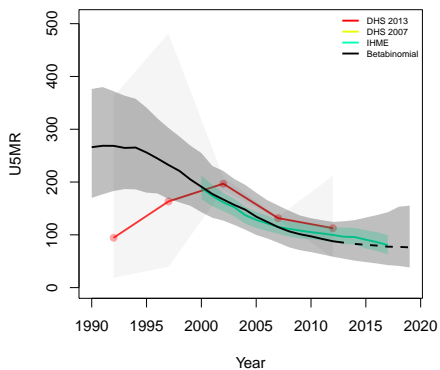
**Zorzor, Lofa**



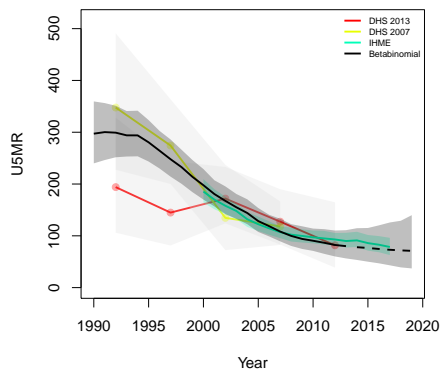
**Firestone, Margibi**



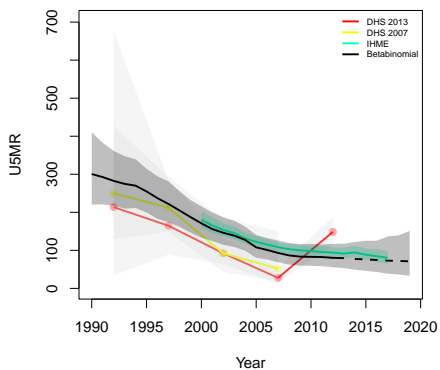
**Gibi, Margibi**



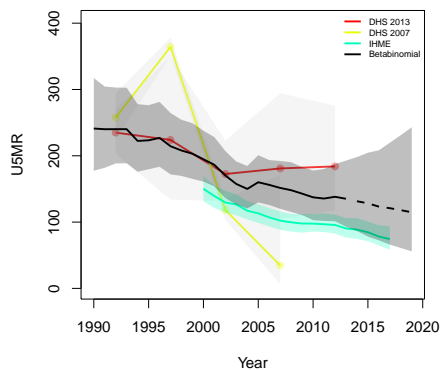
**Kakata, Margibi**



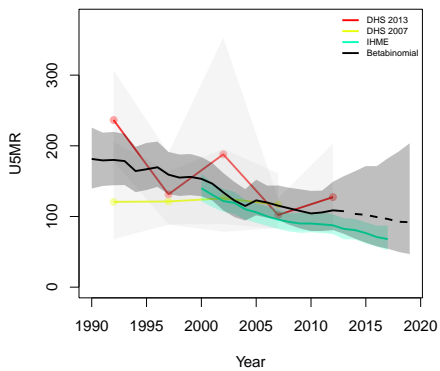
**Mambah-Kaba, Margibi**



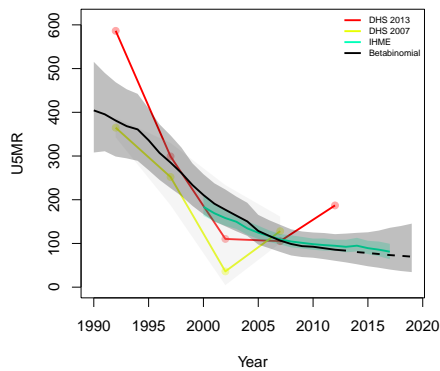
**Barrobo, Maryland**



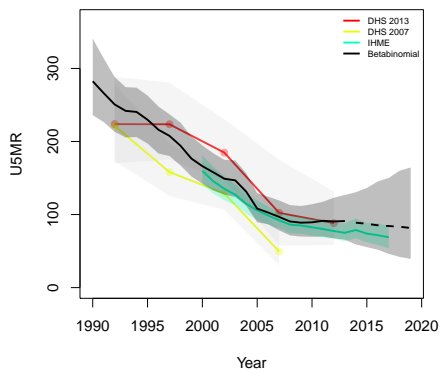
**Pleebo/Sodeken, Maryland**



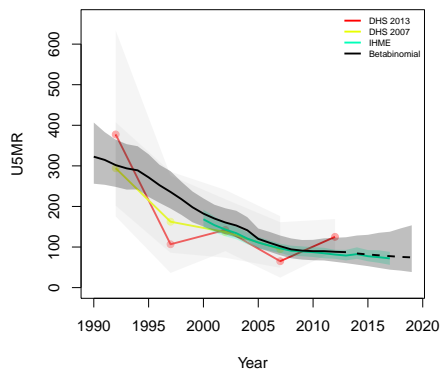
**Careysburg, Montserrat**



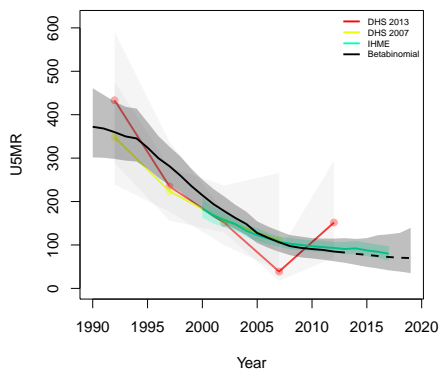
**Greater Monrovia, Montserrat**



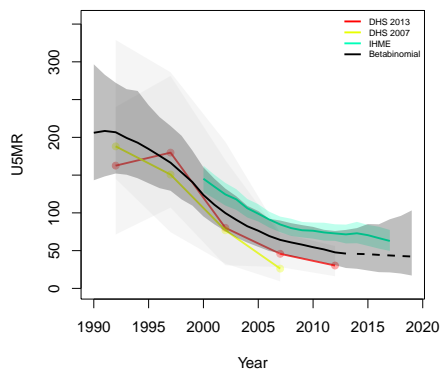
**St Paul River, Montserrat**



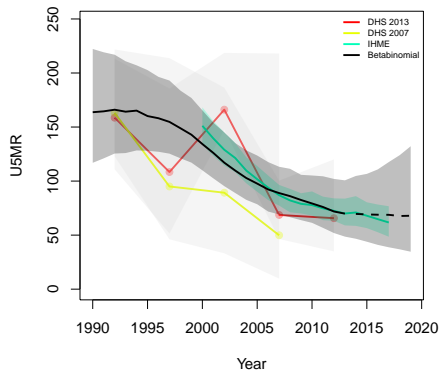
**Todee, Montserrat**



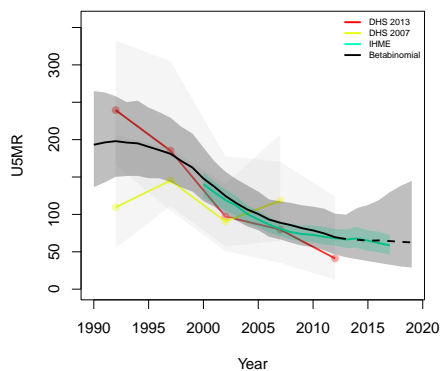
**Gbehlageh, Nimba**



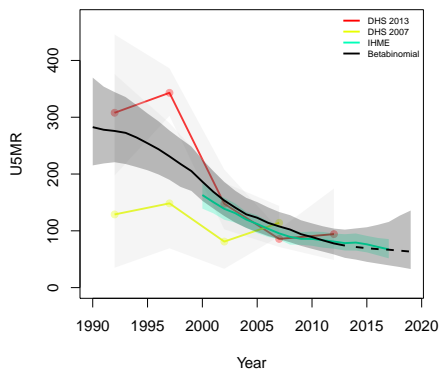
**Saclepea, Nimba**



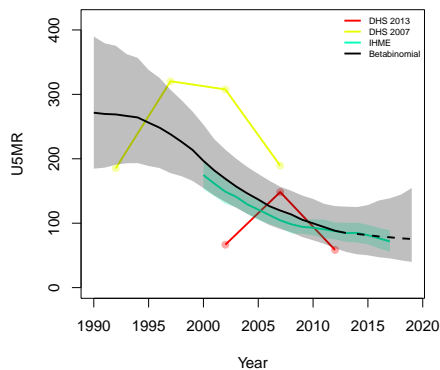
**Sanniqueleh-Mahn, Nimba**



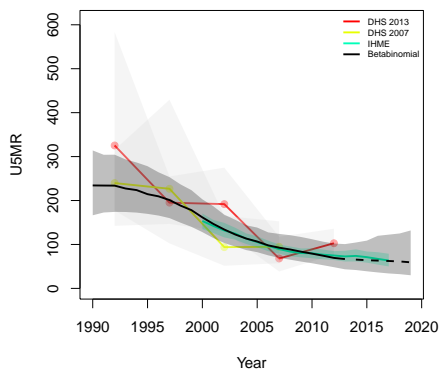
**Tappita, Nimba**



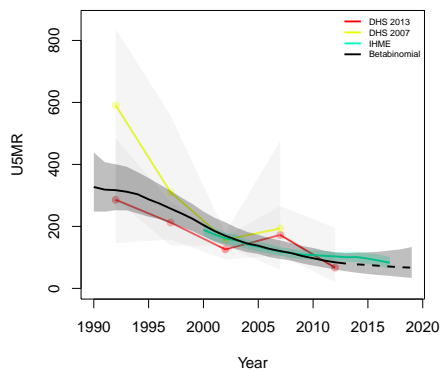
**Yarwein-Mehnsohnne, Nimba**



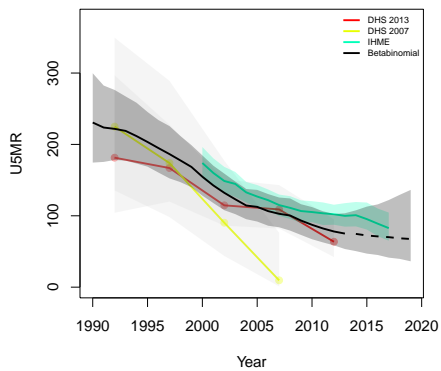
**Zoegeh, Nimba**



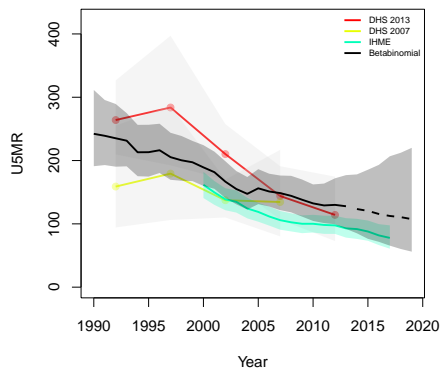
**Morweh, River Cess**



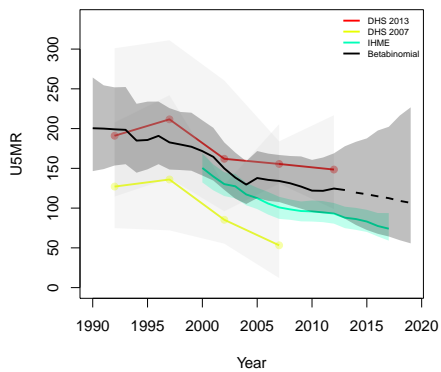
**Timbo, River Cess**



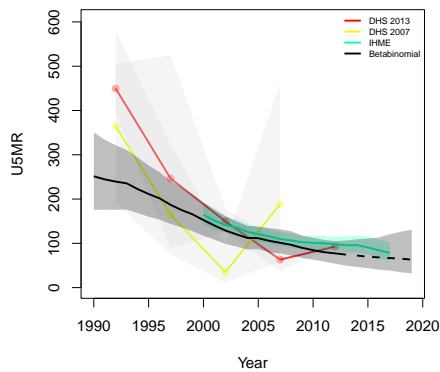
**Gbeapo, River Gee**



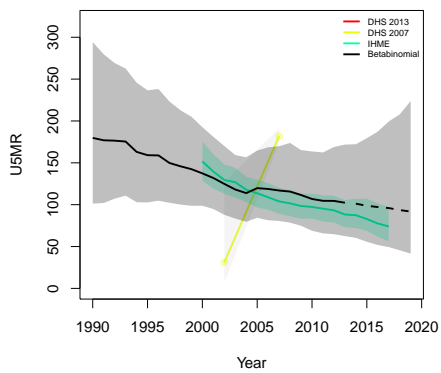
**Webbo, River Gee**



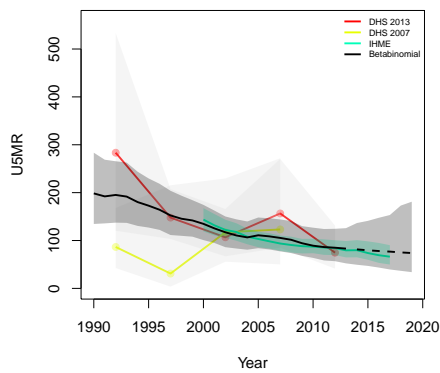
**Butaw, Sinoe**



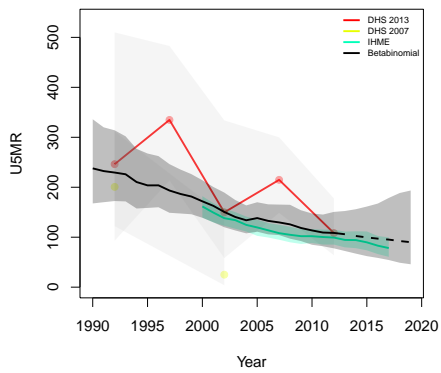
**Dugbe River, Sinoe**



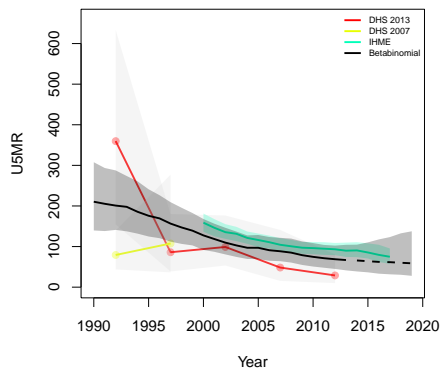
**Greenville, Sinoe**

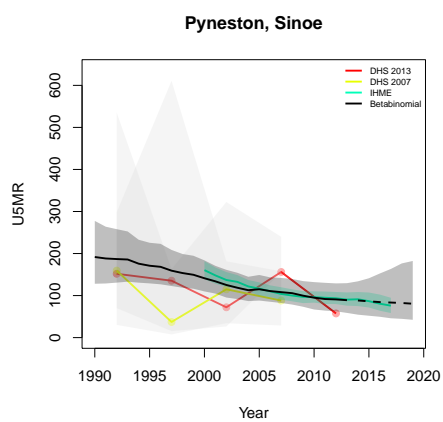
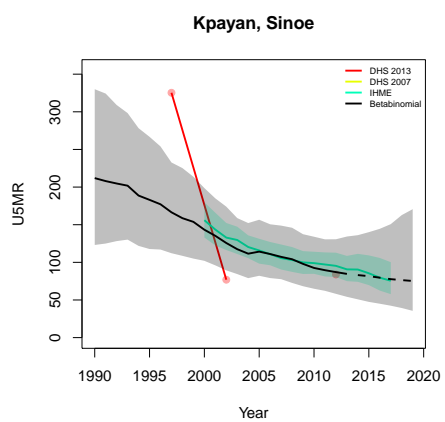


**Jaedae Jaedepo, Sinoe**



**Juarzon, Sinoe**





## B.8 Malawi

Age	Survey	Clusters			Deaths			Agemonths		
		Urban	Rural	Total	Urban	Rural	Total	Urban	Rural	Total
0	2000	110	449	559	142	940	1082	4047	19234	23281
	2004	65	455	520	70	995	1065	2793	23877	26670
	2010	151	676	827	181	2111	2292	5895	56218	62113
	2015	173	677	850	242	1674	1916	10412	55159	65571
1-11	2000	110	449	559	184	1287	1471	41107	189223	230330
	2004	65	455	520	94	1420	1514	28661	237446	266107
	2010	151	676	827	218	2459	2677	60651	569615	630266
	2015	173	677	850	211	1584	1795	108212	567100	675312
12-23	2000	110	449	559	98	693	791	40484	182586	223070
	2004	65	455	520	51	647	698	28512	228976	257488
	2010	151	676	827	98	1432	1530	61170	566159	627329
	2015	173	677	850	113	832	945	110534	574886	685420
24-35	2000	110	449	559	61	461	522	37666	168664	206330
	2004	65	455	520	43	433	476	26204	212076	238280
	2010	151	676	827	66	915	981	57355	525899	583254
	2015	173	677	850	65	642	707	104223	539039	643262
36-47	2000	110	449	559	38	244	282	35222	155735	190957
	2004	65	455	520	21	264	285	24295	197009	221304
	2010	151	676	827	41	628	669	54022	485992	540014
	2015	173	677	850	38	381	419	97785	502959	600744
48-59	2000	110	449	559	8	119	127	33225	145910	179135
	2004	65	455	520	14	117	131	22835	182900	205735
	2010	151	676	827	37	287	324	50494	451106	501600
	2015	173	677	850	18	203	221	91407	468558	559965

Table B.8: **Data summary for Malawi.** Total numbers of clusters (Columns 3–5) with observations in each age group by survey in urban and rural areas and combined. Numbers of deaths (Columns 6–8) and number of agemonths (Columns 9–10) observed in each age group by survey in urban and rural areas and combined.

*B.8.1 Admin-2*

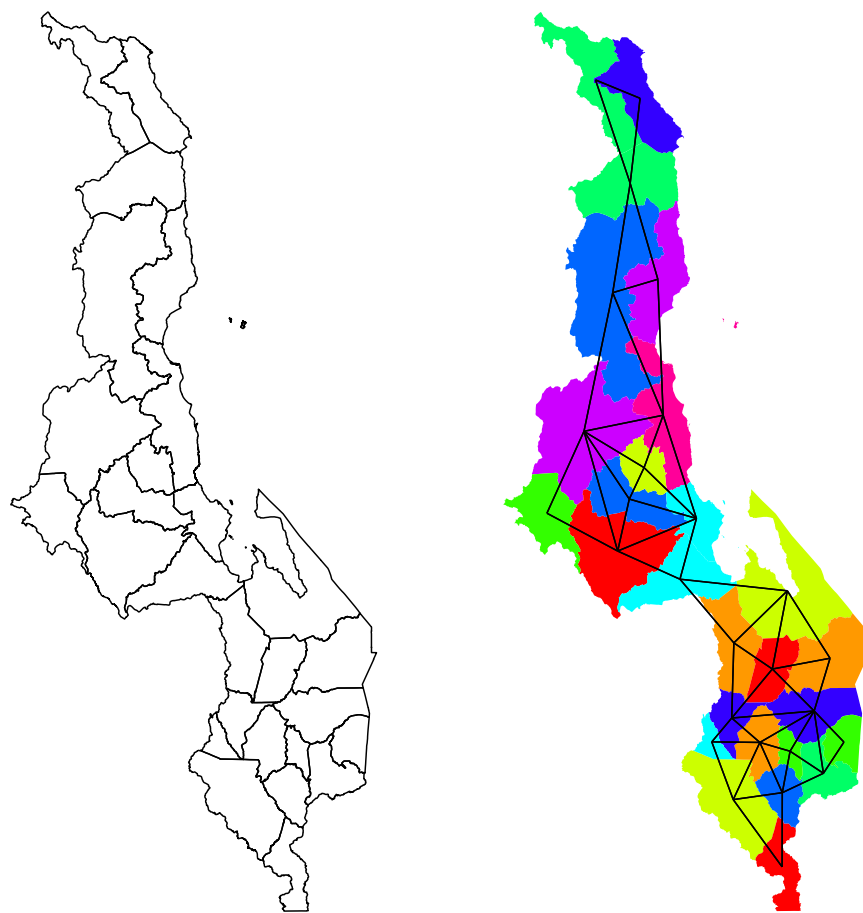


Figure B.37: **Left:** The names of the 28 Admin-2 areas of Malawi . **Right:** The neighborhood structure of Admin-2 areas in Malawi .

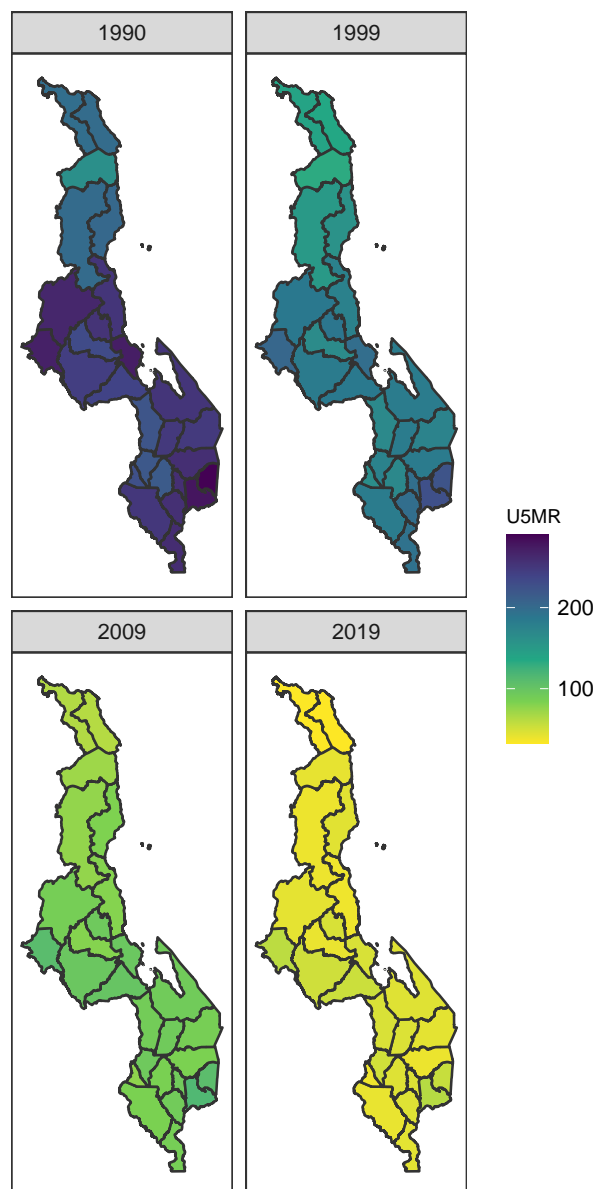


Figure B.38: Median U5MR estimates for years 1990, 1999, 2009, 2019 for Admin-2 areas in Malawi .

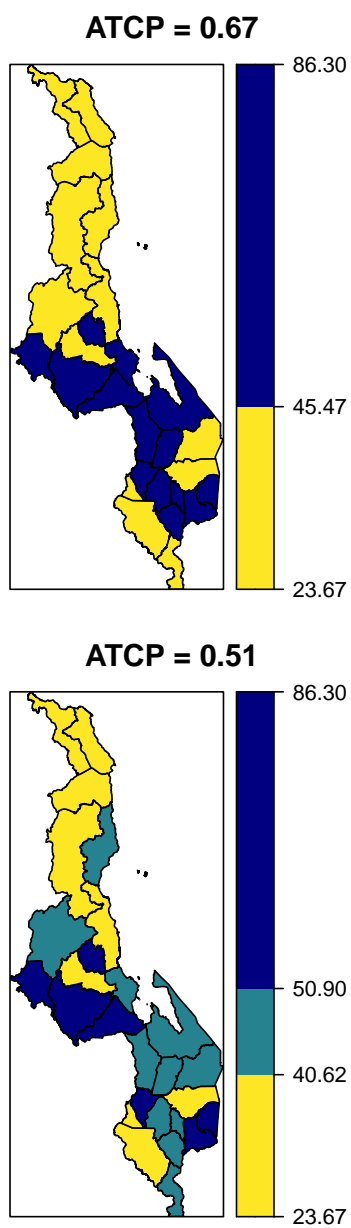
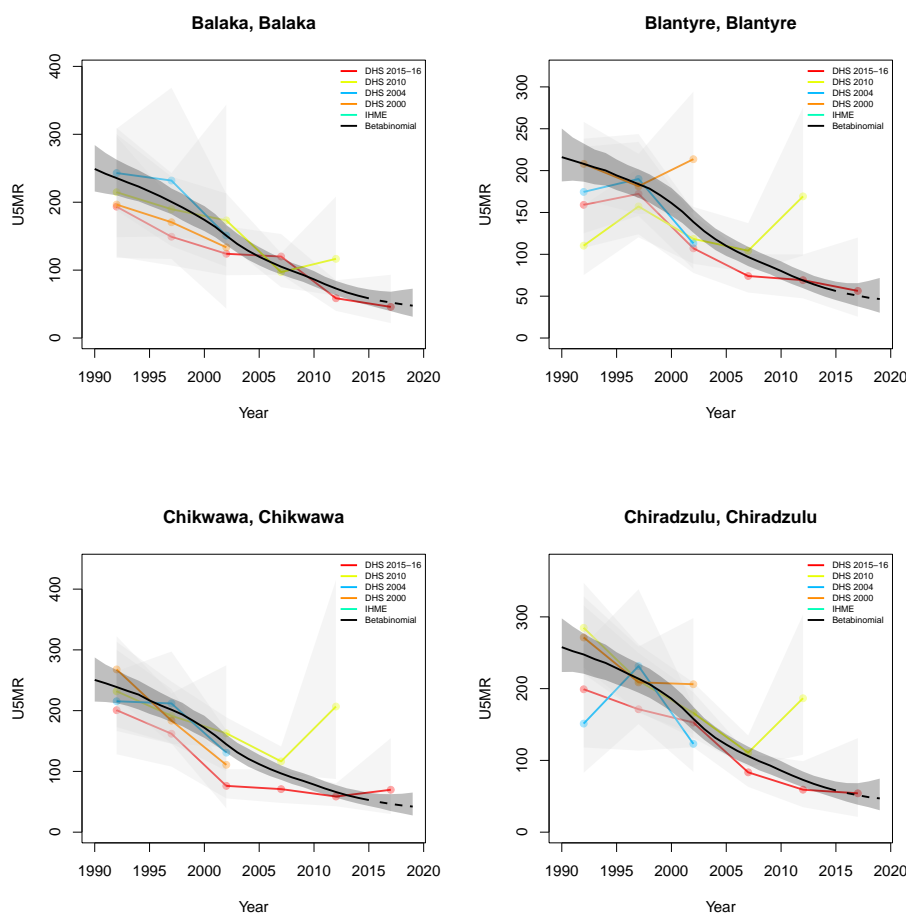


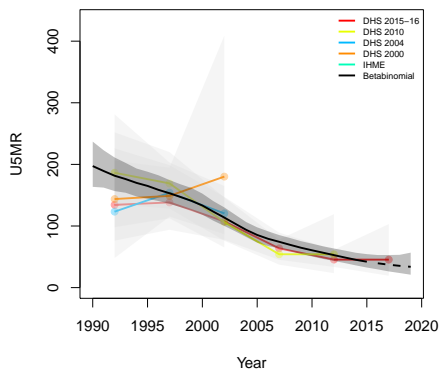
Figure B.39: Expression of uncertainty of U5MR (deaths per 1000 children) estimates for Admin-1 areas based on the average true classification probability (ATCP) in 2019 using  $K = 2, 3$  colors.

*Data and estimates over time by area*

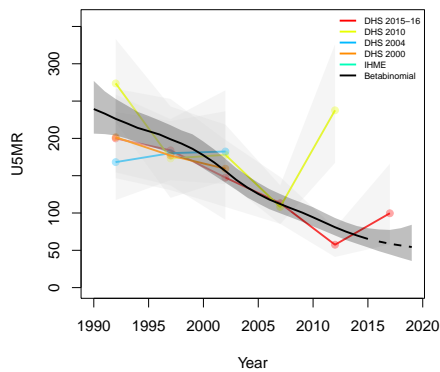
Colored lines with circular points and light grey uncertainty bands are 5-year survey-weighted estimates of U5MR for years 1990–1994 up to 2015–2019 depending on survey timing. For a survey that ends in the middle of a 5-year period, we plot the estimates at the mid-point of the years in that interval for which the survey provides data. Black lines and corresponding intervals represent posterior medians and 95% uncertainty intervals respectively for the betabinomial model. IHME’s estimates and corresponding intervals, where we can compare, are in aquamarine.



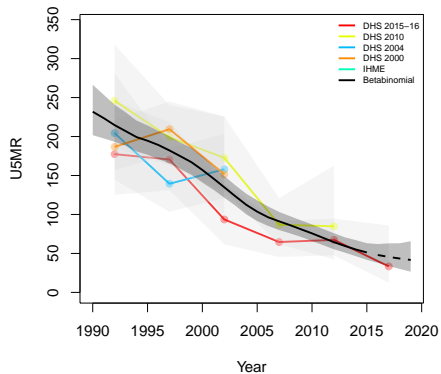
**Chitipa, Chitipa**



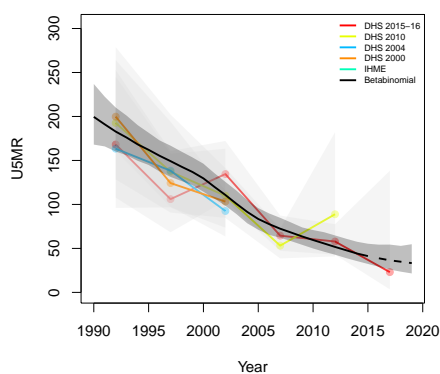
**Dedza, Dedza**



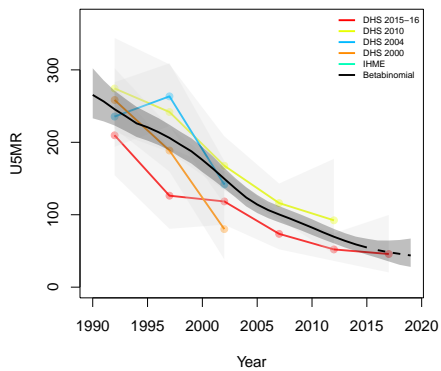
**Dowa, Dowa**



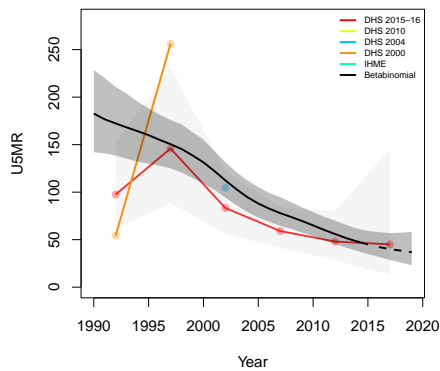
**Karonga, Karonga**



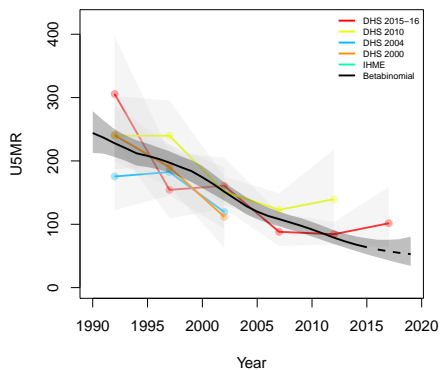
**Kasungu, Kasungu**



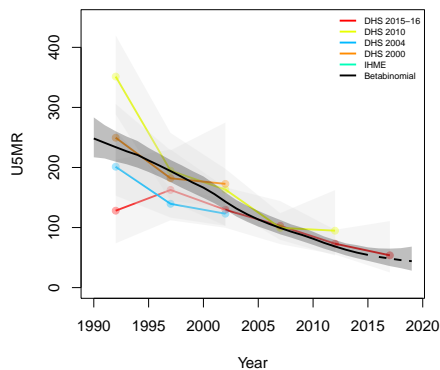
**Likoma, Likoma**



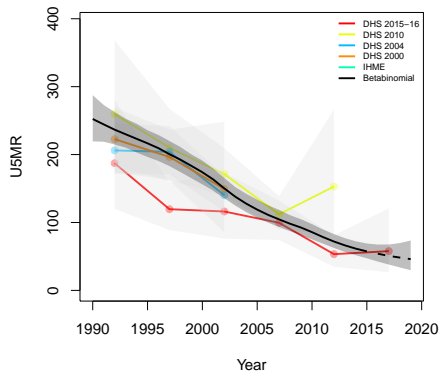
**Lilongwe, Lilongwe**



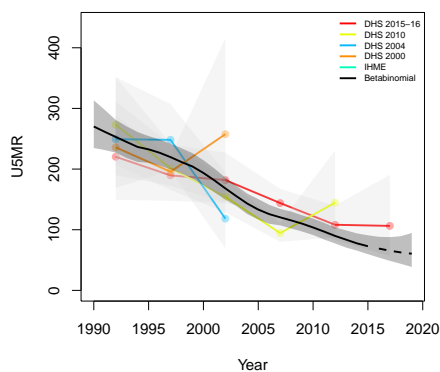
**Machinga, Machinga**



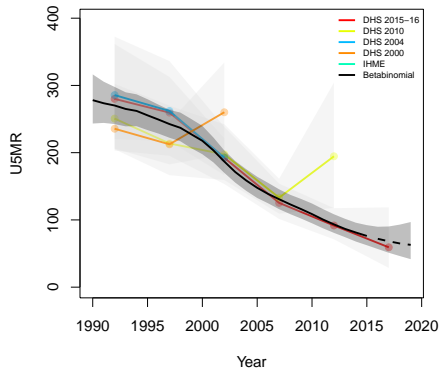
**Mangochi, Mangochi**



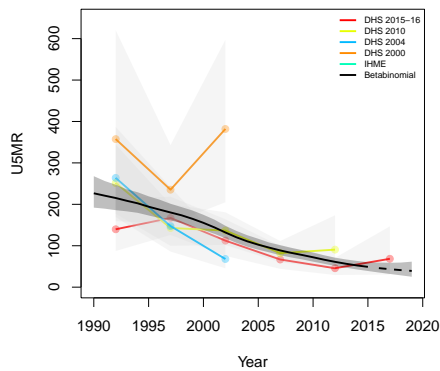
**Mchinji, Mchinji**



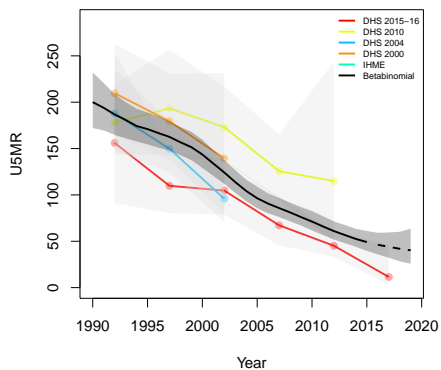
**Mulanje, Mulanje**



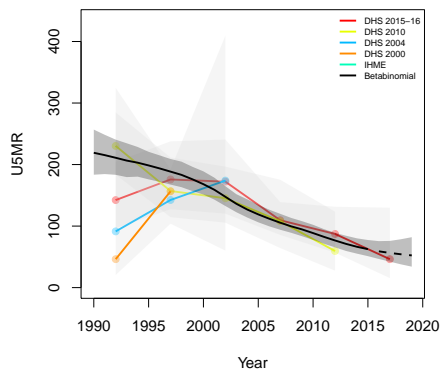
**Mwanza, Mwanza**



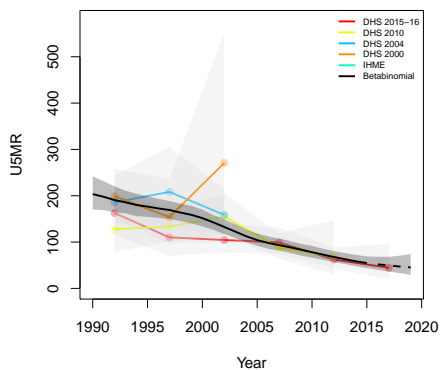
**Mzimba, Mzimba**



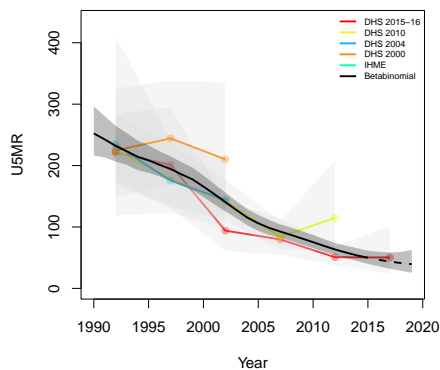
**Neno, Neno**



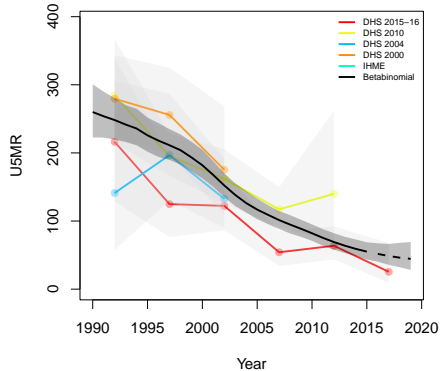
**Nkhata Bay, Nkhata Bay**



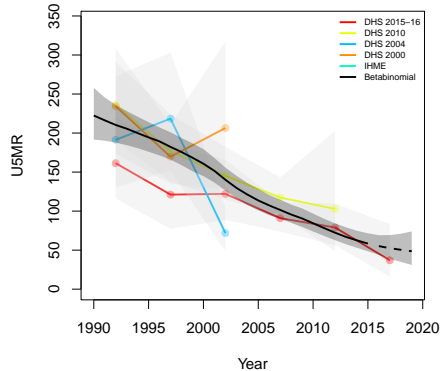
**Nkhotakota, Nkhotakota**

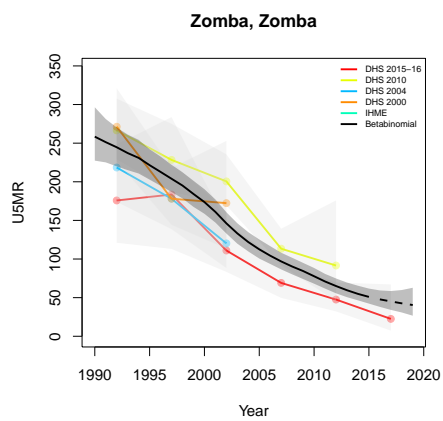
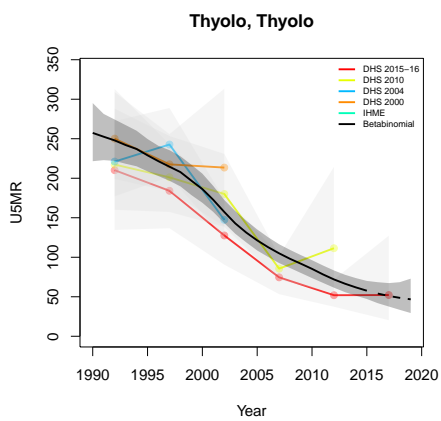
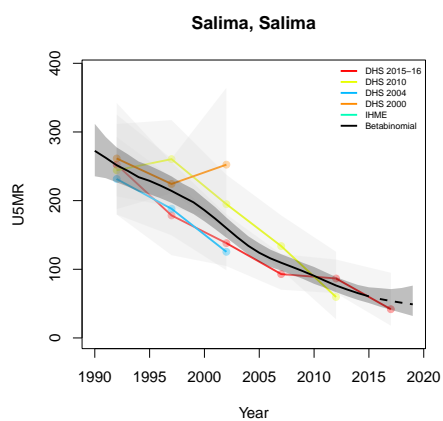
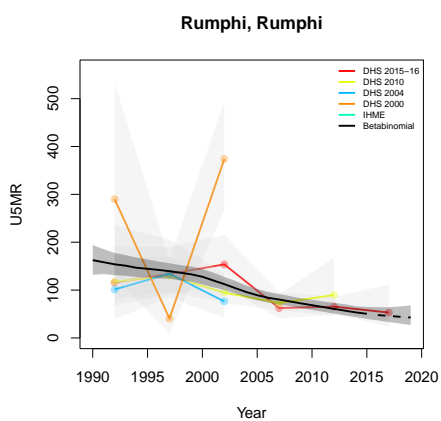
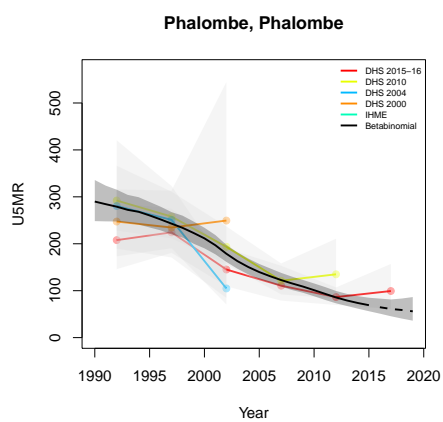
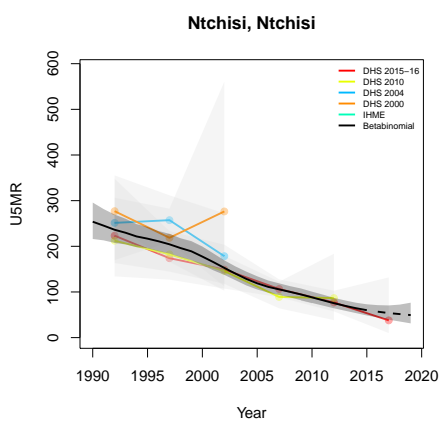


**Nsanje, Nsanje**



**Ntcheu, Ntcheu**





## B.9 Mali

Age	Survey	Clusters			Deaths			Agemonths		
		Urban	Rural	Total	Urban	Rural	Total	Urban	Rural	Total
0	1996	118	182	300	168	605	773	3792	8926	12718
	2001	129	270	399	304	1683	1987	5639	23265	28904
	2006	145	260	405	563	1810	2373	11351	30024	41375
	2012	116	297	413	255	1063	1318	7784	24352	32136
	2018	94	234	328	196	1023	1219	7591	24286	31877
1-11	1996	118	182	300	182	613	795	38328	87988	126316
	2001	129	270	399	248	1348	1596	56003	224372	280375
	2006	145	260	405	463	1720	2183	113398	294201	407599
	2012	116	297	413	157	762	919	79247	243672	322919
	2018	94	234	328	136	825	961	78155	244226	322381
12-23	1996	118	182	300	142	358	500	39058	88509	127567
	2001	129	270	399	136	965	1101	56090	219969	276059
	2006	145	260	405	255	1126	1381	112502	287137	399639
	2012	116	297	413	47	277	324	79866	245142	325008
	2018	94	234	328	56	466	522	78627	241933	320560
24-35	1996	118	182	300	125	349	474	37437	86355	123792
	2001	129	270	399	97	850	947	53493	209423	262916
	2006	145	260	405	224	1082	1306	105048	266177	371225
	2012	116	297	413	52	466	518	74576	227703	302279
	2018	94	234	328	77	522	599	72583	221578	294161
36-47	1996	118	182	300	63	224	287	35429	81383	116812
	2001	129	270	399	72	483	555	50605	195288	245893
	2006	145	260	405	158	528	686	96841	242810	339651
	2012	116	297	413	27	230	257	69134	208360	277494
	2018	94	234	328	36	265	301	66861	201915	268776
48-59	1996	118	182	300	33	85	118	33779	77638	111417
	2001	129	270	399	42	217	259	47931	182304	230235
	2006	145	260	405	75	259	334	90138	225555	315693
	2012	116	297	413	16	93	109	63774	190156	253930

*B.9.1 Admin-1*

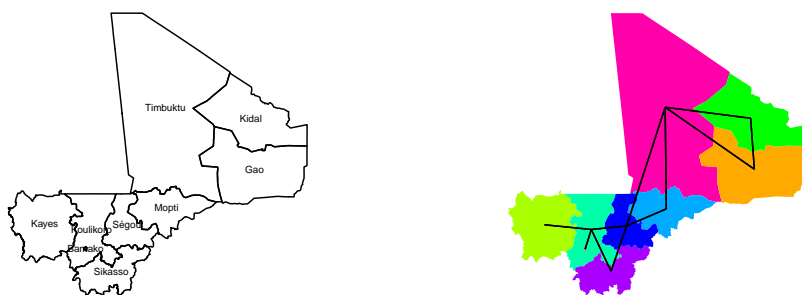


Figure B.40: **Left:** The names of the 9 Admin-1 areas of Mali . **Right:** The neighborhood structure of Admin-1 areas in Mali .

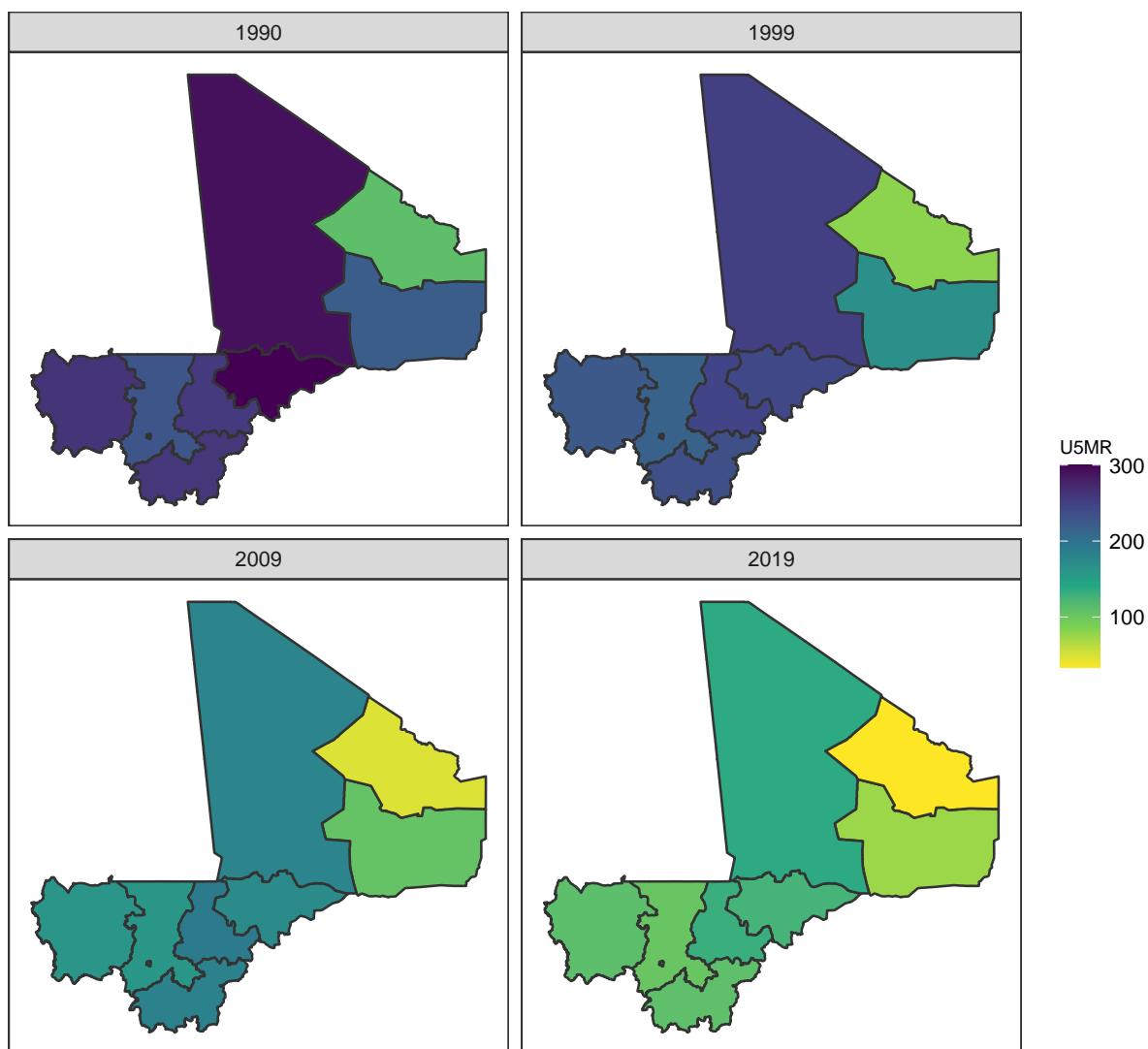


Figure B.41: Median U5MR estimates for years 1990, 1999, 2009, 2019 for Admin-1 areas in Mali .

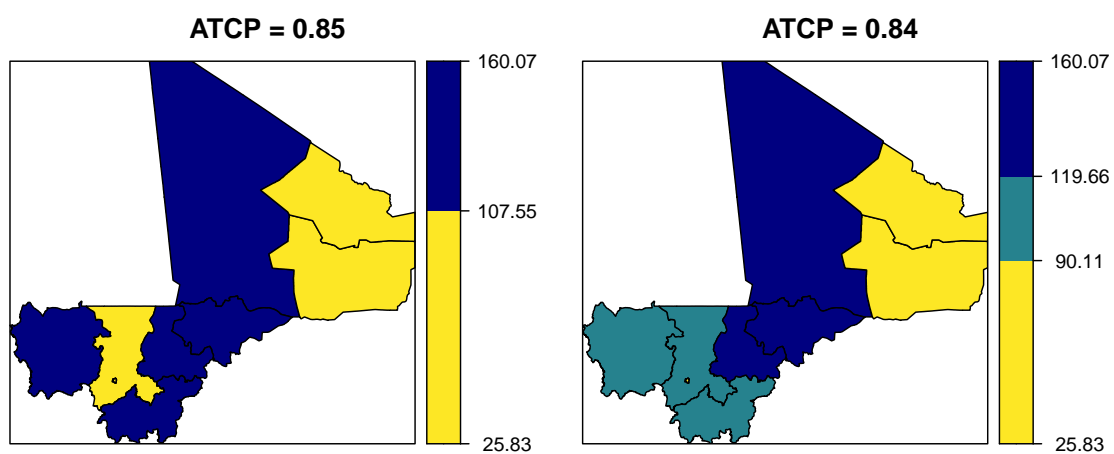
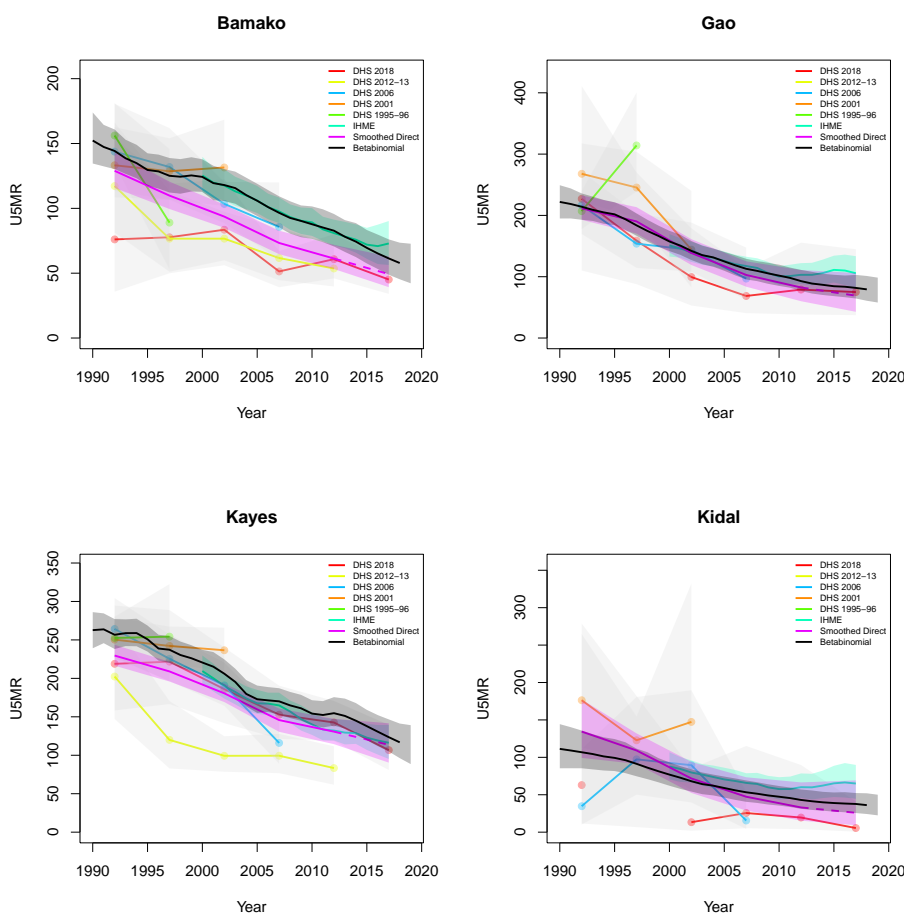
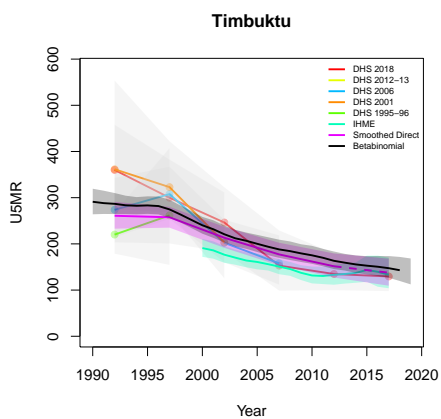
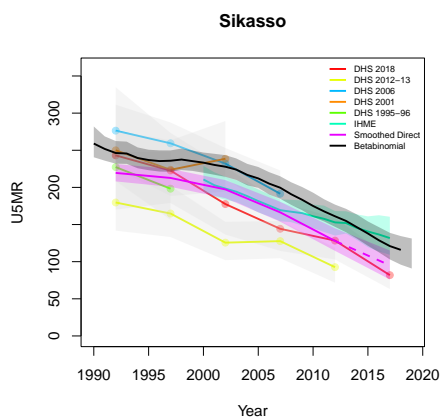
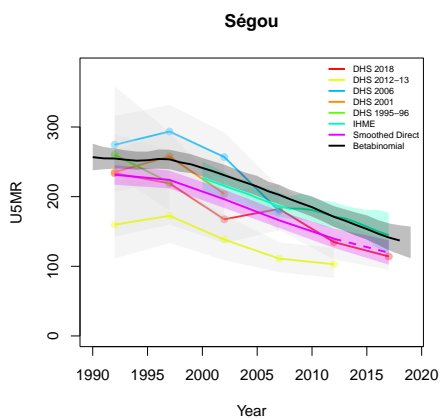
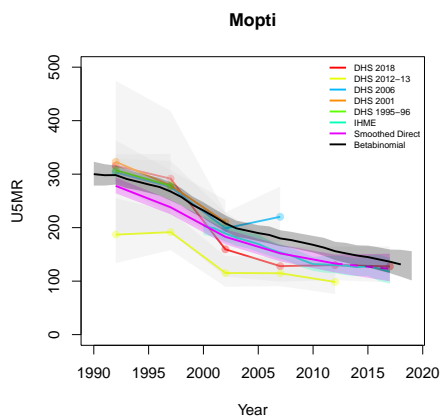
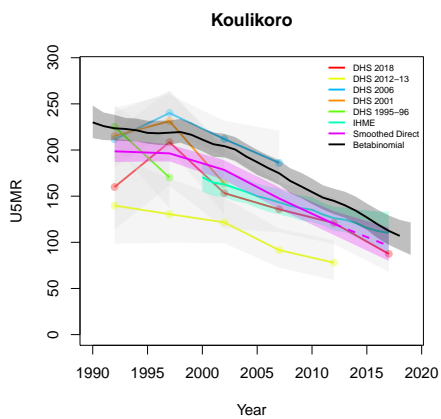


Figure B.42: Expression of uncertainty of U5MR (deaths per 1000 children) estimates for Admin-1 areas based on the average true classification probability (ATCP) in 2019 using  $K = 2, 3$  colors.

*Data and estimates over time by area*

Colored lines with circular points and light grey uncertainty bands are 5-year survey-weighted estimates of U5MR for years 1990–1994 up to 2015–2019 depending on survey timing. For a survey that ends in the middle of a 5-year period, we plot the estimates at the mid-point of the years in that interval for which the survey provides data. Black lines and corresponding intervals represent posterior medians and 95% uncertainty intervals respectively for the betabinomial model. IHME’s estimates and corresponding intervals, where we can compare, are in aquamarine.





*B.9.2 Admin-2*

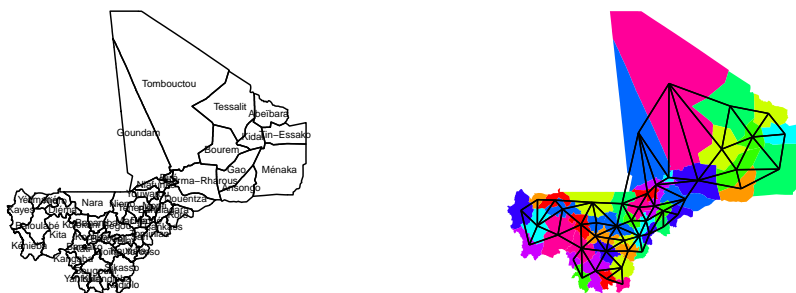


Figure B.43: **Left:** The names of the 50 Admin-2 areas of Mali . **Right:** The neighborhood structure of Admin-2 areas in Mali .

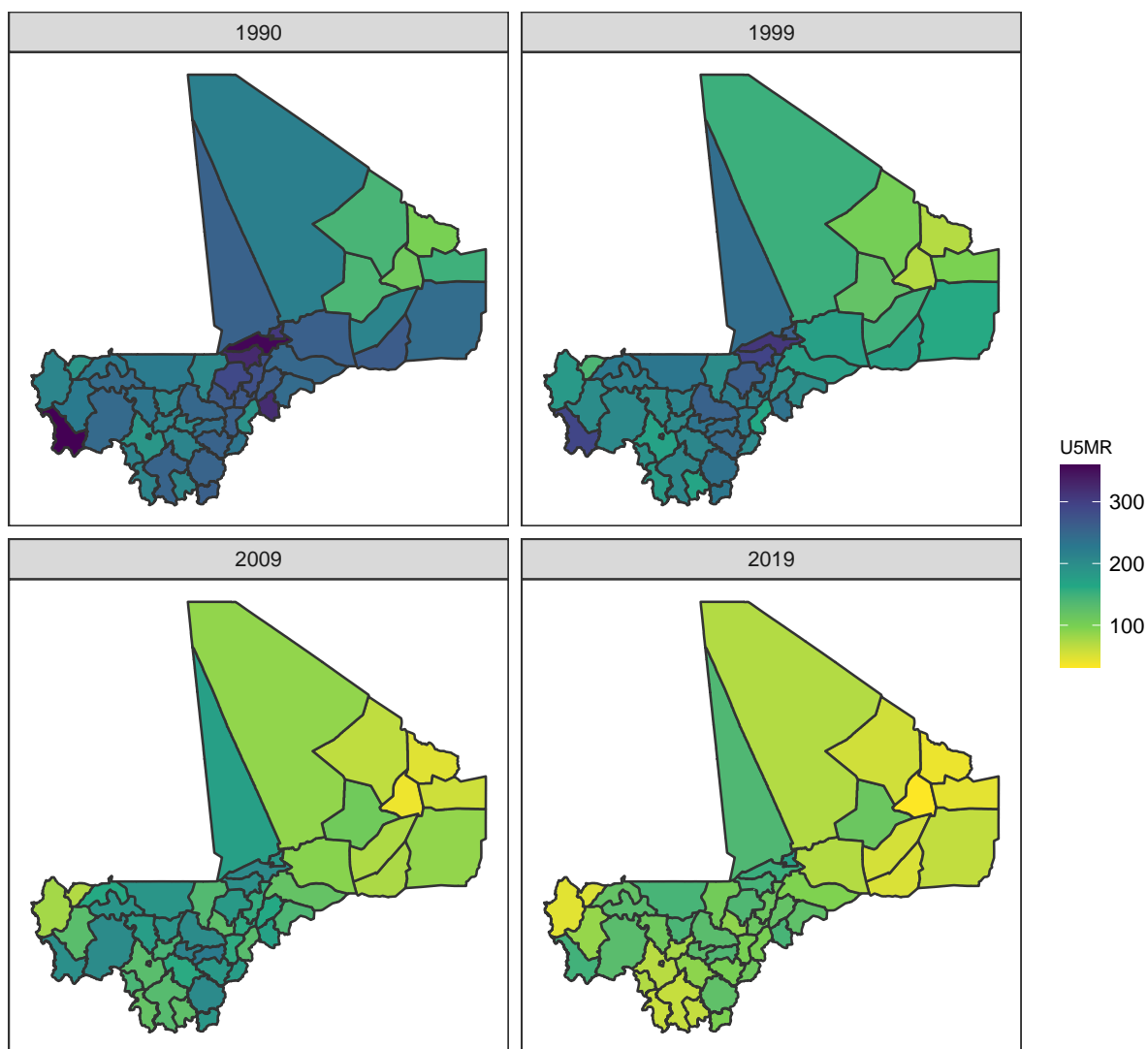


Figure B.44: Median U5MR estimates for years 1990, 1999, 2009, 2019 for Admin-2 areas in Mali .

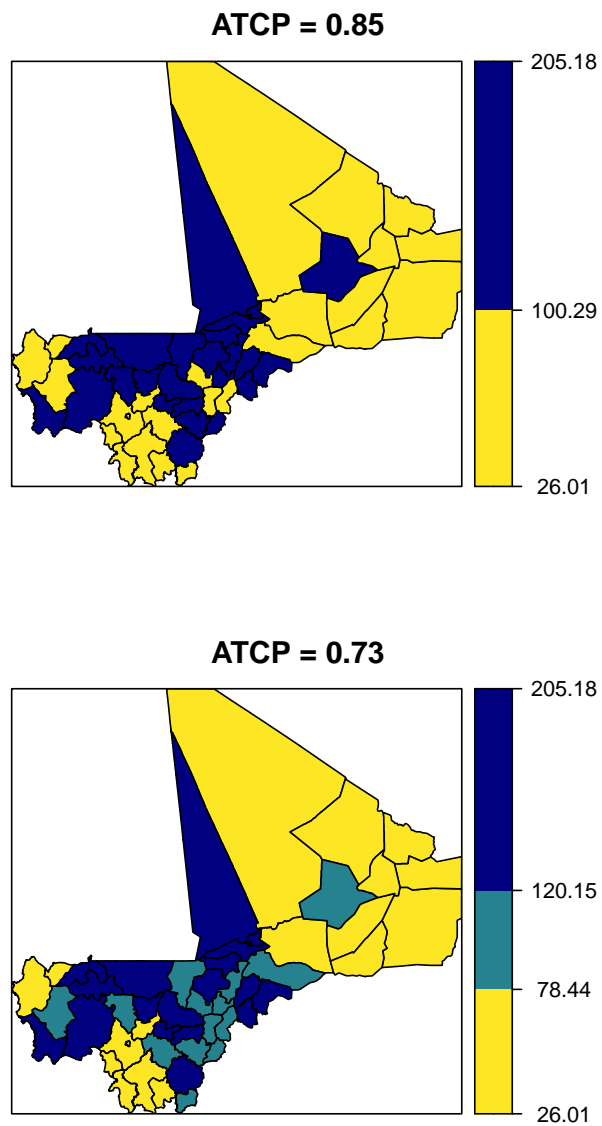
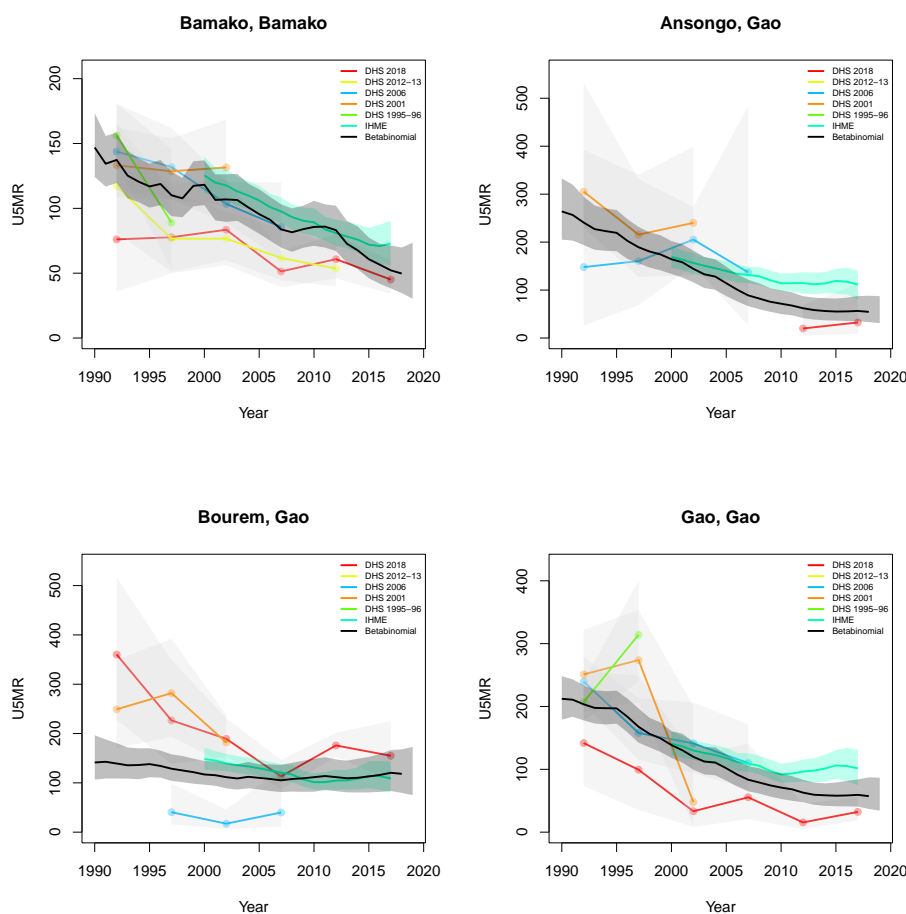


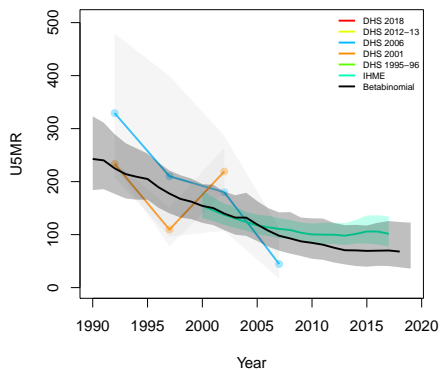
Figure B.45: Expression of uncertainty of U5MR (deaths per 1000 children) estimates for Admin-1 areas based on the average true classification probability (ATCP) in 2019 using  $K = 2, 3$  colors.

*Data and estimates over time by area*

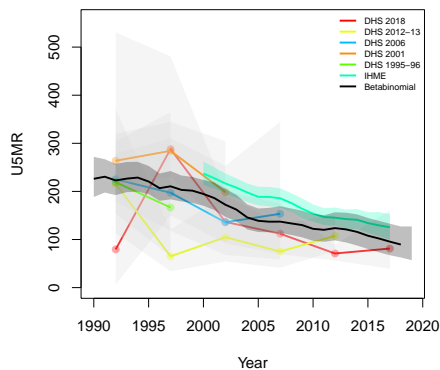
Colored lines with circular points and light grey uncertainty bands are 5-year survey-weighted estimates of U5MR for years 1990–1994 up to 2015–2019 depending on survey timing. For a survey that ends in the middle of a 5-year period, we plot the estimates at the mid-point of the years in that interval for which the survey provides data. Black lines and corresponding intervals represent posterior medians and 95% uncertainty intervals respectively for the betabinomial model. IHME’s estimates and corresponding intervals, where we can compare, are in aquamarine.



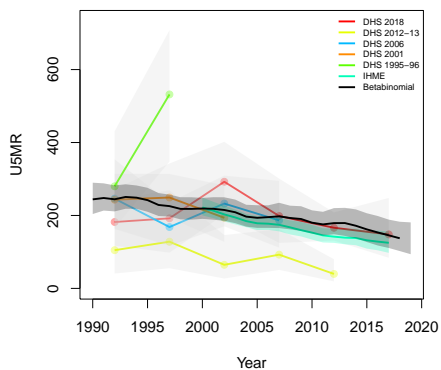
**Ménaka, Gao**



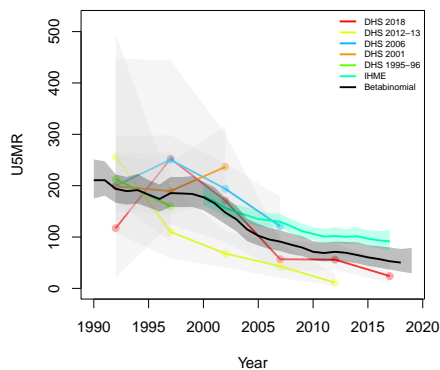
**Bafoulabé, Kayes**



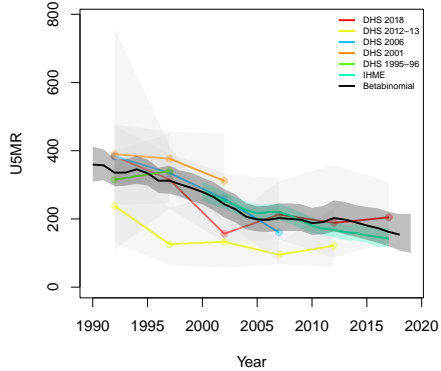
**Diéma, Kayes**



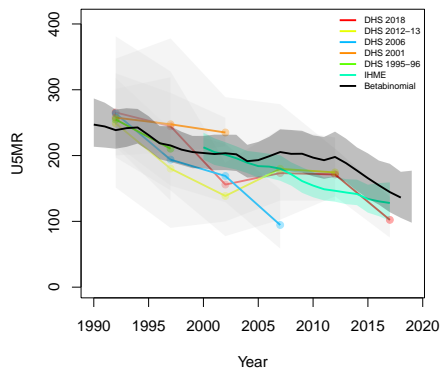
**Kayes, Kayes**



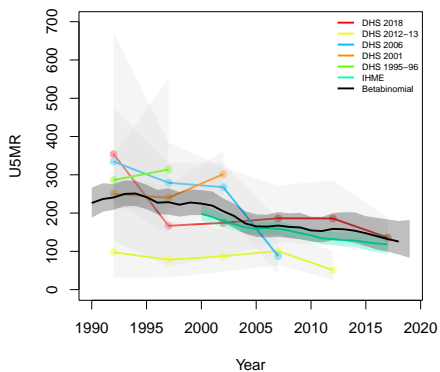
**Kéniéba, Kayes**



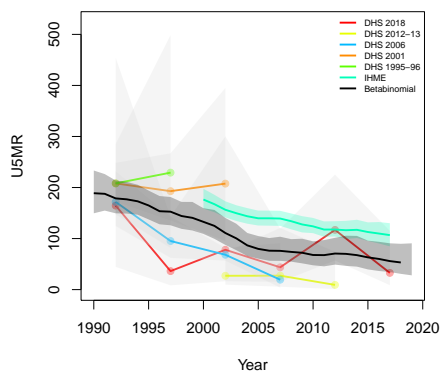
**Kita, Kayes**



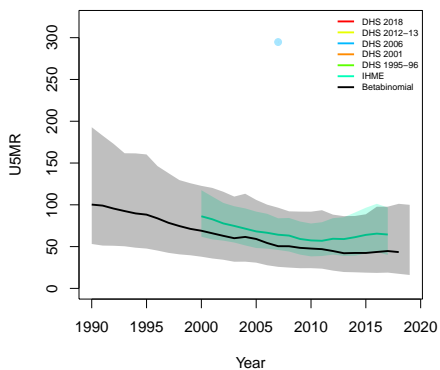
**Nioro, Kayes**



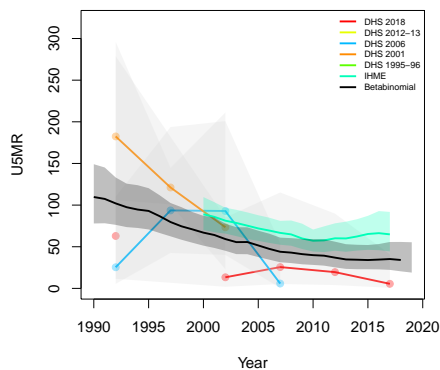
**Yélimané, Kayes**



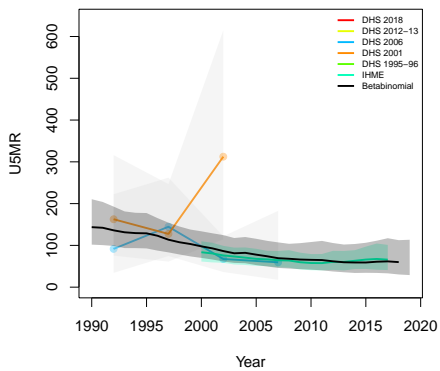
**Abeïbara, Kidal**



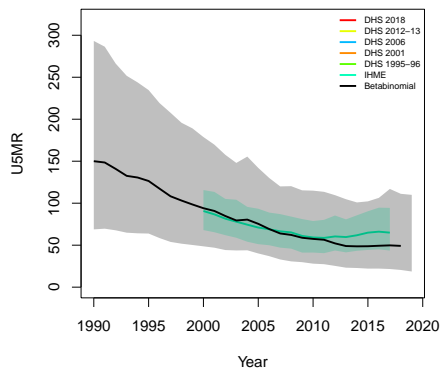
**Kidal, Kidal**



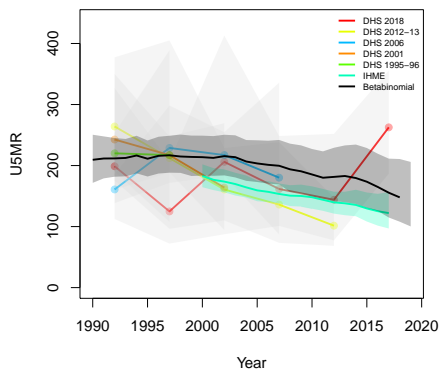
**Tessalit, Kidal**



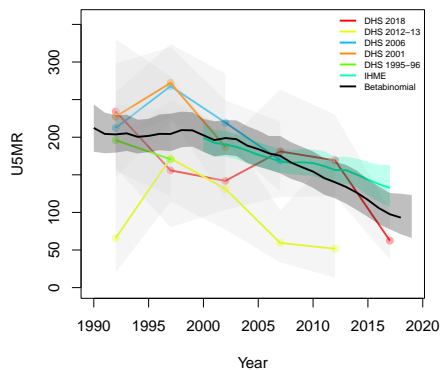
**Tin-Essako, Kidal**



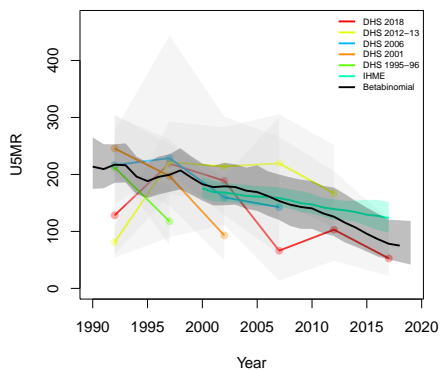
**Banamba, Koulikoro**



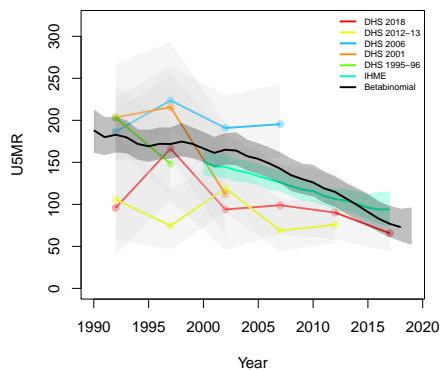
**Dioïla, Koulikoro**



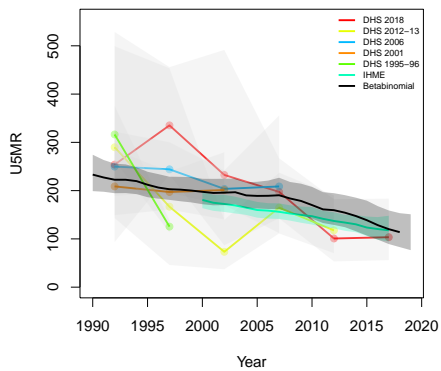
**Kangaba, Koulikoro**



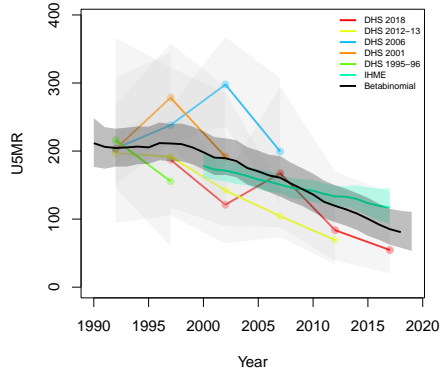
**Kati, Koulikoro**



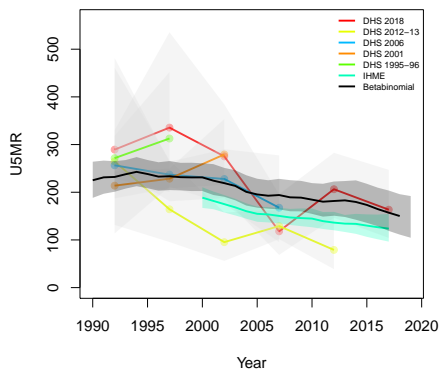
**Kolokani, Koulikoro**



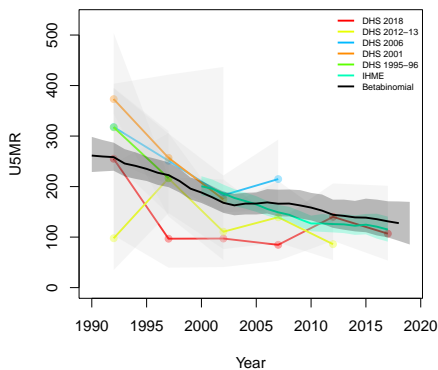
**Koulikoro, Koulikoro**



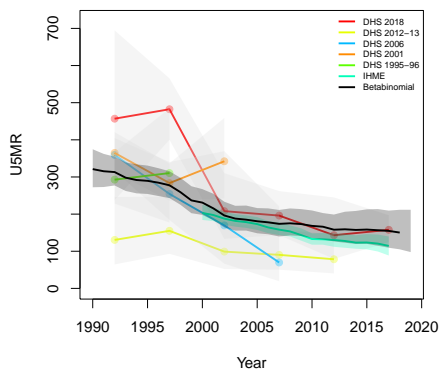
**Nara, Koulikoro**



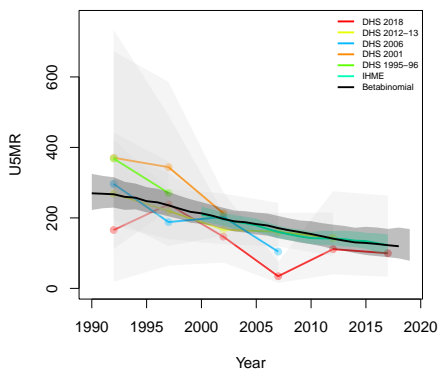
**Bandiagara, Mopti**



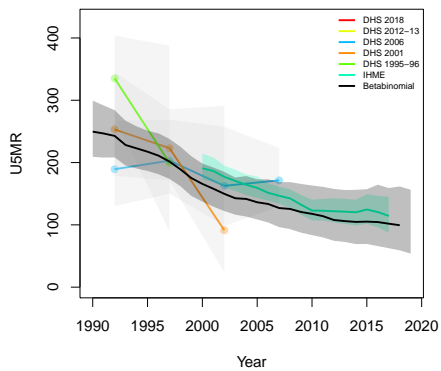
**Bankass, Mopti**



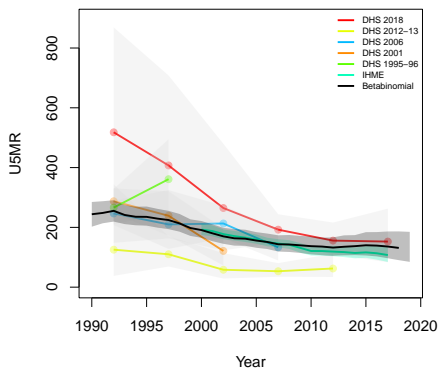
**Djenné, Mopti**



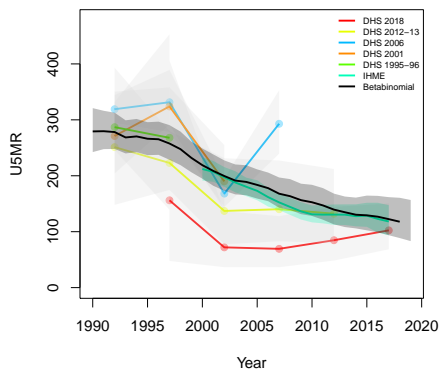
**Douentza, Mopti**



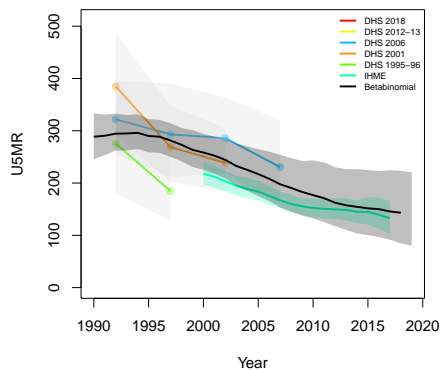
**Koro, Mopti**



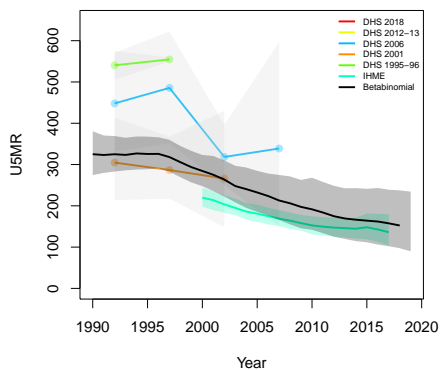
**Mopti, Mopti**



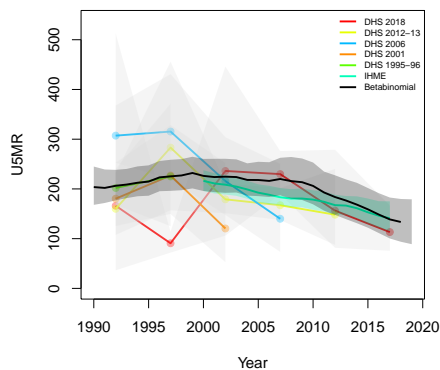
**Ténenkou, Mopti**



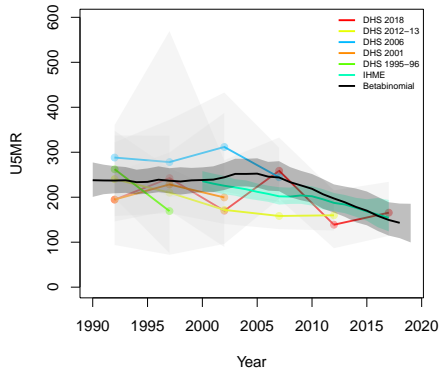
**Youwarou, Mopti**



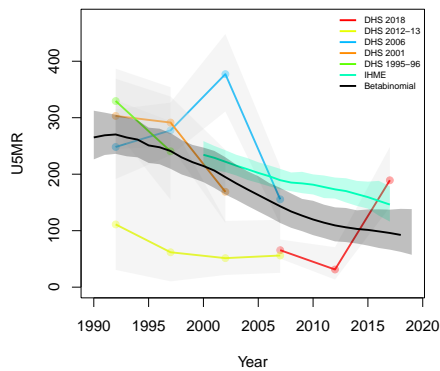
**Barouéli, Ségou**



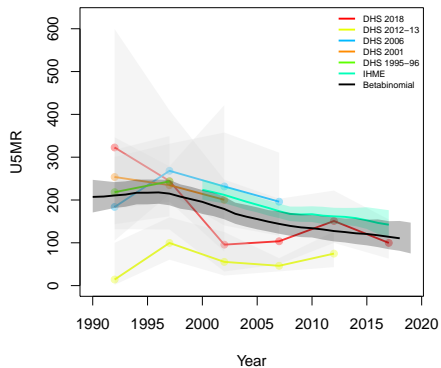
**Bla, Ségou**



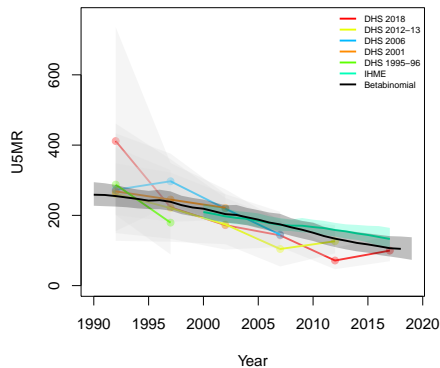
**Macina, Ségou**



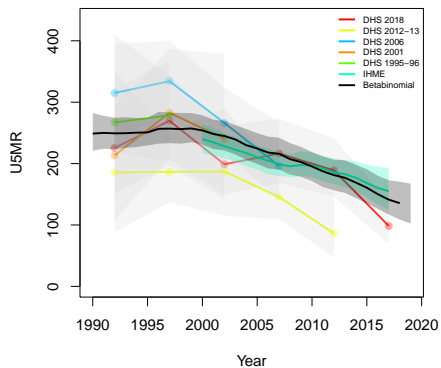
**Niono, Ségou**



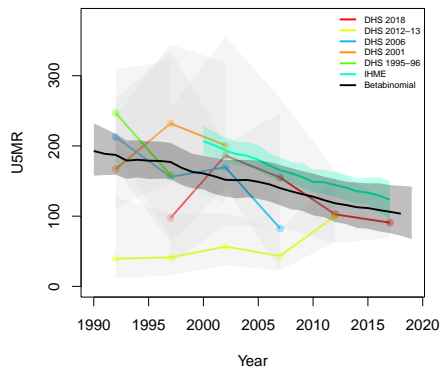
**San, Ségou**



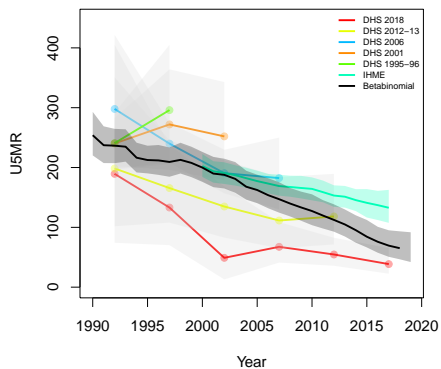
**Ségou, Ségou**



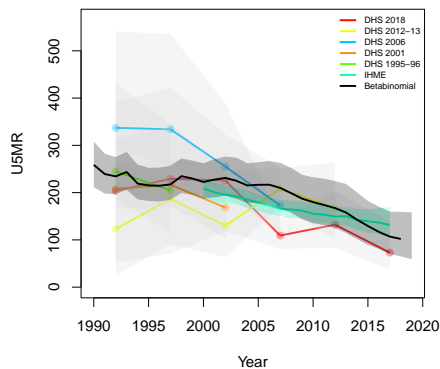
**Tominian, Ségou**



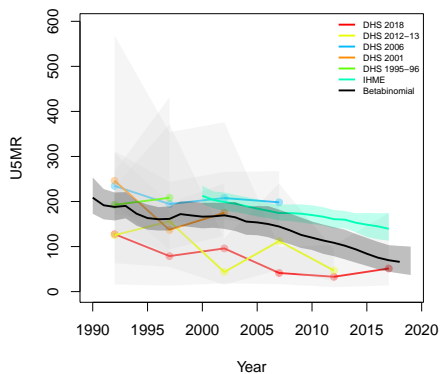
**Bougouni, Sikasso**



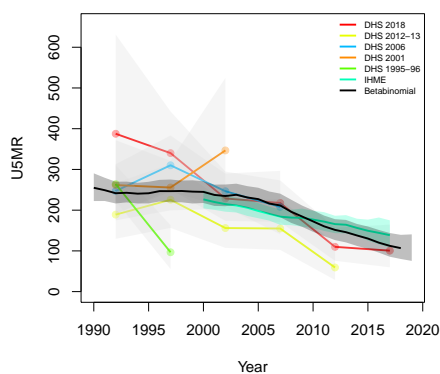
**Kadiolo, Sikasso**



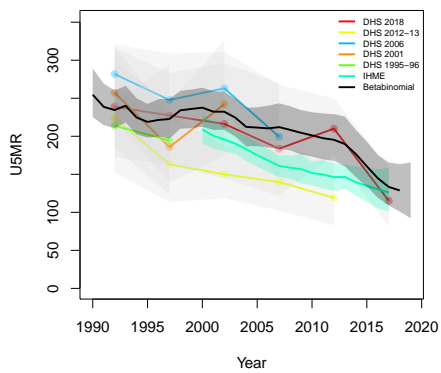
**Kolondiéba, Sikasso**



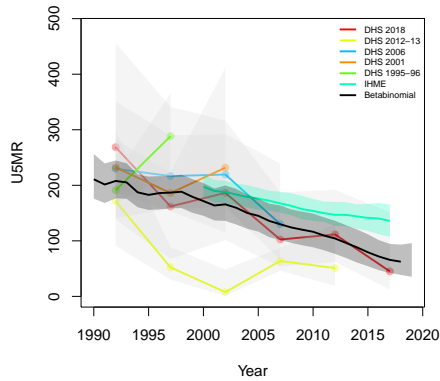
**Koutiala, Sikasso**



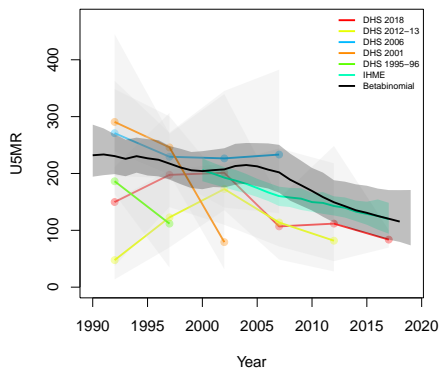
**Sikasso, Sikasso**



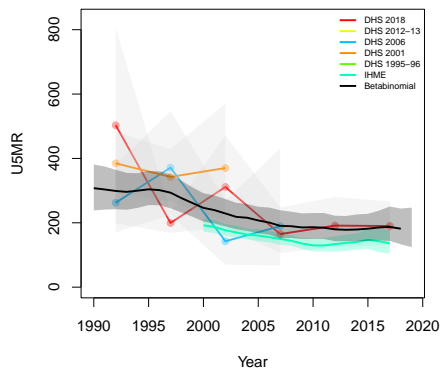
**Yanfolila, Sikasso**

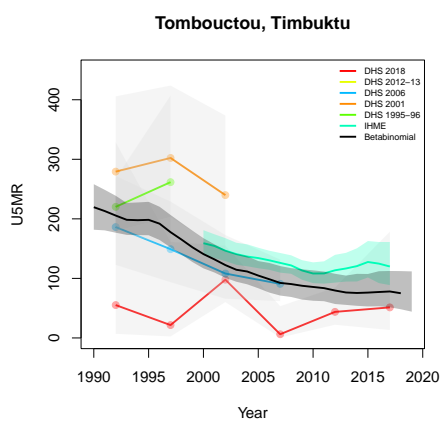
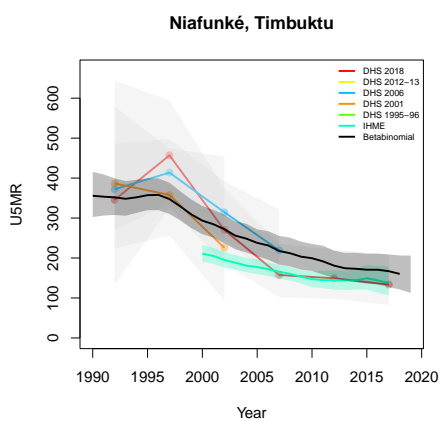
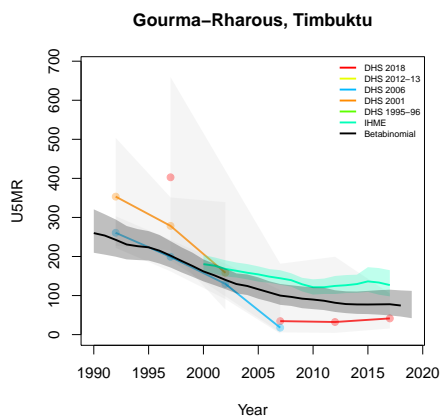
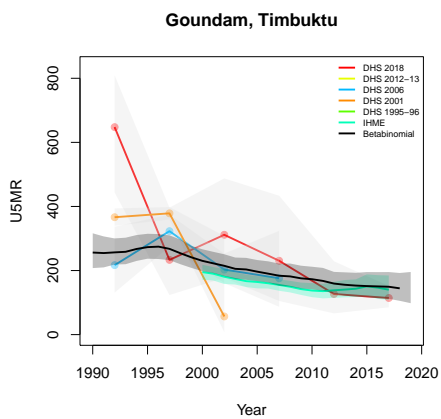


**Yorosso, Sikasso**



**Diré, Timbuktu**





**B.10 Myanmar**

Age 0	Survey	Clusters			Deaths			Agemonths		
		Urban	Rural	Total	Urban	Rural	Total	Urban	Rural	Total
1-11	2016	122	319	441	110	732	842	4682	17281	21963
12-23	2016	122	319	441	93	669	762	48669	173875	222544
24-35	2016	122	319	441	13	146	159	50842	179154	229996
36-47	2016	122	319	441	12	88	100	48952	171041	219993
48-59	2016	122	319	441	12	88	100	46678	162225	208903
	2016	122	319	441	3	49	52	44633	153212	197845

Table B.10: **Data summary for Myanmar.** Total numbers of clusters (Columns 3–5) with observations in each age group by survey in urban and rural areas and combined. Numbers of deaths (Columns 6–8) and number of agemonths (Columns 9–10) observed in each age group by survey in urban and rural areas and combined.

### B.10.1 Admin-1

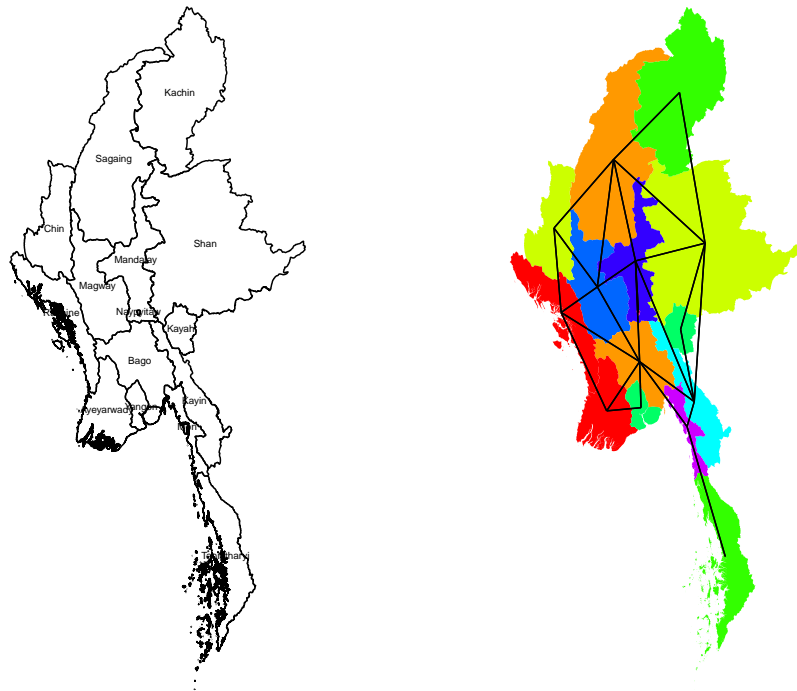


Figure B.46: **Left:** The names of the 15 Admin-1 areas of Myanmar . **Right:** The neighborhood structure of Admin-1 areas in Myanmar .

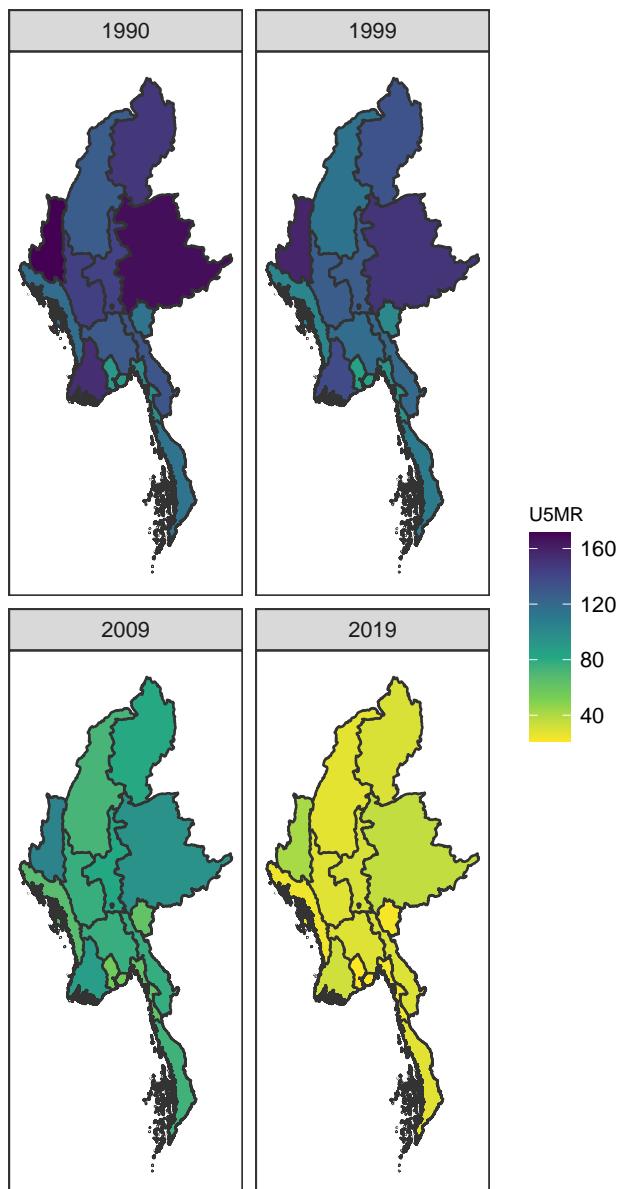


Figure B.47: Median U5MR estimates for years 1990, 1999, 2009, 2019 for Admin-1 areas in Myanmar .

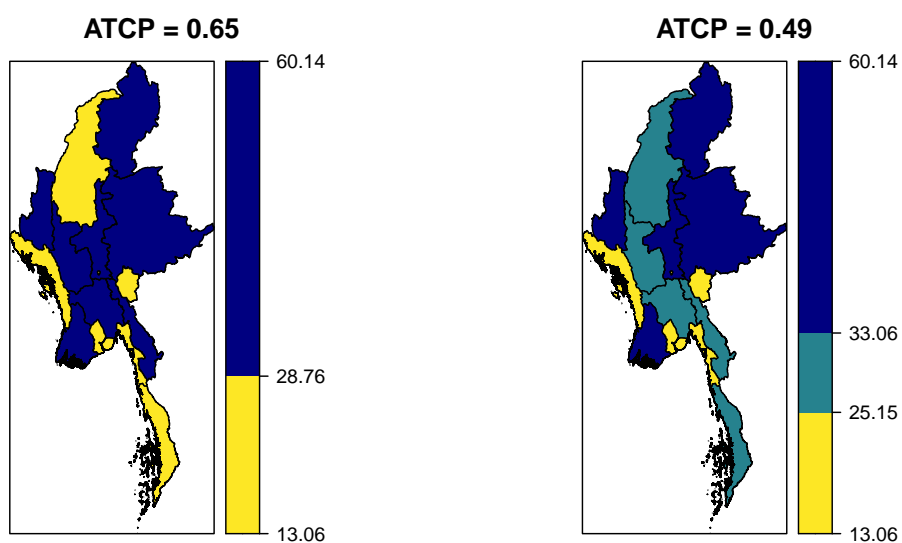
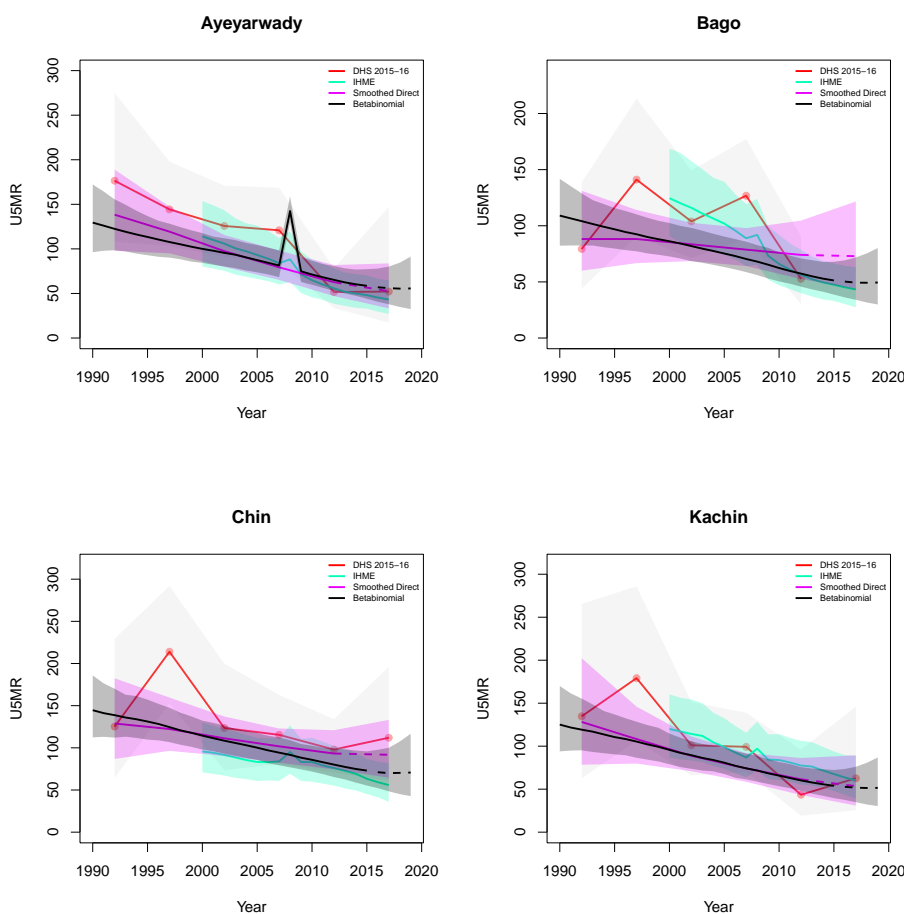
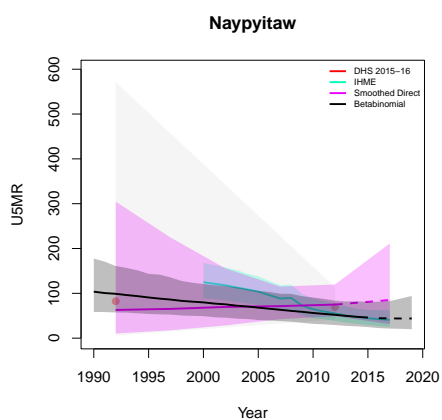
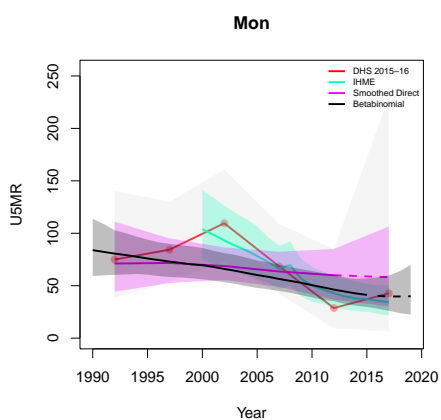
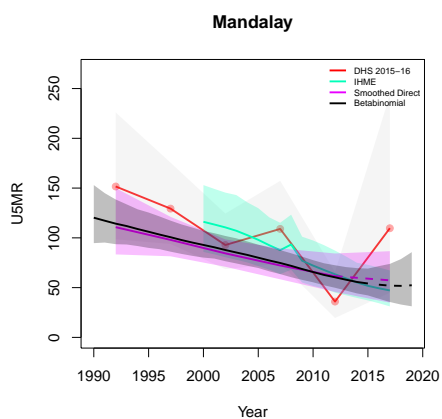
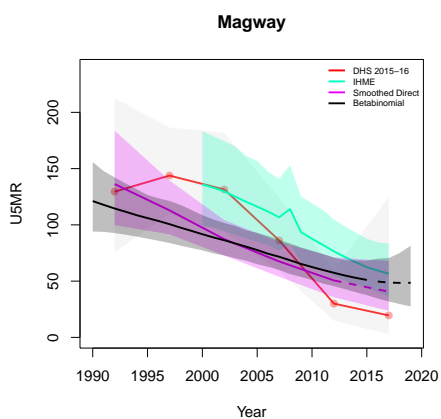
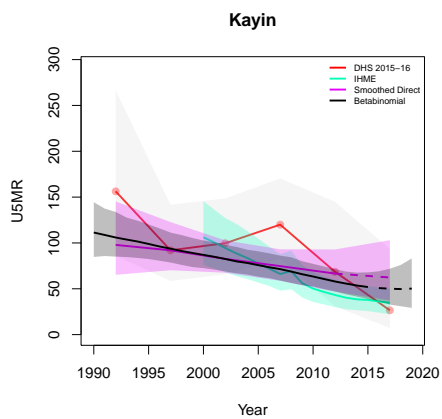
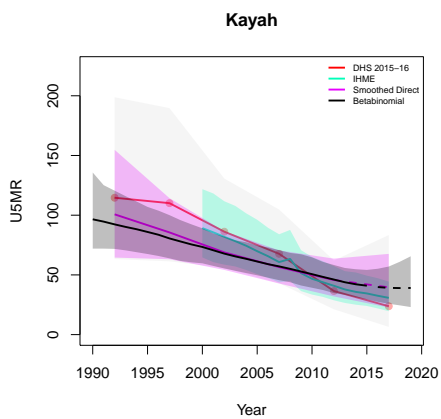


Figure B.48: Expression of uncertainty of U5MR (deaths per 1000 children) estimates for Admin-1 areas based on the average true classification probability (ATCP) in 2019 using  $K = 2, 3$  colors.

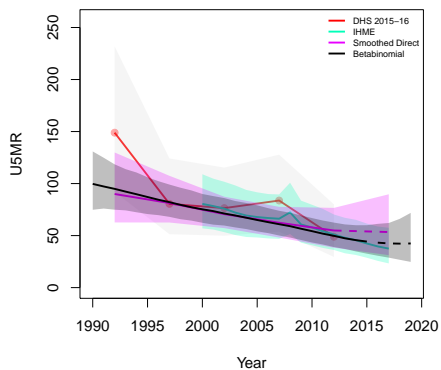
*Data and estimates over time by area*

Colored lines with circular points and light grey uncertainty bands are 5-year survey-weighted estimates of U5MR for years 1990–1994 up to 2015–2019 depending on survey timing. For a survey that ends in the middle of a 5-year period, we plot the estimates at the mid-point of the years in that interval for which the survey provides data. Black lines and corresponding intervals represent posterior medians and 95% uncertainty intervals respectively for the betabinomial model. IHME’s estimates and corresponding intervals, where we can compare, are in aquamarine.

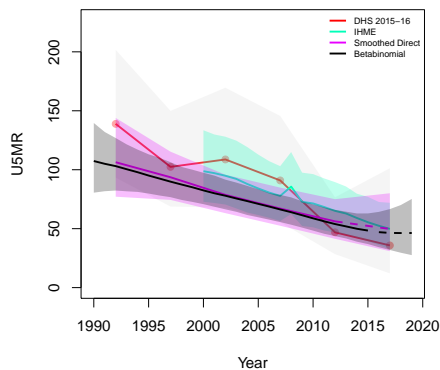




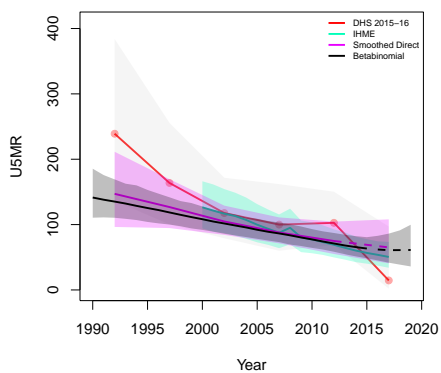
**Rakhine**



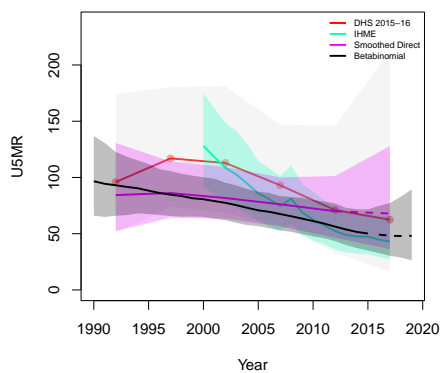
**Sagaing**



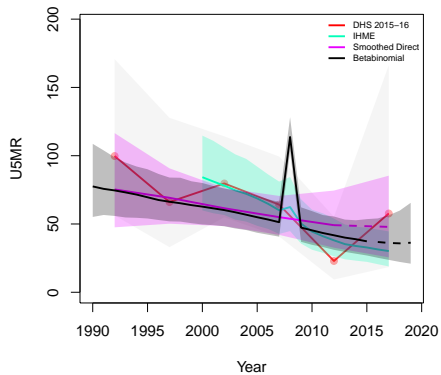
**Shan**



**Tanintharyi**



**Yangon**



*B.10.2 Admin-2*



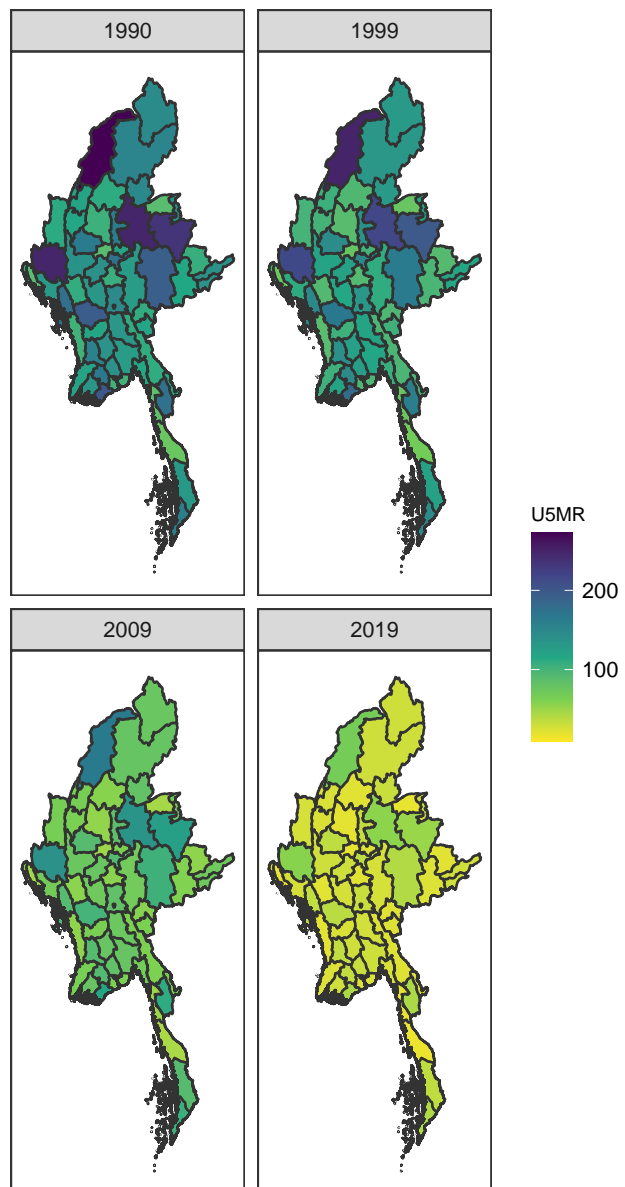


Figure B.50: Median U5MR estimates for years 1990, 1999, 2009, 2019 for Admin-2 areas in Myanmar .

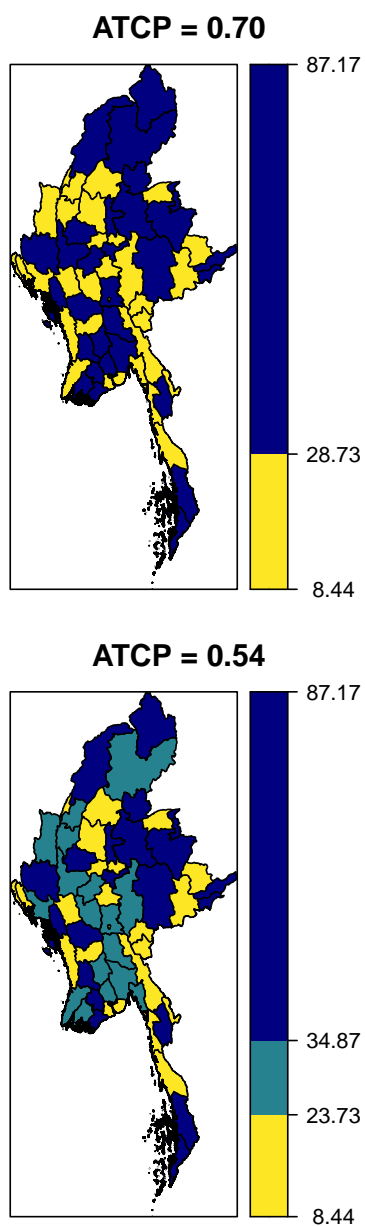
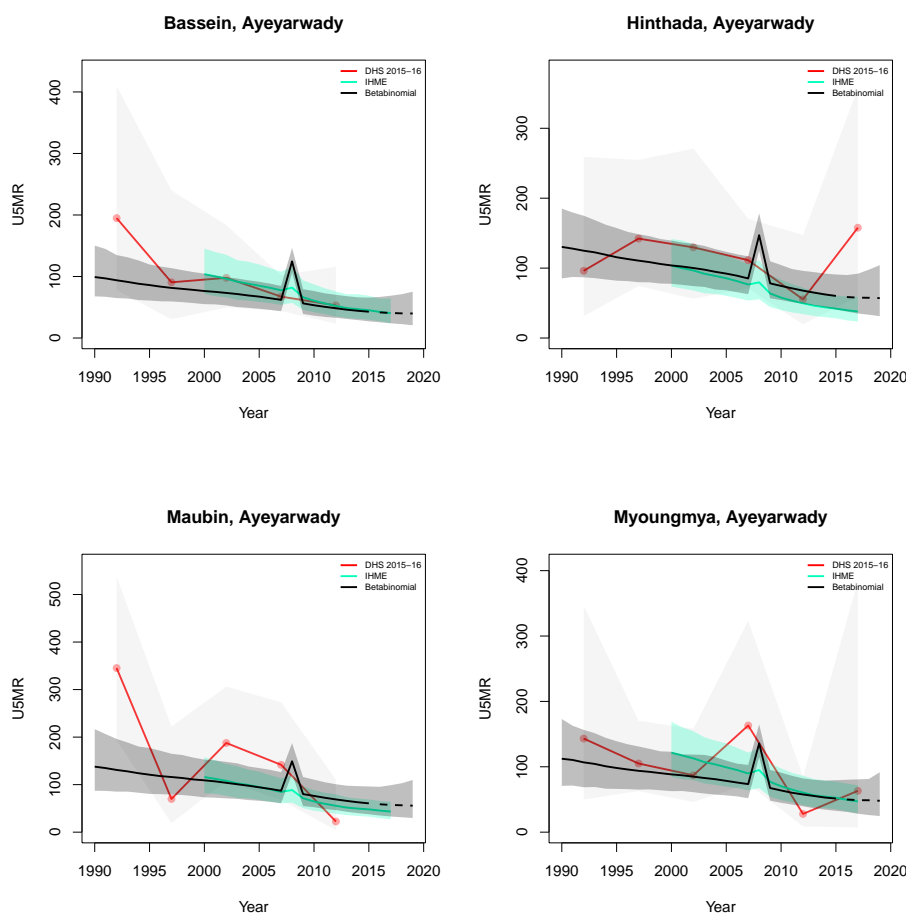
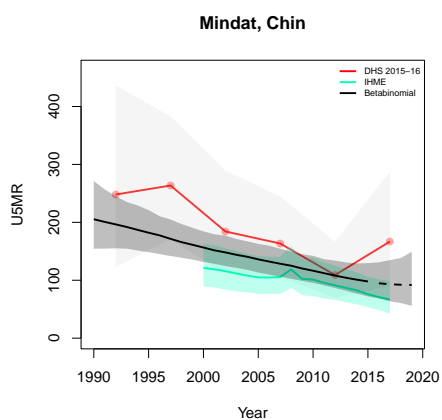
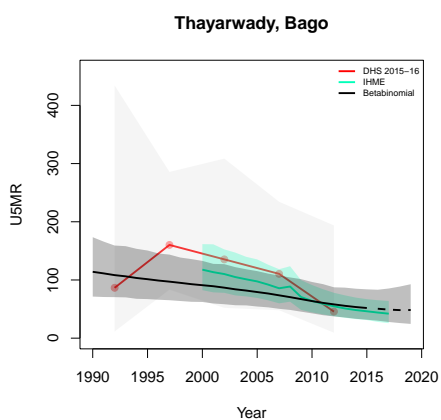
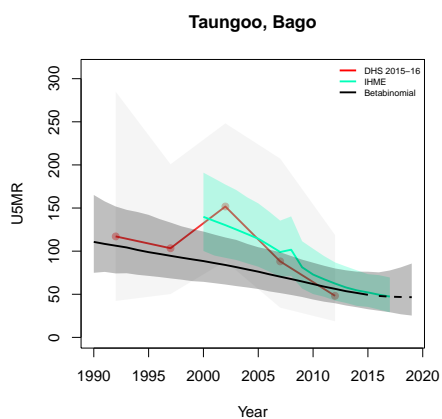
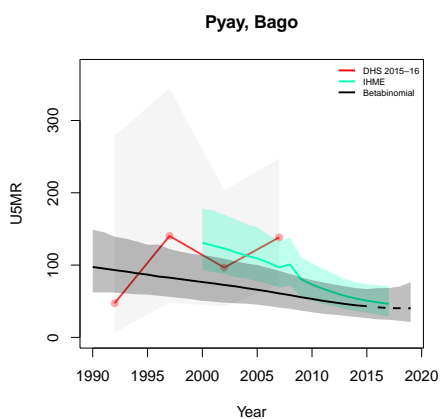
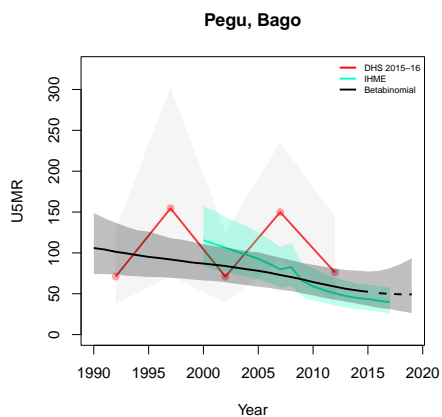
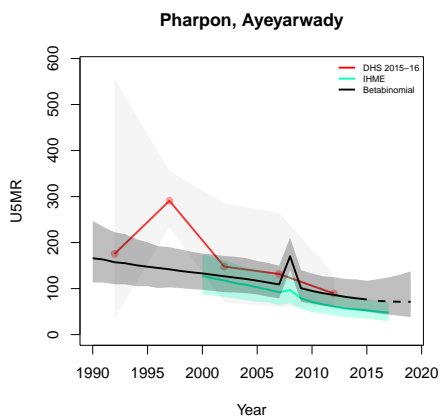


Figure B.51: Expression of uncertainty of U5MR (deaths per 1000 children) estimates for Admin-1 areas based on the average true classification probability (ATCP) in 2019 using  $K = 2, 3$  colors.

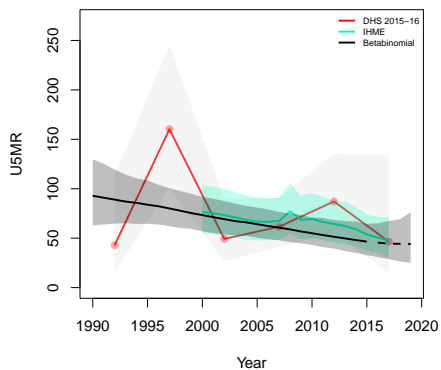
*Data and estimates over time by area*

Colored lines with circular points and light grey uncertainty bands are 5-year survey-weighted estimates of U5MR for years 1990–1994 up to 2015–2019 depending on survey timing. For a survey that ends in the middle of a 5-year period, we plot the estimates at the mid-point of the years in that interval for which the survey provides data. Black lines and corresponding intervals represent posterior medians and 95% uncertainty intervals respectively for the betabinomial model. IHME’s estimates and corresponding intervals, where we can compare, are in aquamarine.





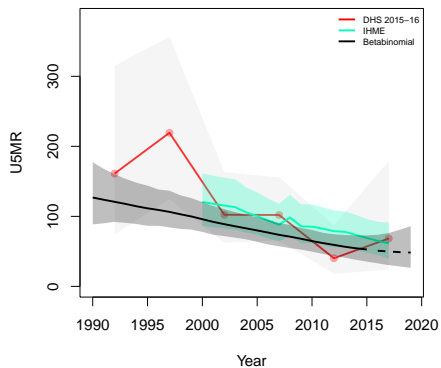
**Palam, Chin**



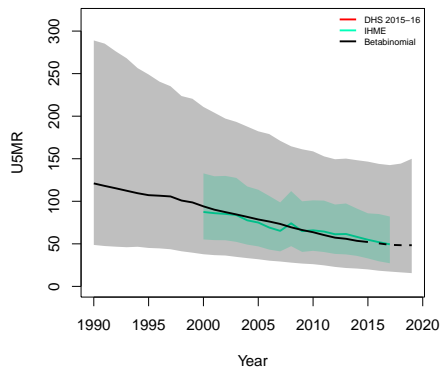
**Bhamo, Kachin**



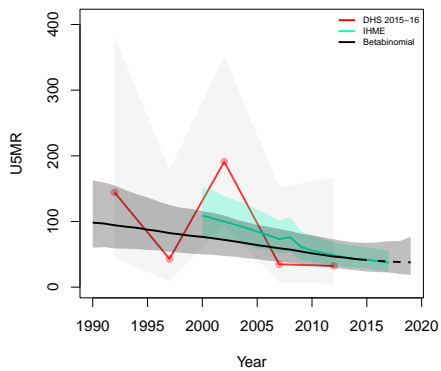
**Myitkyina, Kachin**



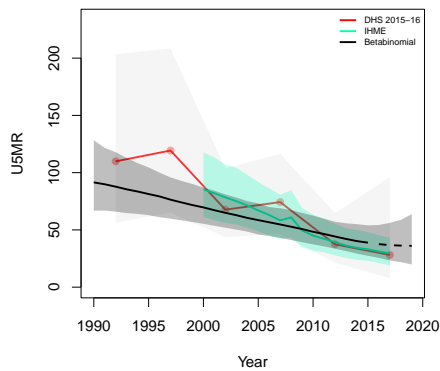
**Putao, Kachin**



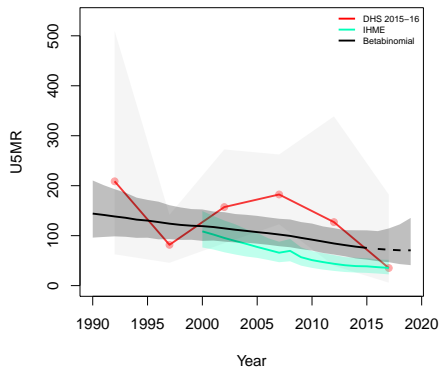
**Bawlake, Kayah**



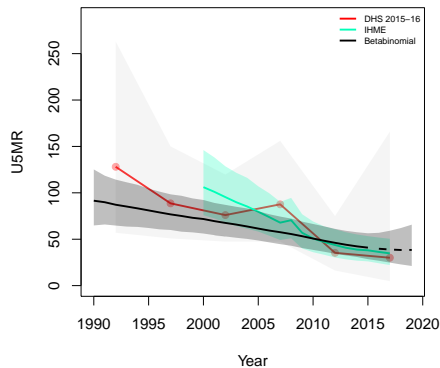
**Loikaw, Kayah**



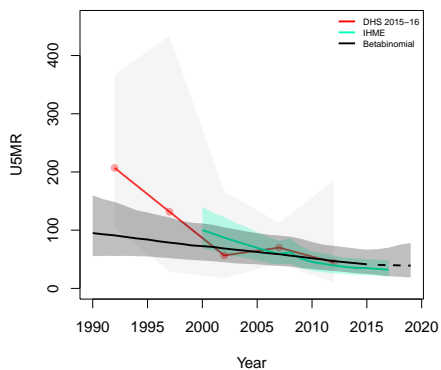
**Hpa-an, Kayin**



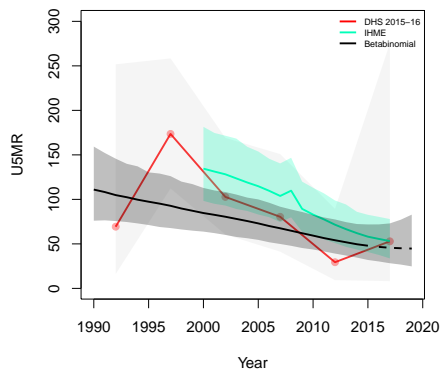
**Kawkareik, Kayin**



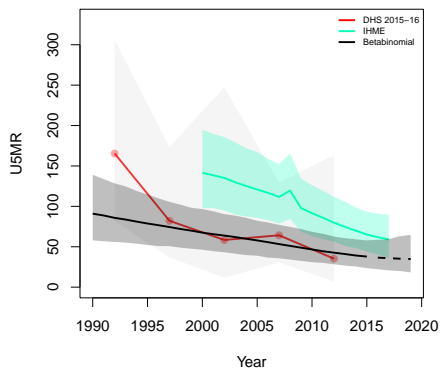
**Myawady, Kayin**



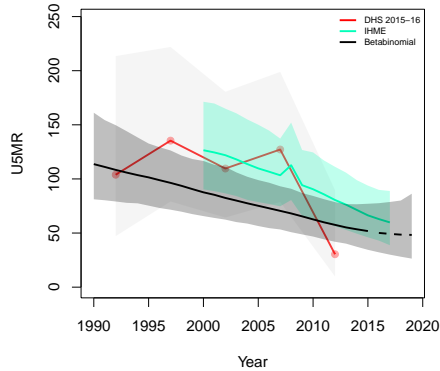
**Magwe Minbu, Magway**



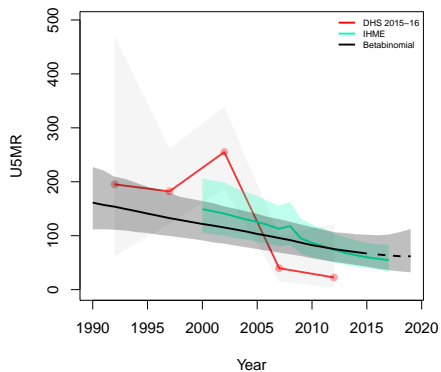
**Minbu, Magway**



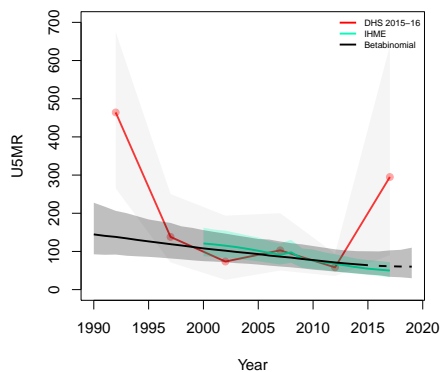
**Pakokku, Magway**



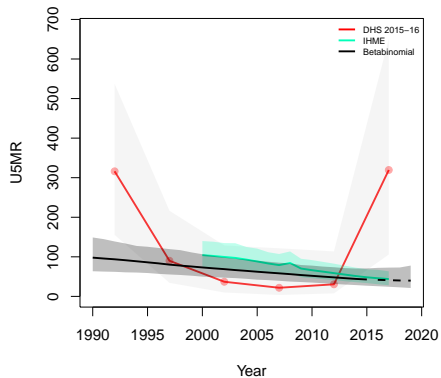
**Thayetmyo, Magway**



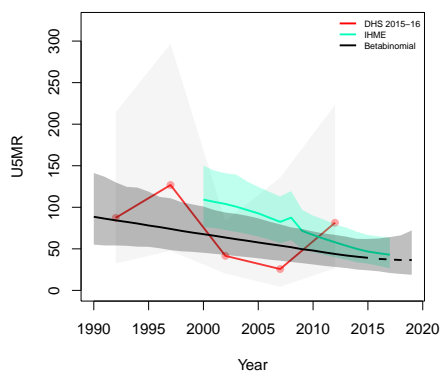
**Kyaukse, Mandalay**



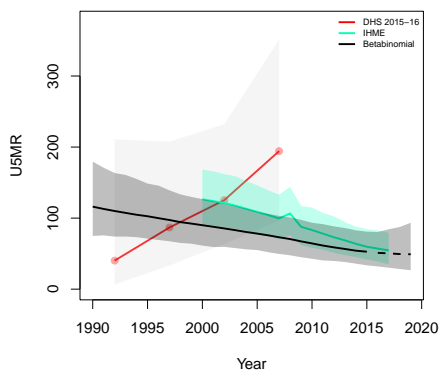
**Mandalay, Mandalay**



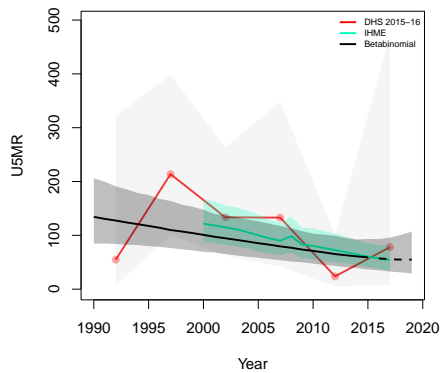
**Meiktila, Mandalay**



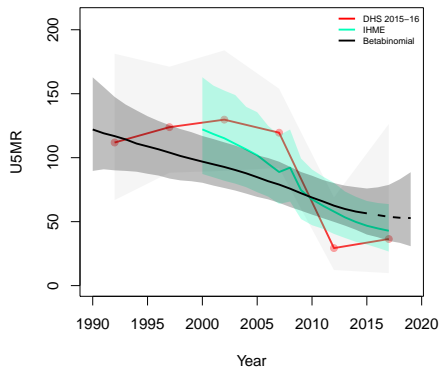
**Myingyan, Mandalay**



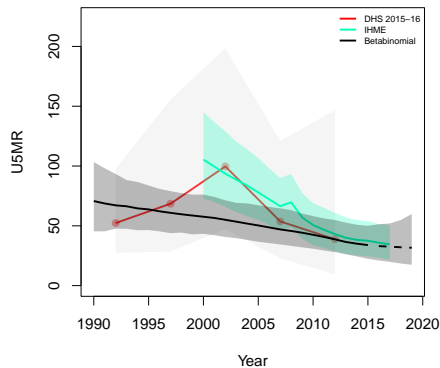
**Pyin-Oo-Lwin, Mandalay**



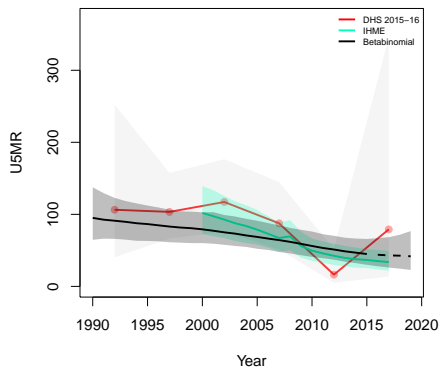
**Yamethin, Mandalay**



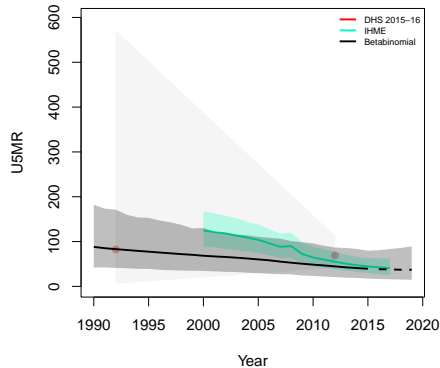
**Mawlamyine, Mon**



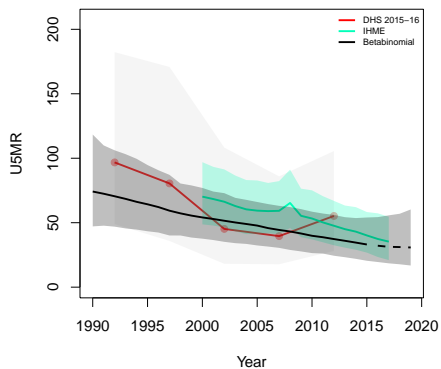
**Thaton, Mon**



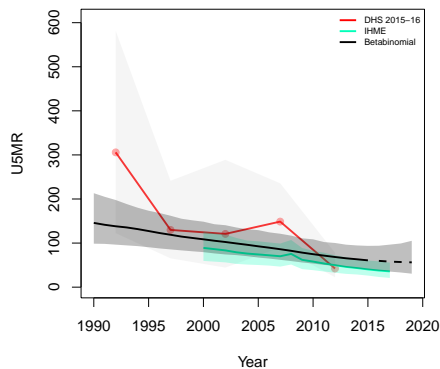
**Naypyitaw, Naypyitaw**



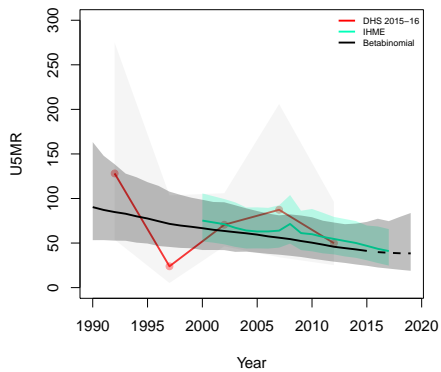
**Buthidaung, Rakhine**



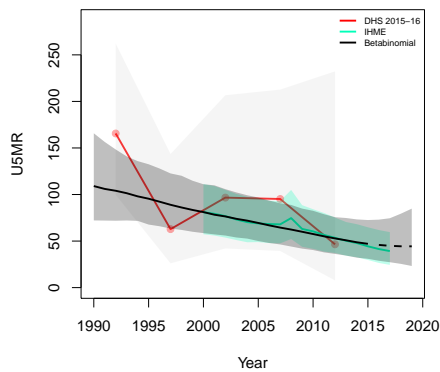
**Kyaunkpyu, Rakhine**



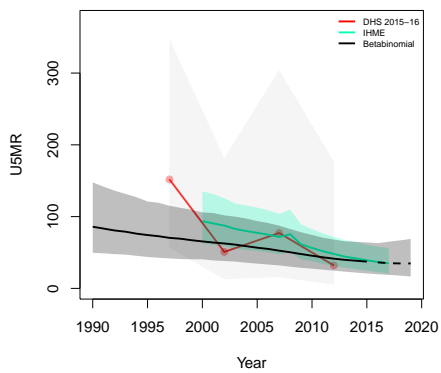
**Maungtaw, Rakhine**



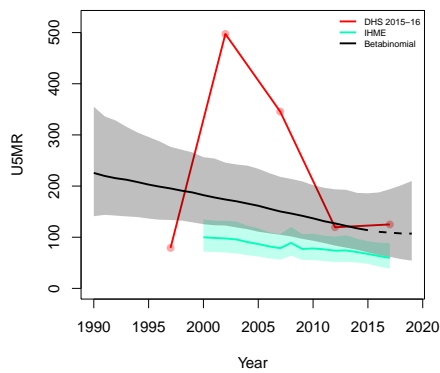
**Sittwe, Rakhine**



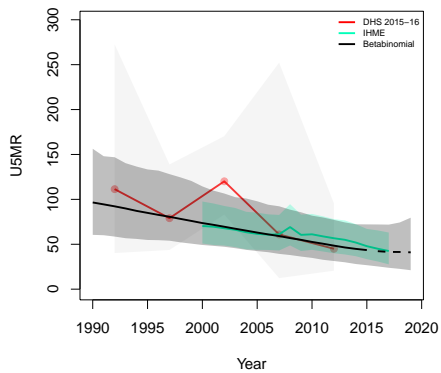
**Thandwe, Rakhine**



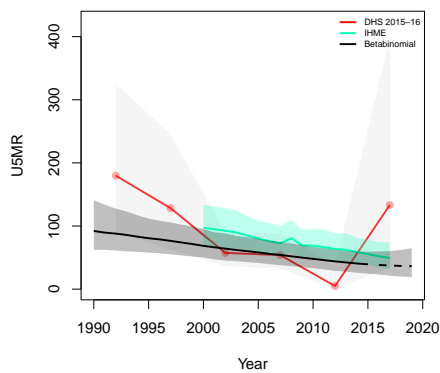
**Hkamti, Sagaing**



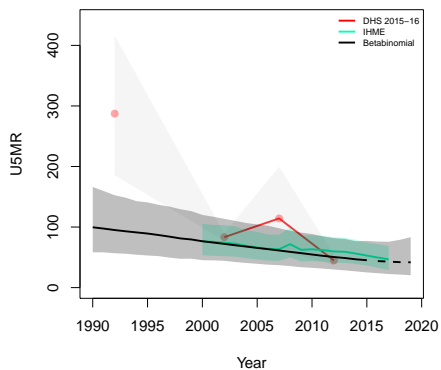
**Kalemyo, Sagaing**



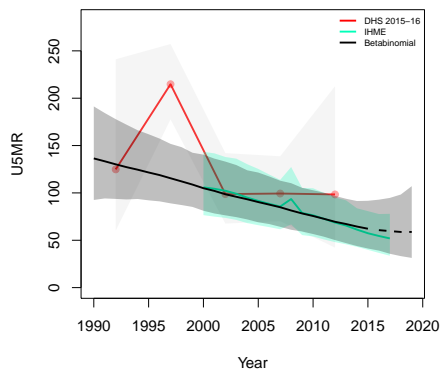
**Katha, Sagaing**



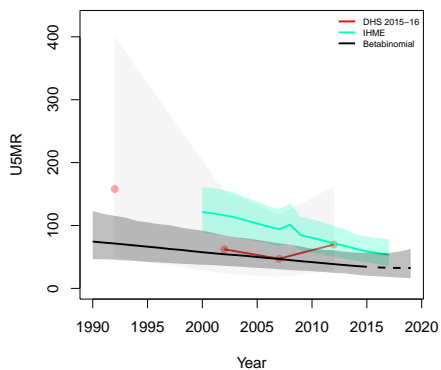
**Mawleik, Sagaing**



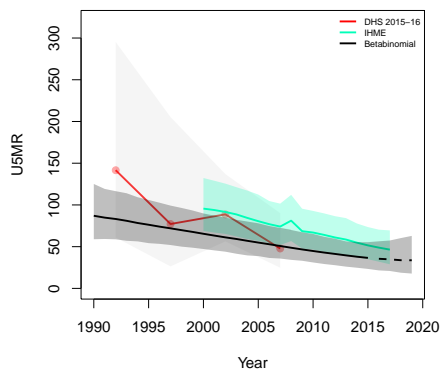
**Monywa, Sagaing**



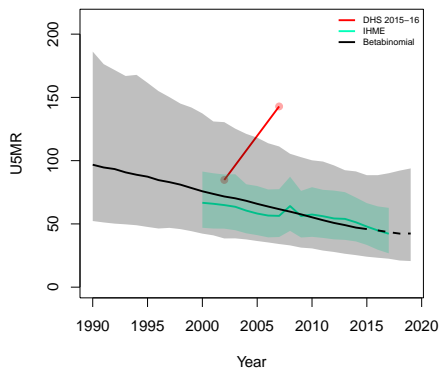
**Sagaing, Sagaing**



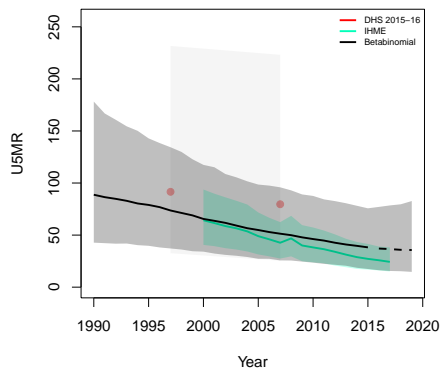
**Shwebo, Sagaing**



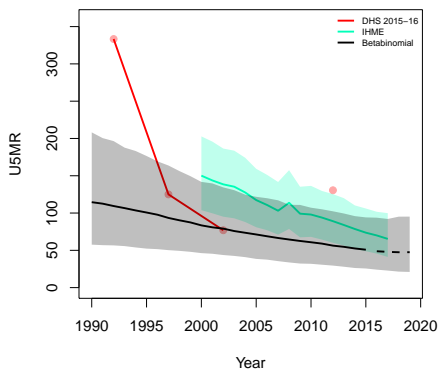
**Tamu, Sagaing**



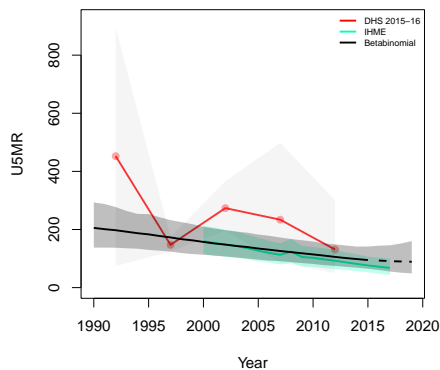
**Kengtung, Shan**



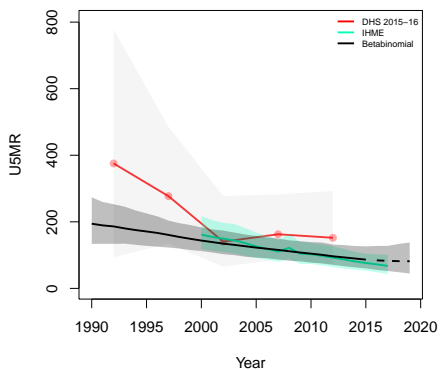
**Kunlong, Shan**



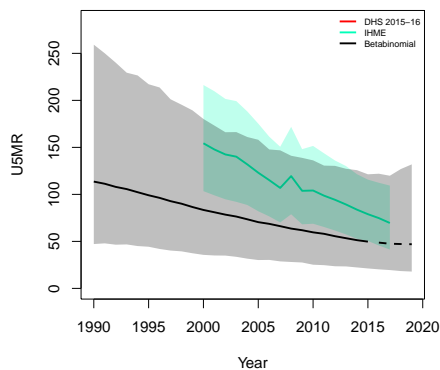
**Kyaukme, Shan**



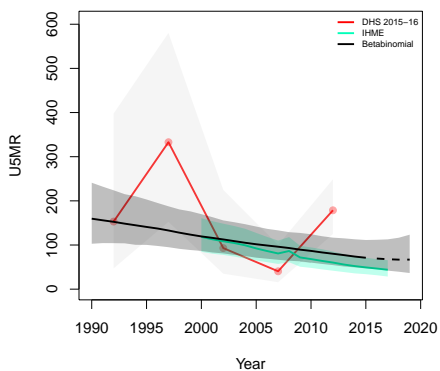
**Lasho, Shan**



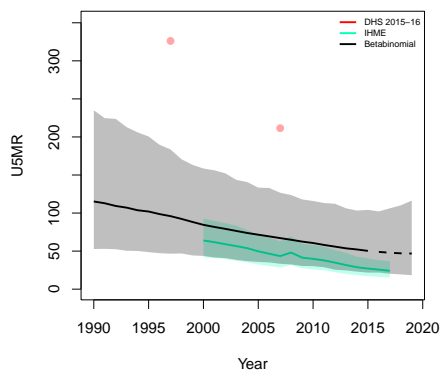
**Lauking, Shan**



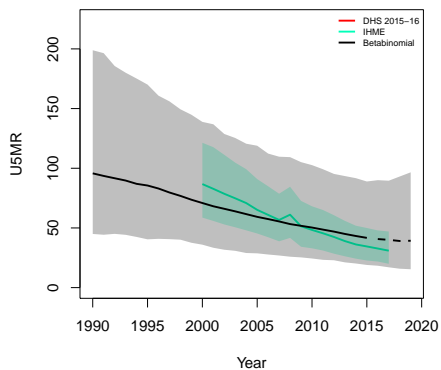
**Loilen, Shan**



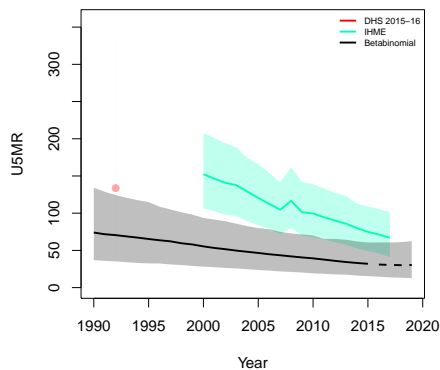
**Mongphat, Shan**



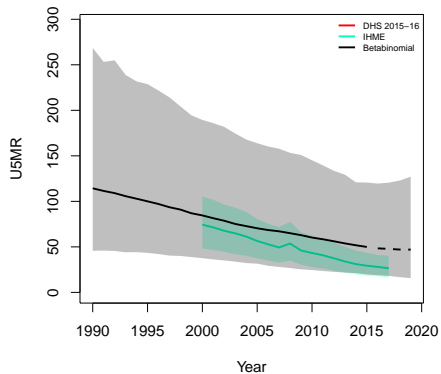
**Mongsat, Shan**



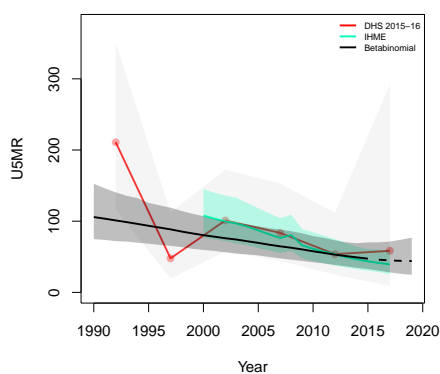
**Muse, Shan**



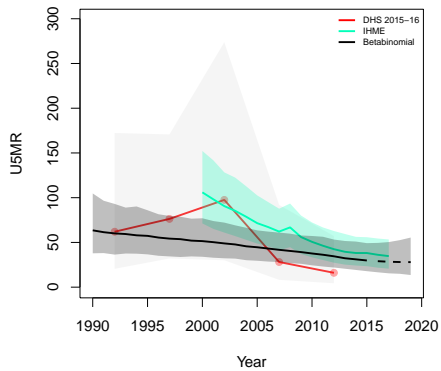
**Tarchilaik, Shan**



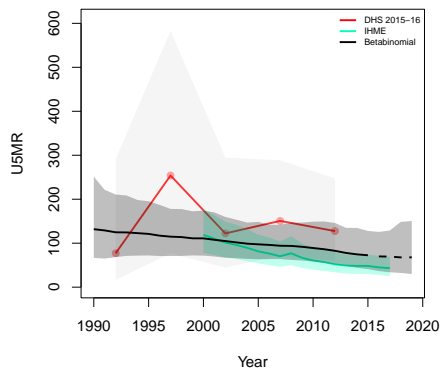
**Taunggye, Shan**

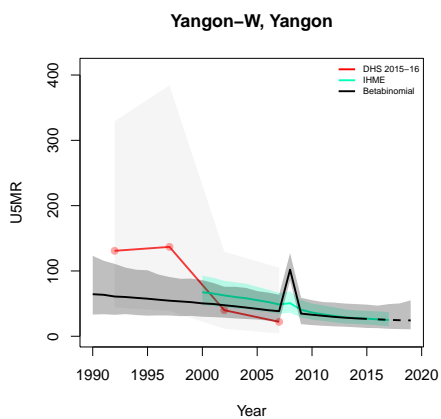
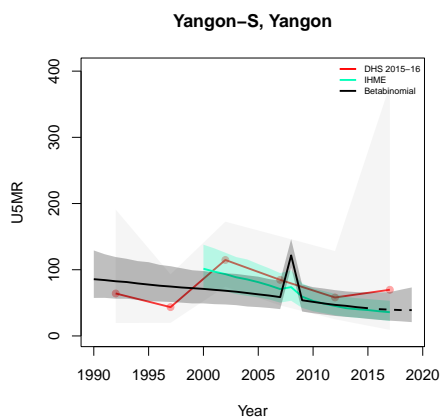
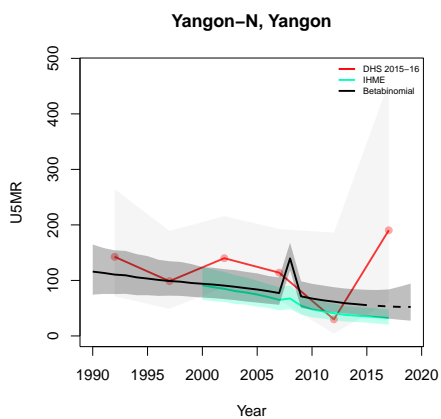
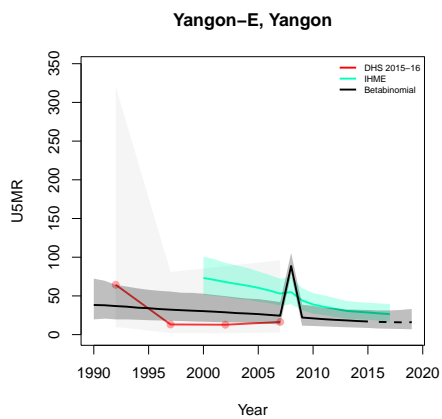
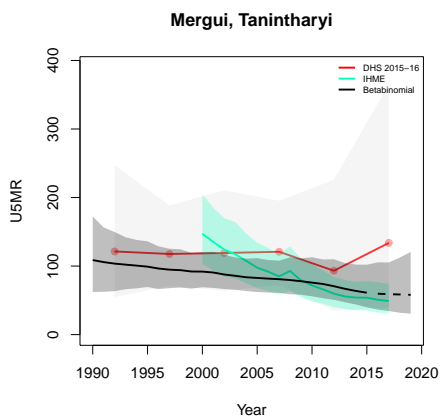


**Dawei, Tanintharyi**



**Kawthoung, Tanintharyi**





**B.11 Namibia**

Age	Survey	Clusters			Deaths			Agemonths		
		Urban	Rural	Total	Urban	Rural	Total	Urban	Rural	Total
0	2000	106	153	259	57	144	201	3322	5127	8449
	2006	208	283	491	129	232	361	5891	9055	14946
	2013	267	282	549	134	215	349	7697	9257	16954
1-11	2000	106	153	259	69	89	158	35150	53167	88317
	2006	208	283	491	109	225	334	61781	93593	155374
	2013	267	282	549	130	202	332	80579	95450	176029
12-23	2000	106	153	259	30	38	68	36680	54826	91506
	2006	208	283	491	49	87	136	63823	95664	159487
	2013	267	282	549	57	70	127	82589	97141	179730
24-35	2000	106	153	259	12	30	42	35796	52746	88542
	2006	208	283	491	29	49	78	61207	91733	152940
	2013	267	282	549	22	32	54	77994	91641	169635
36-47	2000	106	153	259	8	16	24	34689	50798	85487
	2006	208	283	491	16	30	46	58253	87357	145610
	2013	267	282	549	20	28	48	73272	85882	159154
48-59	2000	106	153	259	6	11	17	33696	48714	82410
	2006	208	283	491	6	23	29	55500	83110	138610
	2013	267	282	549	10	10	20	68687	81086	149773

Table B.11: **Data summary for Namibia.** Total numbers of clusters (Columns 3–5) with observations in each age group by survey in urban and rural areas and combined. Numbers of deaths (Columns 6–8) and number of agemonths (Columns 9–10) observed in each age group by survey in urban and rural areas and combined.

*B.11.1 Admin-1*

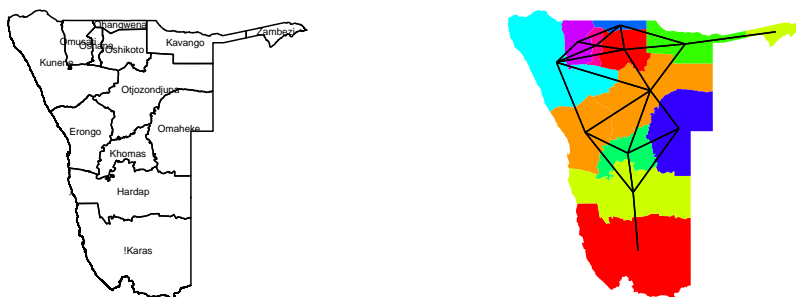


Figure B.52: **Left:** The names of the 13 Admin-1 areas of Namibia . **Right:** The neighborhood structure of Admin-1 areas in Namibia .

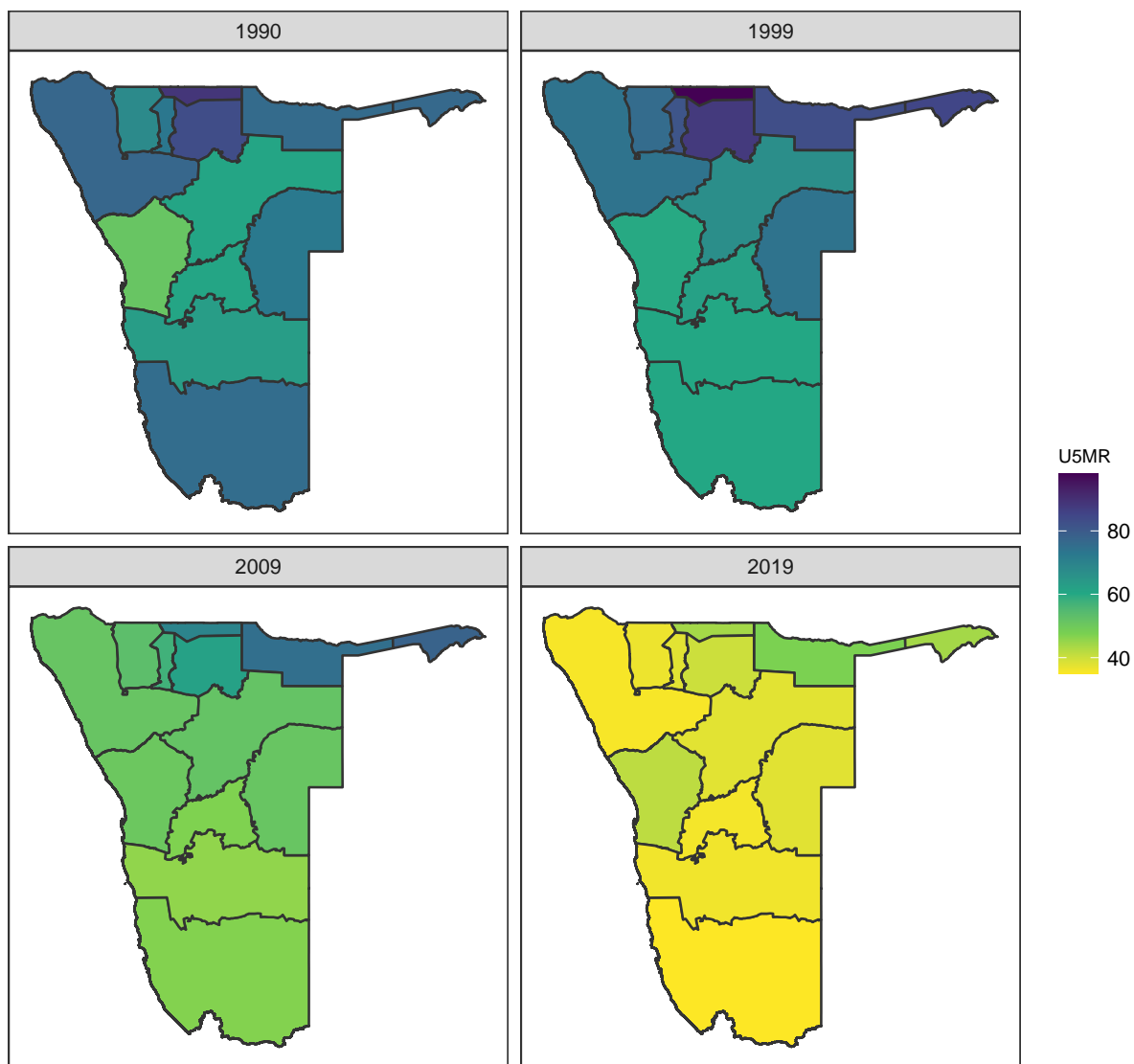


Figure B.53: Median U5MR estimates for years 1990, 1999, 2009, 2019 for Admin-1 areas in Namibia .

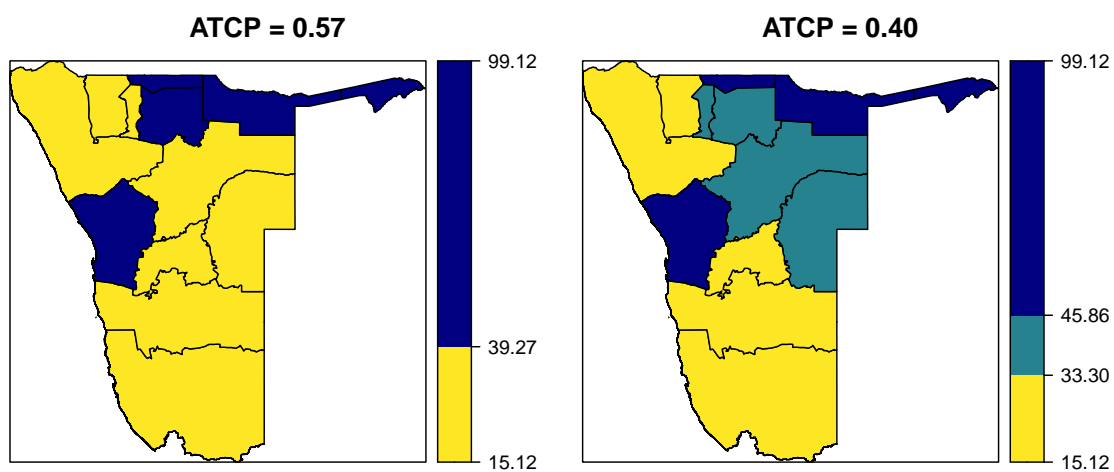
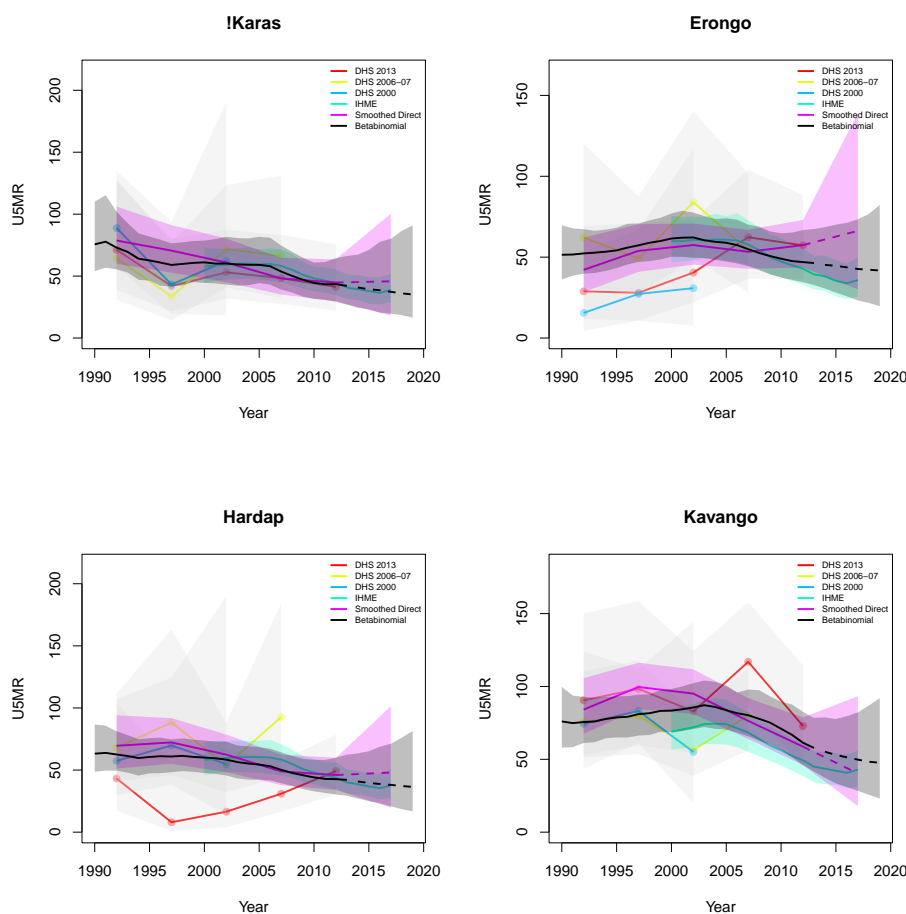
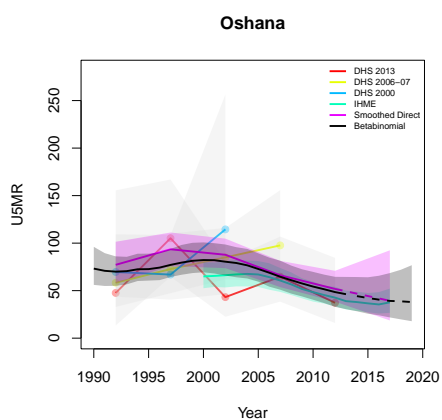
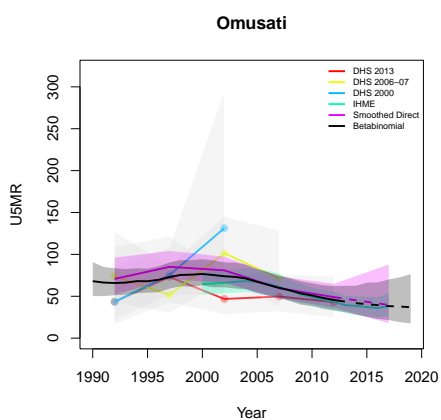
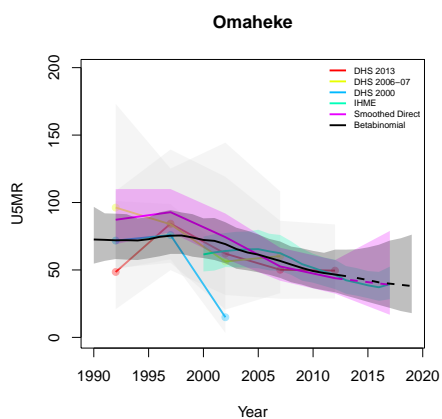
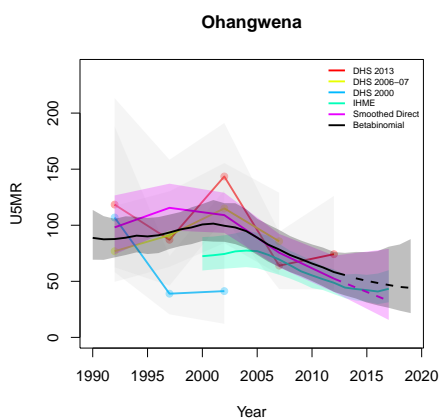
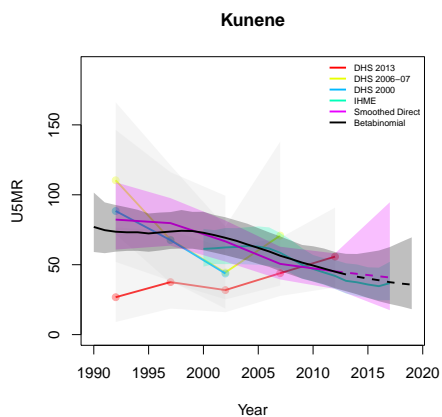
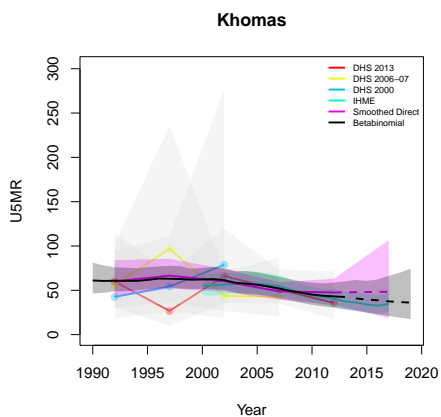


Figure B.54: Expression of uncertainty of U5MR (deaths per 1000 children) estimates for Admin-1 areas based on the average true classification probability (ATCP) in 2019 using  $K = 2, 3$  colors.

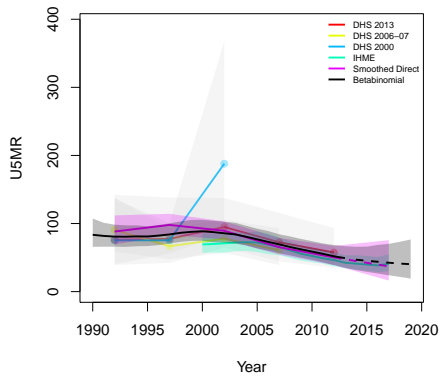
*Data and estimates over time by area*

Colored lines with circular points and light grey uncertainty bands are 5-year survey-weighted estimates of U5MR for years 1990–1994 up to 2015–2019 depending on survey timing. For a survey that ends in the middle of a 5-year period, we plot the estimates at the mid-point of the years in that interval for which the survey provides data. Black lines and corresponding intervals represent posterior medians and 95% uncertainty intervals respectively for the betabinomial model. IHME’s estimates and corresponding intervals, where we can compare, are in aquamarine.

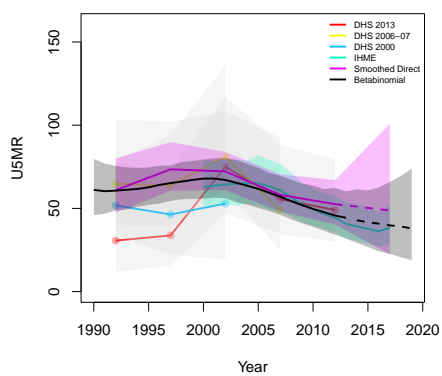




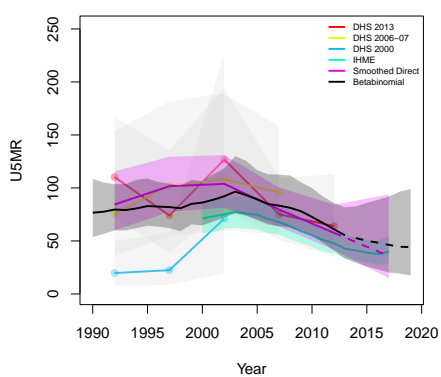
**Oshikoto**



**Otjozondjupa**



**Zambezi**



*B.11.2 Admin-2*



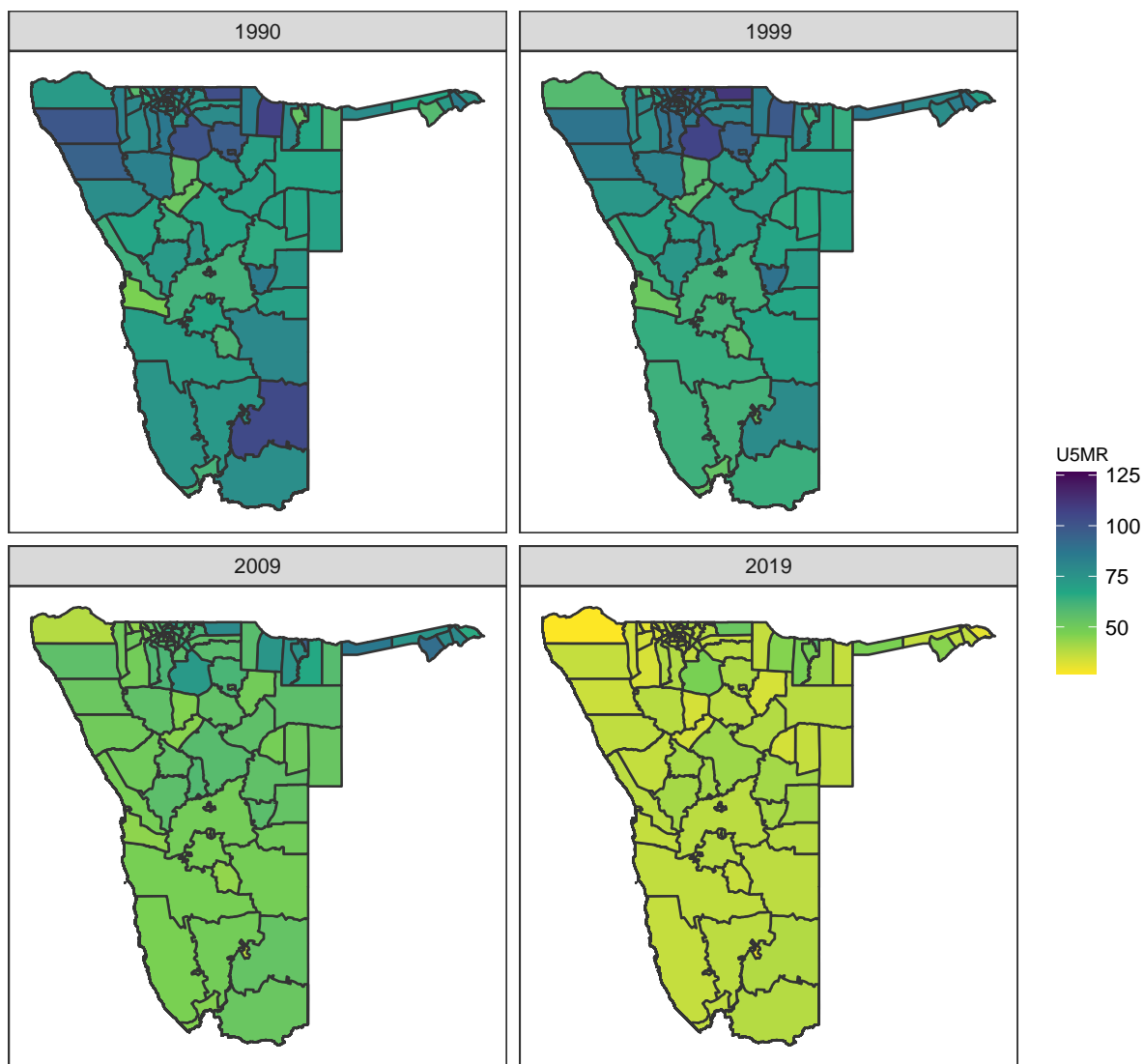


Figure B.56: Median U5MR estimates for years 1990, 1999, 2009, 2019 for Admin-2 areas in Namibia .

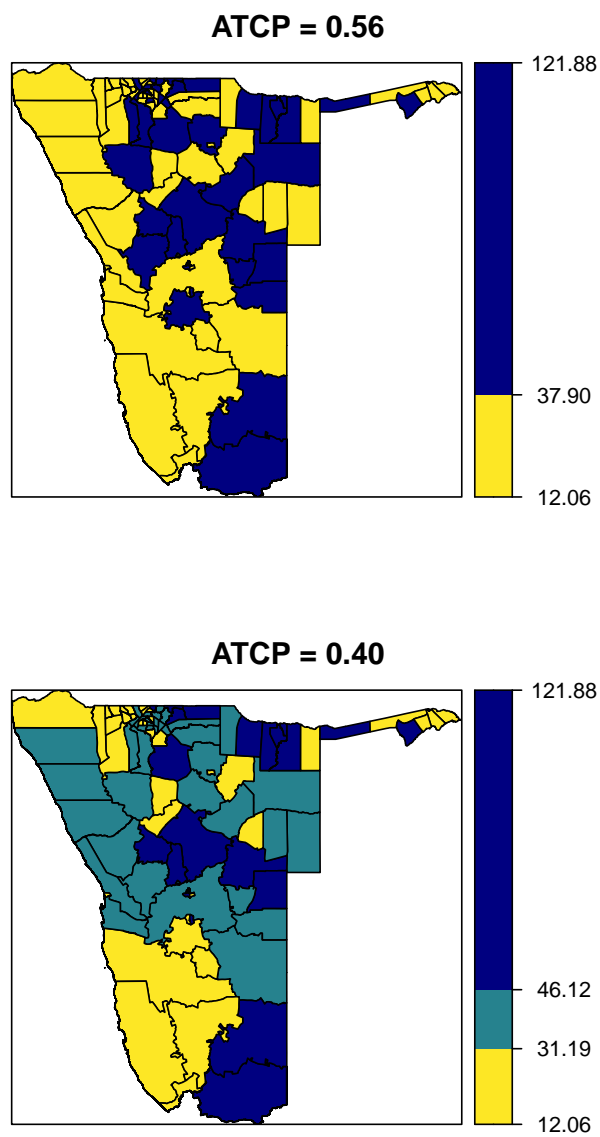
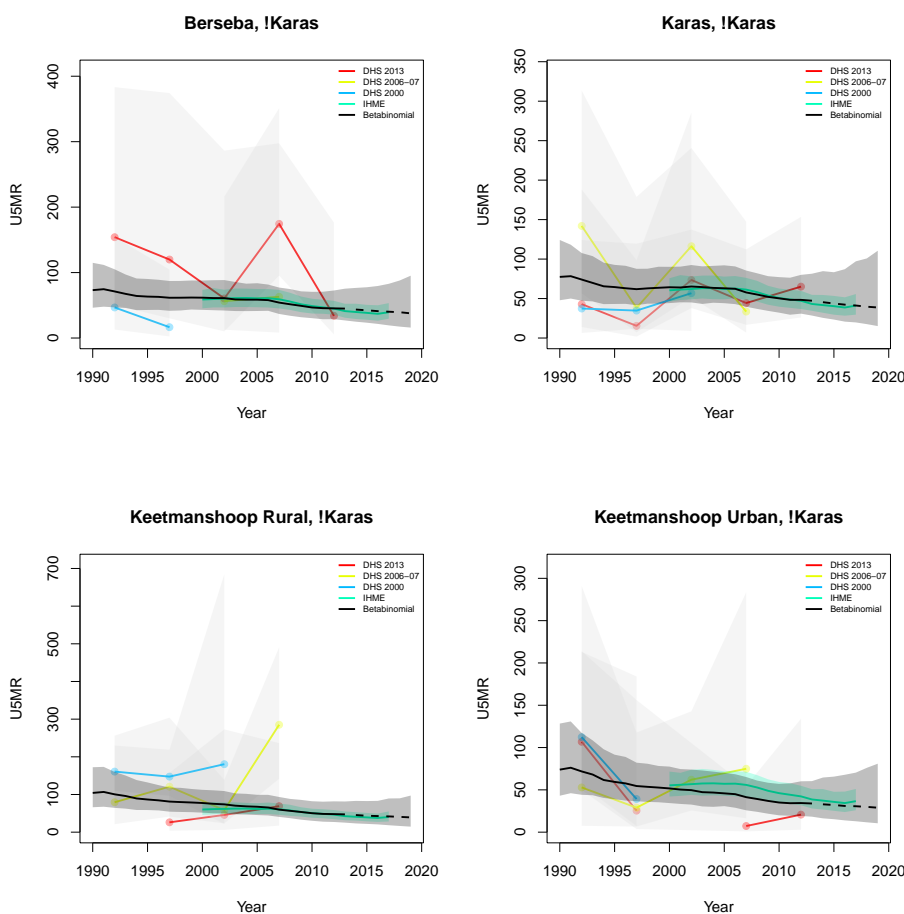


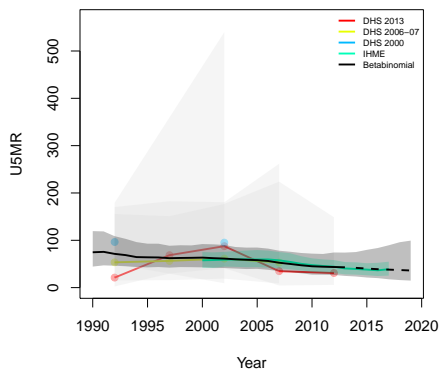
Figure B.57: Expression of uncertainty of U5MR (deaths per 1000 children) estimates for Admin-1 areas based on the average true classification probability (ATCP) in 2019 using  $K = 2, 3$  colors.

*Data and estimates over time by area*

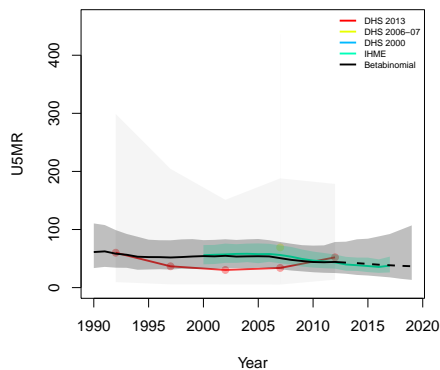
Colored lines with circular points and light grey uncertainty bands are 5-year survey-weighted estimates of U5MR for years 1990–1994 up to 2015–2019 depending on survey timing. For a survey that ends in the middle of a 5-year period, we plot the estimates at the mid-point of the years in that interval for which the survey provides data. Black lines and corresponding intervals represent posterior medians and 95% uncertainty intervals respectively for the betabinomial model. IHME’s estimates and corresponding intervals, where we can compare, are in aquamarine.



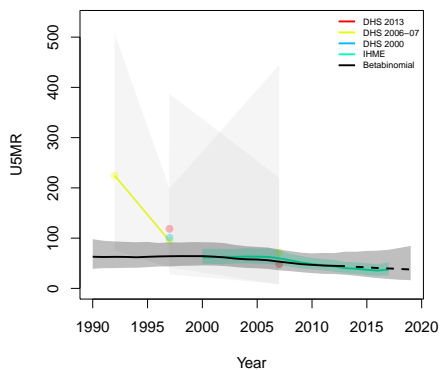
**Luderitz, !Karas**



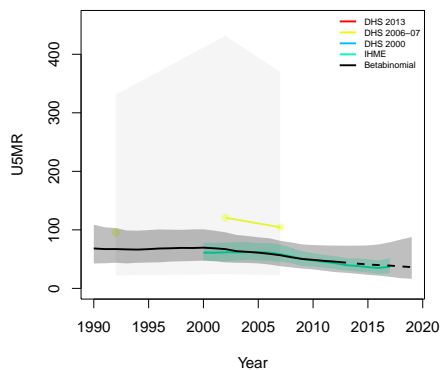
**Oranjemund, !Karas**



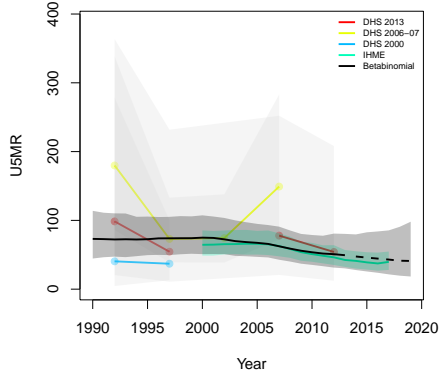
**Arandis, Erongo**



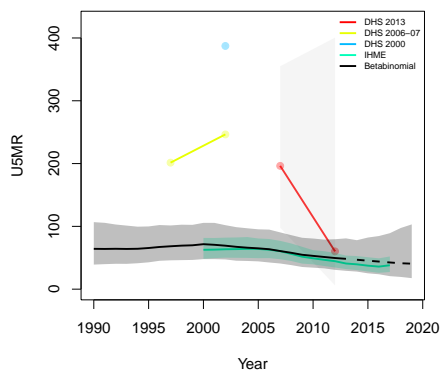
**Daures, Erongo**



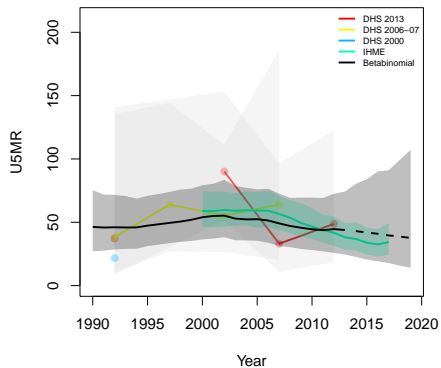
**Karibib, Erongo**



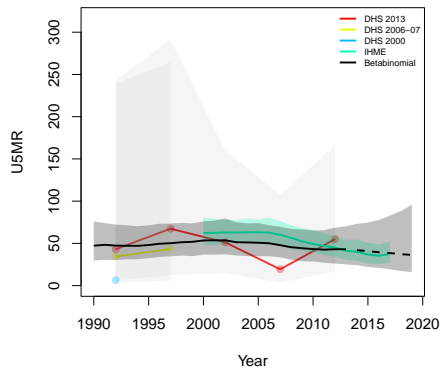
**Omaruru, Erongo**



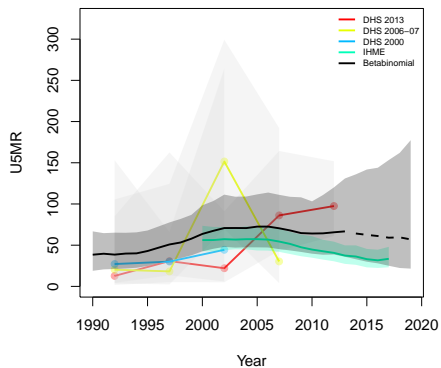
**Swakopmund, Erongo**



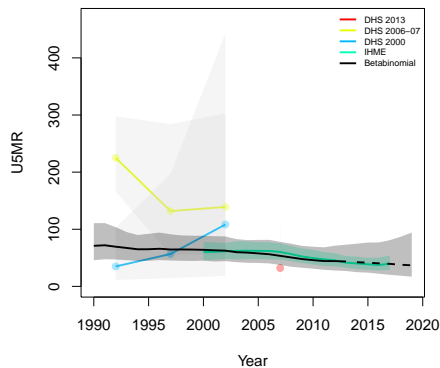
**Walvisbay Rural, Erongo**



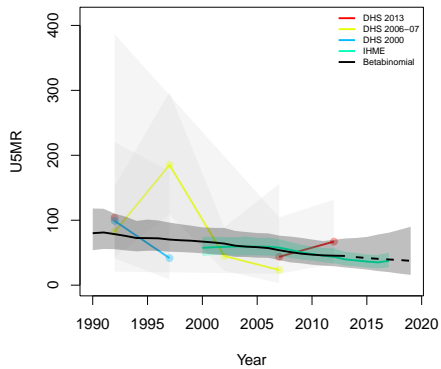
**Walvisbay Urban, Erongo**



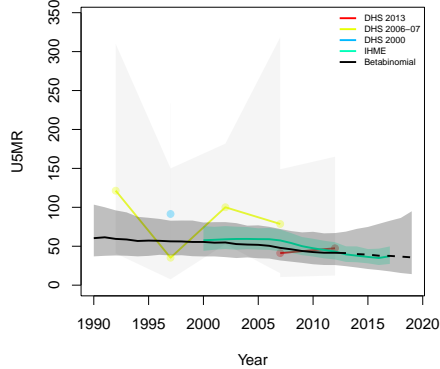
**Gibeon, Hardap**



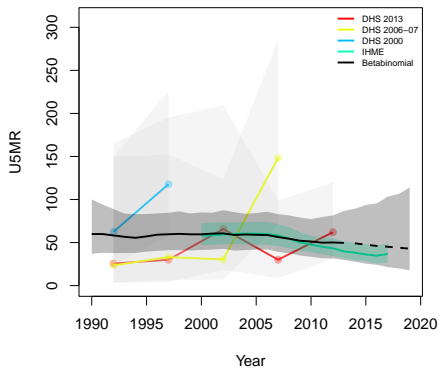
**Mariental Rural, Hardap**



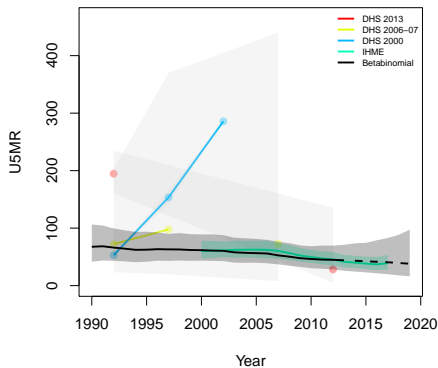
**Mariental Urban, Hardap**



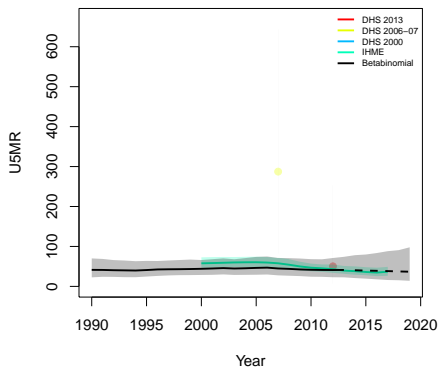
**Rehoboth East, Hardap**



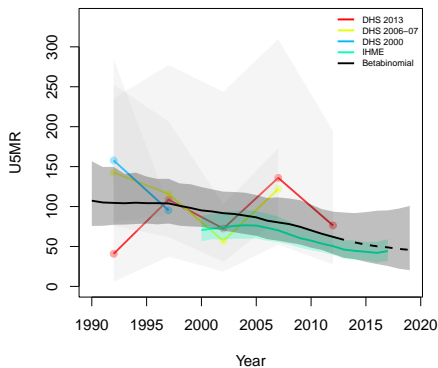
**Rehoboth Rural, Hardap**



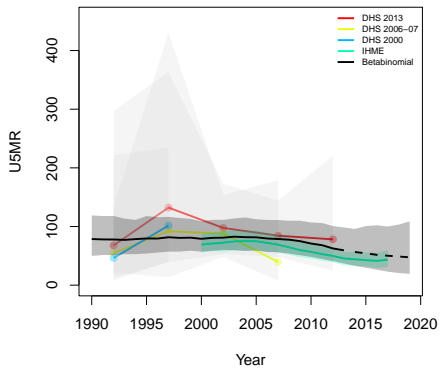
**Rehoboth West, Hardap**



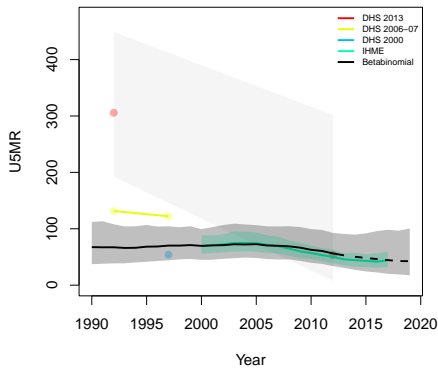
**Kahenge, Kavango**



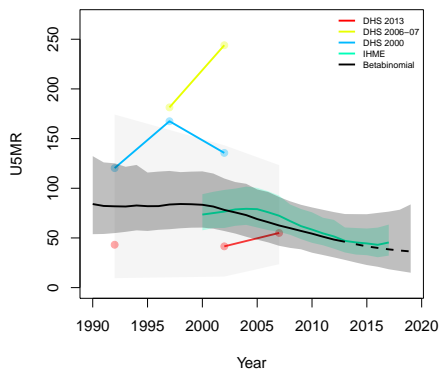
**Kapako, Kavango**



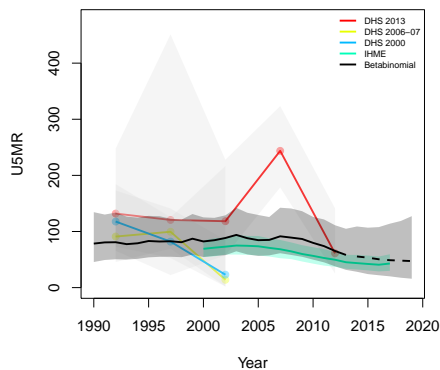
**Mashare, Kavango**



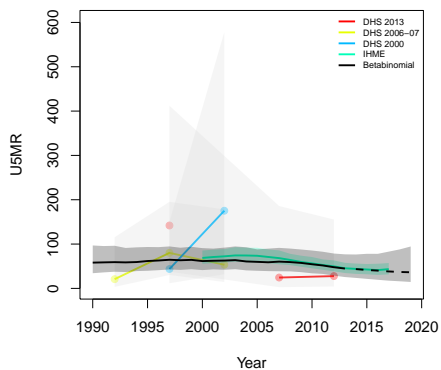
**Mpungu, Kavango**



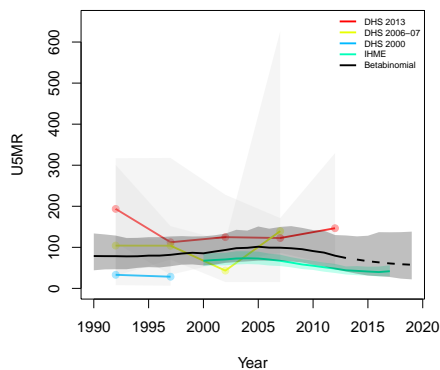
**Mukwe, Kavango**



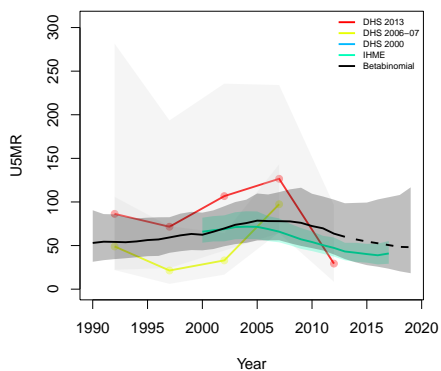
**Ndiyona, Kavango**



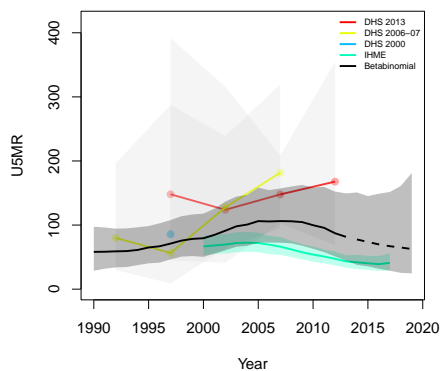
**Rundu Rural East, Kavango**



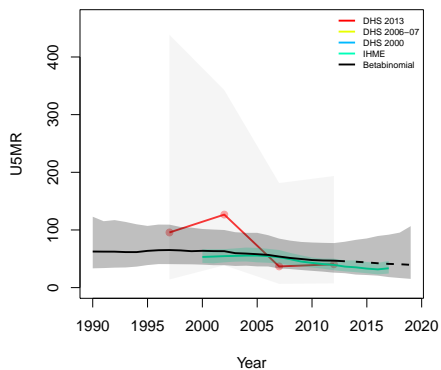
**Rundu Rural West, Kavango**



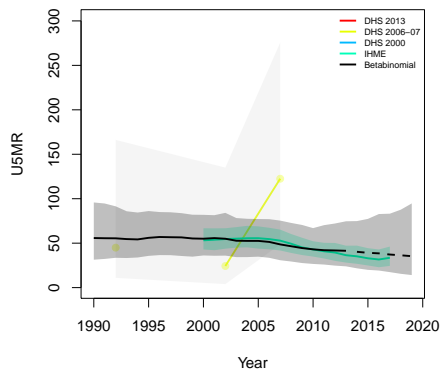
**Rundu Urban, Kavango**



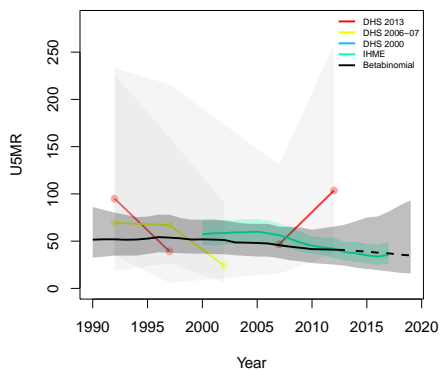
**Katutura Central, Khomas**



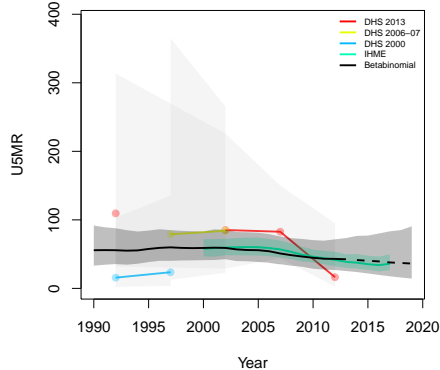
**Katutura East, Khomas**



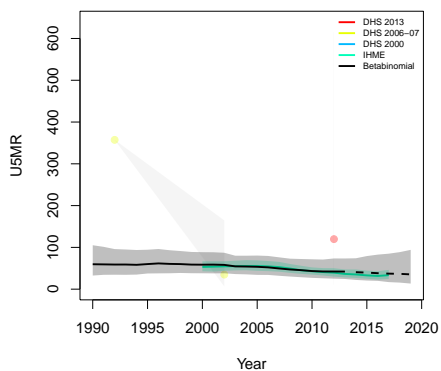
**Khomasdal North, Khomas**



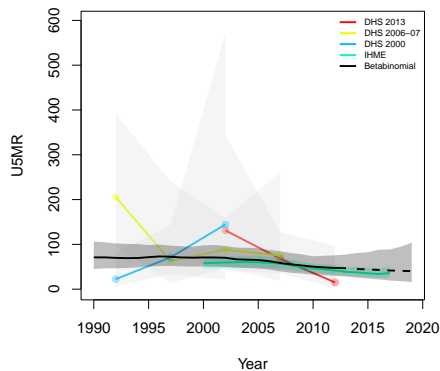
**Moses Garoeb, Khomas**



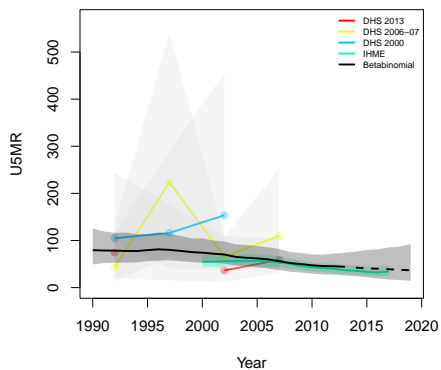
**Soweto, Khomas**



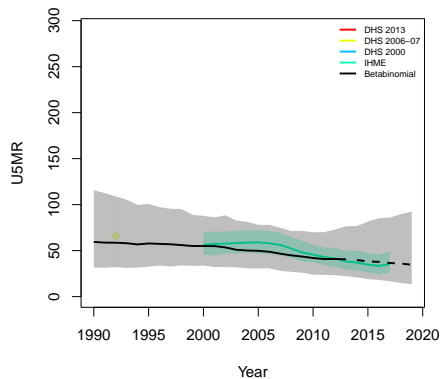
**Tobias Haiyeko, Khomas**



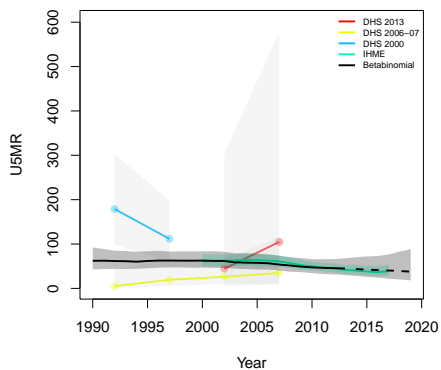
**Wanaheda, Khomas**



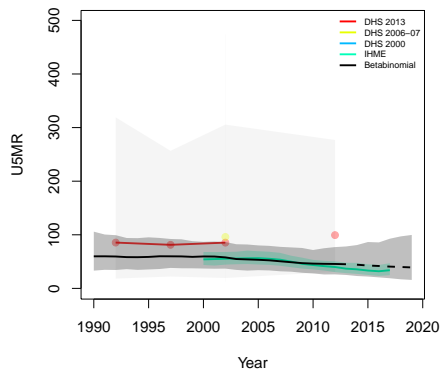
**Windhoek East, Khomas**



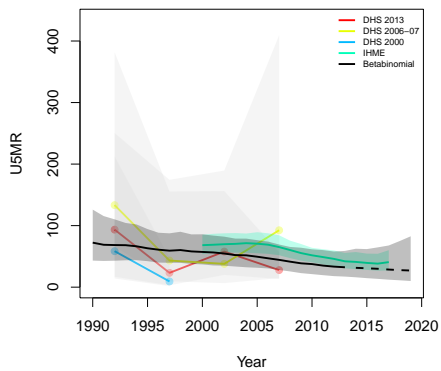
**Windhoek Rural, Khomas**



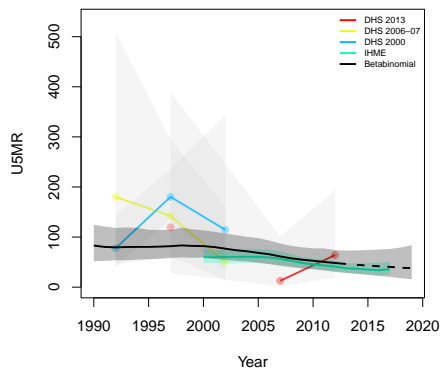
**Windhoek West, Khomas**



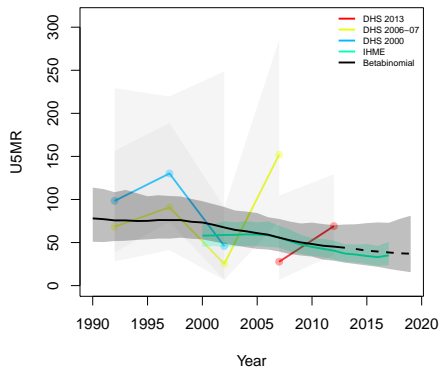
**Epupa, Kunene**



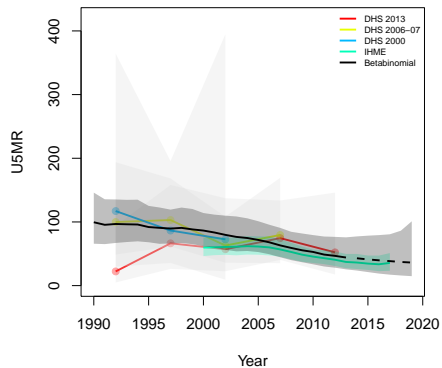
**Kamanjab, Kunene**



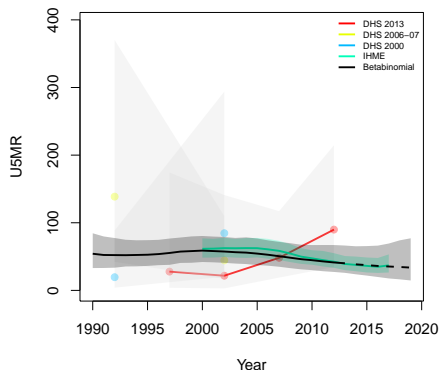
**Khorixas, Kunene**



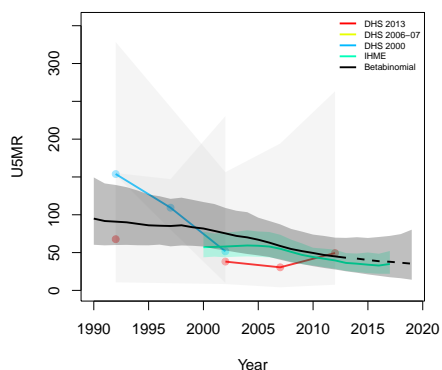
**Opuwo, Kunene**



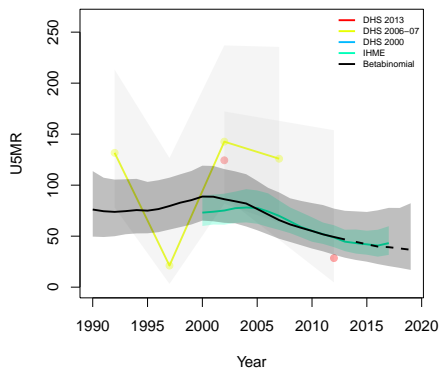
**Outjo, Kunene**



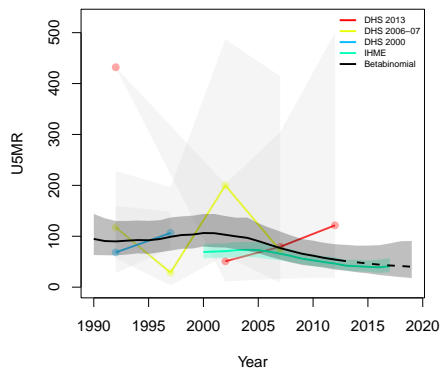
**Sesfontein, Kunene**



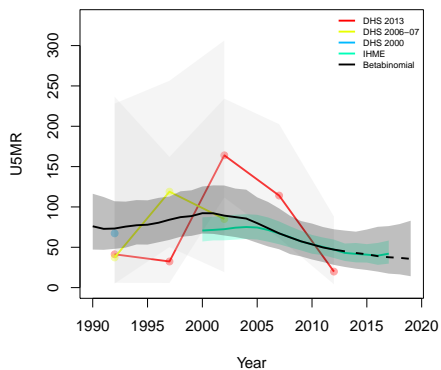
**Eenhana, Oshana**



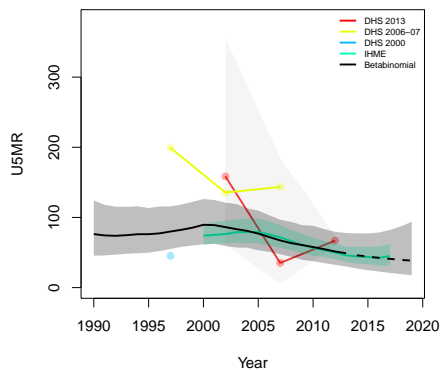
**Endola, Oshana**



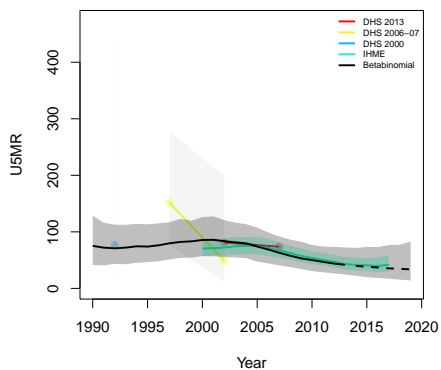
**Engela, Ohangwena**



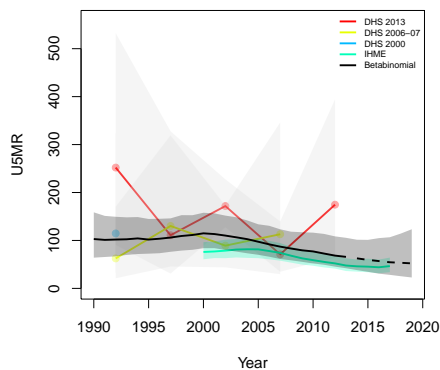
**Epembe, Ohangwena**



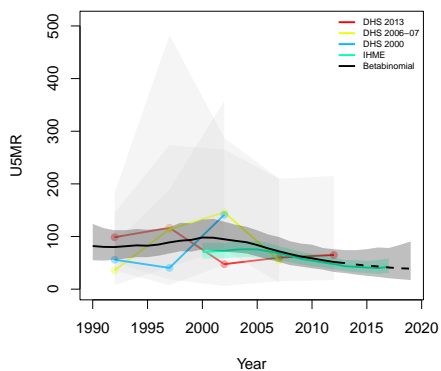
**Ohangwena, Ohangwena**



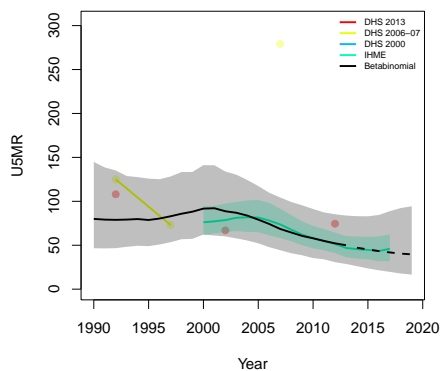
**Okongo, Ohangwena**



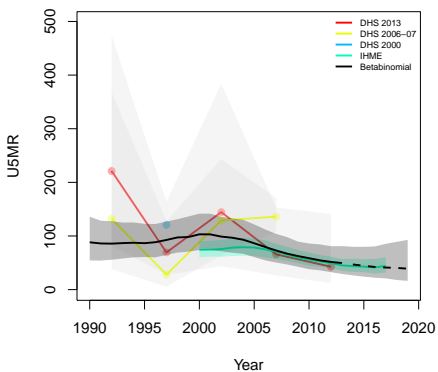
**Omulonga, Ohangwena**



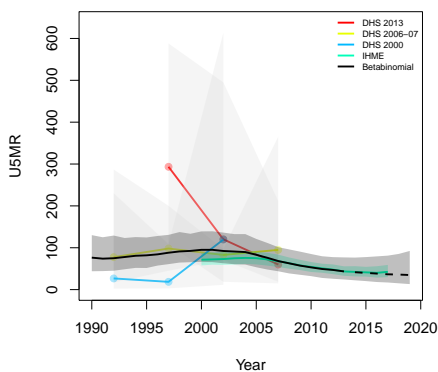
**Omundaungilo, Ohangwena**



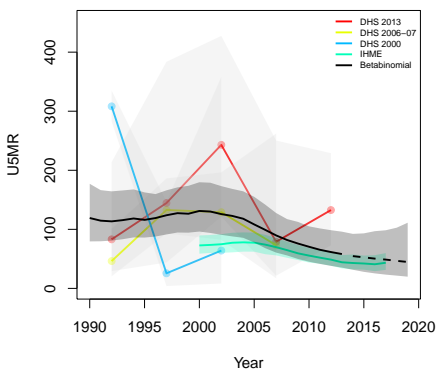
**Ondobe, Ohangwena**



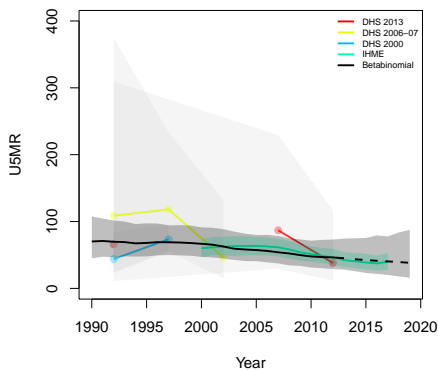
**Ongenga, Ohangwena**



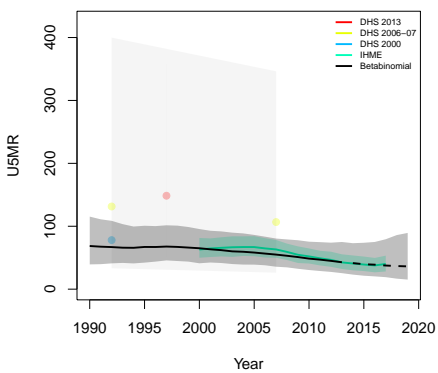
**Oshikango, Ohangwena**



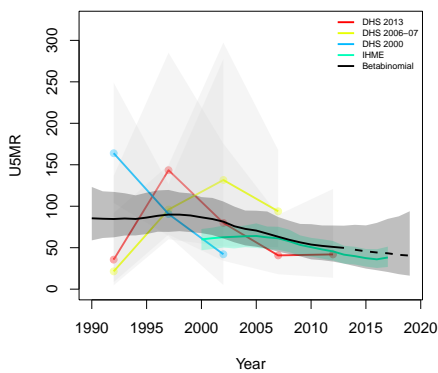
**Aminius, Omaheke**



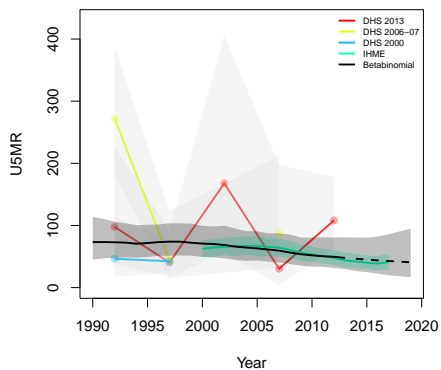
**Epukiro, Omaheke**



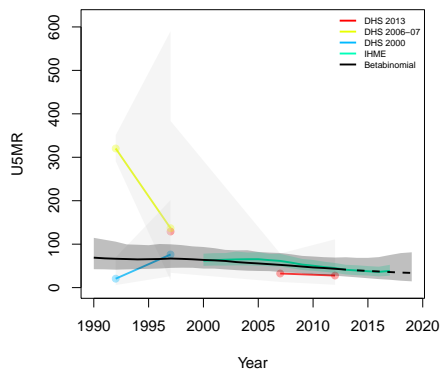
**Gobabis, Omaheke**



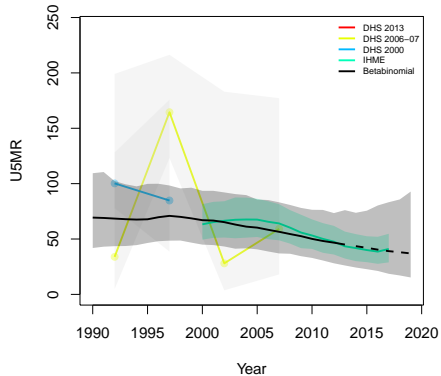
**Kalahari, Omaheke**



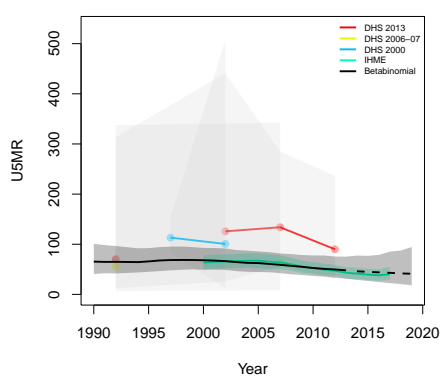
**Otjinene, Omaheke**



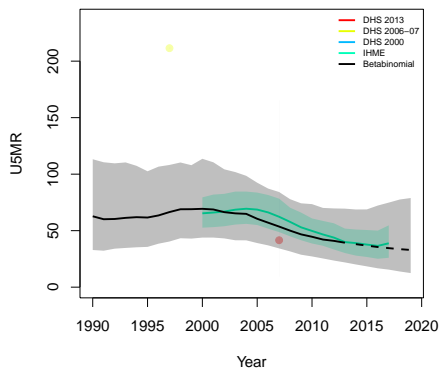
**Otjombinde, Omaheke**



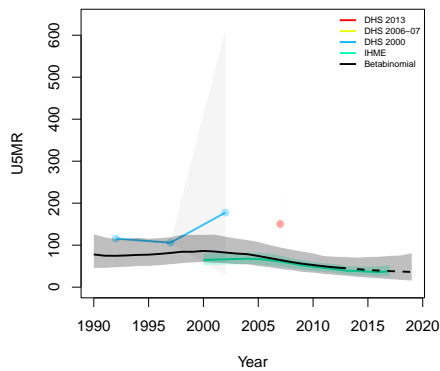
**Steinhausen, Omaheke**

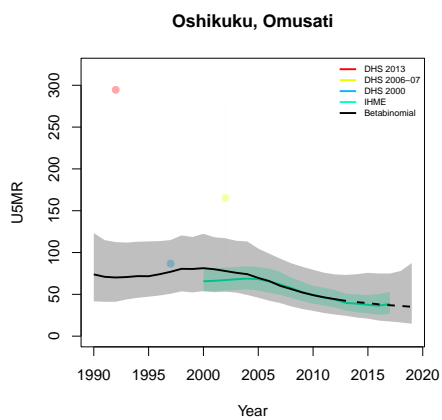
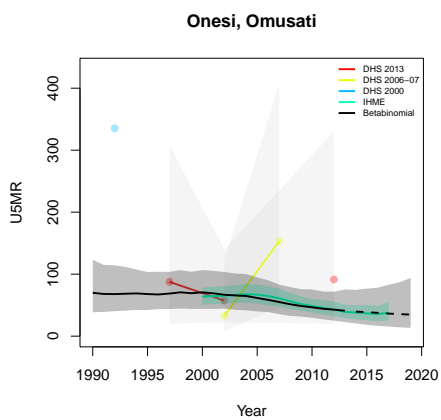
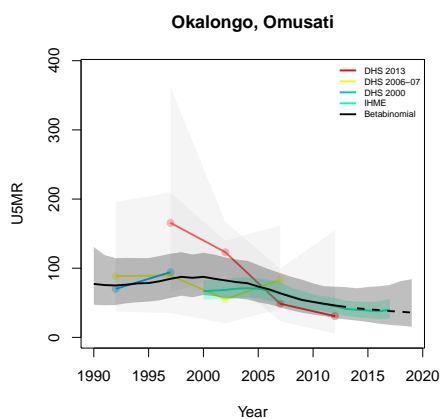
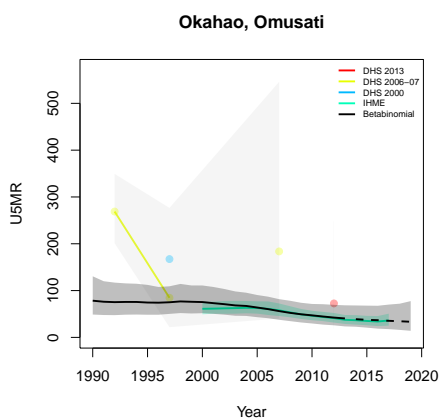
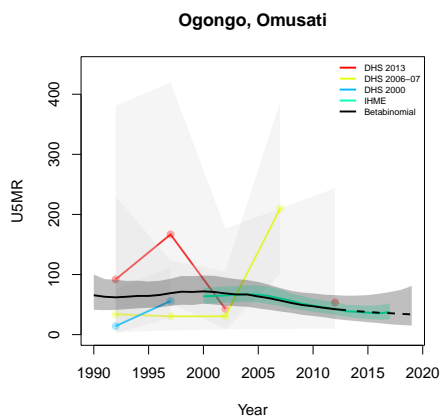
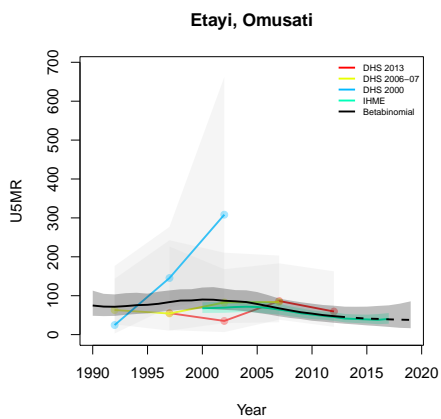


**Anamulenge, Omusati**

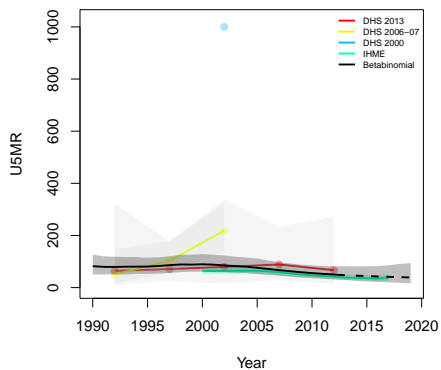


**Eiim, Omusati**

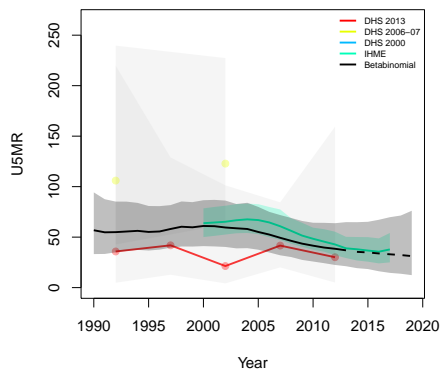




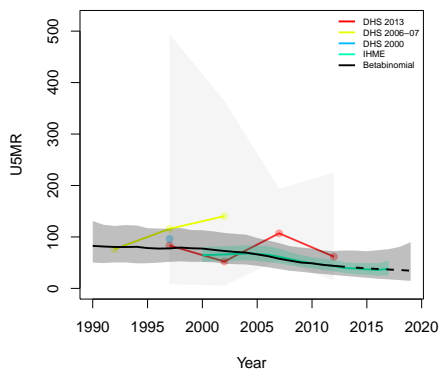
**Otamanzi, Omusati**



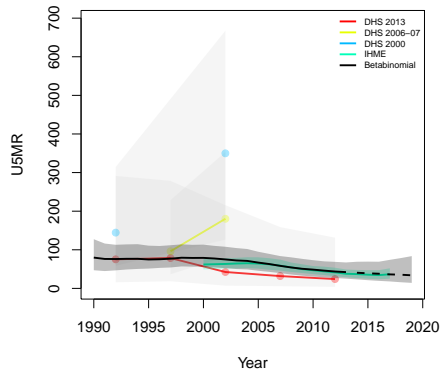
**Outapi, Omusati**



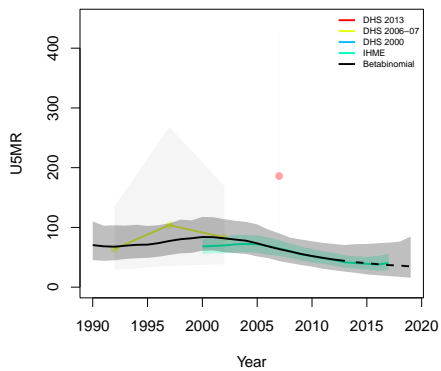
**Ruacana, Omusati**



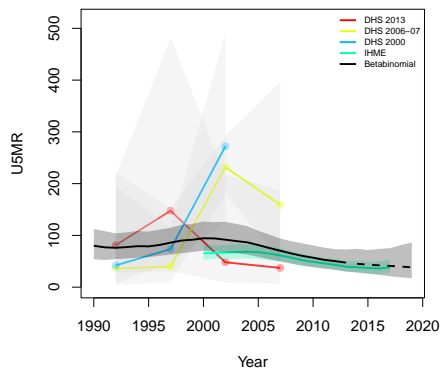
**Tsandi, Omusati**



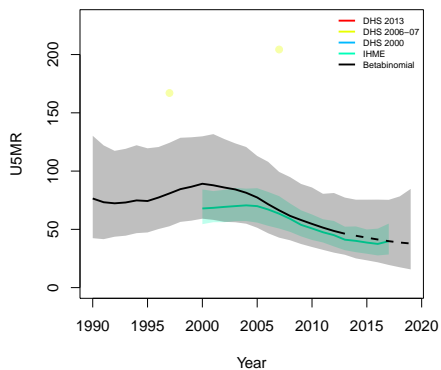
**Okaku, Oshana**



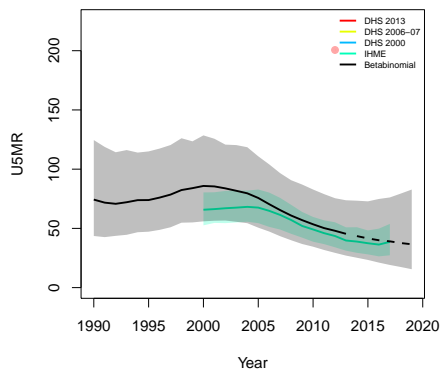
**Okatana, Oshana**



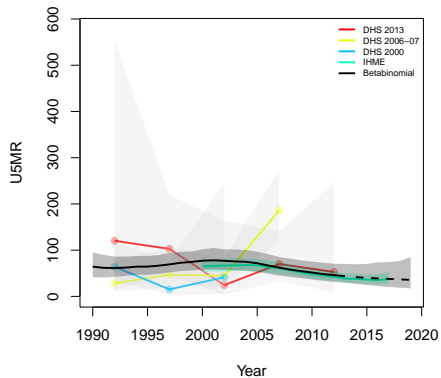
**Okatyali, Oshana**



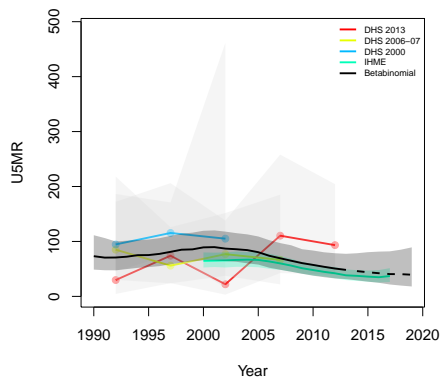
**Ompundja, Oshana**



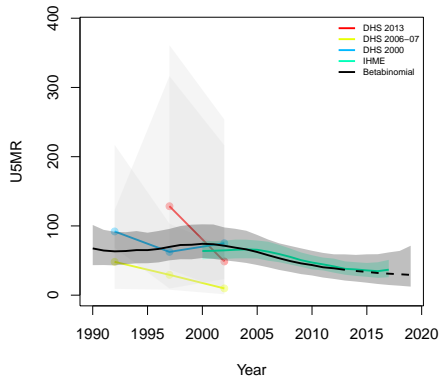
**Ondangwa, Oshana**



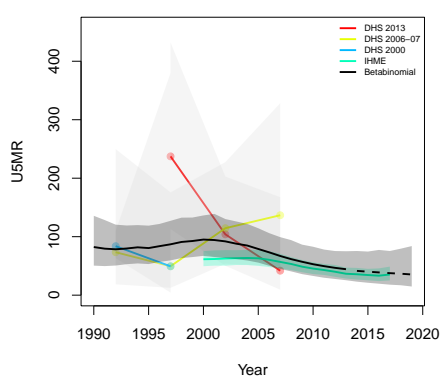
**Ongwediva, Oshana**



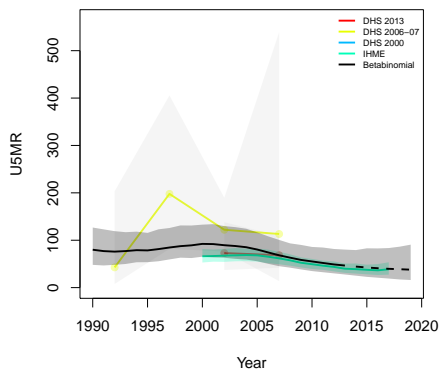
**Oshakati East, Oshana**



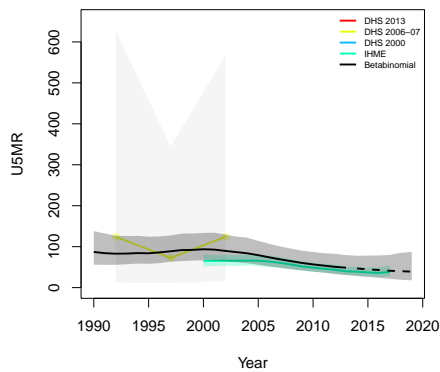
**Oshakati West, Oshana**



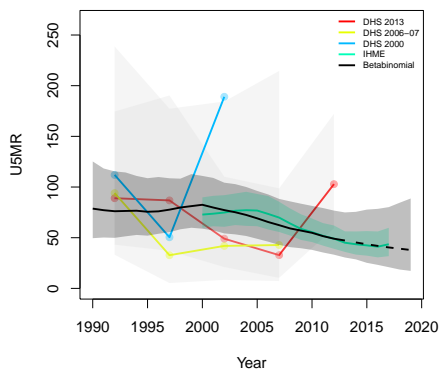
**Uukwiyu, Oshana**



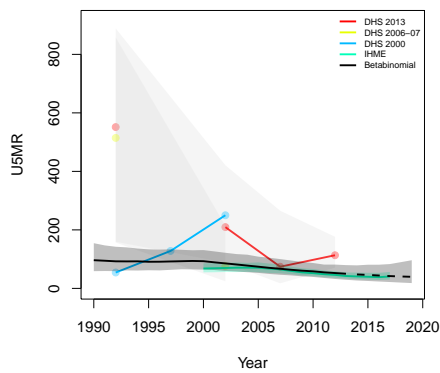
**Uuvudhiya, Oshana**



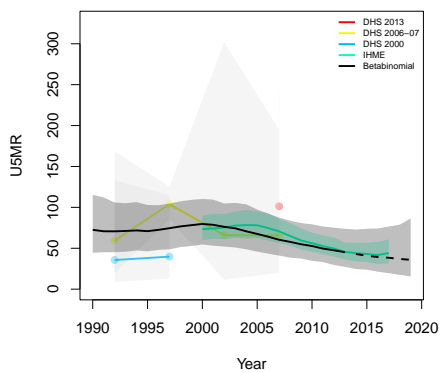
**Engodi, Oshikoto**



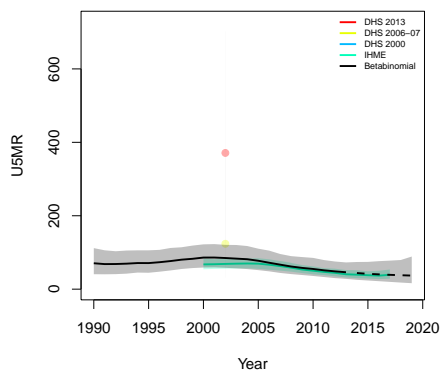
**Guinas, Oshikoto**



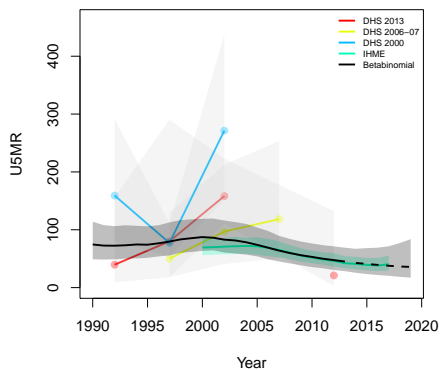
**Okankolo, Oshikoto**



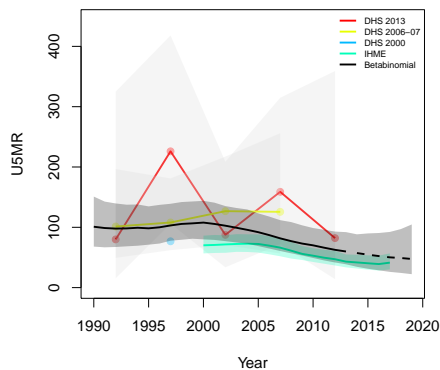
**Olukonda, Oshikoto**



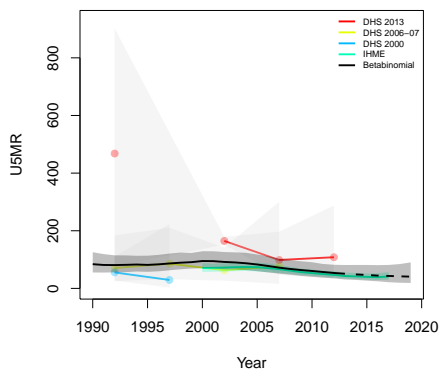
**Omuntele, Oshikoto**



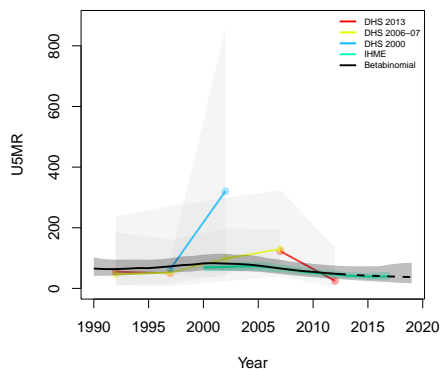
**Omuthiyagwipundi, Oshikoto**



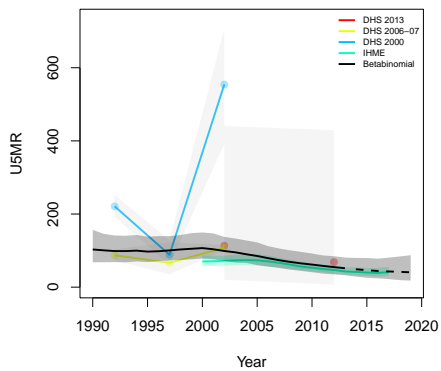
**Onayena, Oshikoto**



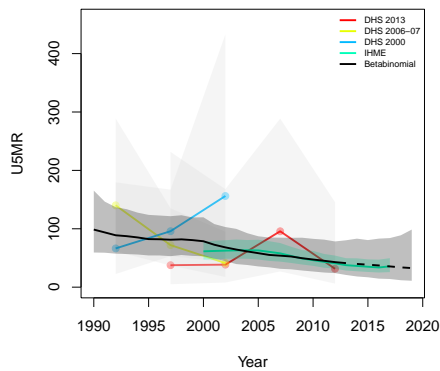
**Oniipa, Oshikoto**



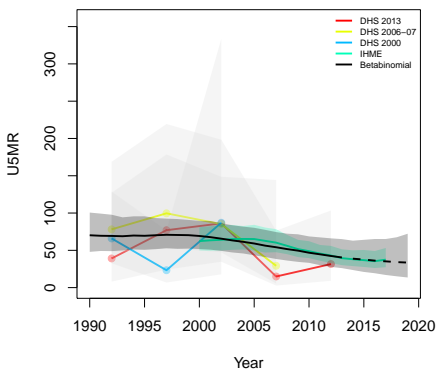
**Onyaanya, Oshikoto**



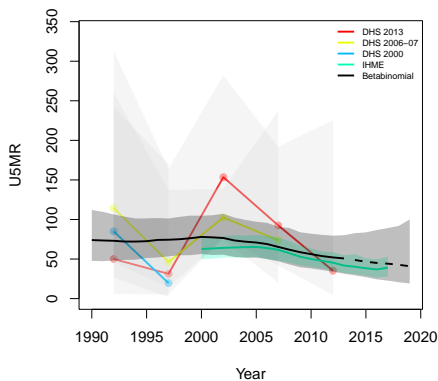
**Tsumeb, Oshikoto**



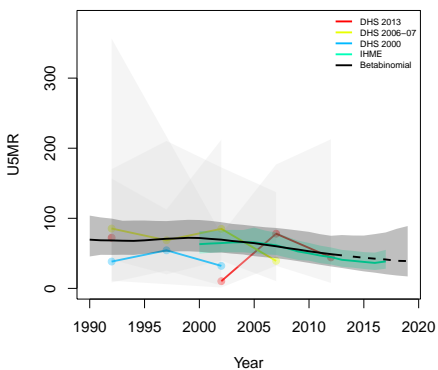
**Grootfontein, Otjozondjupa**



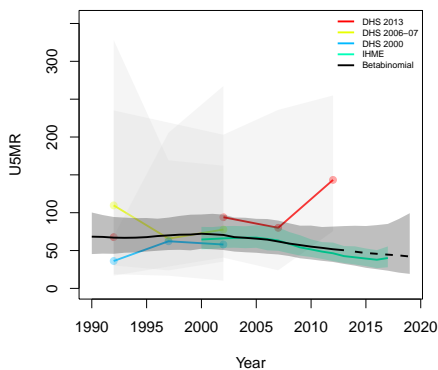
**Okahandja, Otjozondjupa**



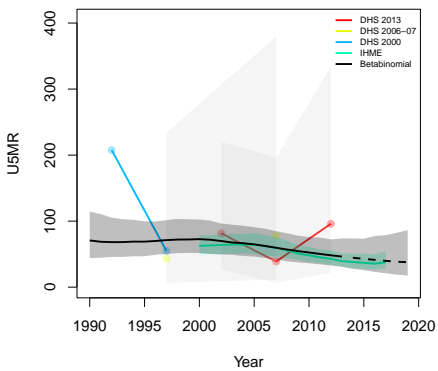
**Okakarara, Otjozondjupa**



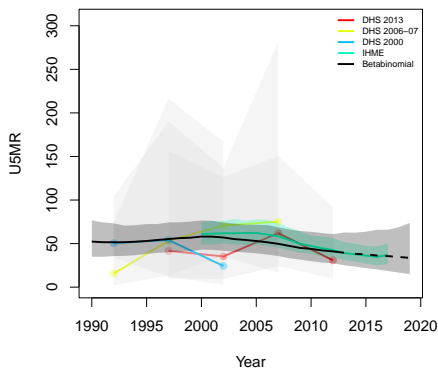
**Omatako, Otjozondjupa**



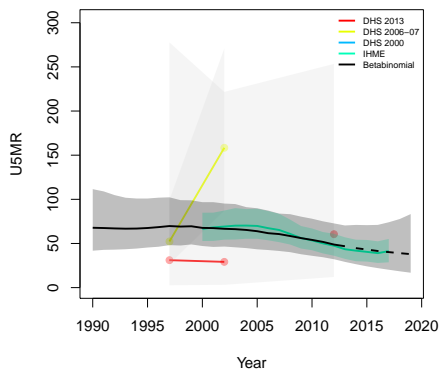
**Otavi, Otjozondjupa**



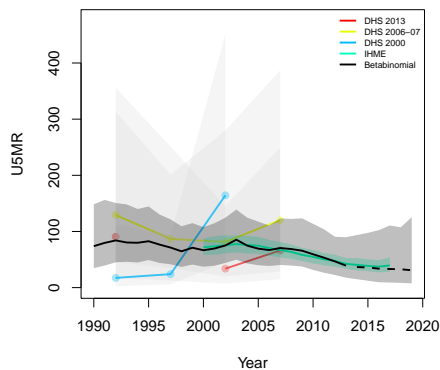
**Otjiwarongo, Otjozondjupa**



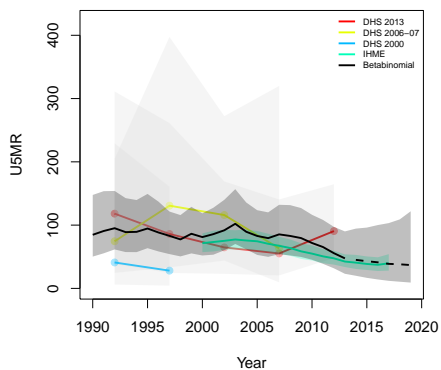
**Tsumkwe, Otjozondjupa**



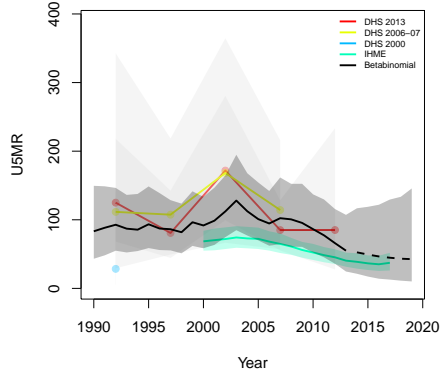
**Kabe, Zambezi**



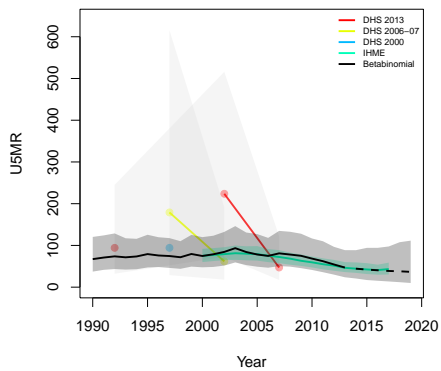
**Katima Muliro Rural, Zambezi**



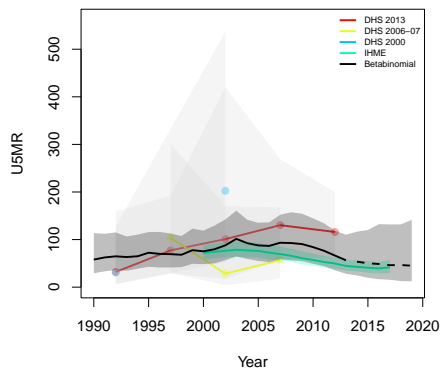
**Katima Muliro Urban, Zambezi**

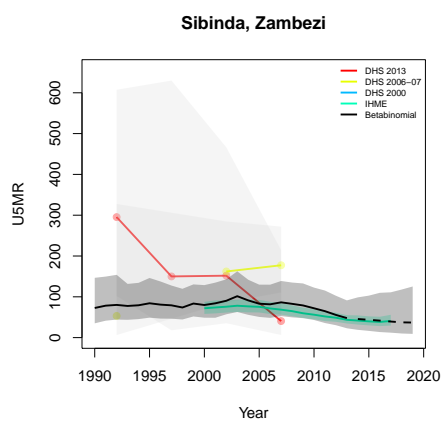


**Kongola, Zambezi**



**Linyandi, Zambezi**





## B.12 Nigeria

Age	Survey	Clusters			Deaths			Agemonths		
		Urban	Rural	Total	Urban	Rural	Total	Urban	Rural	Total
0	1990	131	166	297	15	27	42	306	716	1022
	2003	165	195	360	219	560	779	5159	9471	14630
	2008	279	607	886	864	3200	4064	22007	65081	87088
	2013	369	520	889	1282	3481	4763	35162	73177	108339
	2018	570	812	1382	1356	3088	4444	42786	81726	124512
1-11	1990	131	166	297	18	35	53	3242	7064	10306
	2003	165	195	360	221	559	780	52345	92849	145194
	2008	279	607	886	716	3088	3804	222827	645904	868731
	2013	369	520	889	1009	3375	4384	359290	731288	1090578
	2018	570	812	1382	974	2981	3955	440296	829030	1269326
12-23	1990	131	166	297	6	22	28	2859	6277	9136
	2003	165	195	360	160	439	599	52473	90645	143118
	2008	279	607	886	509	2435	2944	221772	632164	853936
	2013	369	520	889	701	2637	3338	361724	719011	1080735
	2018	570	812	1382	660	2134	2794	446860	825360	1272220
24-35	1990	131	166	297	5	27	32	2922	6285	9207
	2003	165	195	360	123	297	420	49277	84683	133960
	2008	279	607	886	374	2024	2398	208243	590384	798627
	2013	369	520	889	448	1943	2391	338879	664145	1003024
	2018	570	812	1382	547	2040	2587	417741	762377	1180118
36-47	1990	131	166	297	3	23	26	3053	5946	8999
	2003	165	195	360	65	158	223	46045	78387	124432
	2008	279	607	886	175	956	1131	193367	540737	734104
	2013	369	520	889	246	964	1210	316619	610921	927540
	2018	570	812	1382	258	1036	1294	388169	699903	1088072
48-59	1990	131	166	297	2	6	8	2829	5498	8327
	2003	165	195	360	25	86	111	43967	73118	117085
	2008	279	607	886	115	429	544	180439	501884	682323
	2013	369	520	889	131	456	587	295930	566309	862239

*B.12.1 Admin-1*

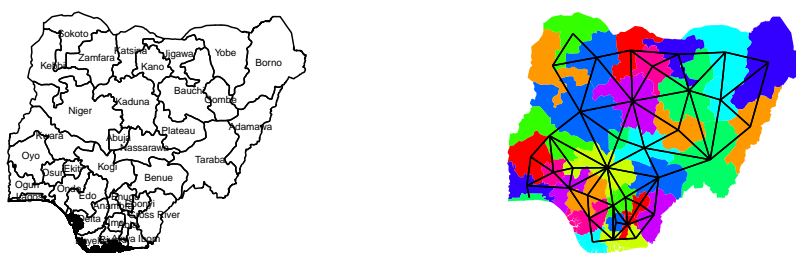


Figure B.58: **Left:** The names of the 37 Admin-1 areas of Nigeria . **Right:** The neighborhood structure of Admin-1 areas in Nigeria .

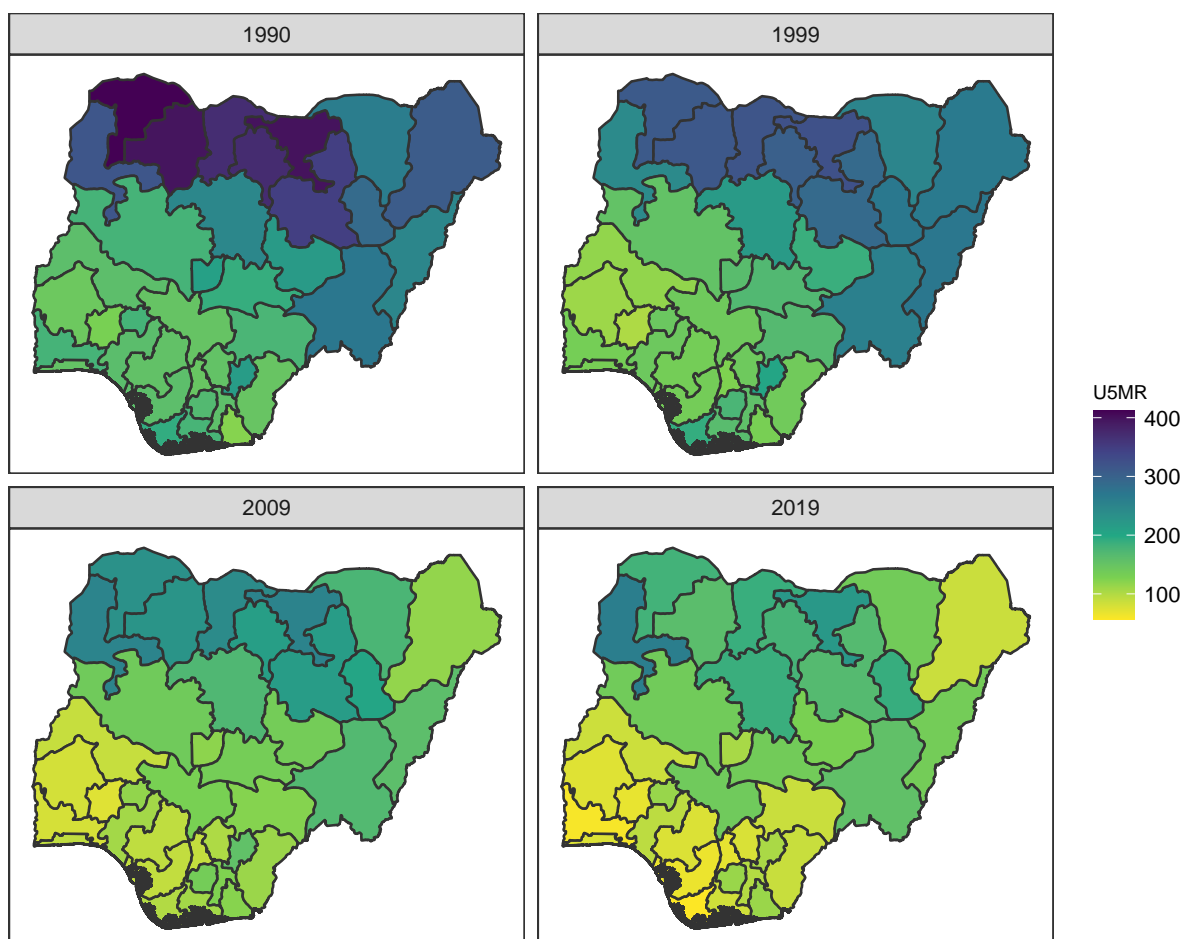


Figure B.59: Median U5MR estimates for years 1990, 1999, 2009, 2019 for Admin-1 areas in Nigeria .

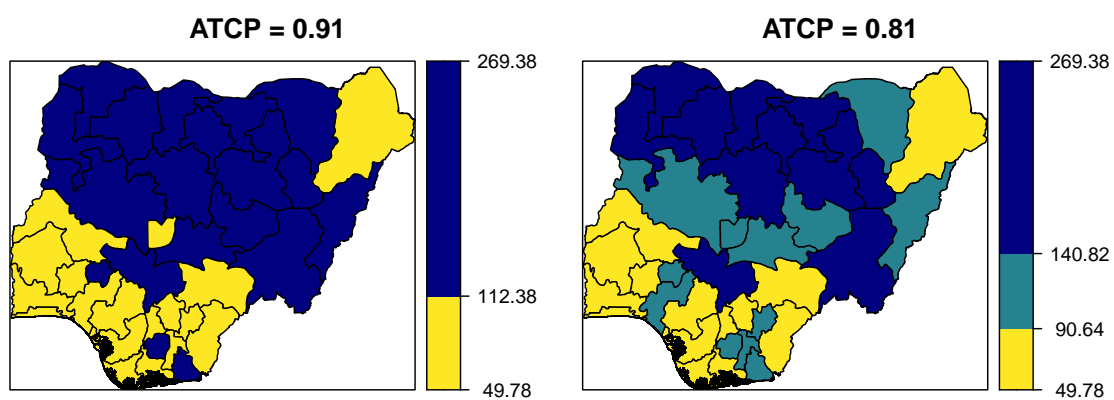
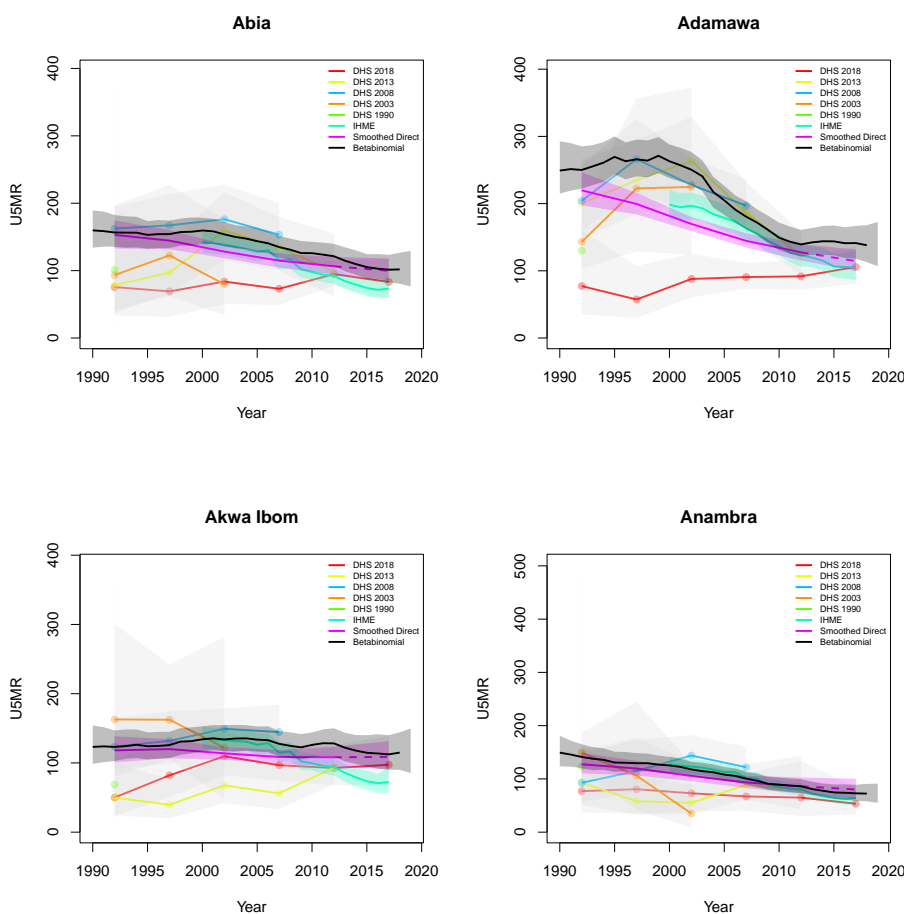
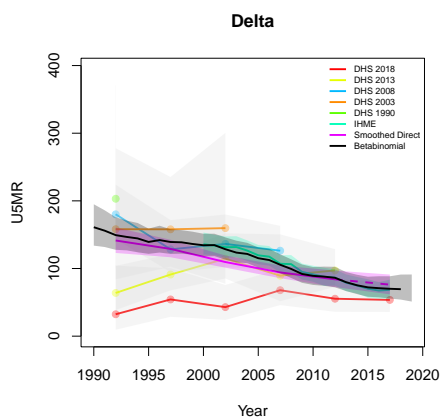
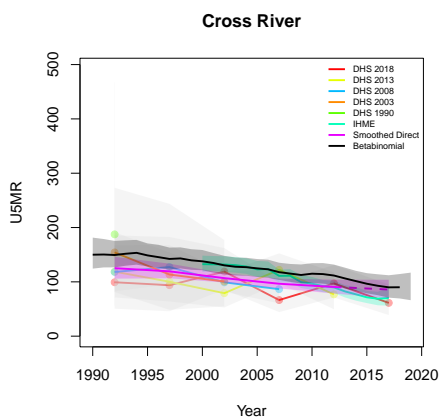
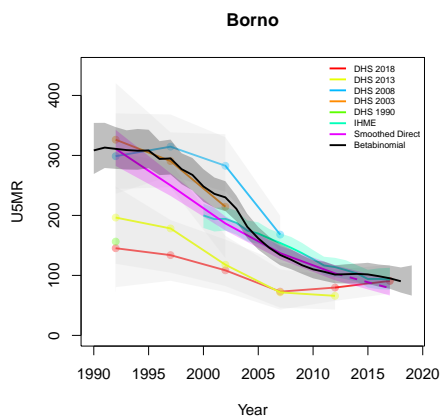
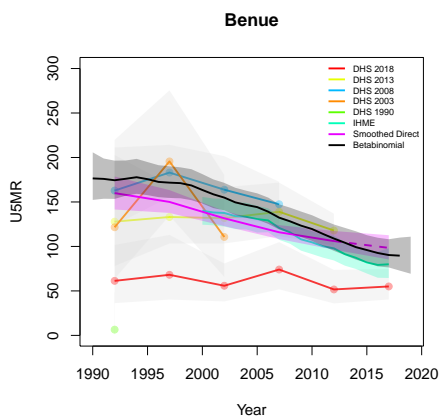
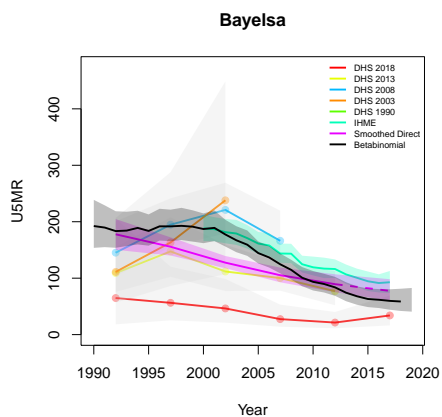
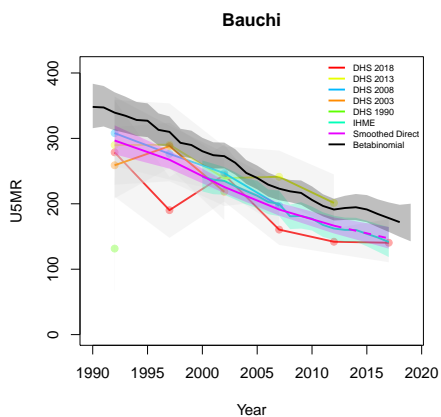


Figure B.60: Expression of uncertainty of U5MR (deaths per 1000 children) estimates for Admin-1 areas based on the average true classification probability (ATCP) in 2019 using  $K = 2, 3$  colors.

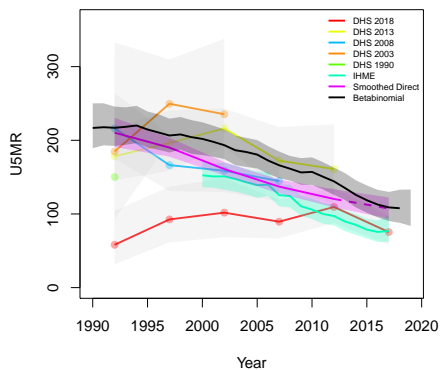
*Data and estimates over time by area*

Colored lines with circular points and light grey uncertainty bands are 5-year survey-weighted estimates of U5MR for years 1990–1994 up to 2015–2019 depending on survey timing. For a survey that ends in the middle of a 5-year period, we plot the estimates at the mid-point of the years in that interval for which the survey provides data. Black lines and corresponding intervals represent posterior medians and 95% uncertainty intervals respectively for the betabinomial model. IHME’s estimates and corresponding intervals, where we can compare, are in aquamarine.

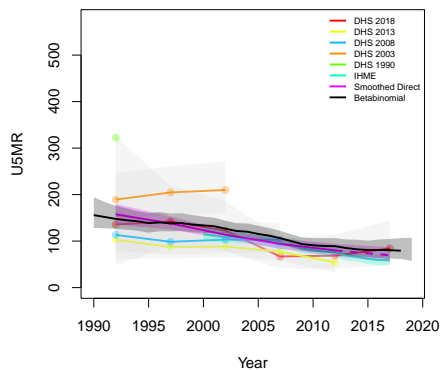




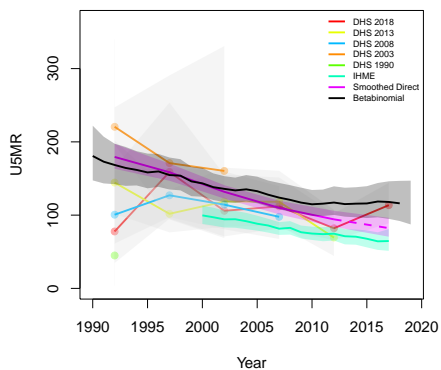
**Ebonyi**



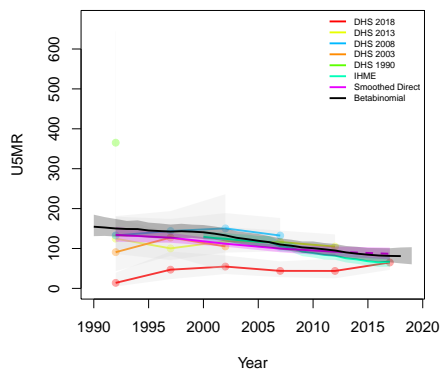
**Edo**



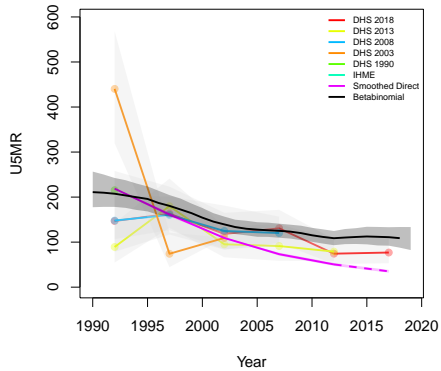
**Ekiti**



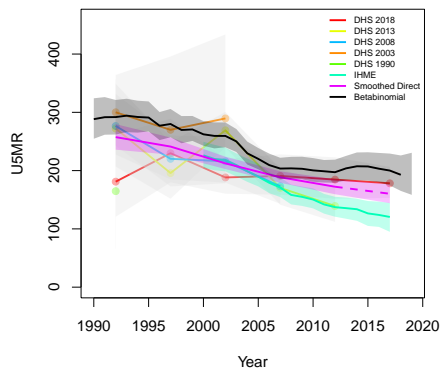
**Enugu**

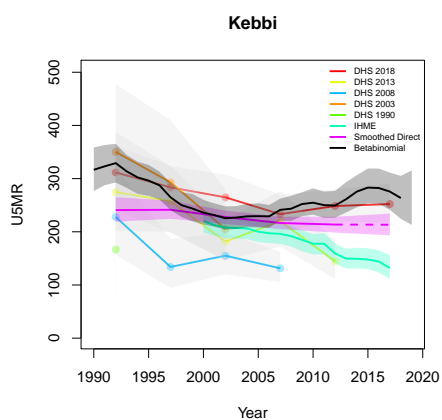
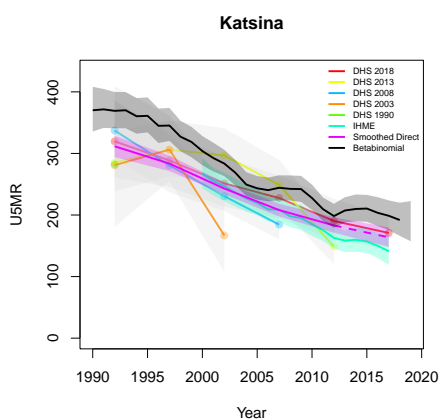
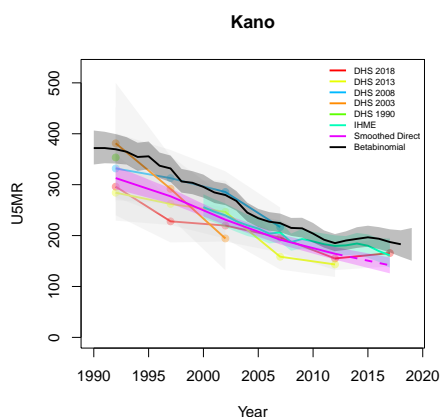
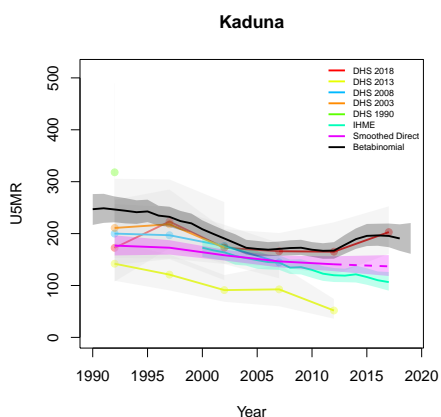
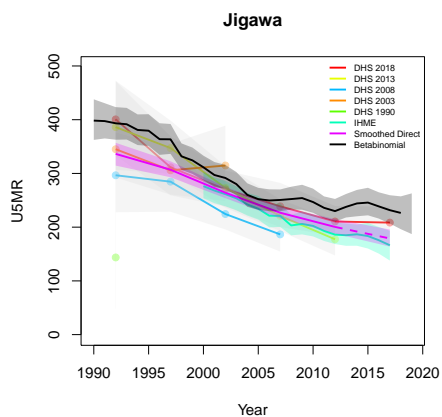
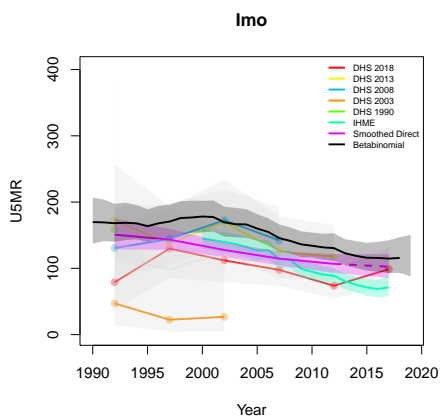


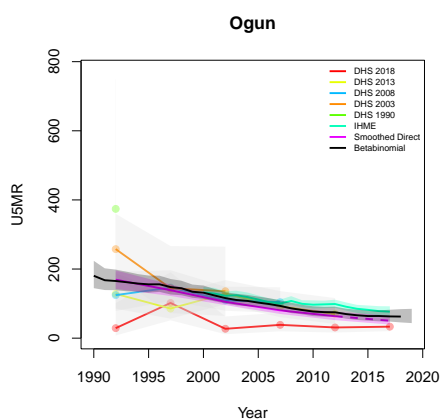
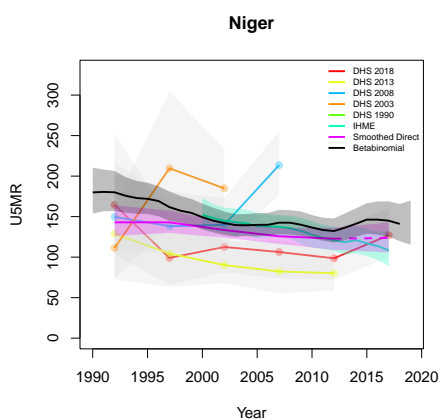
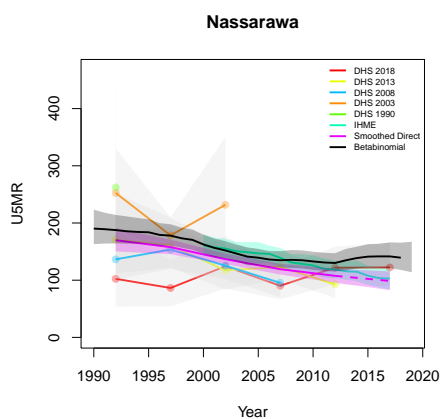
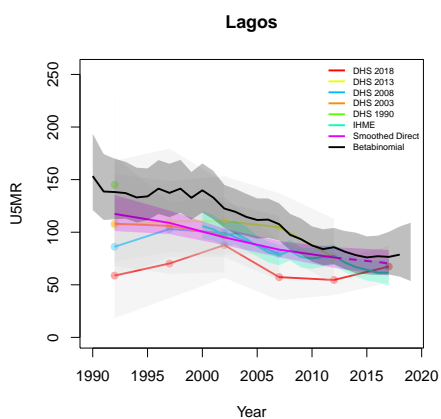
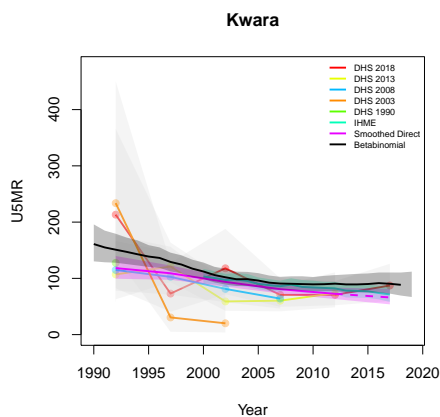
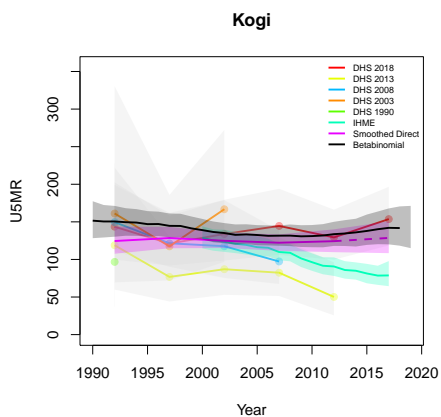
**Abuja**

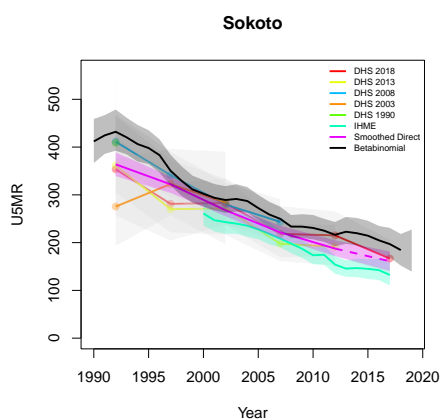
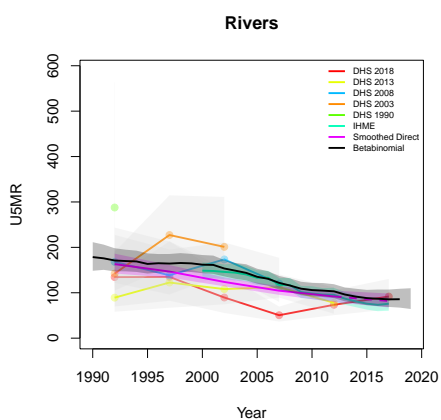
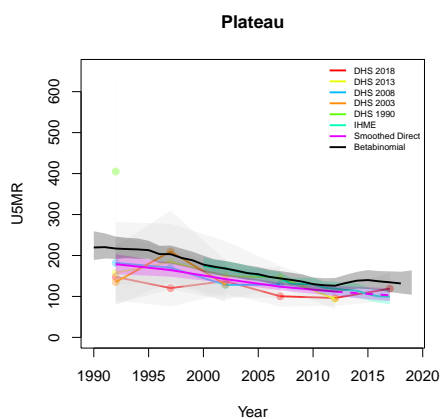
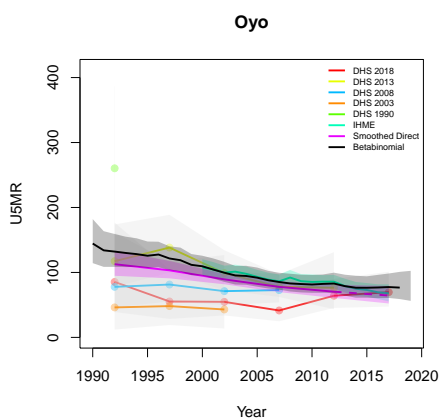
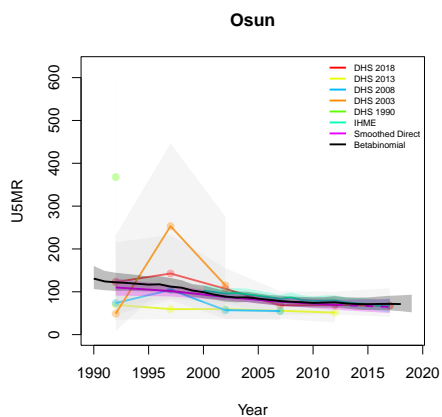
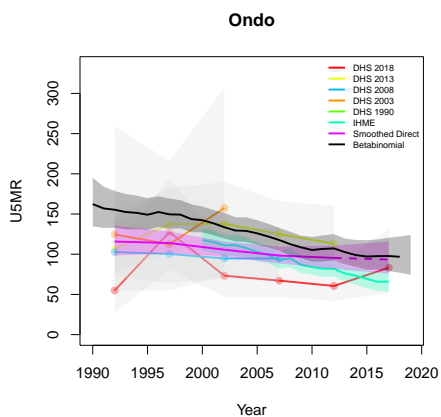


**Gombe**

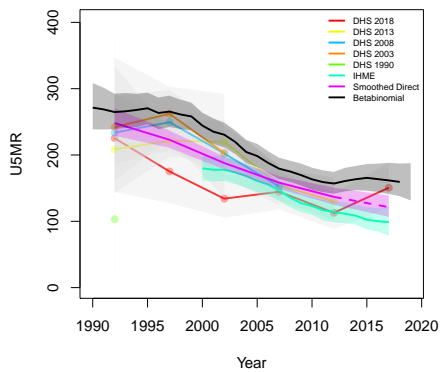




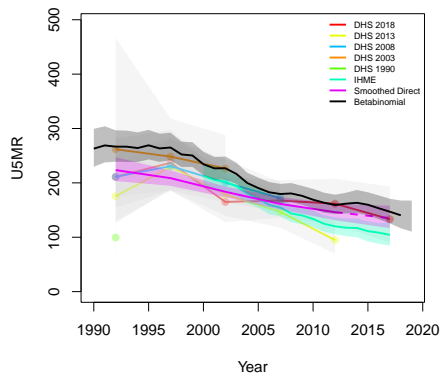




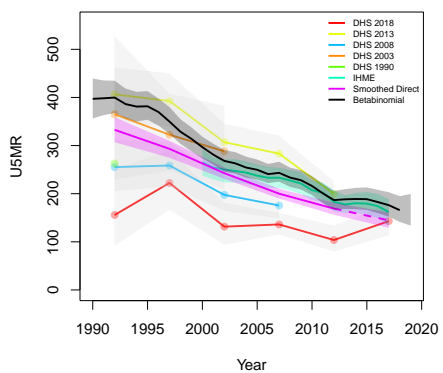
**Taraba**



**Yobe**



**Zamfara**



**B.13 Pakistan**

Age	Survey	Clusters			Deaths			Agemonths		
		Urban	Rural	Total	Urban	Rural	Total	Urban	Rural	Total
0	2006	384	571	955	453	1066	1519	10584	19057	29641
	2017	282	278	560	740	1134	1874	22648	26946	49594
1-11	2006	384	571	955	219	547	766	108964	191113	300077
	2017	282	278	560	416	655	1071	232509	272215	504724
12-23	2006	384	571	955	64	148	212	115765	199812	315577
	2017	282	278	560	78	155	233	239698	278683	518381
24-35	2006	384	571	955	31	84	115	113880	194165	308045
	2017	282	278	560	46	88	134	227670	262804	490474
36-47	2006	384	571	955	24	63	87	110907	185437	296344
	2017	282	278	560	32	63	95	215105	246491	461596
48-59	2006	384	571	955	11	28	39	108046	177436	285482
	2017	282	278	560	19	29	48	202205	230176	432381

Table B.13: **Data summary for Pakistan.** Total numbers of clusters (Columns 3–5) with observations in each age group by survey in urban and rural areas and combined. Numbers of deaths (Columns 6–8) and number of agemonths (Columns 9–10) observed in each age group by survey in urban and rural areas and combined.

**B.13.1 Admin-1**



Figure B.61: **Left:** The names of the 8 Admin-1 areas of Pakistan . **Right:** The neighborhood structure of Admin-1 areas in Pakistan .

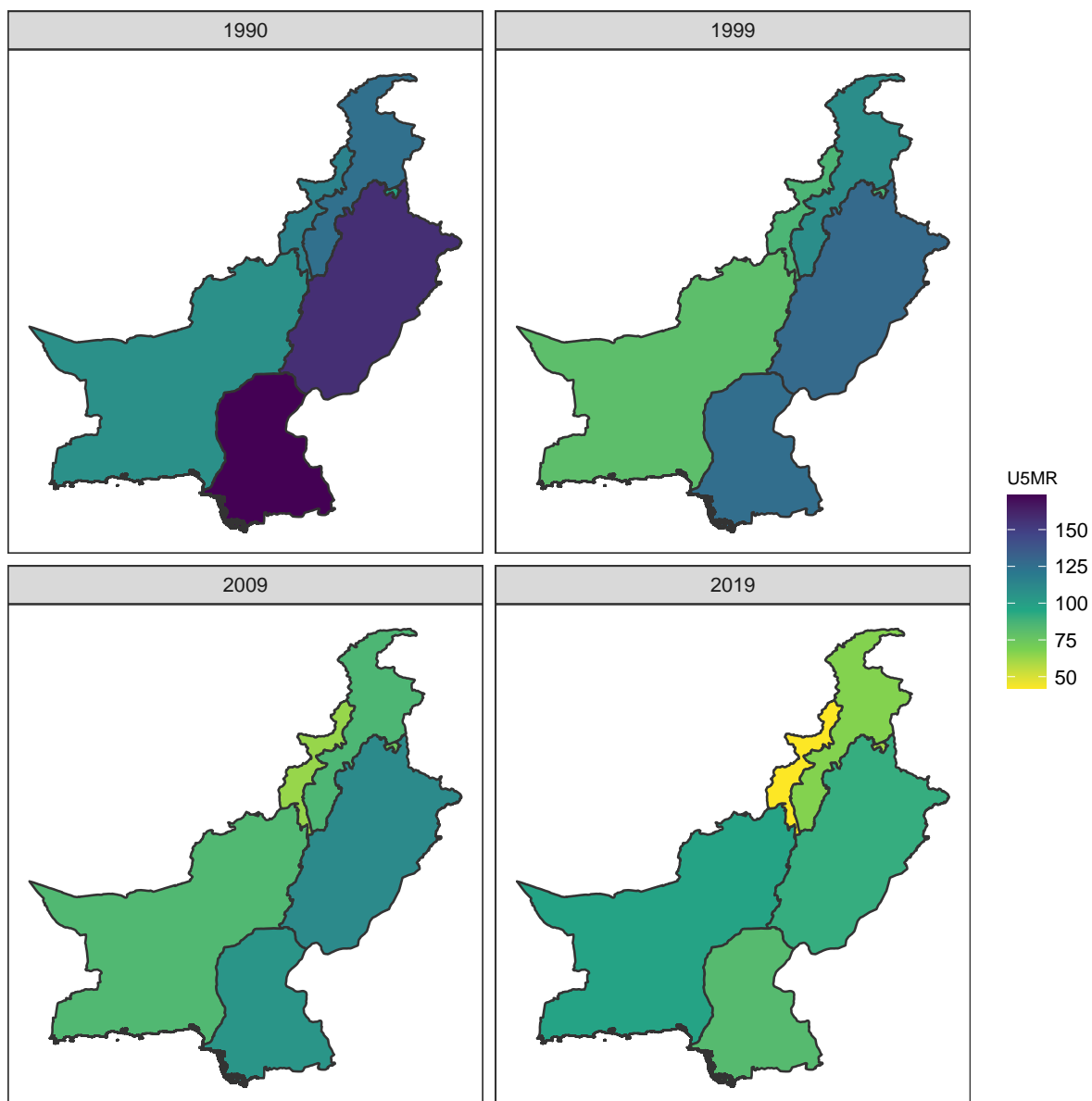


Figure B.62: Median U5MR estimates for years 1990, 1999, 2009, 2019 for Admin-1 areas in Pakistan .

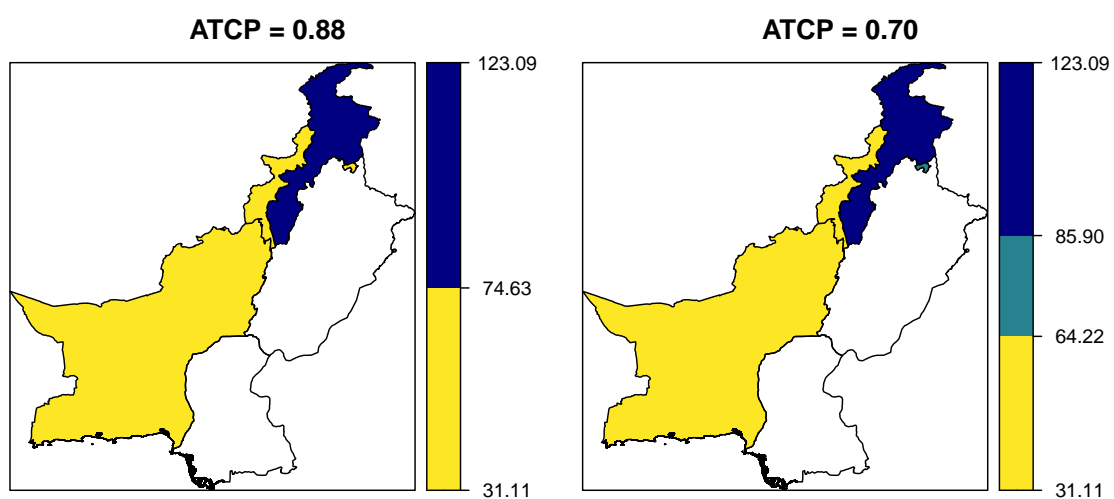
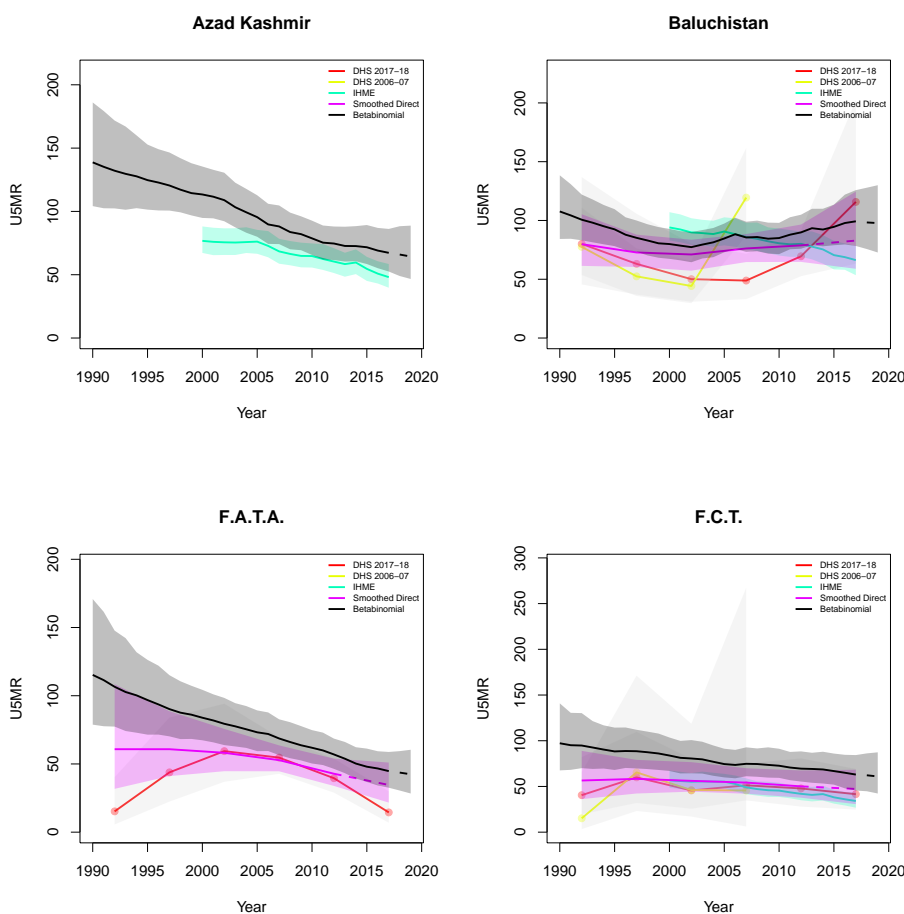


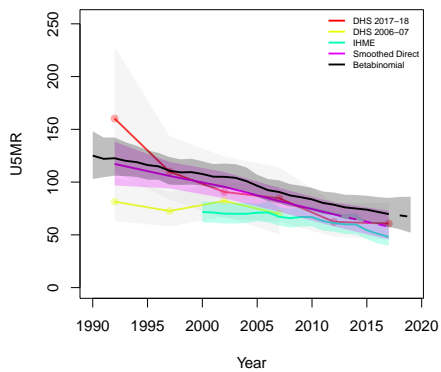
Figure B.63: Expression of uncertainty of U5MR (deaths per 1000 children) estimates for Admin-1 areas based on the average true classification probability (ATCP) in 2019 using  $K = 2, 3$  colors.

*Data and estimates over time by area*

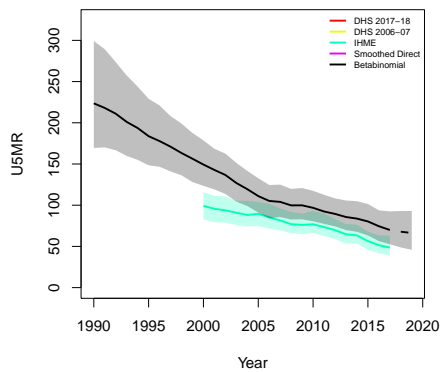
Colored lines with circular points and light grey uncertainty bands are 5-year survey-weighted estimates of U5MR for years 1990–1994 up to 2015–2019 depending on survey timing. For a survey that ends in the middle of a 5-year period, we plot the estimates at the mid-point of the years in that interval for which the survey provides data. Black lines and corresponding intervals represent posterior medians and 95% uncertainty intervals respectively for the betabinomial model. IHME’s estimates and corresponding intervals, where we can compare, are in aquamarine.



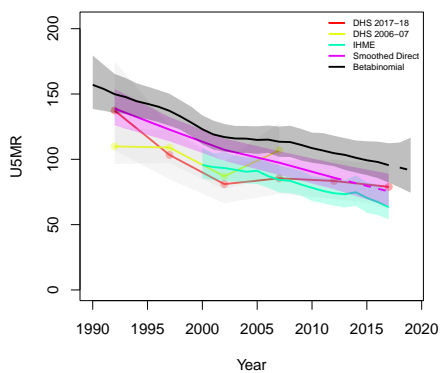
**N.W.F.P.**



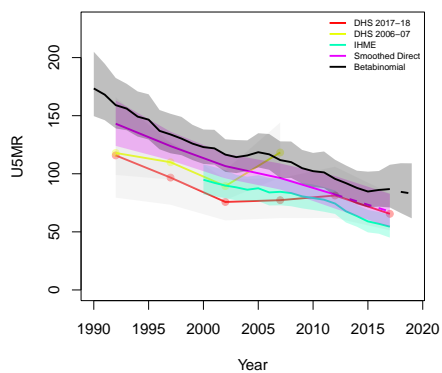
**Northern Areas**



**Punjab**



**Sind**



**B.14 Rwanda**

Age	Survey	Clusters			Deaths			Agemonths		
		Urban	Rural	Total	Urban	Rural	Total	Urban	Rural	Total
0	2005	108	348	456	129	912	1041	4235	18296	22531
	2008	61	185	246	70	437	507	3408	12057	15465
	2010	79	413	492	106	949	1055	3864	25635	29499
	2015	113	379	492	156	740	896	5911	23171	29082
1-11	2005	108	348	456	170	1075	1245	43108	180641	223749
	2008	61	185	246	99	593	692	35206	121922	157128
	2010	79	413	492	154	1063	1217	39589	260719	300308
	2015	113	379	492	145	824	969	60594	236058	296652
12-23	2005	108	348	456	104	544	648	43049	178774	221823
	2008	61	185	246	66	247	313	35270	121517	156787
	2010	79	413	492	56	583	639	39970	262901	302871
	2015	113	379	492	65	412	477	60915	236826	297741
24-35	2005	108	348	456	51	350	401	40045	166771	206816
	2008	61	185	246	35	176	211	32967	113377	146344
	2010	79	413	492	22	355	377	37489	244888	282377
	2015	113	379	492	36	271	307	56555	220885	277440
36-47	2005	108	348	456	31	268	299	36896	154765	191661
	2008	61	185	246	30	138	168	30563	105454	136017
	2010	79	413	492	20	255	275	34759	227302	262061
	2015	113	379	492	41	200	241	52111	204503	256614
48-59	2005	108	348	456	22	188	210	34621	145821	180442
	2008	61	185	246	19	83	102	28179	97477	125656
	2010	79	413	492	14	187	201	32023	209414	241437
	2015	113	379	492	17	124	141	48408	190260	238668

Table B.14: **Data summary for Rwanda.** Total numbers of clusters (Columns 3–5) with observations in each age group by survey in urban and rural areas and combined. Numbers of deaths (Columns 6–8) and number of agemonths (Columns 9–10) observed in each age group by survey in urban and rural areas and combined.

*B.14.1 Admin-1*

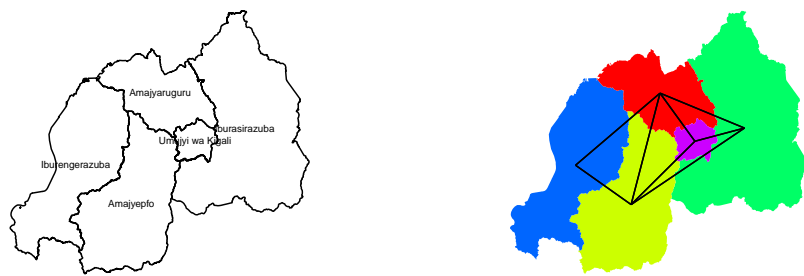


Figure B.64: **Left:** The names of the 5 Admin-1 areas of Rwanda . **Right:** The neighborhood structure of Admin-1 areas in Rwanda .

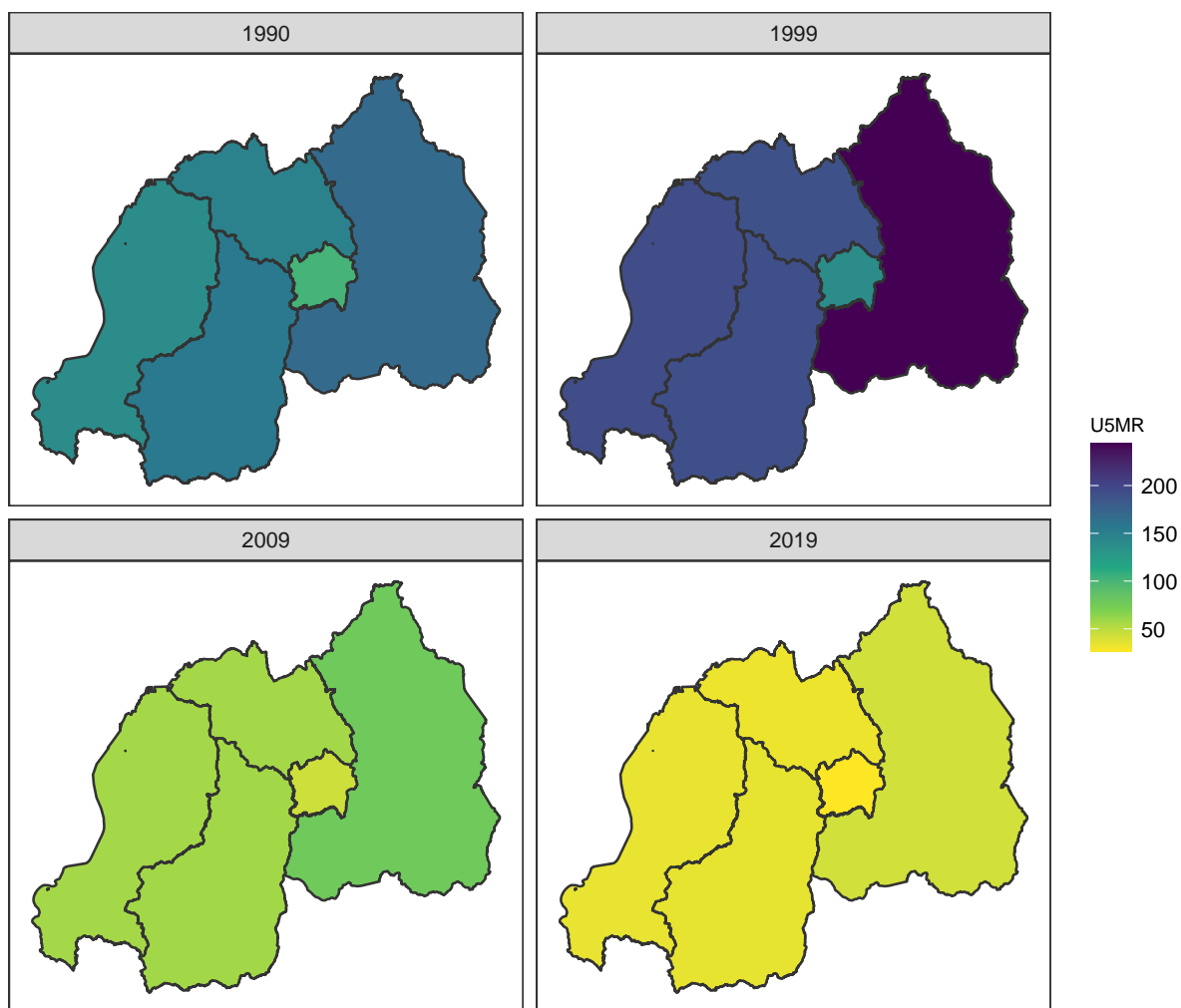


Figure B.65: Median U5MR estimates for years 1990, 1999, 2009, 2019 for Admin-1 areas in Rwanda .

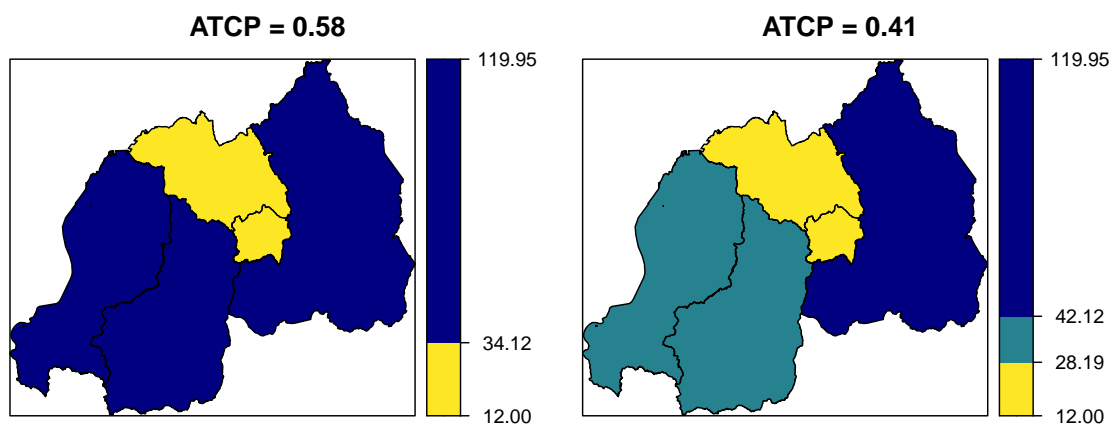
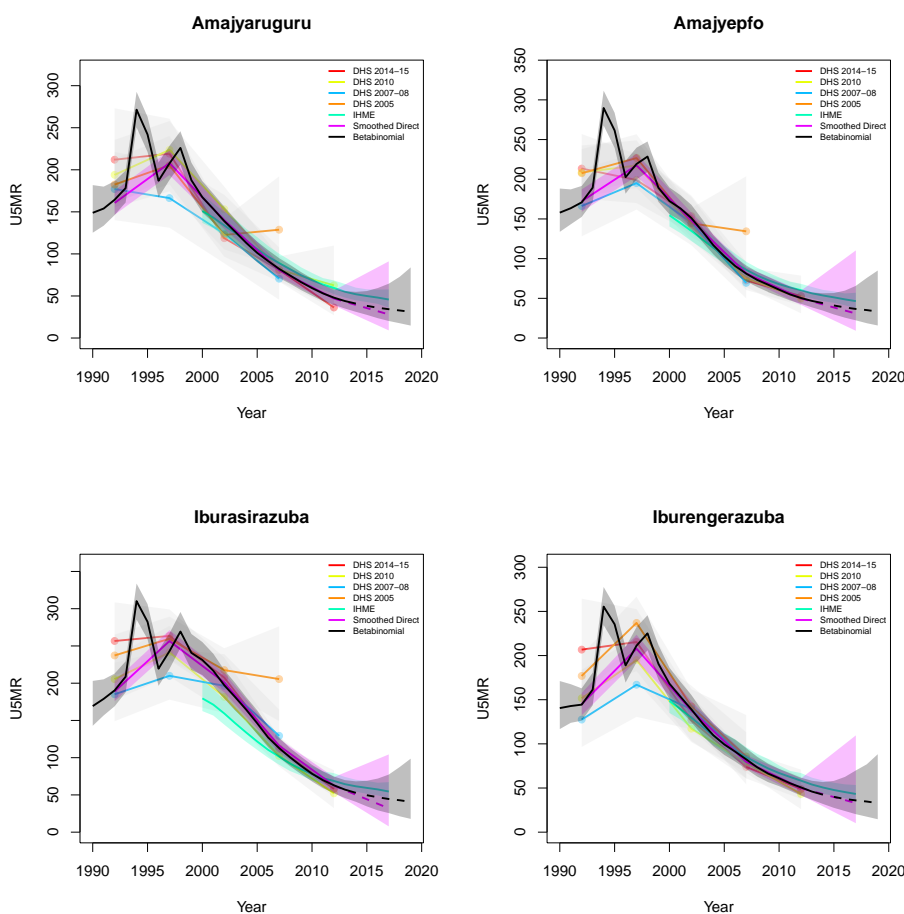
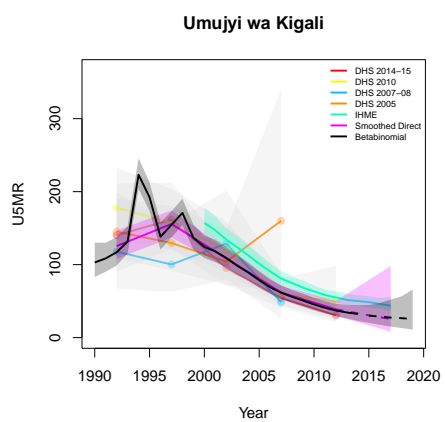


Figure B.66: Expression of uncertainty of U5MR (deaths per 1000 children) estimates for Admin-1 areas based on the average true classification probability (ATCP) in 2019 using  $K = 2, 3$  colors.

*Data and estimates over time by area*

Colored lines with circular points and light grey uncertainty bands are 5-year survey-weighted estimates of U5MR for years 1990–1994 up to 2015–2019 depending on survey timing. For a survey that ends in the middle of a 5-year period, we plot the estimates at the mid-point of the years in that interval for which the survey provides data. Black lines and corresponding intervals represent posterior medians and 95% uncertainty intervals respectively for the betabinomial model. IHME’s estimates and corresponding intervals, where we can compare, are in aquamarine.





*B.14.2 Admin-2*

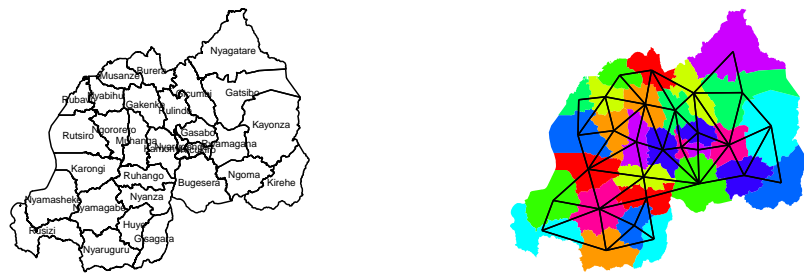


Figure B.67: **Left:** The names of the 30 Admin-2 areas of Rwanda . **Right:** The neighborhood structure of Admin-2 areas in Rwanda .

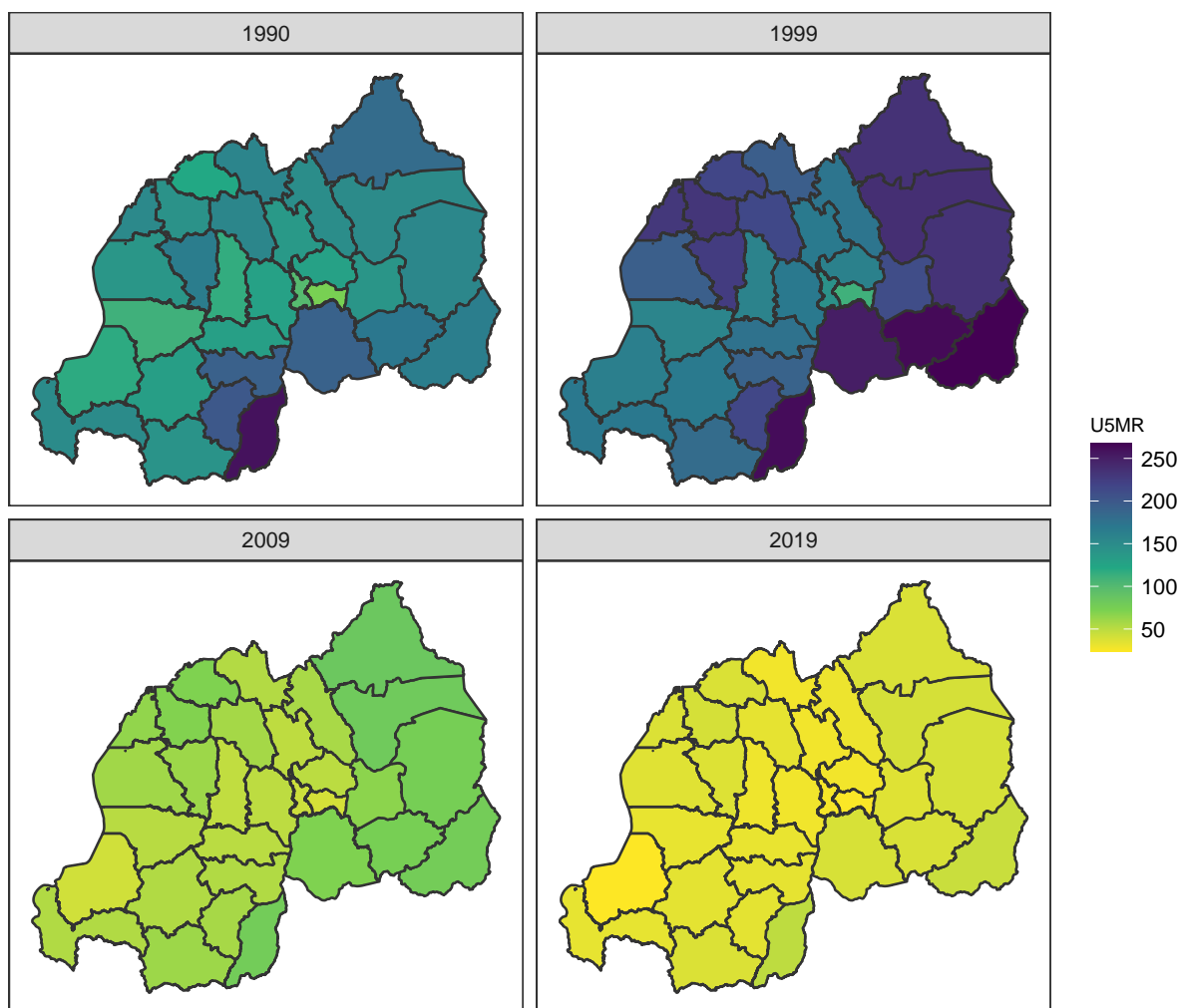


Figure B.68: Median U5MR estimates for years 1990, 1999, 2009, 2019 for Admin-2 areas in Rwanda .

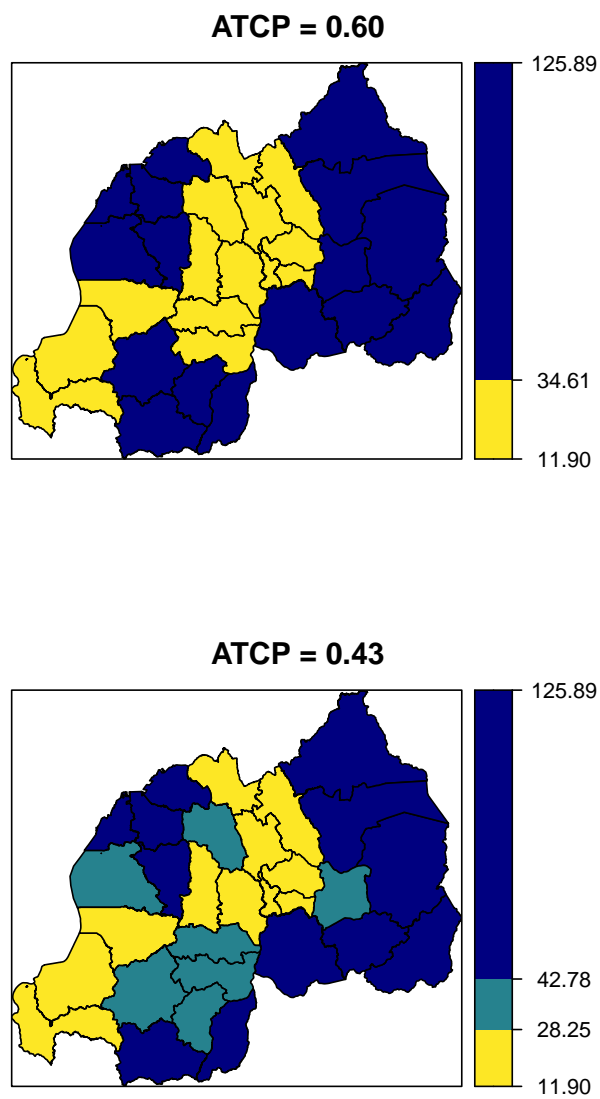
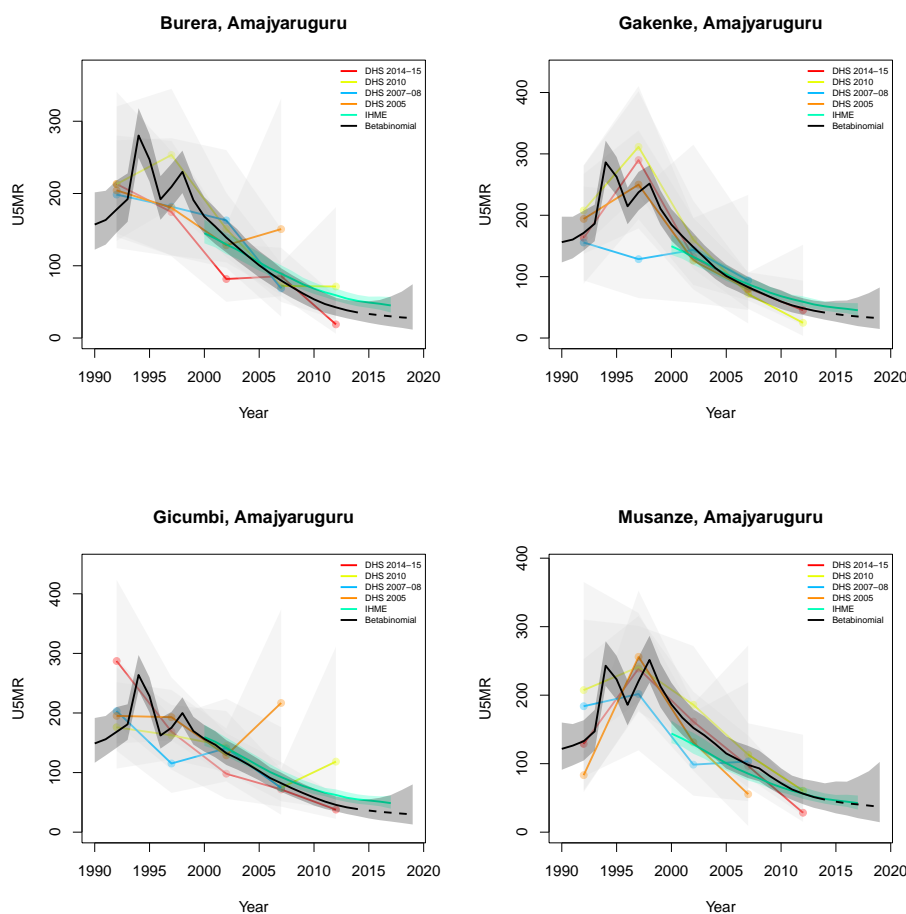


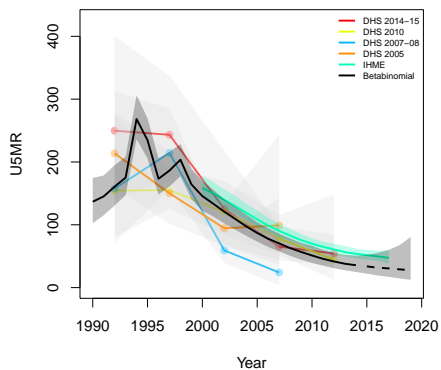
Figure B.69: Expression of uncertainty of U5MR (deaths per 1000 children) estimates for Admin-1 areas based on the average true classification probability (ATCP) in 2019 using  $K = 2, 3$  colors.

*Data and estimates over time by area*

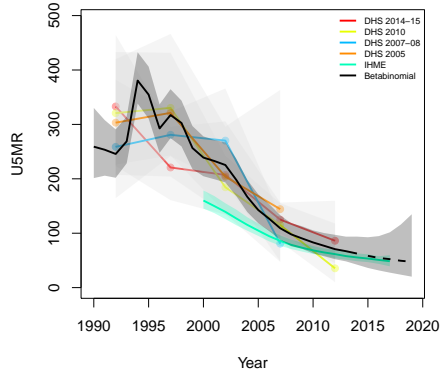
Colored lines with circular points and light grey uncertainty bands are 5-year survey-weighted estimates of U5MR for years 1990–1994 up to 2015–2019 depending on survey timing. For a survey that ends in the middle of a 5-year period, we plot the estimates at the mid-point of the years in that interval for which the survey provides data. Black lines and corresponding intervals represent posterior medians and 95% uncertainty intervals respectively for the betabinomial model. IHME’s estimates and corresponding intervals, where we can compare, are in aquamarine.



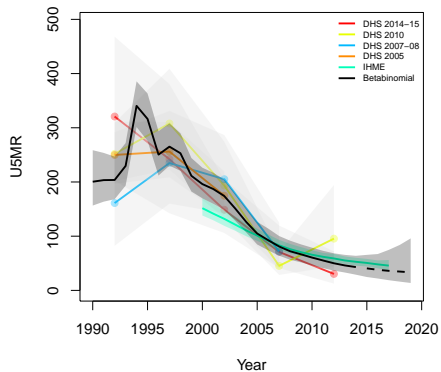
**Rulindo, Amajyaruguru**



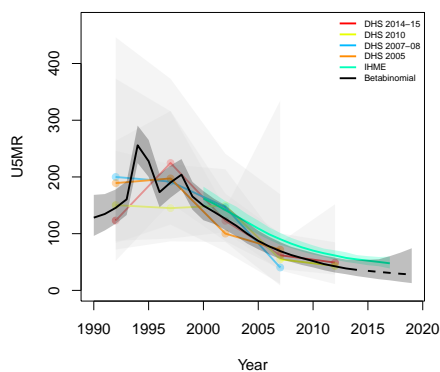
**Gisagara, Amajyepfo**



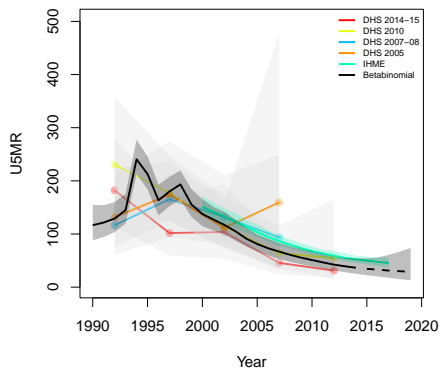
**Huye, Amajyepfo**



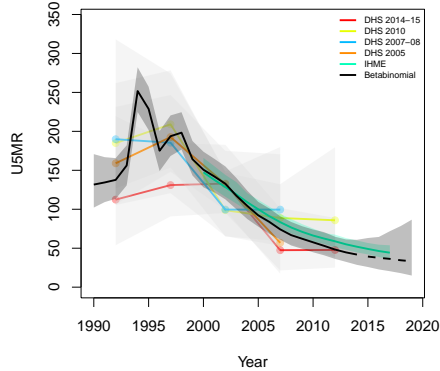
**Kamonyi, Amajyepfo**



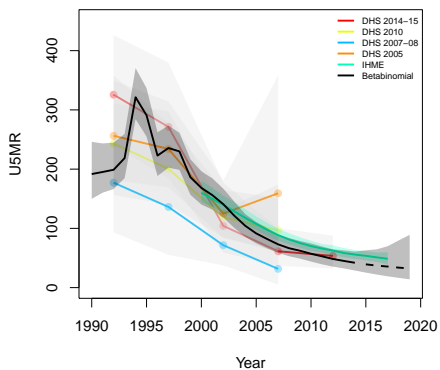
**Muhanga, Amajyepfo**



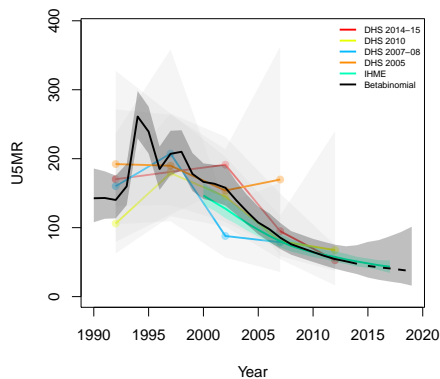
**Nyamagabe, Amajyepfo**



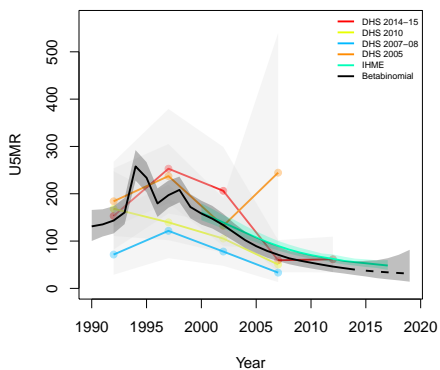
**Nyanza, Amajepfo**



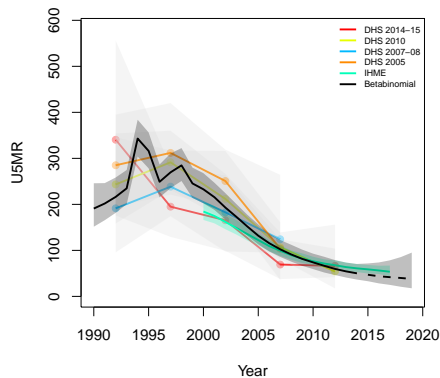
**Nyaruguru, Amajepfo**



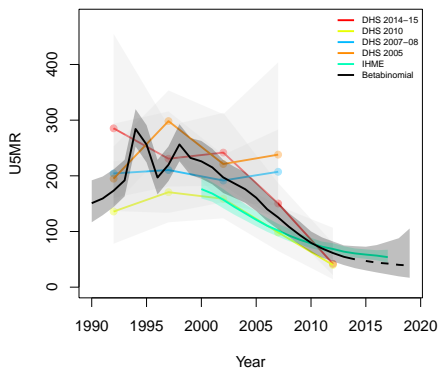
**Ruhango, Amajepfo**



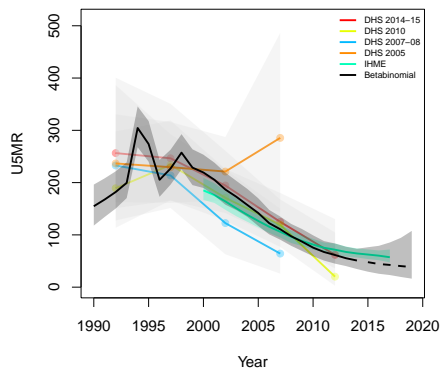
**Bugesera, Iburasirazuba**



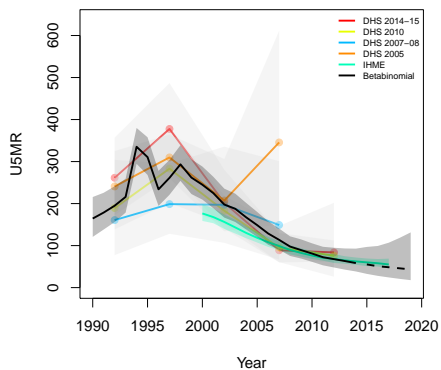
**Gatsibo, Iburasirazuba**



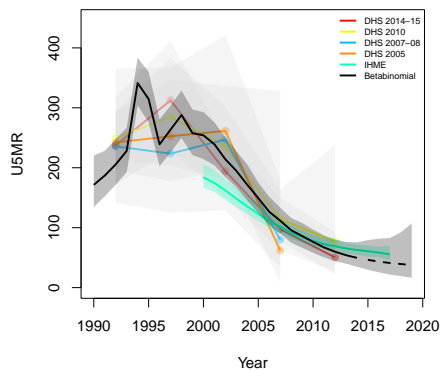
**Kayonza, Iburasirazuba**



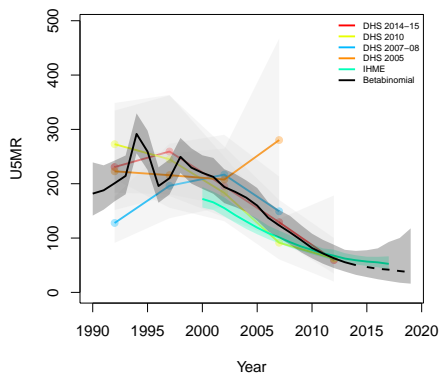
**Kirehe, Iburasirazuba**



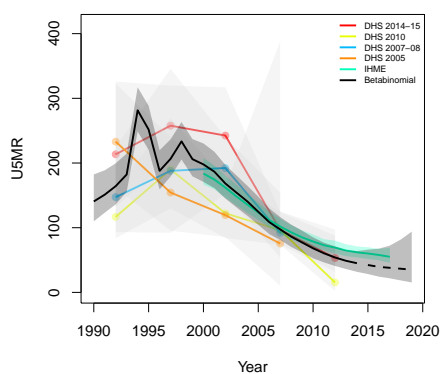
**Ngoma, Iburasirazuba**



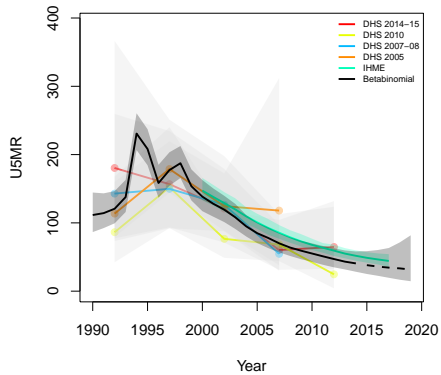
**Nyagatare, Iburasirazuba**



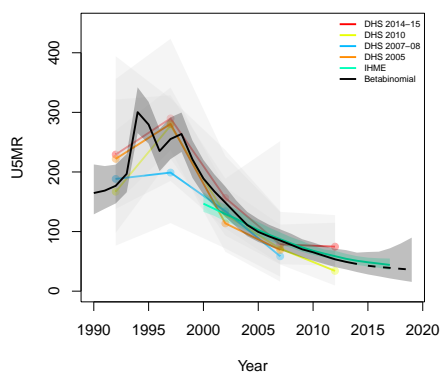
**Rwamagana, Iburasirazuba**



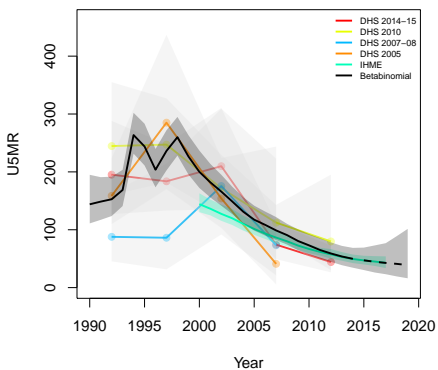
**Karongi, Iburengerazuba**



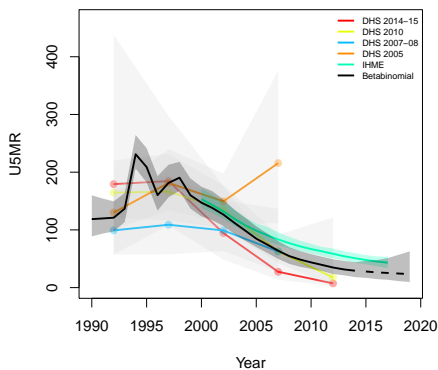
**Ngororero, Iburengerazuba**



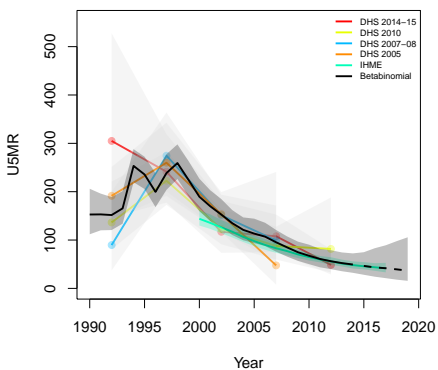
**Nyabihu, Iburengerazuba**



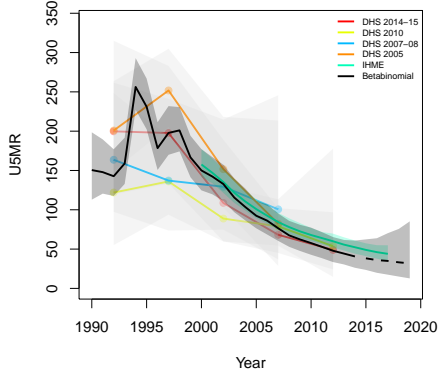
**Nyamasheke, Iburengerazuba**



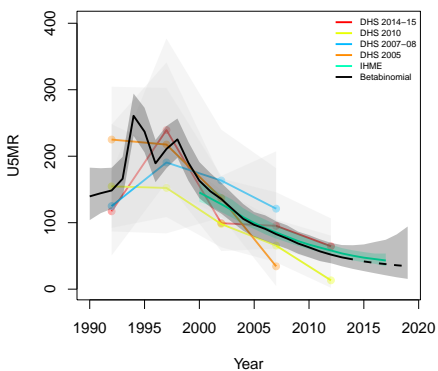
**Rubavu, Iburengerazuba**



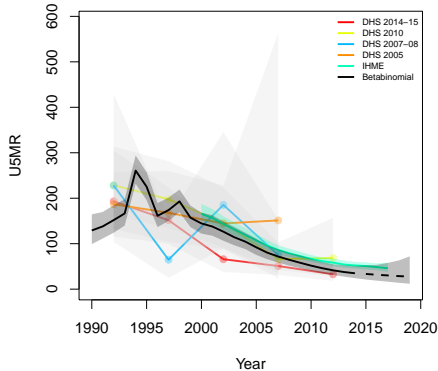
**Rusizi, Iburengerazuba**

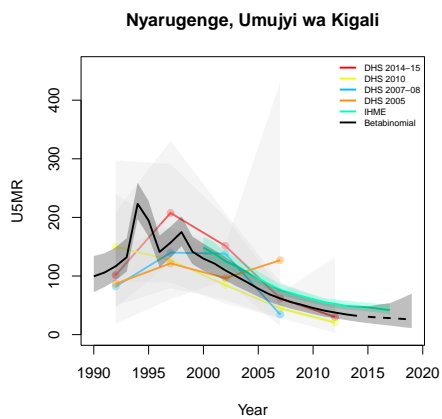
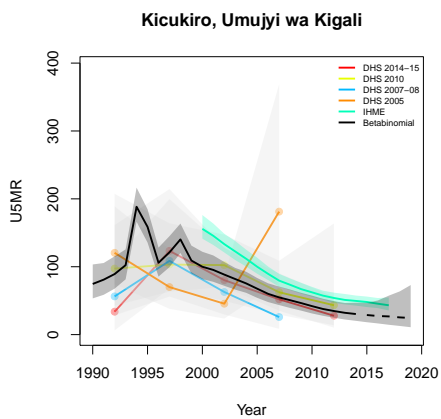


**Rutsiro, Iburengerazuba**



**Gasabo, Umujyi wa Kigali**





**B.15 Senegal**

Age	Survey	Clusters			Deaths			Agemonths		
		Urban	Rural	Total	Urban	Rural	Total	Urban	Rural	Total
0	1993	132	126	258	35	79	114	1221	2297	3518
	1997	100	219	319	80	324	404	2798	7904	10702
	2005	158	208	366	316	840	1156	9572	18861	28433
	2010	147	238	385	373	1002	1375	11161	26025	37186
	2012	79	121	200	171	486	657	6234	14584	20818
	2014	77	120	197	150	455	605	6401	14580	20981
	2015	84	130	214	172	513	685	6490	16096	22586
	2016	84	130	214	151	478	629	6411	15817	22228
	2019	84	130	214	144	454	598	6473	14977	21450
1-11	1993	132	126	258	25	95	120	12586	23703	36289
	1997	100	219	319	71	276	347	29511	81002	110513
	2005	158	208	366	239	614	853	98628	189999	288627
	2010	147	238	385	203	675	878	114204	263259	377463
	2012	79	121	200	97	438	535	64414	148160	212574
	2014	77	120	197	117	410	527	66470	149241	215711
	2015	84	130	214	140	491	631	66897	164349	231246
	2016	84	130	214	114	381	495	66599	161935	228534
	2019	84	130	214	93	351	444	67391	152898	220289
12-23	1993	132	126	258	18	52	70	13401	24765	38166
	1997	100	219	319	44	232	276	31447	84027	115474
	2005	158	208	366	146	470	616	101234	191567	292801
	2010	147	238	385	104	505	609	116253	264124	380377
	2012	79	121	200	44	217	261	65801	148468	214269
	2014	77	120	197	38	245	283	68005	150152	218157
	2015	84	130	214	51	313	364	68430	163562	231992
	2016	84	130	214	46	221	267	68271	162609	230880
	2019	84	130	214	34	160	194	68687	154239	222926
	1993	132	126	258	12	52	64	12849	24564	37413
	1997	100	219	319	31	215	246	31027	80516	111543

*B.15.1 Admin-1*



Figure B.70: **Left:** The names of the 14 Admin-1 areas of Senegal . **Right:** The neighborhood structure of Admin-1 areas in Senegal .

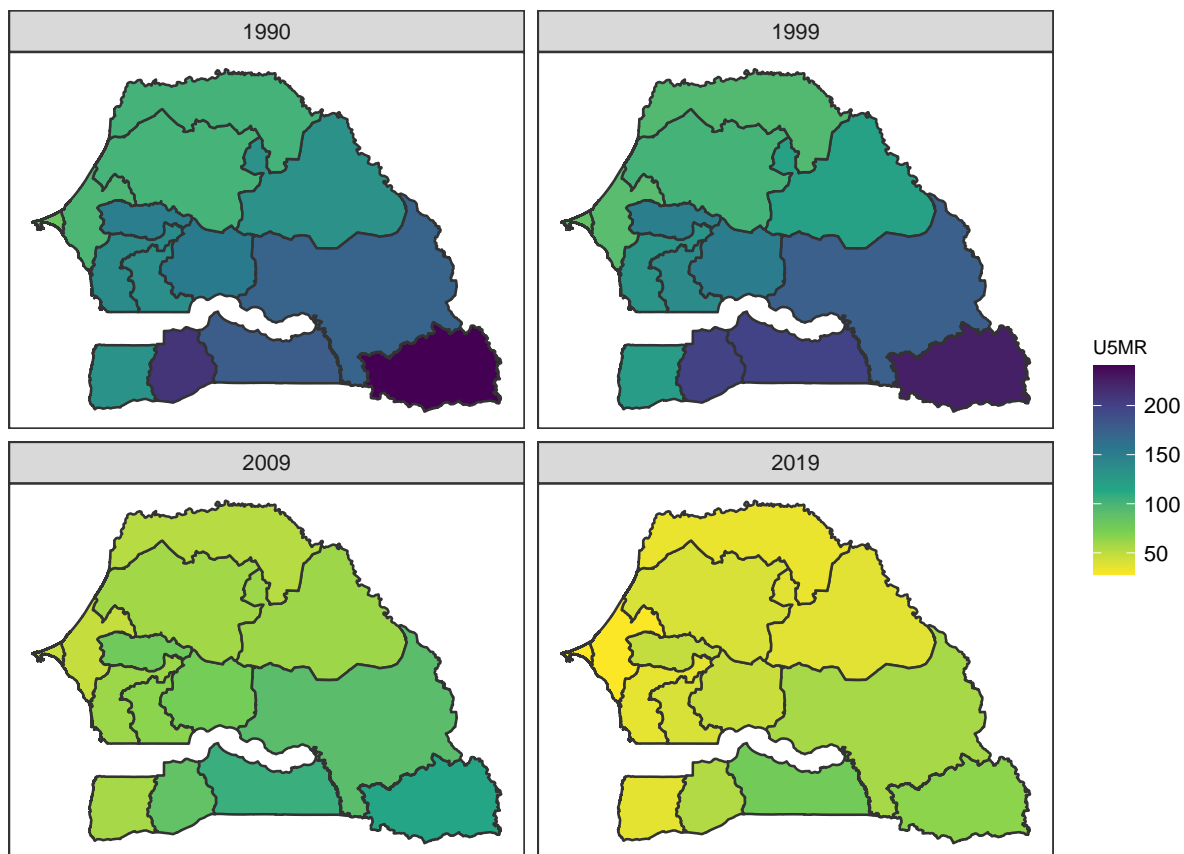


Figure B.71: Median U5MR estimates for years 1990, 1999, 2009, 2019 for Admin-1 areas in Senegal .

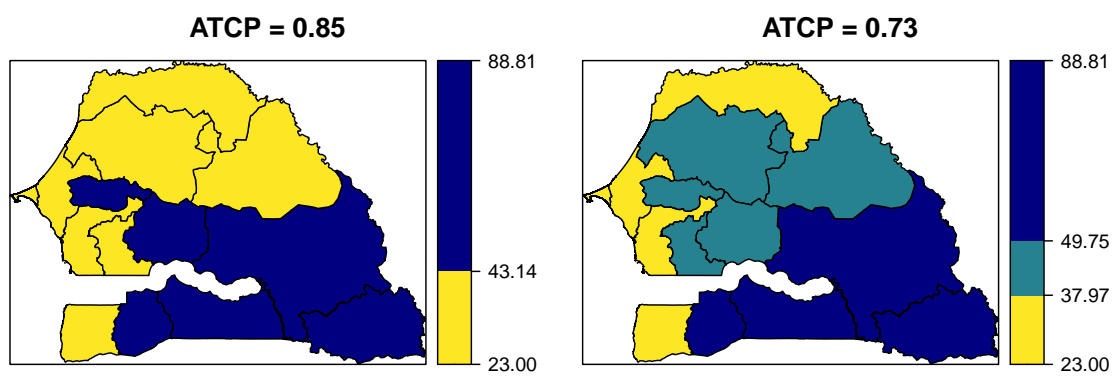
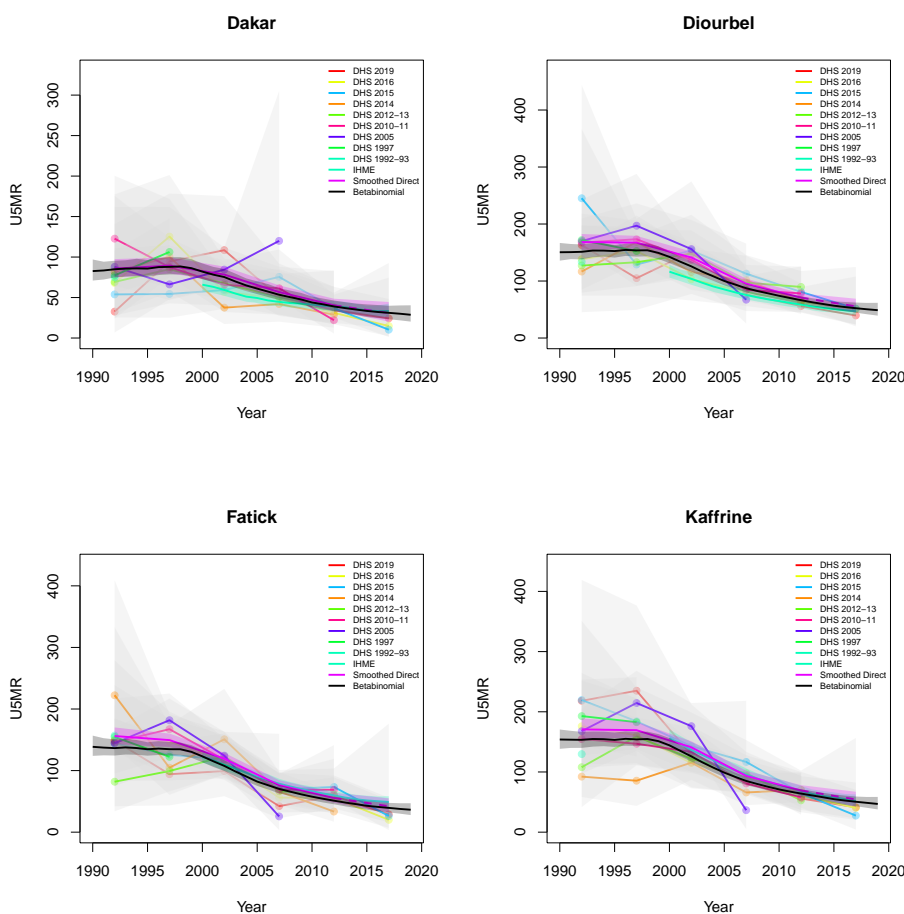


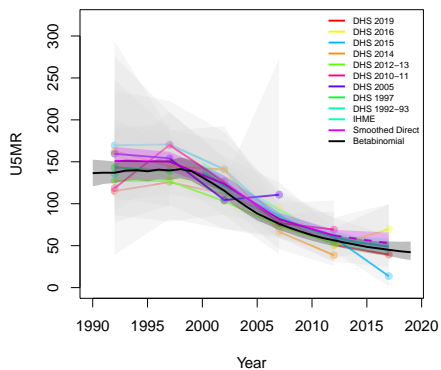
Figure B.72: Expression of uncertainty of U5MR (deaths per 1000 children) estimates for Admin-1 areas based on the average true classification probability (ATCP) in 2019 using  $K = 2, 3$  colors.

*Data and estimates over time by area*

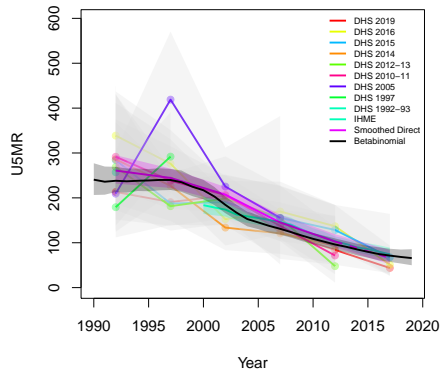
Colored lines with circular points and light grey uncertainty bands are 5-year survey-weighted estimates of U5MR for years 1990–1994 up to 2015–2019 depending on survey timing. For a survey that ends in the middle of a 5-year period, we plot the estimates at the mid-point of the years in that interval for which the survey provides data. Black lines and corresponding intervals represent posterior medians and 95% uncertainty intervals respectively for the betabinomial model. IHME’s estimates and corresponding intervals, where we can compare, are in aquamarine.



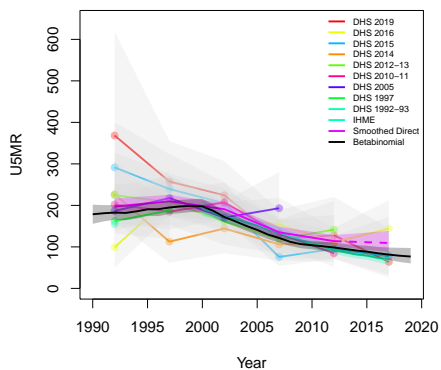
**Kaolack**



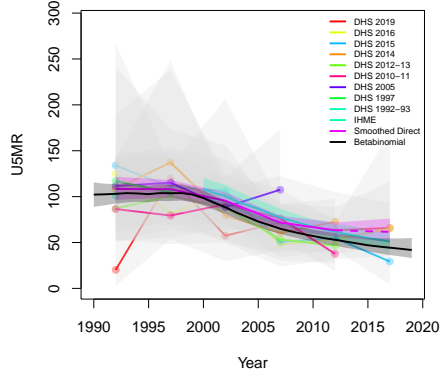
**Kédougou**



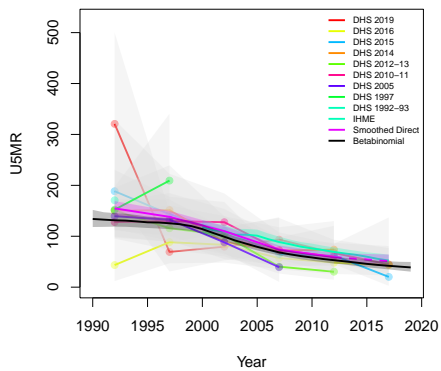
**Kolda**



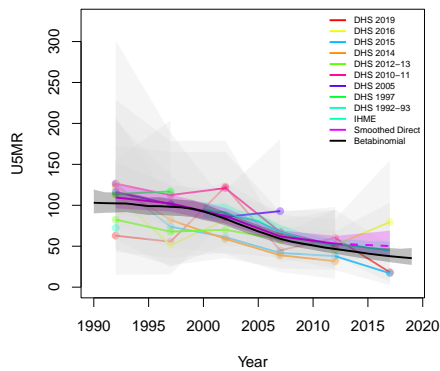
**Louga**

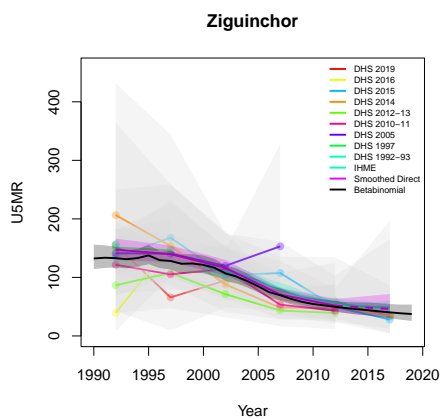
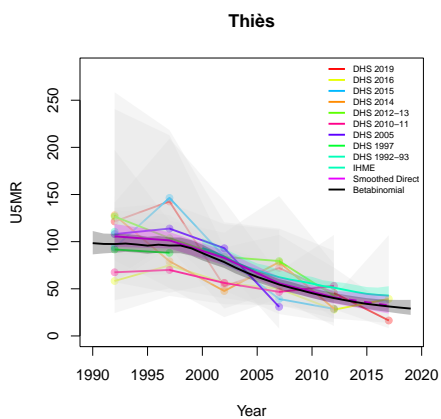
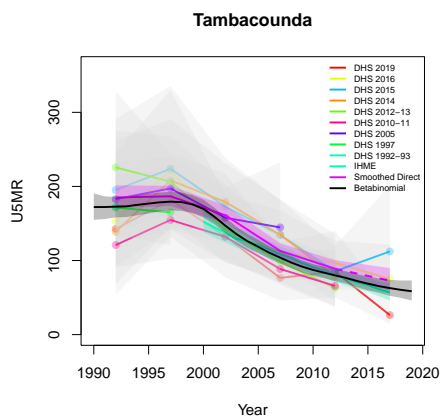
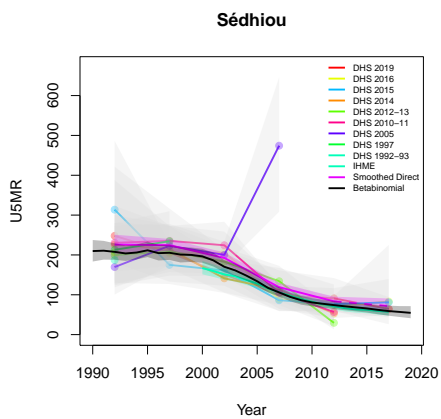


**Matam**



**Saint-Louis**





*B.15.2 Admin-2*



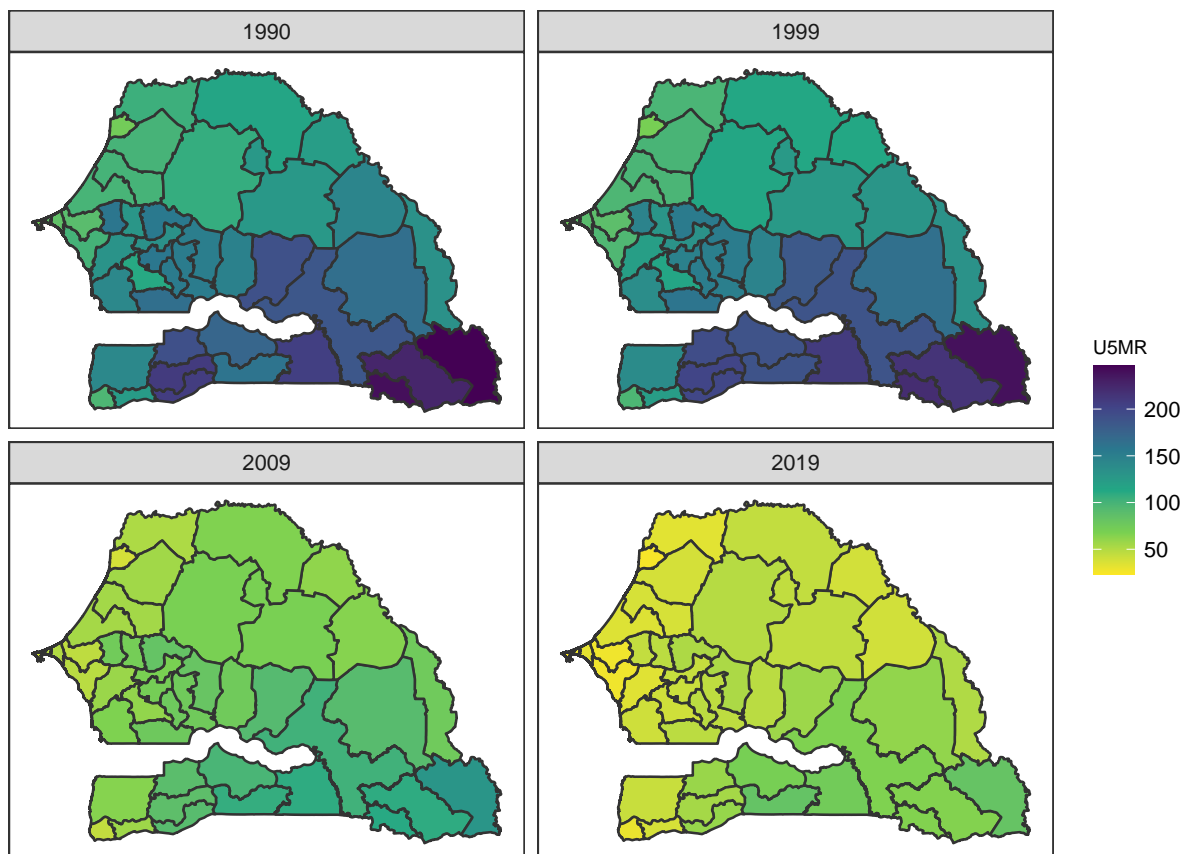


Figure B.74: Median U5MR estimates for years 1990, 1999, 2009, 2019 for Admin-2 areas in Senegal .

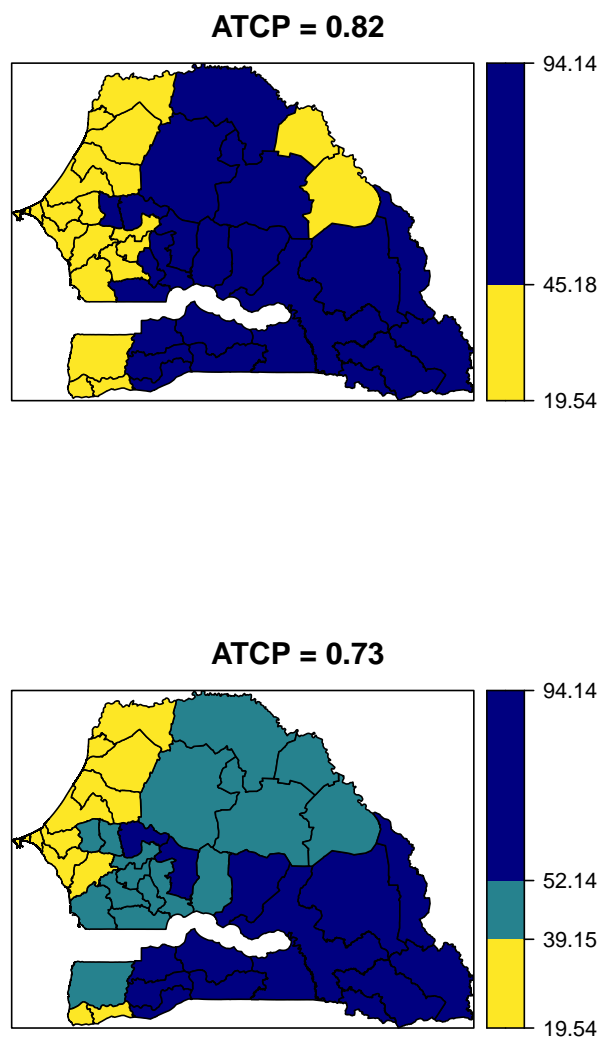
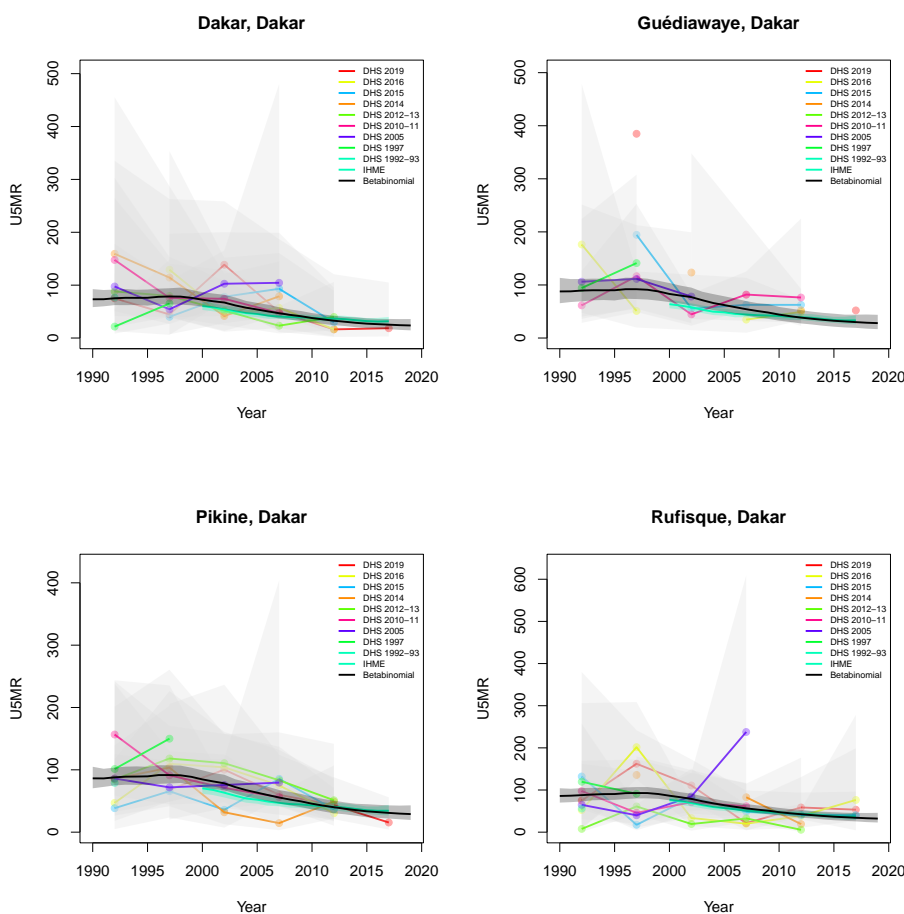


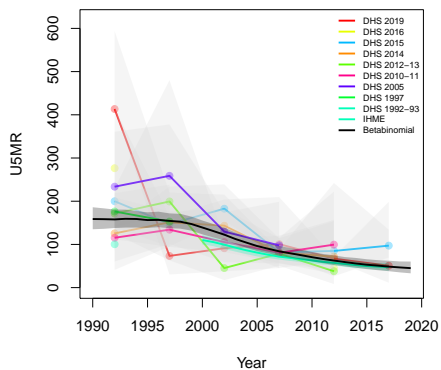
Figure B.75: Expression of uncertainty of U5MR (deaths per 1000 children) estimates for Admin-1 areas based on the average true classification probability (ATCP) in 2019 using  $K = 2, 3$  colors.

*Data and estimates over time by area*

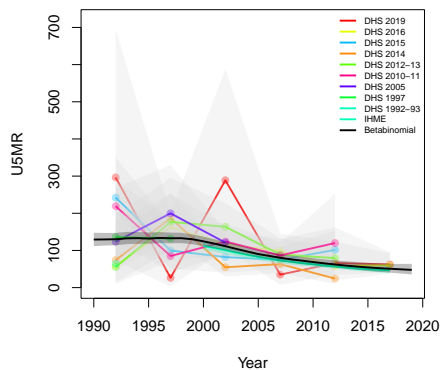
Colored lines with circular points and light grey uncertainty bands are 5-year survey-weighted estimates of U5MR for years 1990–1994 up to 2015–2019 depending on survey timing. For a survey that ends in the middle of a 5-year period, we plot the estimates at the mid-point of the years in that interval for which the survey provides data. Black lines and corresponding intervals represent posterior medians and 95% uncertainty intervals respectively for the betabinomial model. IHME’s estimates and corresponding intervals, where we can compare, are in aquamarine.



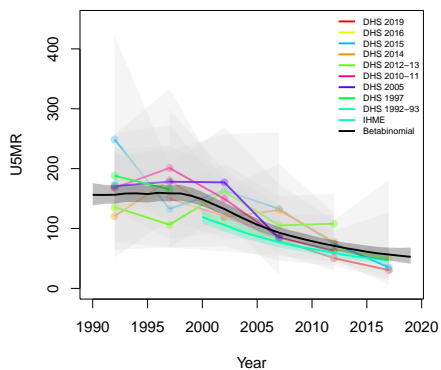
**Bambey, Diourbel**



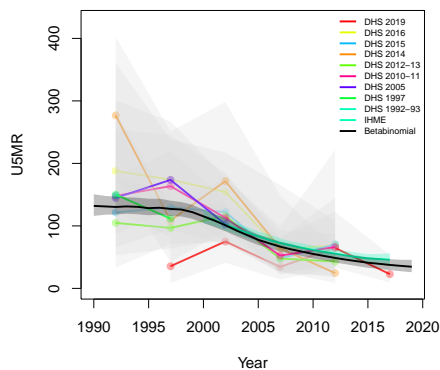
**Diourbel, Diourbel**



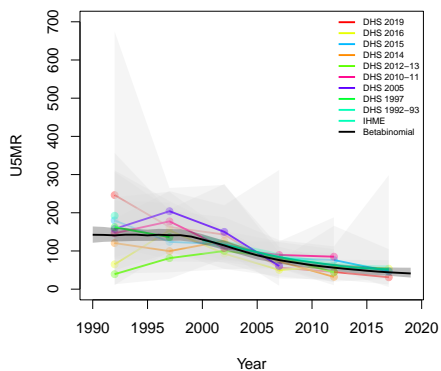
**Mbacké, Diourbel**



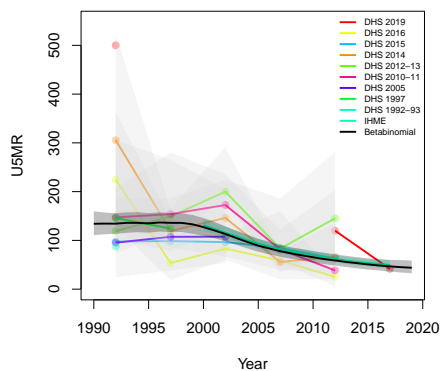
**Fatick, Fatick**



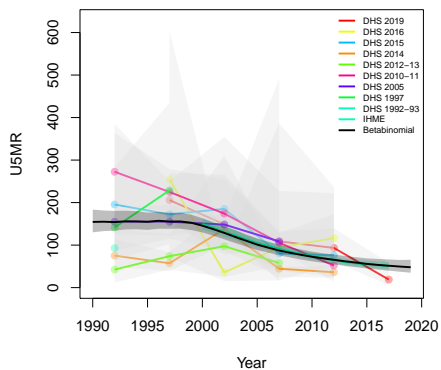
**Foundiougne, Fatick**



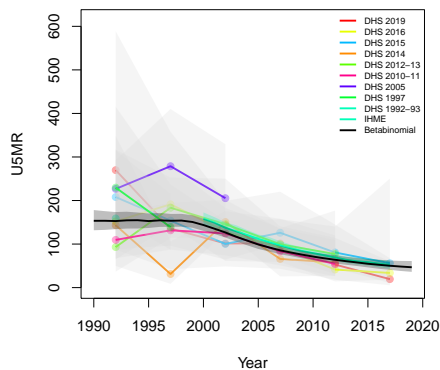
**Gossas, Fatick**



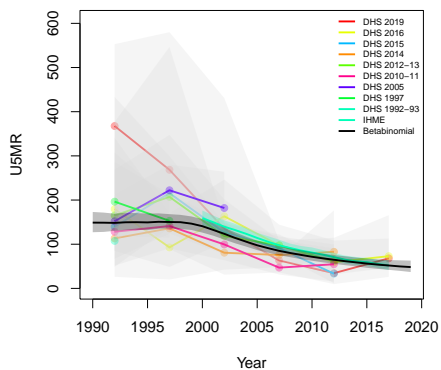
**Birkilane, Kaffrine**



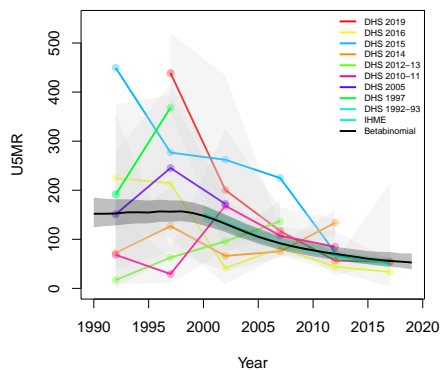
**Kaffrine, Kaffrine**



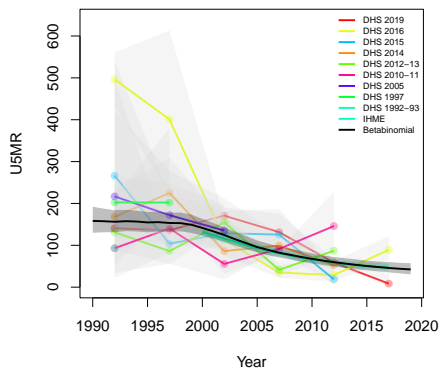
**Koungheul, Kaffrine**



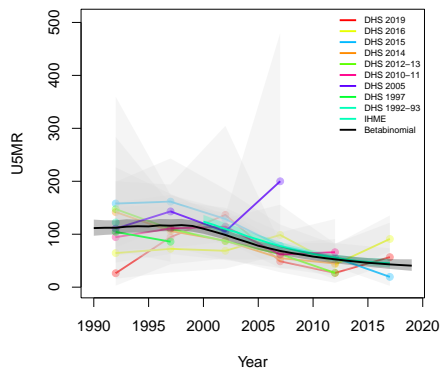
**Malème Hodar, Kaffrine**



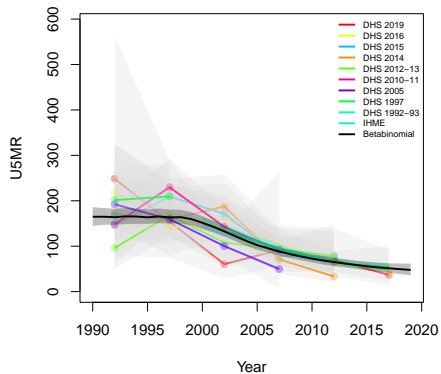
**Guinguinéo, Kaolack**



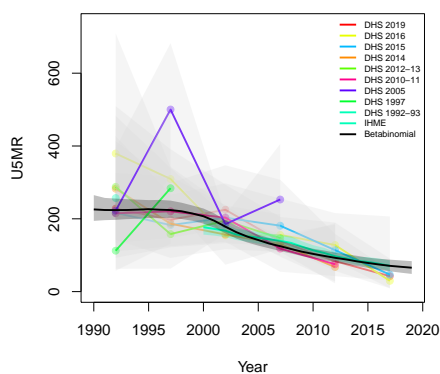
**Kaolack, Kaolack**



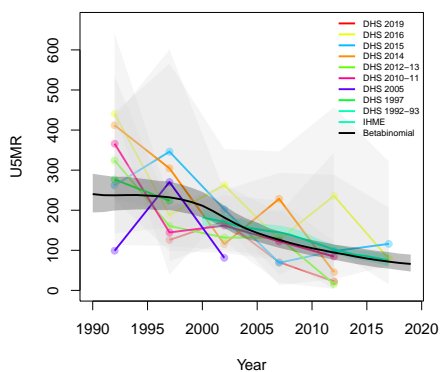
**Nioro du Rip, Kaolack**



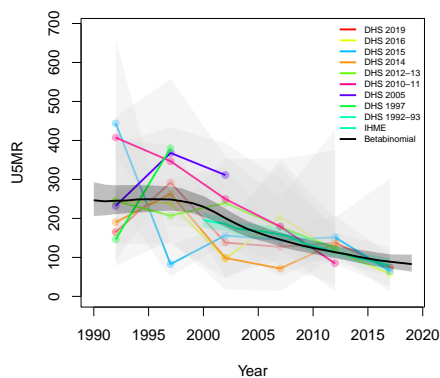
**Kédougou, Kédougou**



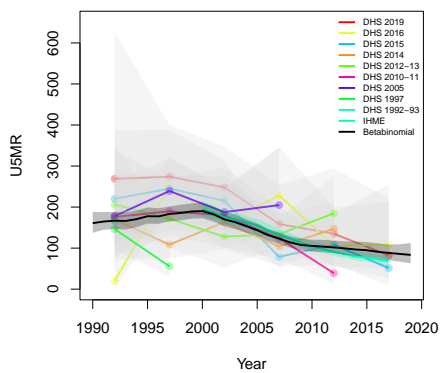
**Salémata, Kédougou**



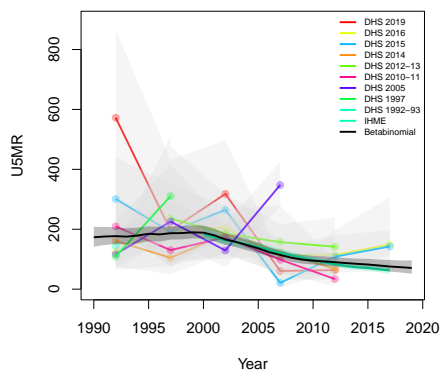
**Saraya, Kédougou**



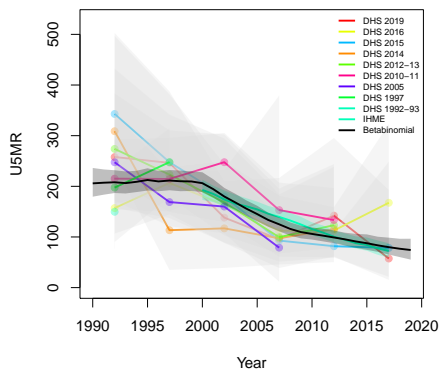
**Kolda, Kolda**



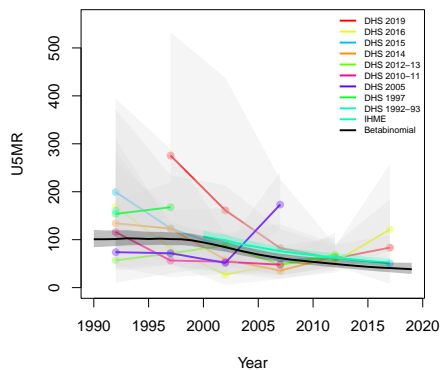
**Médina Yoro Foula, Kolda**



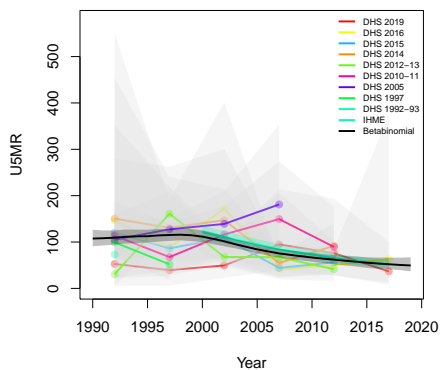
Vélingara, Kolda



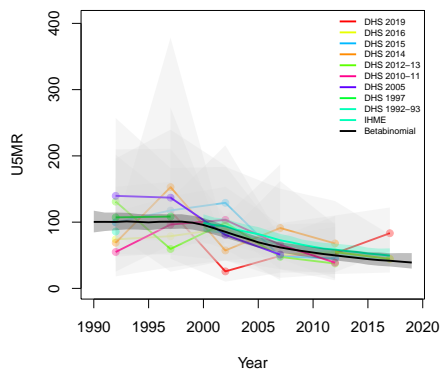
Kébémér, Louga



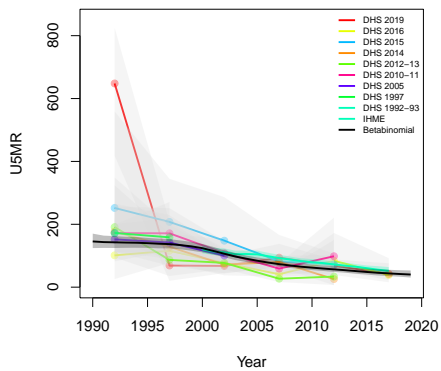
Linguère, Louga



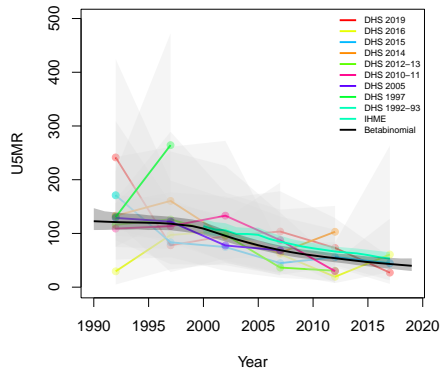
Louga, Louga



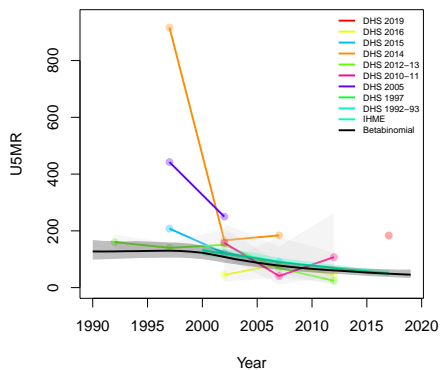
Kanel, Matam



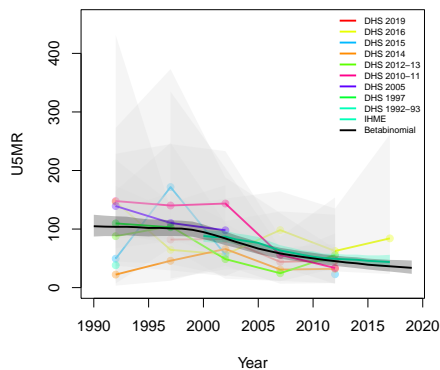
Matam, Matam



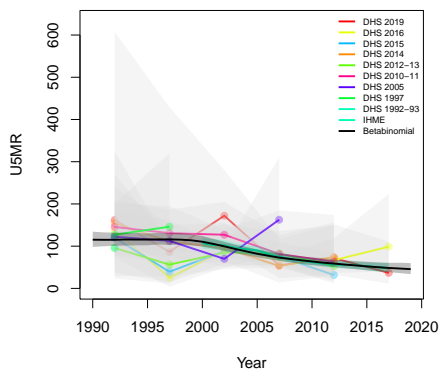
**Ranéroù Ferlo, Matam**



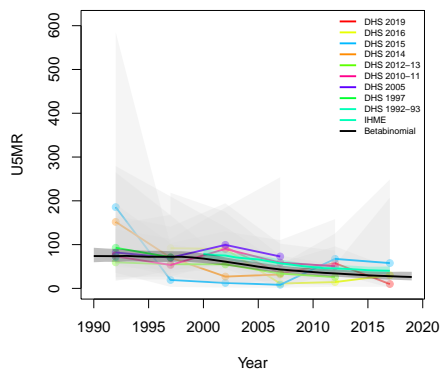
**Dagana, Saint-Louis**



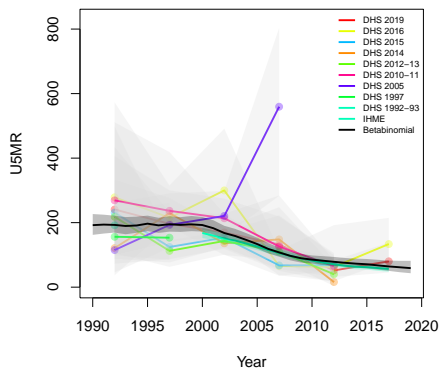
**Podor, Saint-Louis**



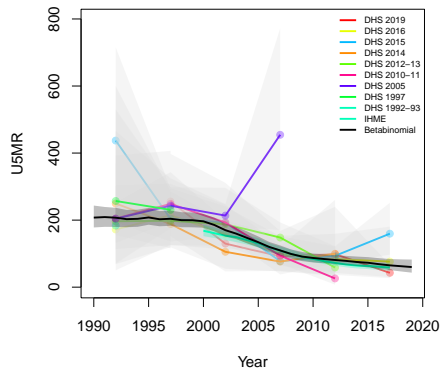
**Saint-Louis, Saint-Louis**



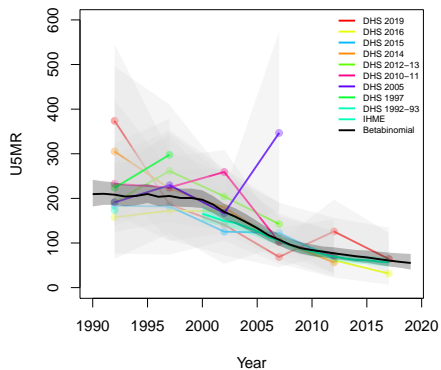
**Boukiling, Sédhiou**



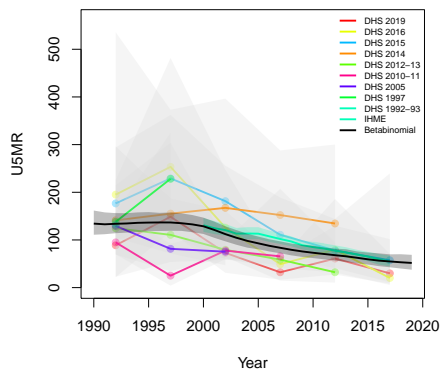
**Goudomp, Sédhiou**



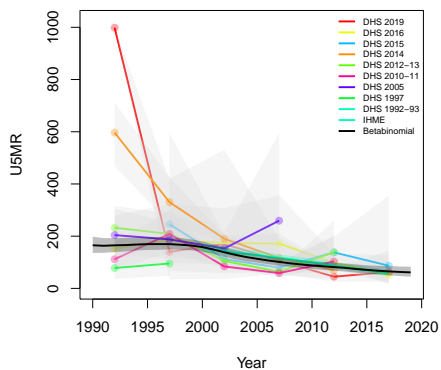
Sédhiou, Sédhiou



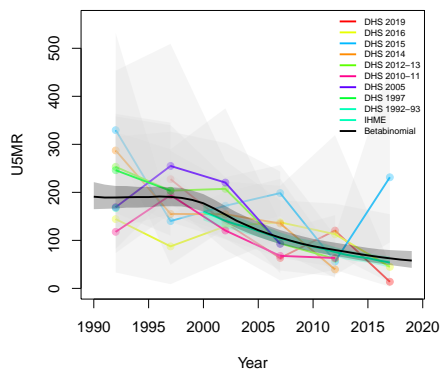
Bakel, Tambacounda



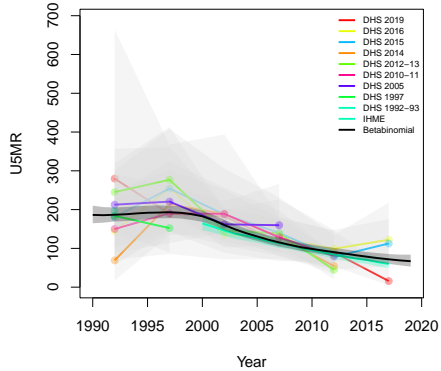
Goudiry, Tambacounda



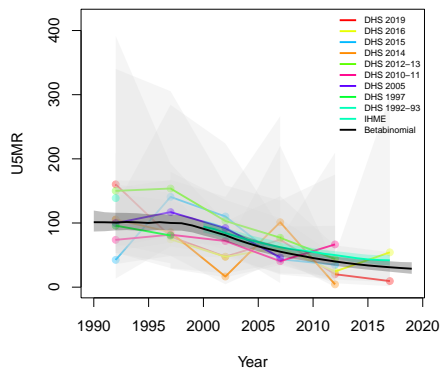
Koupentoum, Tambacounda

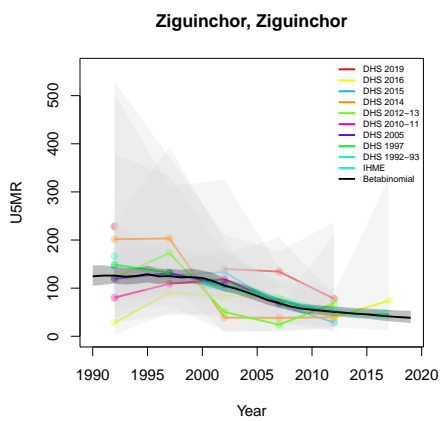
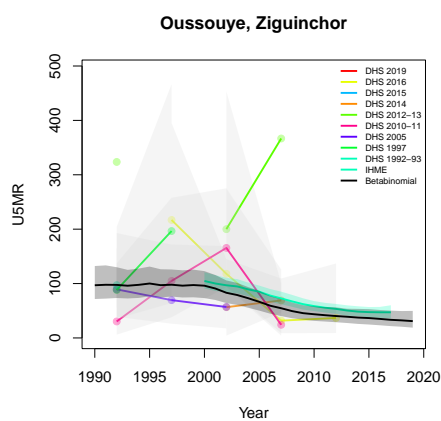
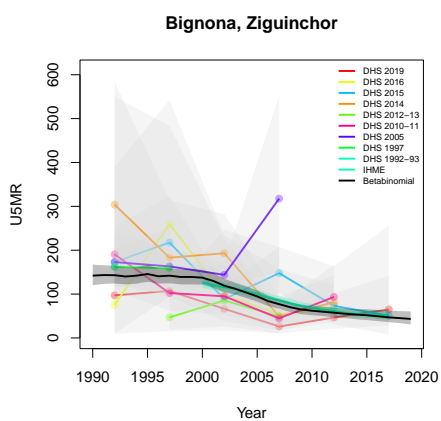
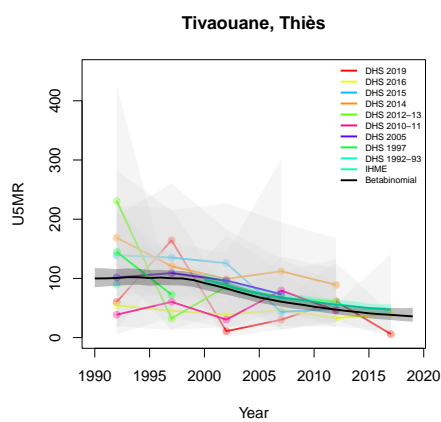
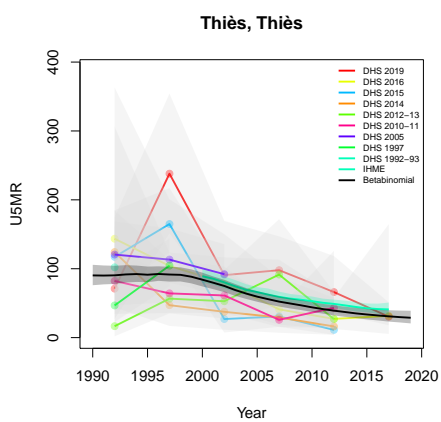


Tambacounda, Tambacounda



Mbour, Thiès





## B.16 Sierra Leone

Age 0	Survey	Clusters			Deaths			Agemonths		
		Urban	Rural	Total	Urban	Rural	Total	Urban	Rural	Total
1-11	2013	158	277	435	622	1330	1952	13622	30526	44148
12-23	2013	158	277	435	846	2514	3360	135599	299349	434948
24-35	2013	158	277	435	351	1031	1382	133386	286874	420260
36-47	2013	158	277	435	213	600	813	125395	267164	392559
48-59	2013	158	277	435	116	286	402	117036	246247	363283
	2013	158	277	435	68	143	211	109222	228618	337840

Table B.16: **Data summary for Sierra Leone.** Total numbers of clusters (Columns 3–5) with observations in each age group by survey in urban and rural areas and combined. Numbers of deaths (Columns 6–8) and number of agemonths (Columns 9–10) observed in each age group by survey in urban and rural areas and combined.

### B.16.1 Admin-1

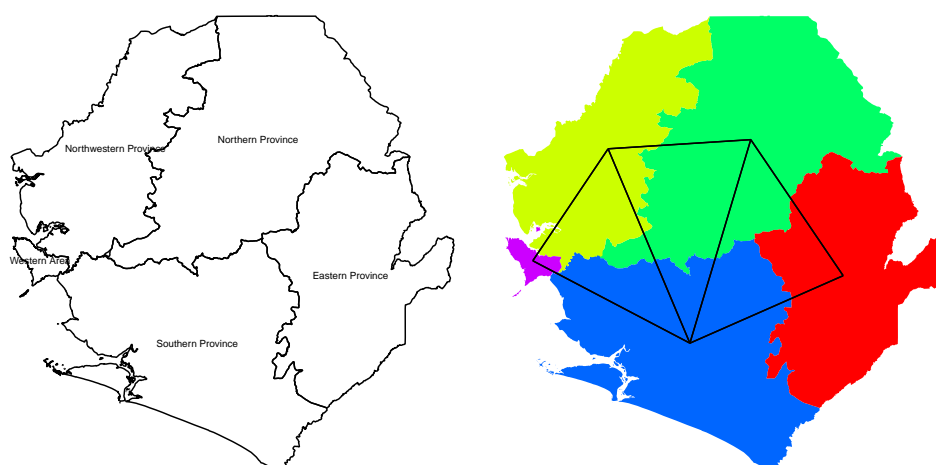


Figure B.76: **Left:** The names of the 5 Admin-1 areas of Sierra Leone . **Right:** The neighborhood structure of Admin-1 areas in Sierra Leone .

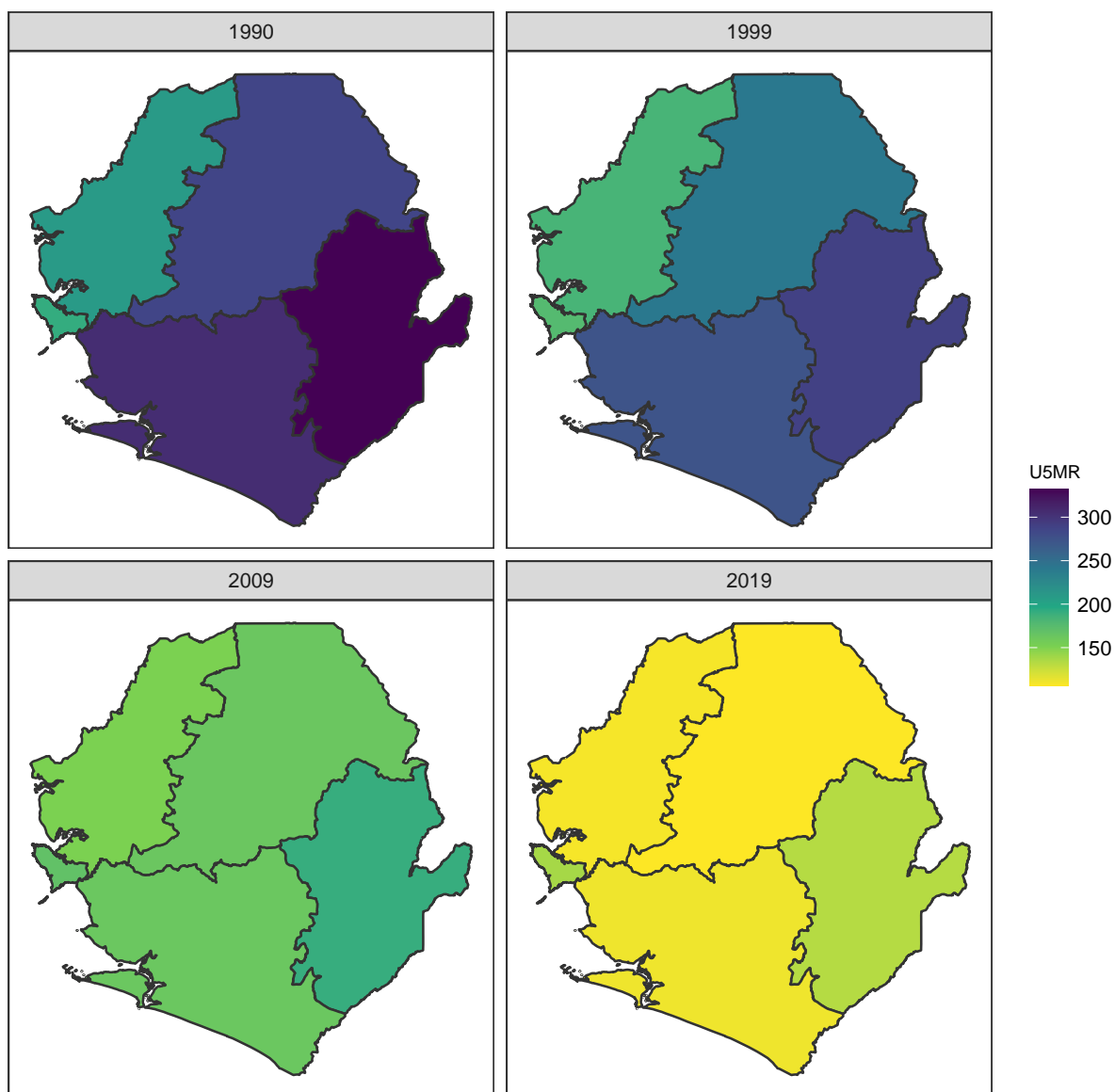


Figure B.77: Median U5MR estimates for years 1990, 1999, 2009, 2019 for Admin-1 areas in Sierra Leone .

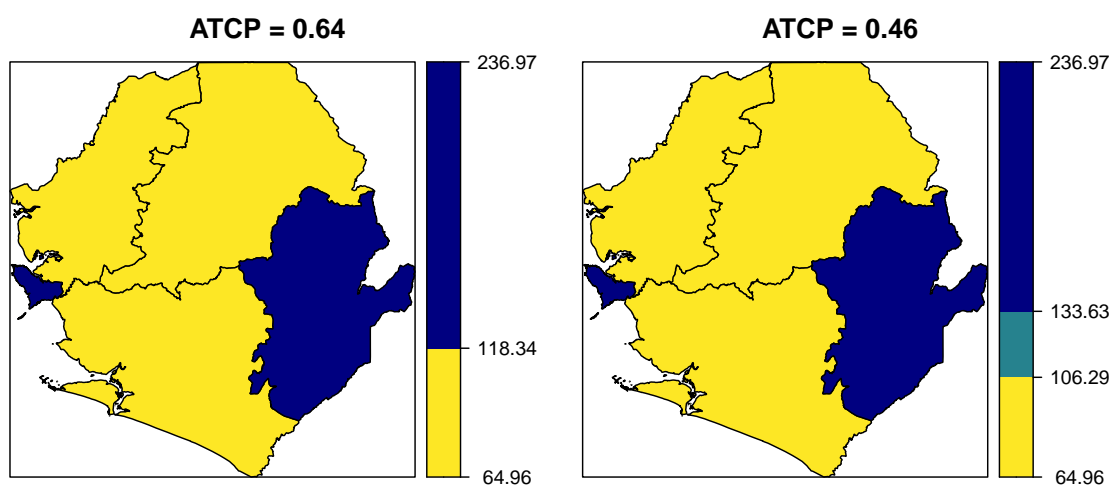
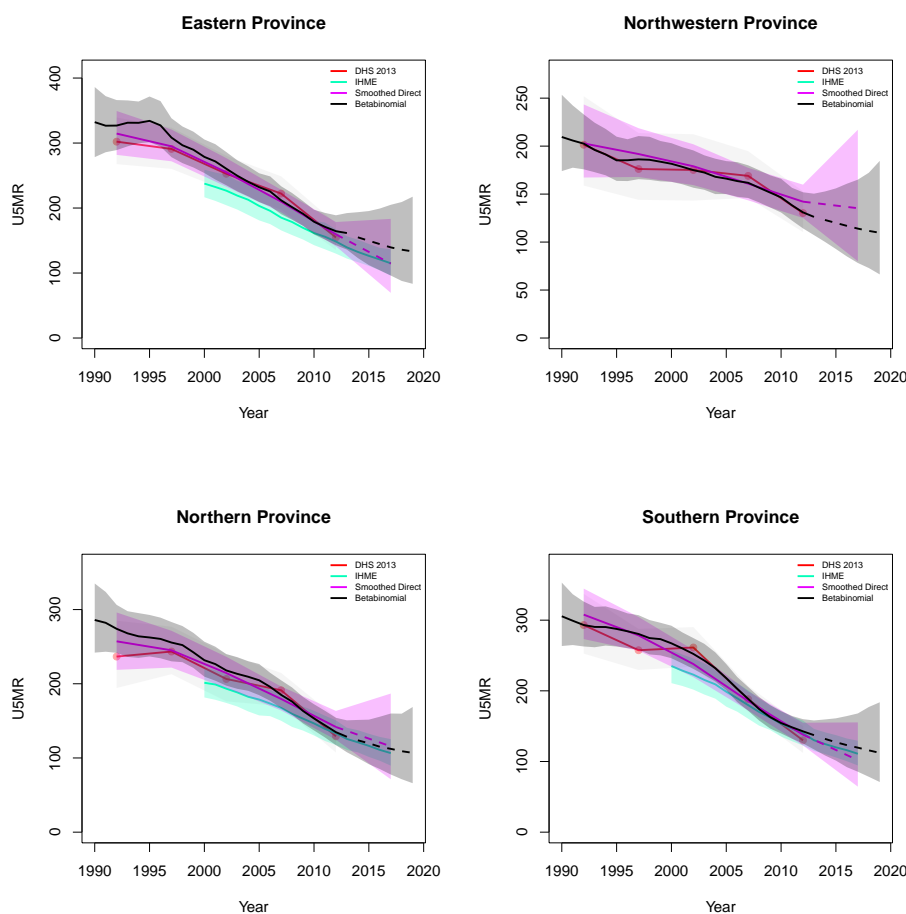
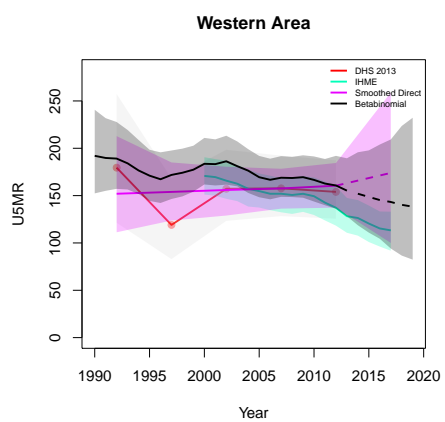


Figure B.78: Expression of uncertainty of U5MR (deaths per 1000 children) estimates for Admin-1 areas based on the average true classification probability (ATCP) in 2019 using  $K = 2, 3$  colors.

*Data and estimates over time by area*

Colored lines with circular points and light grey uncertainty bands are 5-year survey-weighted estimates of U5MR for years 1990–1994 up to 2015–2019 depending on survey timing. For a survey that ends in the middle of a 5-year period, we plot the estimates at the mid-point of the years in that interval for which the survey provides data. Black lines and corresponding intervals represent posterior medians and 95% uncertainty intervals respectively for the betabinomial model. IHME’s estimates and corresponding intervals, where we can compare, are in aquamarine.





*B.16.2 Admin-2*

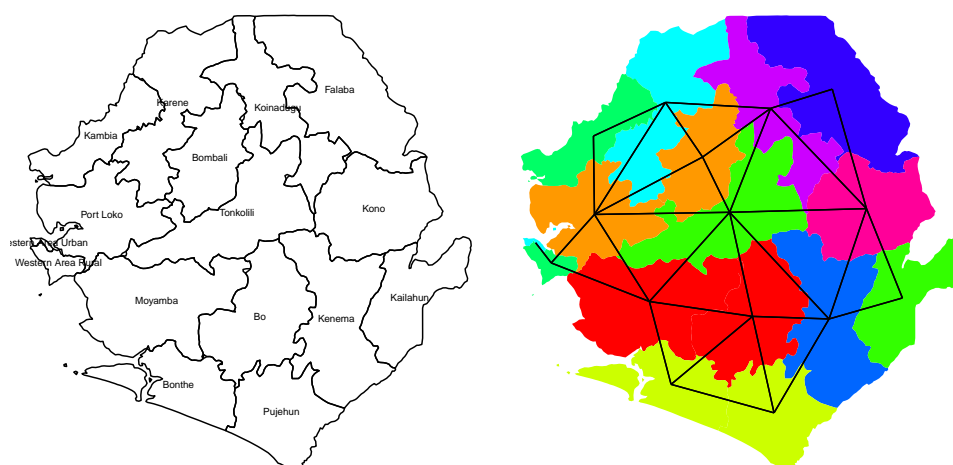


Figure B.79: **Left:** The names of the 16 Admin-2 areas of Sierra Leone . **Right:** The neighborhood structure of Admin-2 areas in Sierra Leone .

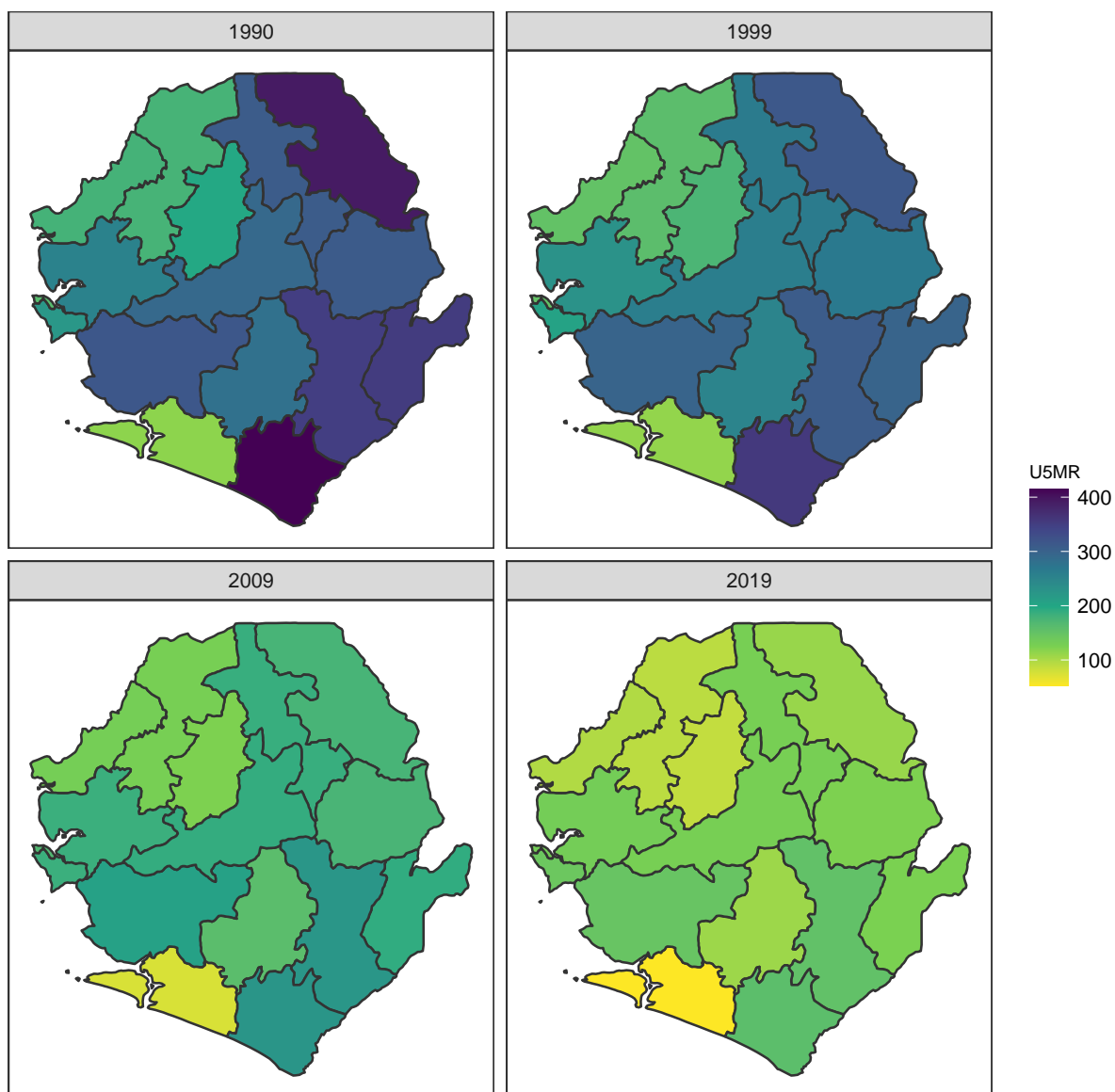


Figure B.80: Median U5MR estimates for years 1990, 1999, 2009, 2019 for Admin-2 areas in Sierra Leone .

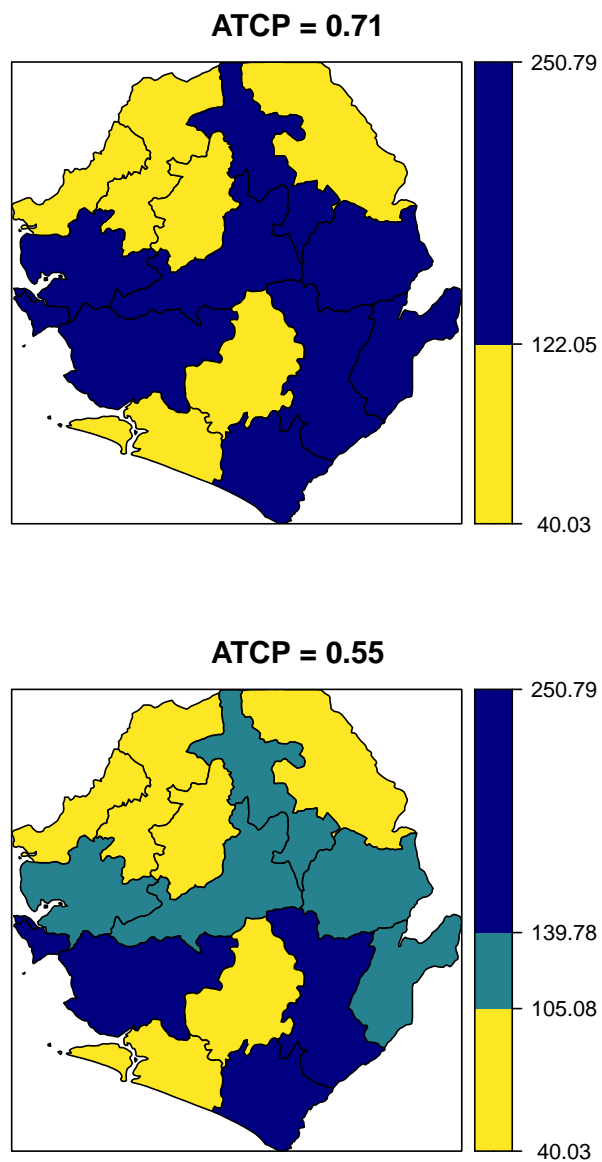
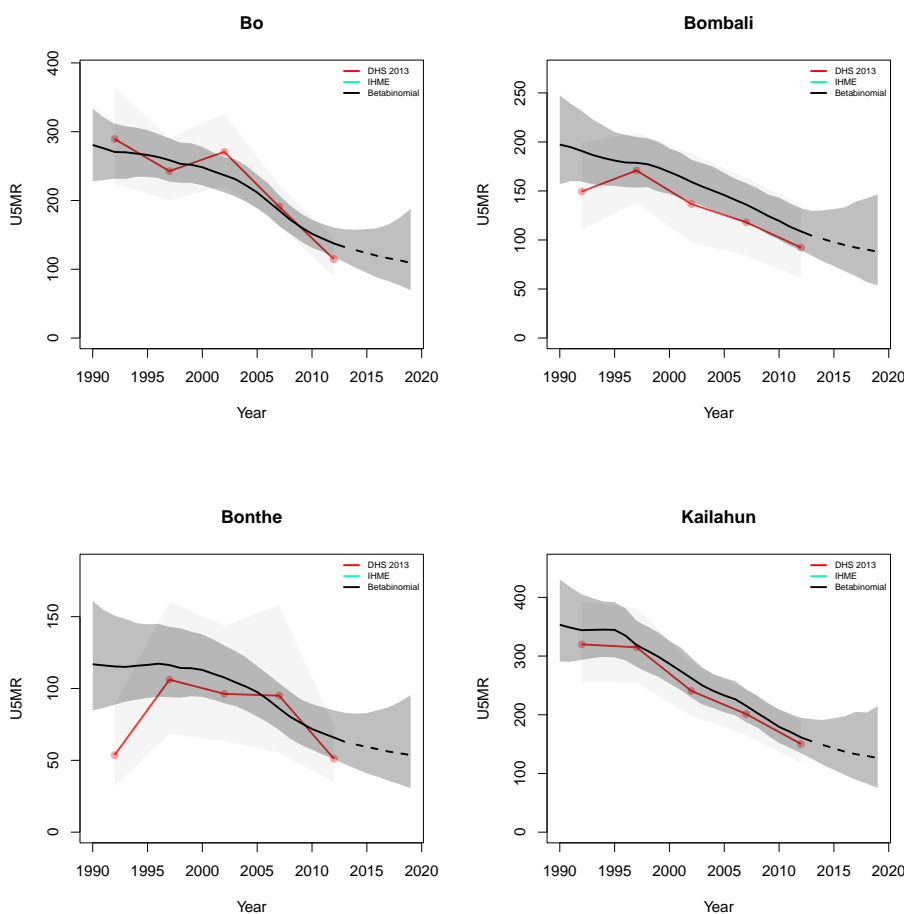
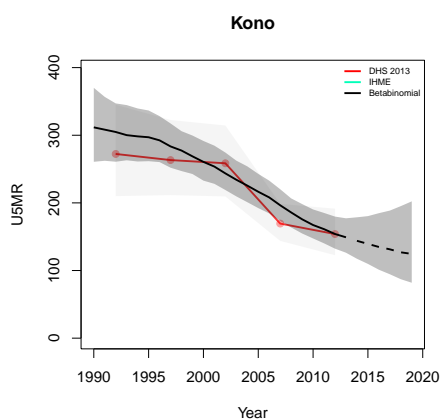
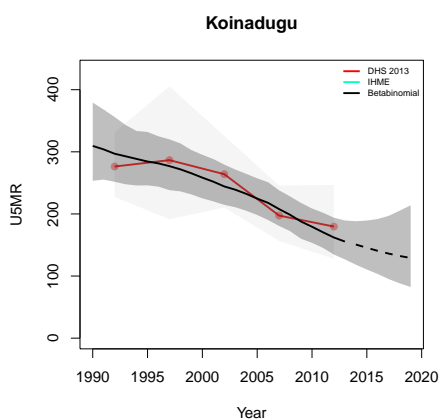
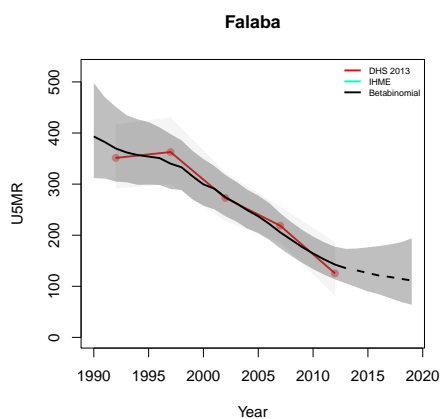
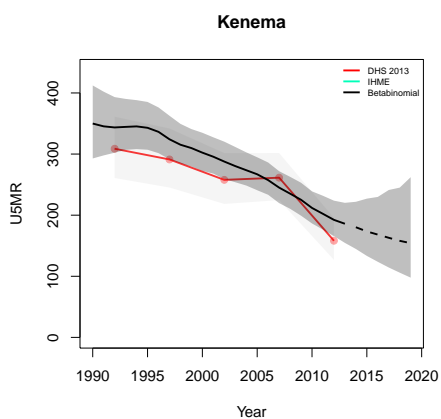
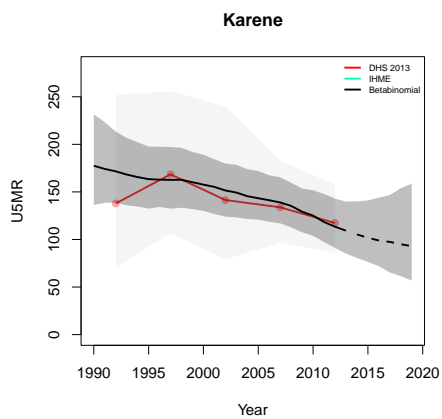
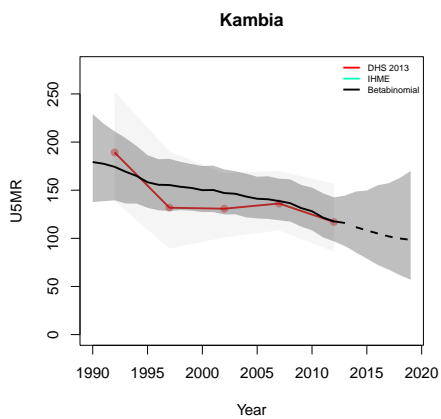


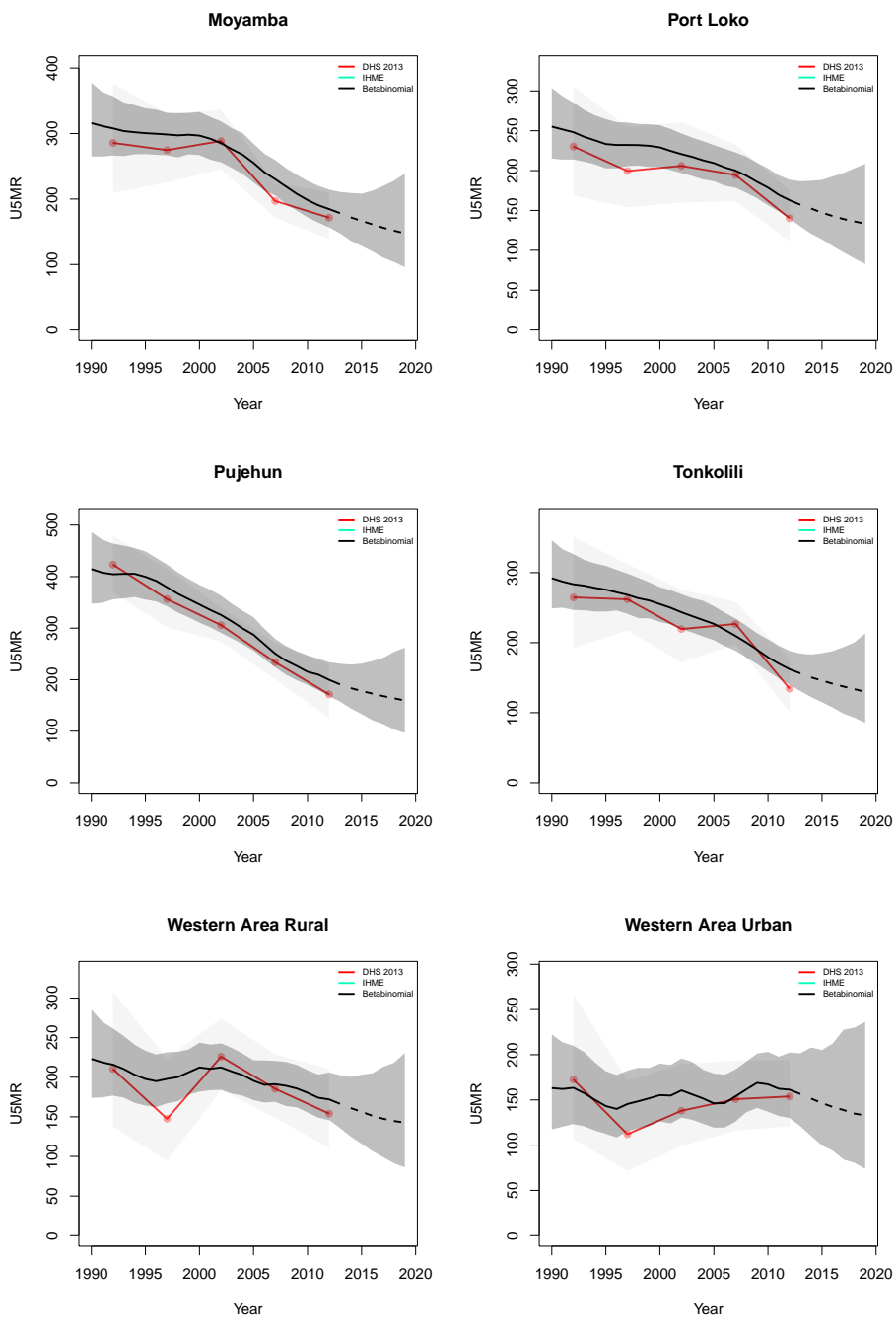
Figure B.81: Expression of uncertainty of U5MR (deaths per 1000 children) estimates for Admin-1 areas based on the average true classification probability (ATCP) in 2019 using  $K = 2, 3$  colors.

*Data and estimates over time by area*

Colored lines with circular points and light grey uncertainty bands are 5-year survey-weighted estimates of U5MR for years 1990–1994 up to 2015–2019 depending on survey timing. For a survey that ends in the middle of a 5-year period, we plot the estimates at the mid-point of the years in that interval for which the survey provides data. Black lines and corresponding intervals represent posterior medians and 95% uncertainty intervals respectively for the betabinomial model. IHME’s estimates and corresponding intervals, where we can compare, are in aquamarine.







**B.17 Tanzania**

Age	Survey	Clusters			Deaths			Agemonths		
		Urban	Rural	Total	Urban	Rural	Total	Urban	Rural	Total
0	1999	58	115	173	66	189	255	1564	4395	5959
	2010	110	348	458	139	621	760	4451	20153	24604
	2015	180	428	608	315	790	1105	8152	27492	35644
1-11	1999	58	115	173	63	290	353	16038	43993	60031
	2010	110	348	458	134	813	947	45908	206247	252155
	2015	180	428	608	216	880	1096	83058	281787	364845
12-23	1999	58	115	173	25	109	134	16514	44198	60712
	2010	110	348	458	50	339	389	46682	208431	255113
	2015	180	428	608	116	403	519	83796	283609	367405
24-35	1999	58	115	173	9	52	61	15752	42235	57987
	2010	110	348	458	25	196	221	44364	196376	240740
	2015	180	428	608	35	194	229	77995	264364	342359
36-47	1999	58	115	173	5	32	37	15128	40421	55549
	2010	110	348	458	13	116	129	41904	184920	226824
	2015	180	428	608	33	122	155	72911	247378	320289
48-59	1999	58	115	173	8	23	31	14506	38401	52907
	2010	110	348	458	12	58	70	39467	173071	212538
	2015	180	428	608	14	66	80	68303	230693	298996

Table B.17: **Data summary for Tanzania.** Total numbers of clusters (Columns 3–5) with observations in each age group by survey in urban and rural areas and combined. Numbers of deaths (Columns 6–8) and number of agemonths (Columns 9–10) observed in each age group by survey in urban and rural areas and combined.

*B.17.1 Admin-1*



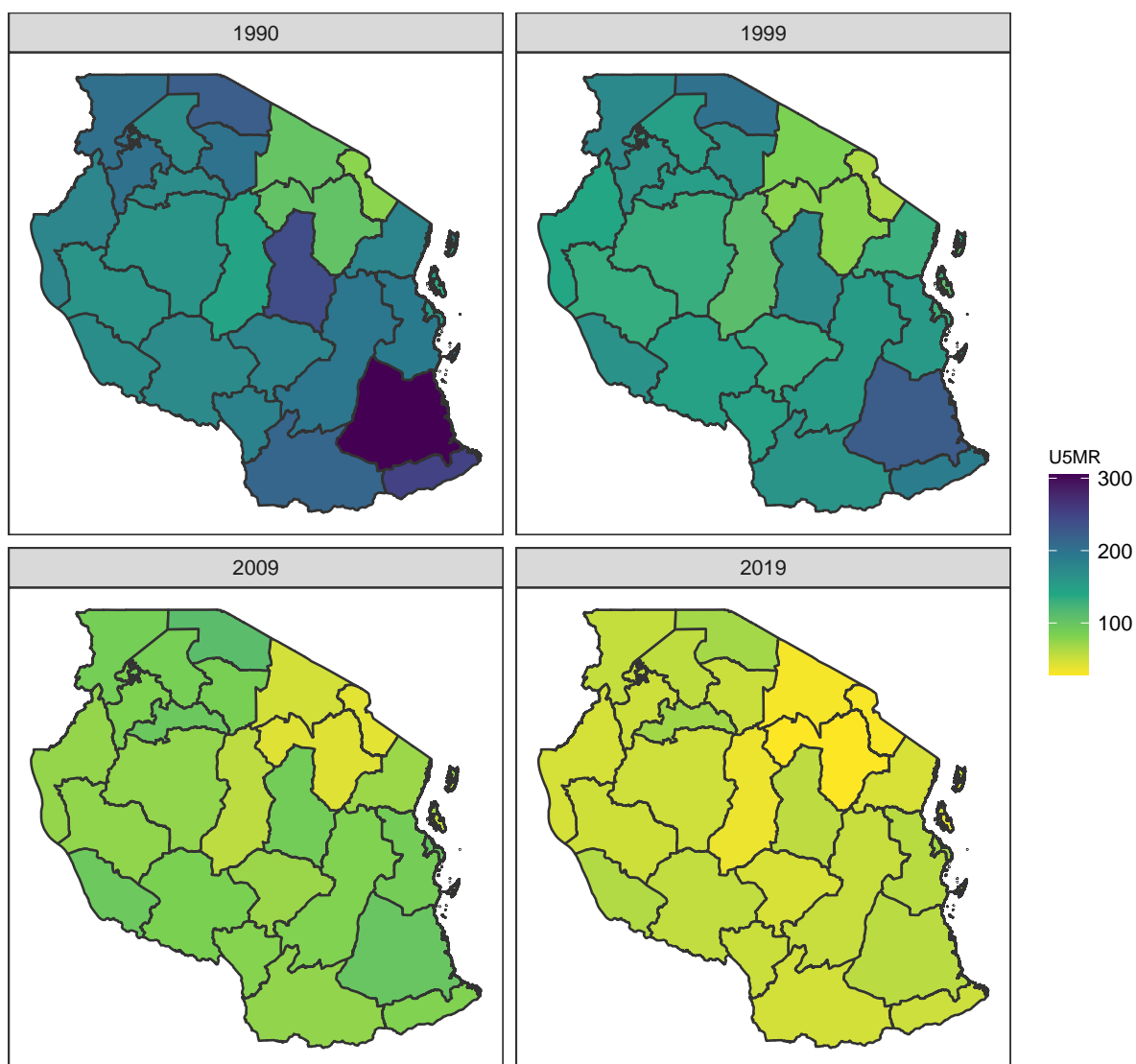


Figure B.83: Median U5MR estimates for years 1990, 1999, 2009, 2019 for Admin-1 areas in Tanzania .

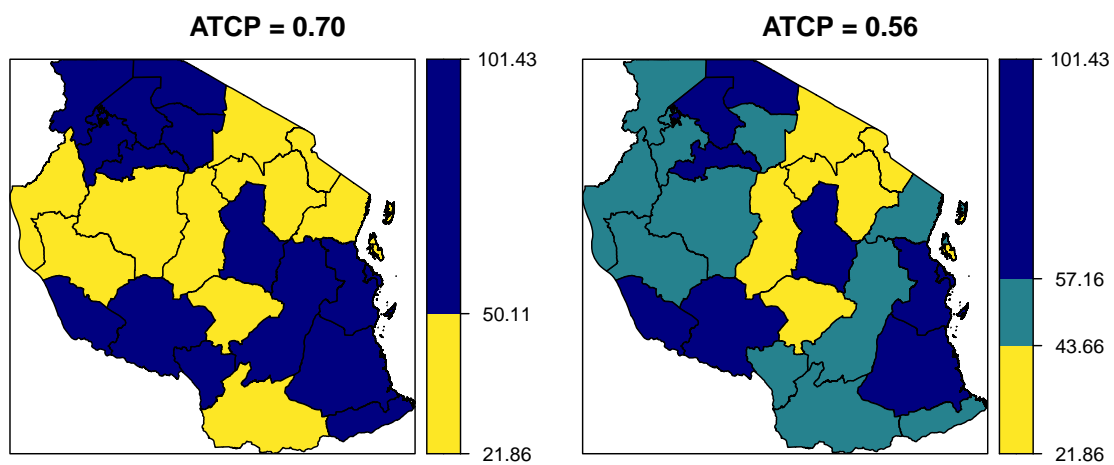
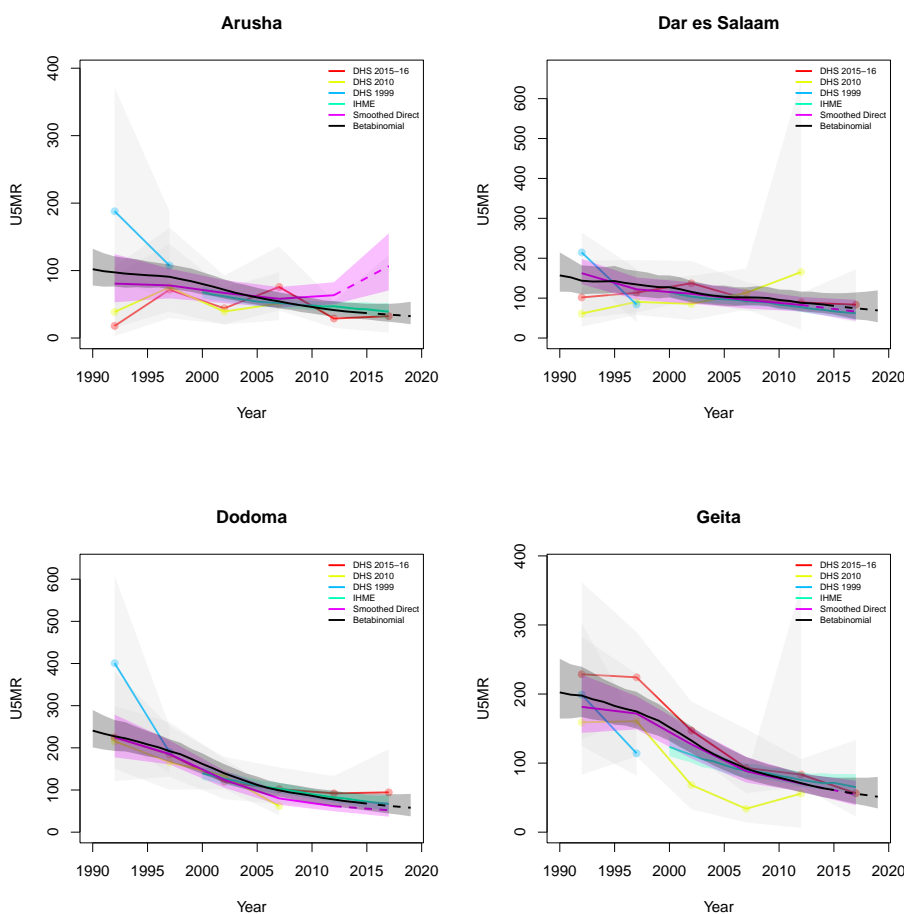


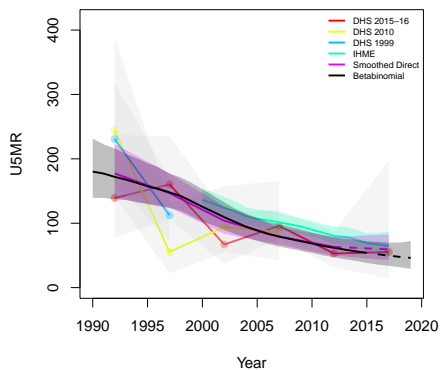
Figure B.84: Expression of uncertainty of U5MR (deaths per 1000 children) estimates for Admin-1 areas based on the average true classification probability (ATCP) in 2019 using  $K = 2, 3$  colors.

*Data and estimates over time by area*

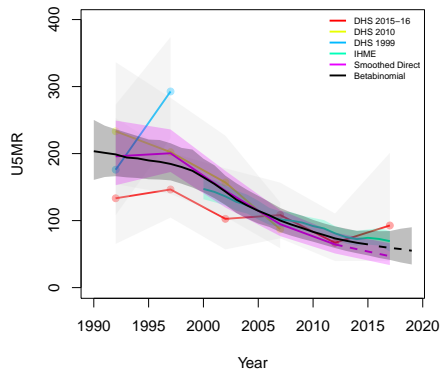
Colored lines with circular points and light grey uncertainty bands are 5-year survey-weighted estimates of U5MR for years 1990–1994 up to 2015–2019 depending on survey timing. For a survey that ends in the middle of a 5-year period, we plot the estimates at the mid-point of the years in that interval for which the survey provides data. Black lines and corresponding intervals represent posterior medians and 95% uncertainty intervals respectively for the betabinomial model. IHME’s estimates and corresponding intervals, where we can compare, are in aquamarine.



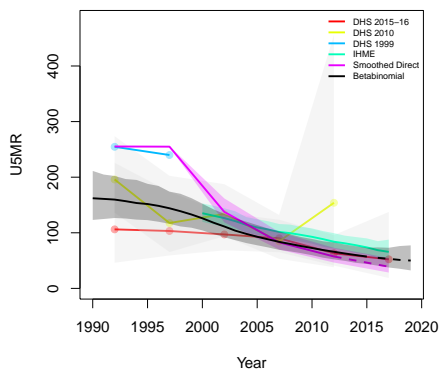
Iringa



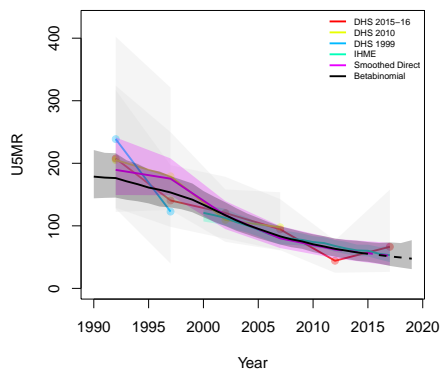
Kagera



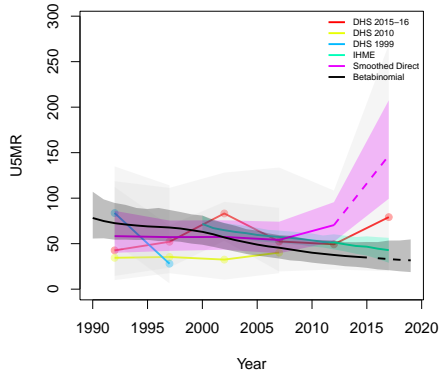
Katavi



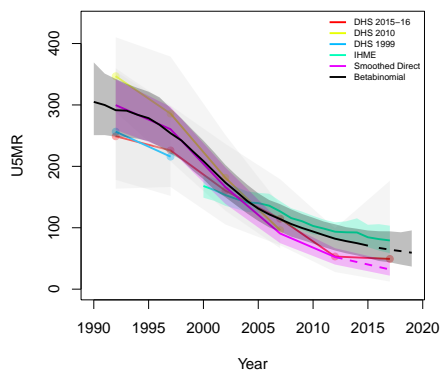
Kigoma

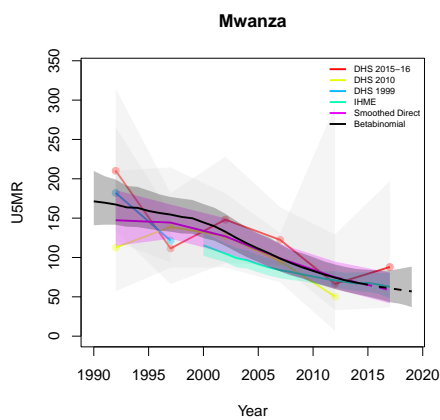
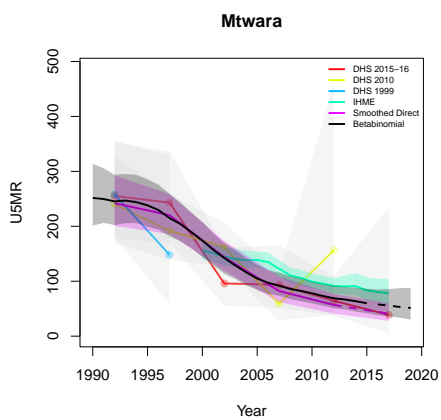
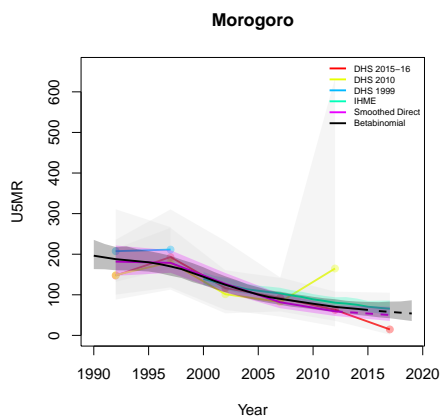
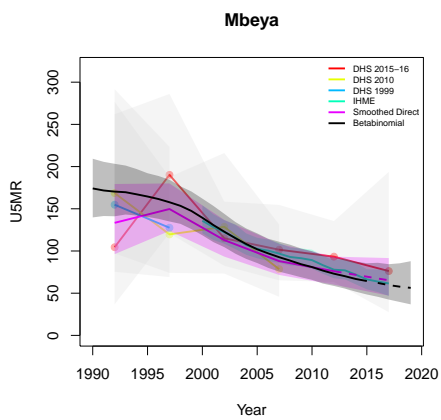
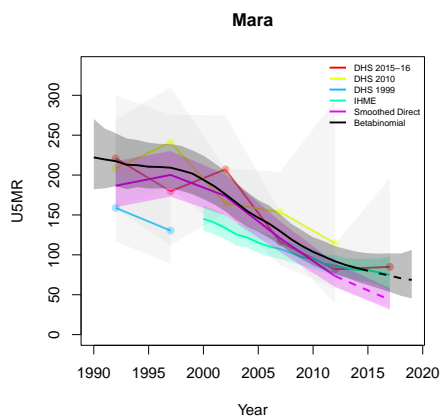
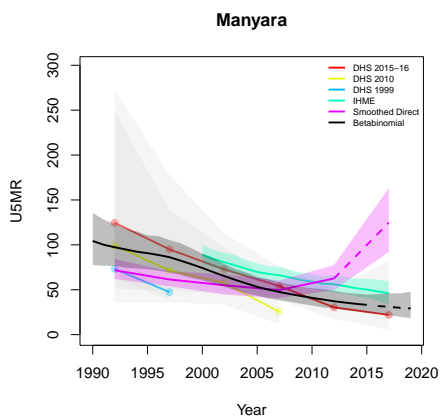


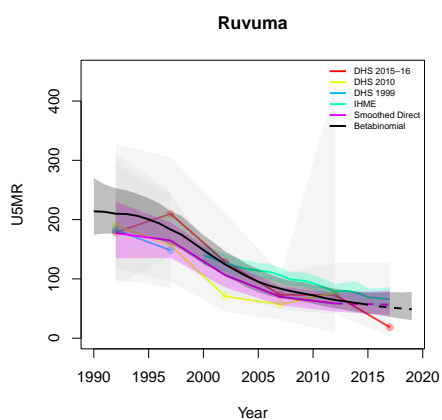
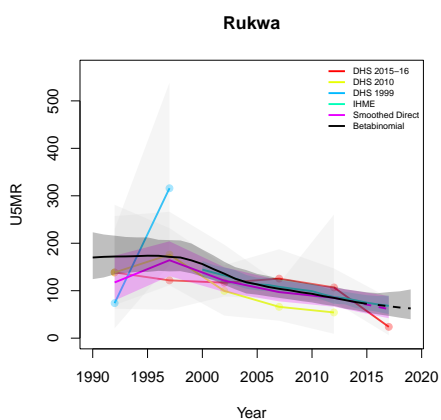
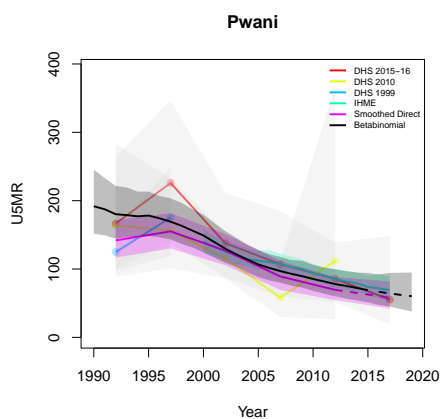
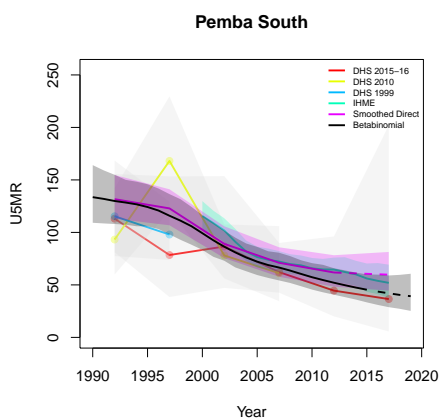
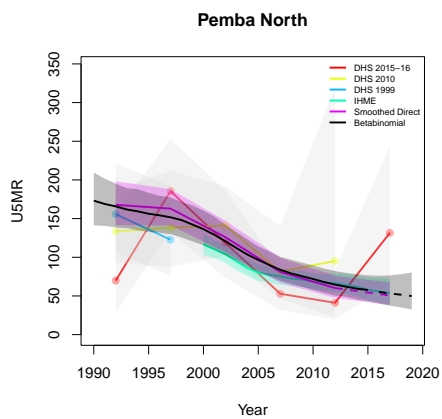
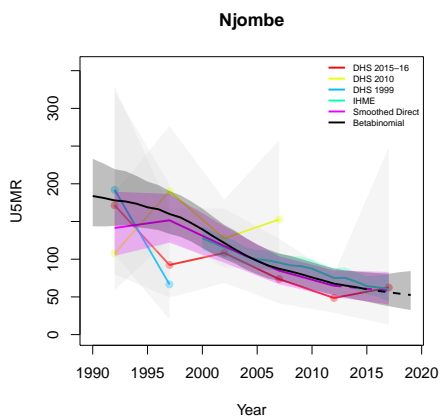
Kilimanjaro

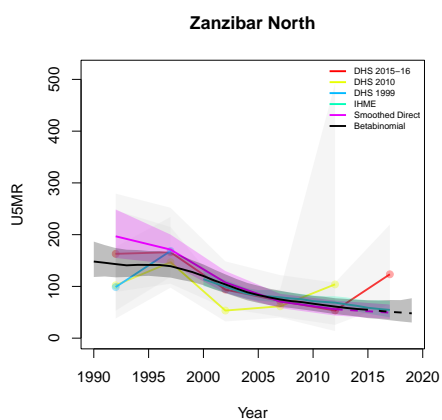
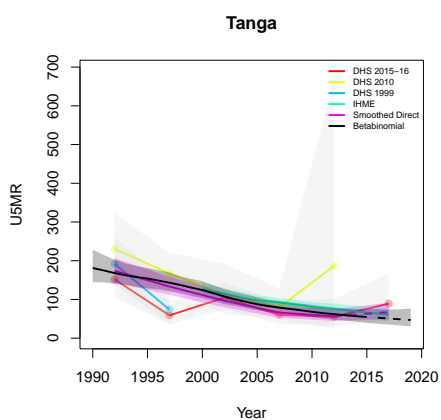
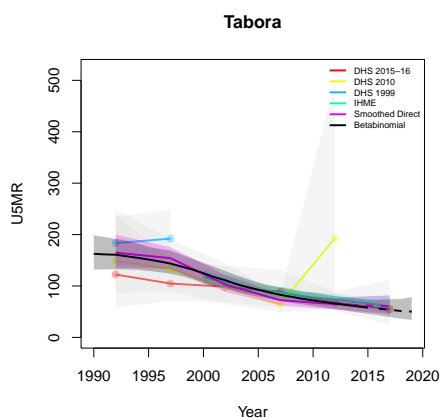
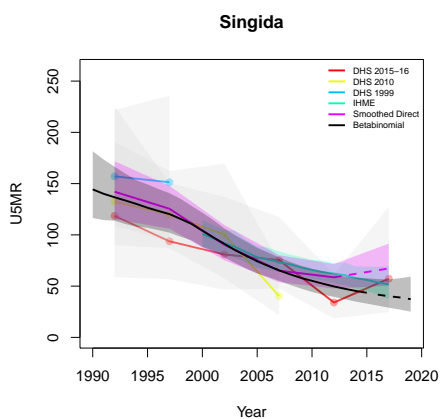
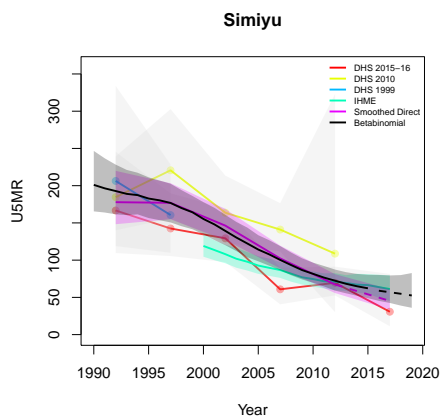
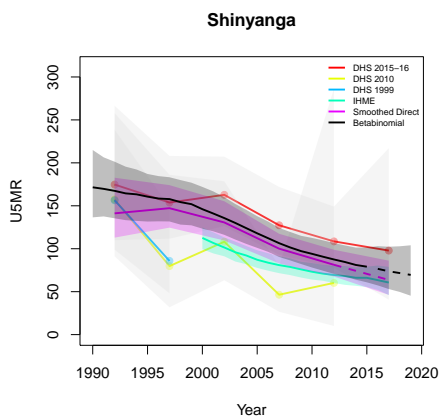


Lindi

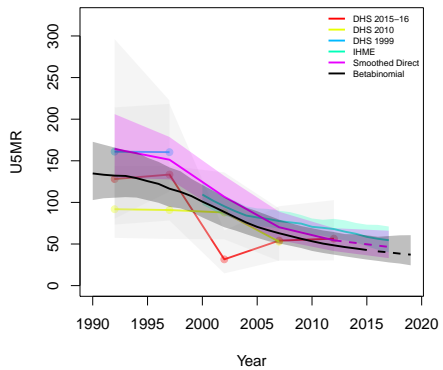




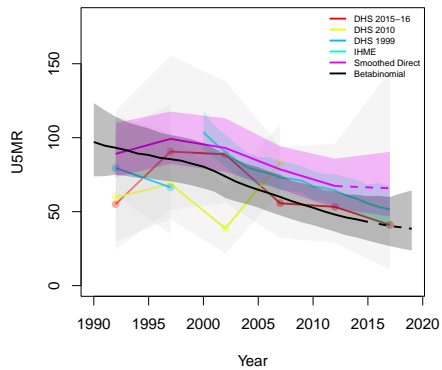




Zanzibar South and Central



Zanzibar West



*B.17.2 Admin-2*

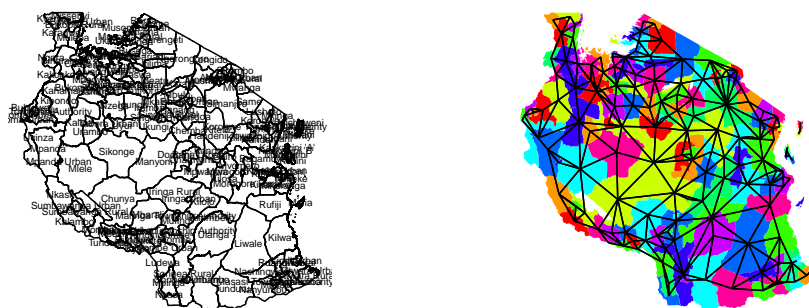


Figure B.85: **Left:** The names of the 169 Admin-2 areas of Tanzania . **Right:** The neighborhood structure of Admin-2 areas in Tanzania .

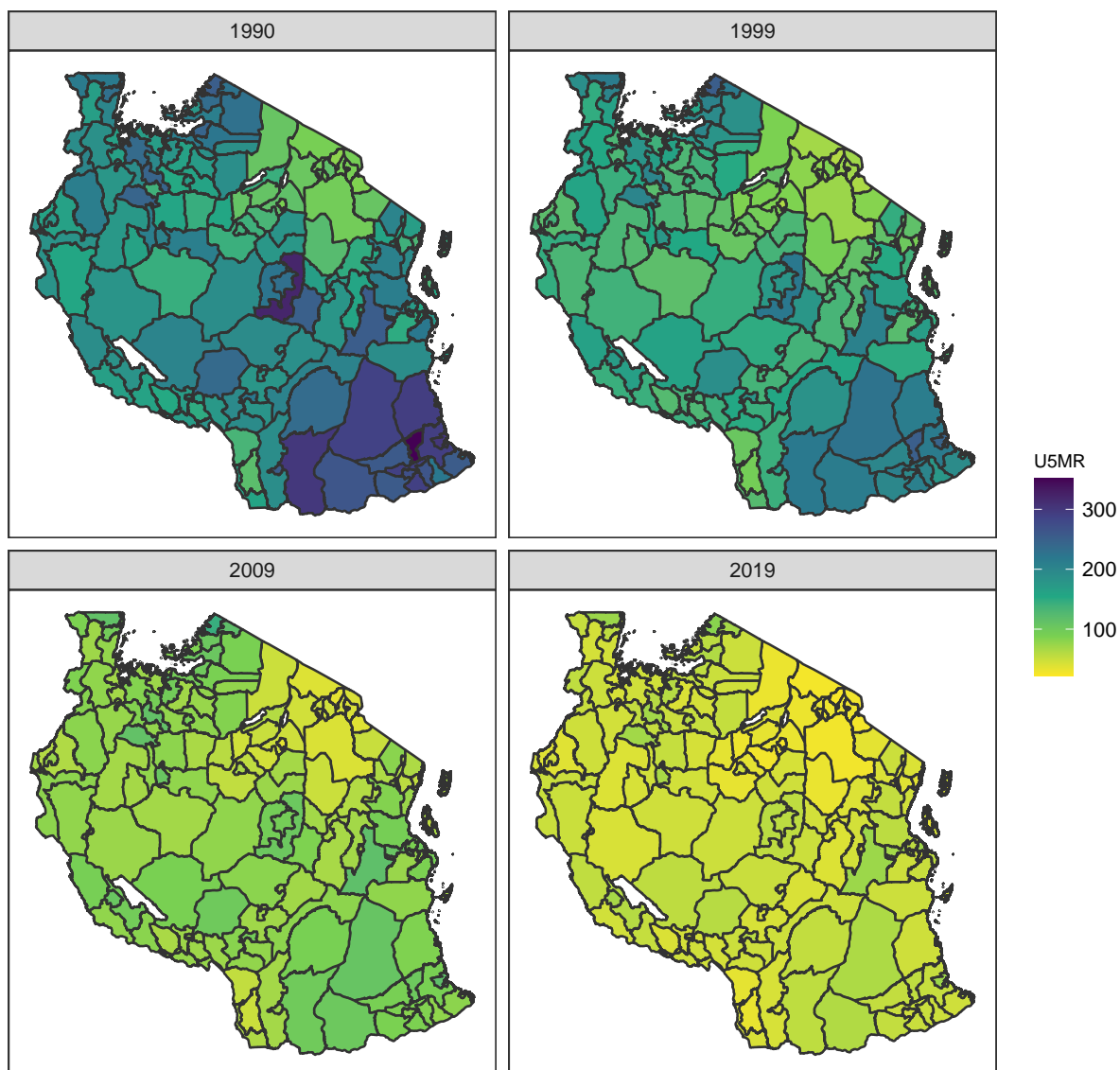


Figure B.86: Median U5MR estimates for years 1990, 1999, 2009, 2019 for Admin-2 areas in Tanzania .

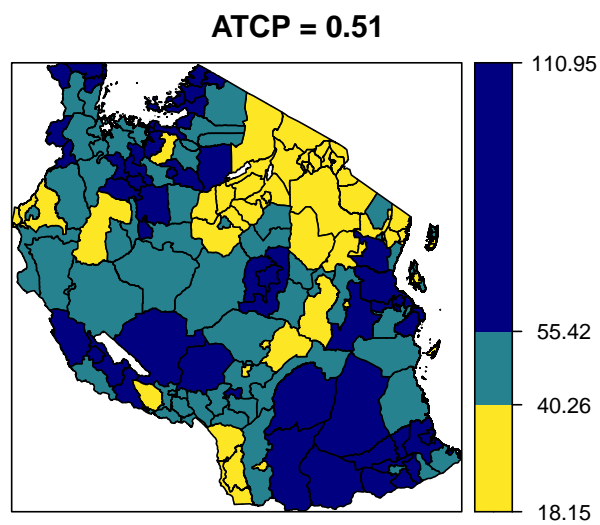
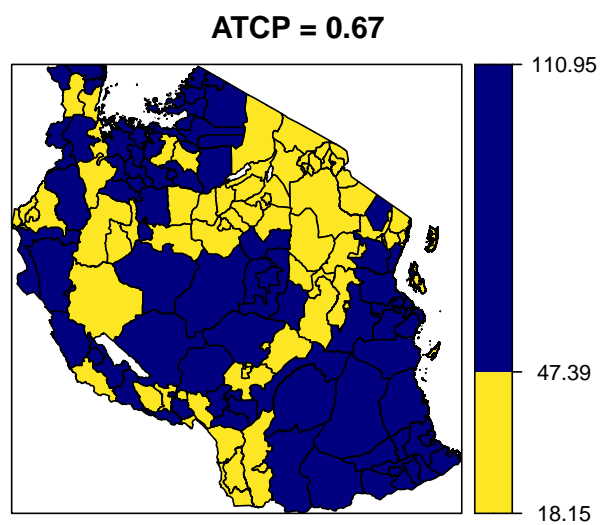
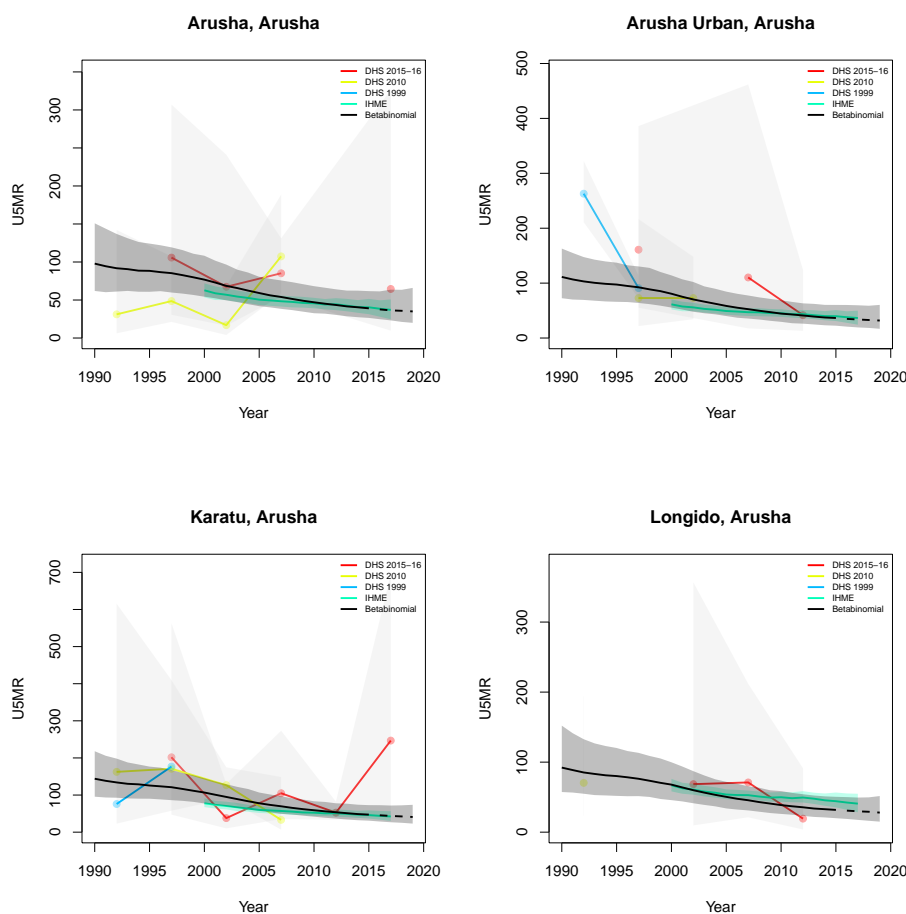


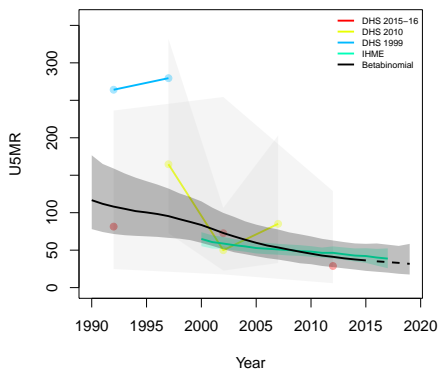
Figure B.87: Expression of uncertainty of U5MR (deaths per 1000 children) estimates for Admin-1 areas based on the average true classification probability (ATCP) in 2019 using  $K = 2, 3$  colors.

*Data and estimates over time by area*

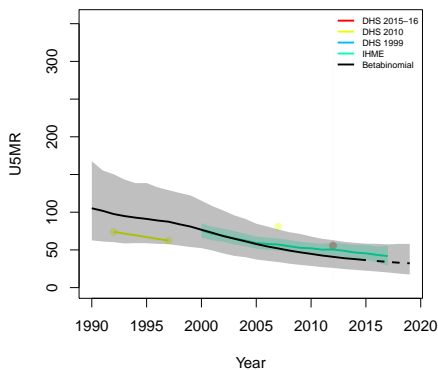
Colored lines with circular points and light grey uncertainty bands are 5-year survey-weighted estimates of U5MR for years 1990–1994 up to 2015–2019 depending on survey timing. For a survey that ends in the middle of a 5-year period, we plot the estimates at the mid-point of the years in that interval for which the survey provides data. Black lines and corresponding intervals represent posterior medians and 95% uncertainty intervals respectively for the betabinomial model. IHME’s estimates and corresponding intervals, where we can compare, are in aquamarine.



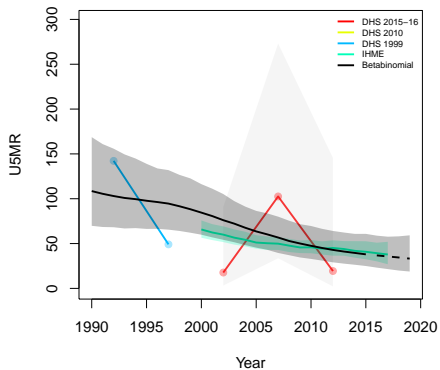
**Meru, Arusha**



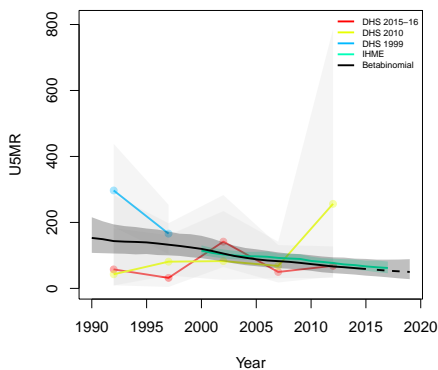
**Monduli, Arusha**



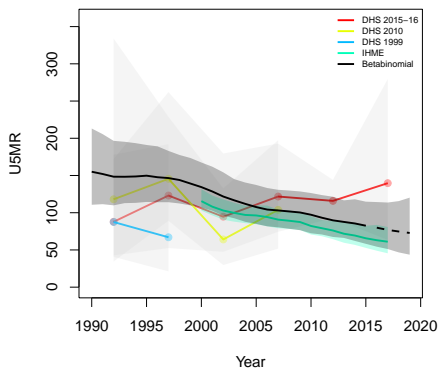
**Ngorongoro, Arusha**



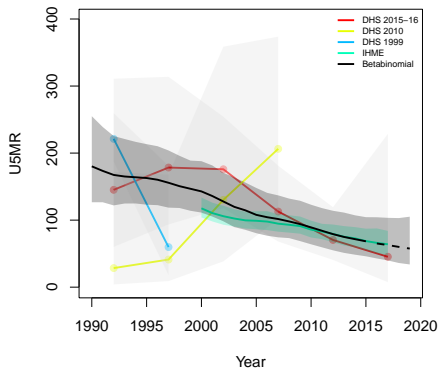
**Ilala, Dar es Salaam**



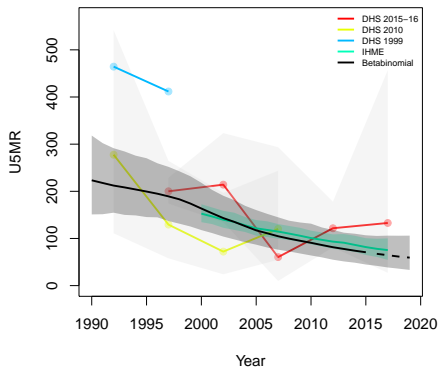
**Kinondoni, Dar es Salaam**



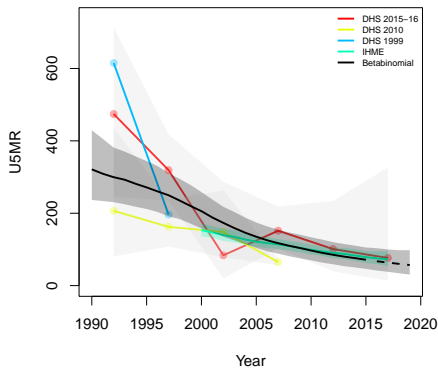
**Temeke, Dar es Salaam**



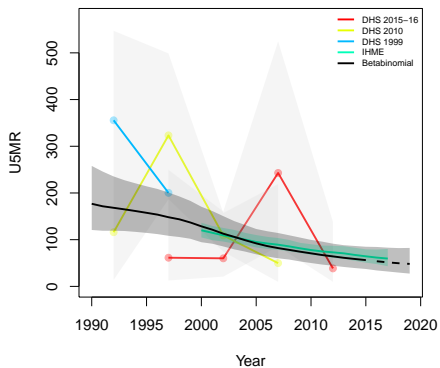
**Bahi, Dodoma**



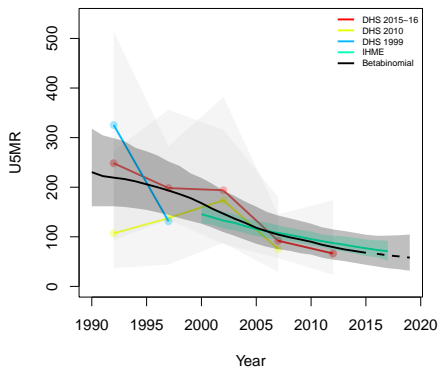
**Chamwino, Dodoma**



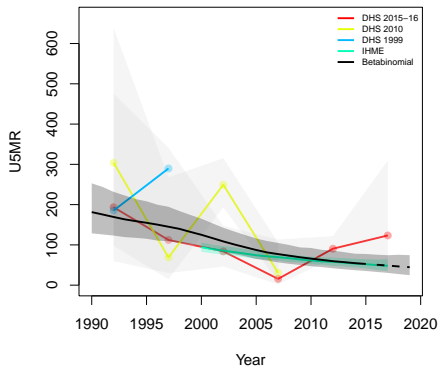
**Chemba, Dodoma**



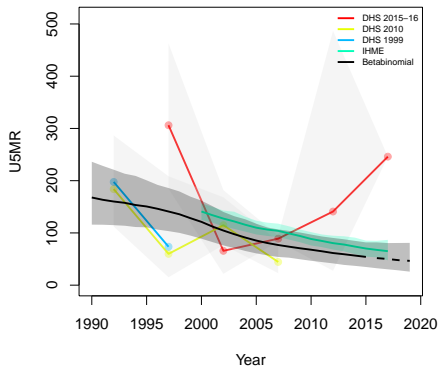
**Dodoma Urban, Dodoma**



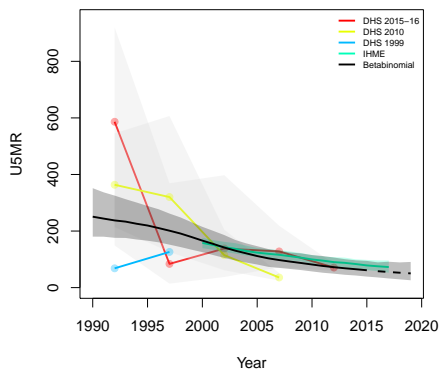
**Kondoa, Dodoma**



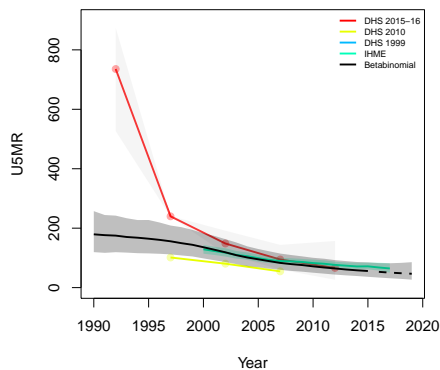
**Kongwa, Dodoma**



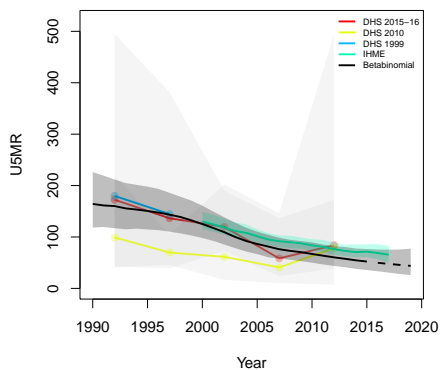
**Mpwapwa, Dodoma**



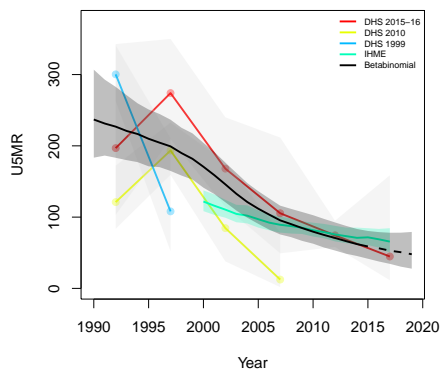
**Bukombe, Geita**



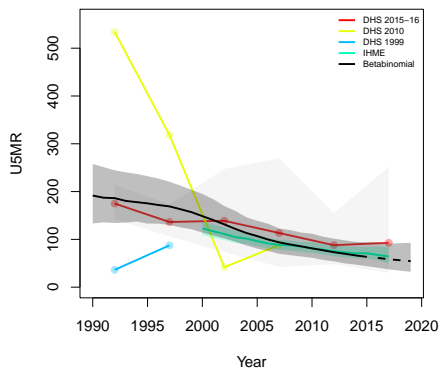
**Chato, Geita**



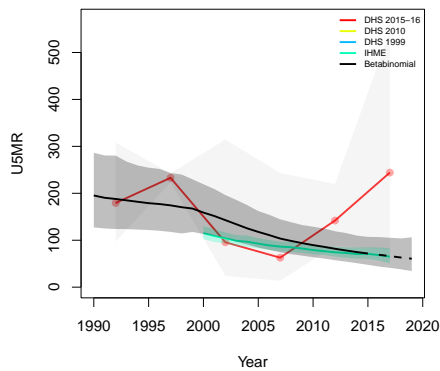
**Geita, Geita**



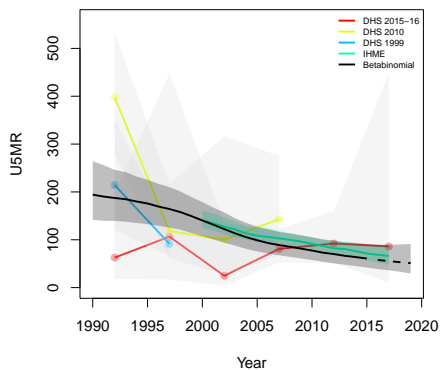
**Mbogwe, Geita**



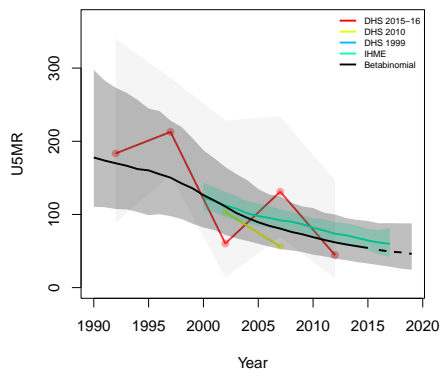
**Nyang'wale, Geita**



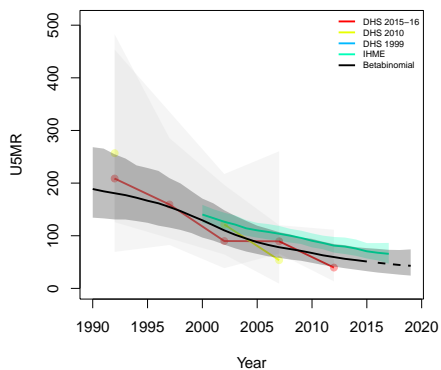
**Iringa Rural, Iringa**



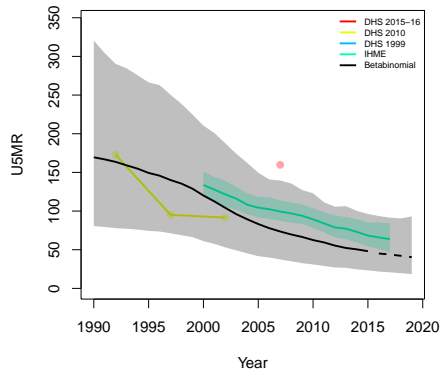
**Iringa Urban, Iringa**



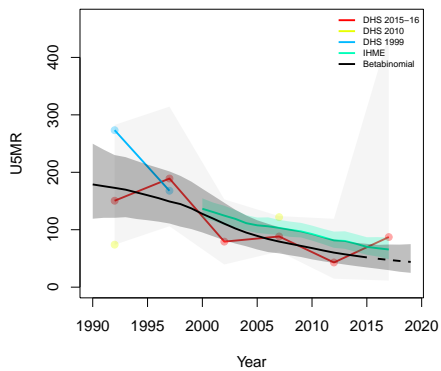
**Kilolo, Iringa**



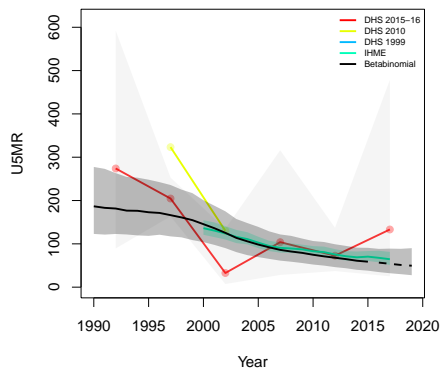
**Mafinga Township Authority, Iringa**



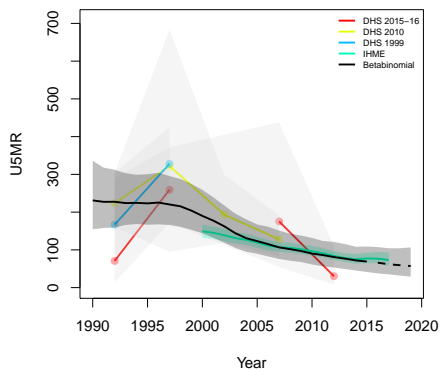
**Mufindi, Iringa**



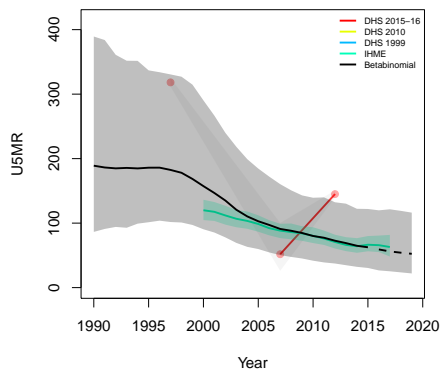
**Biharamulo, Kagera**



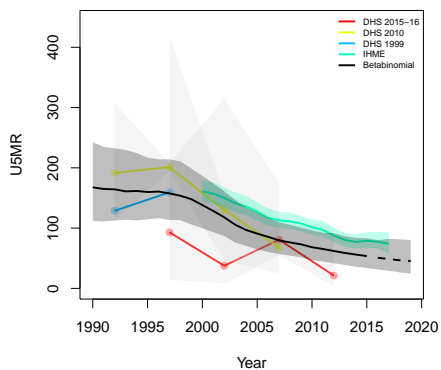
**Bukoba Rural, Kagera**



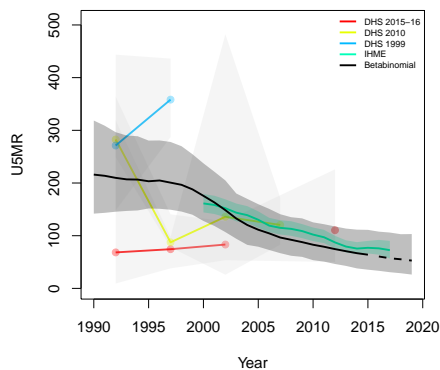
**Bukoba Urban, Kagera**



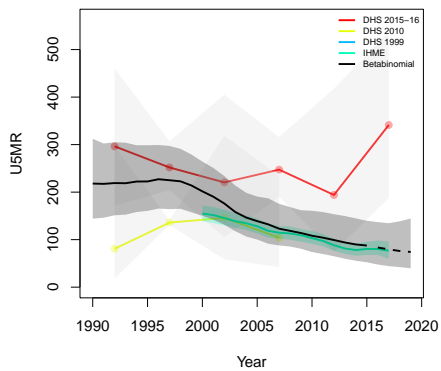
**Karagwe, Kagera**



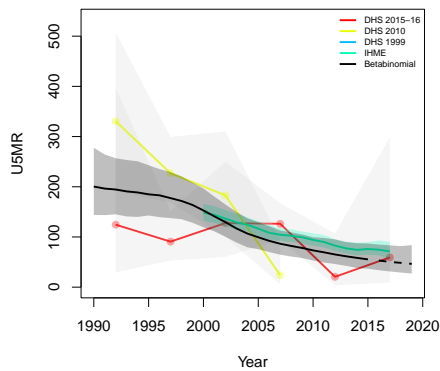
**Kyerwa, Kagera**



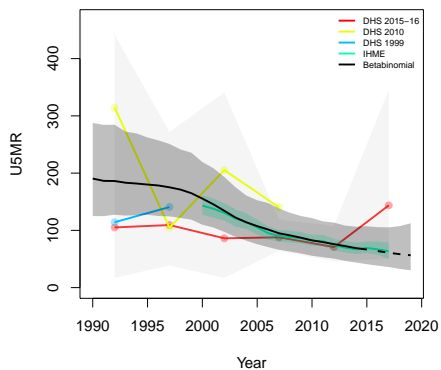
**Missenyi, Kagera**



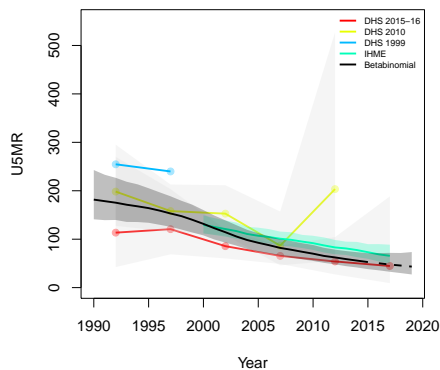
**Muleba, Kagera**



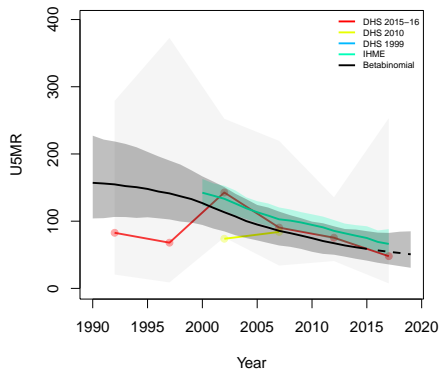
**Ngara, Kagera**



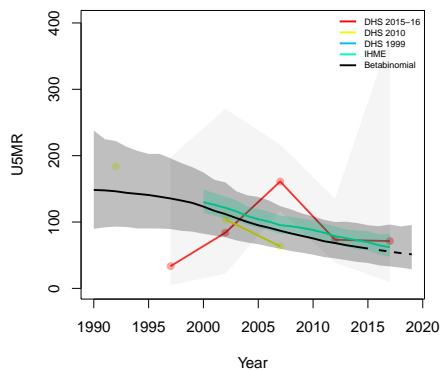
**Mlele, Katavi**



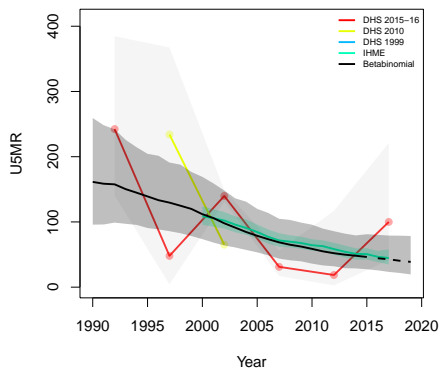
**Mpanda, Katavi**



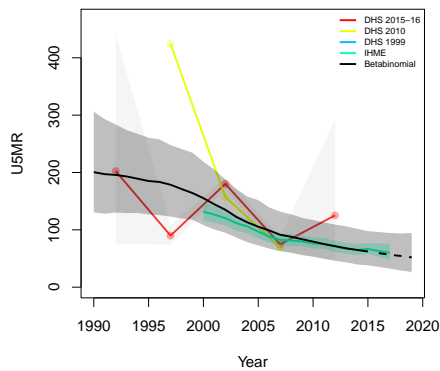
**Mpanda Urban, Katavi**



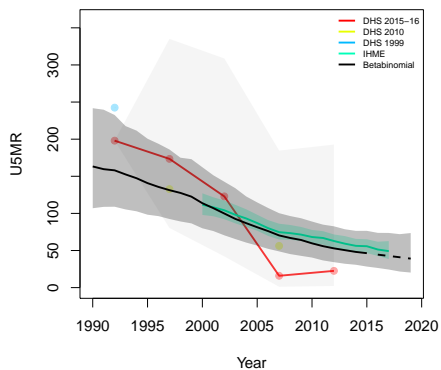
**Buhigwe, Kigoma**



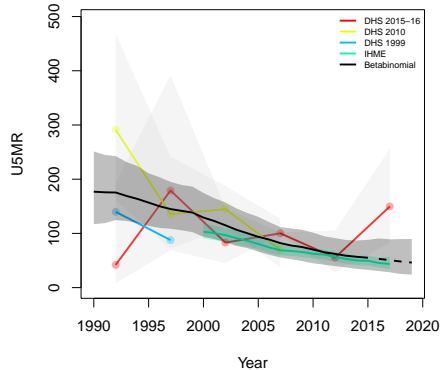
**Kakonko, Kigoma**



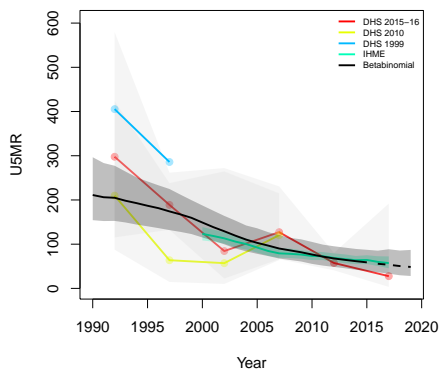
**Kasulu, Kigoma**



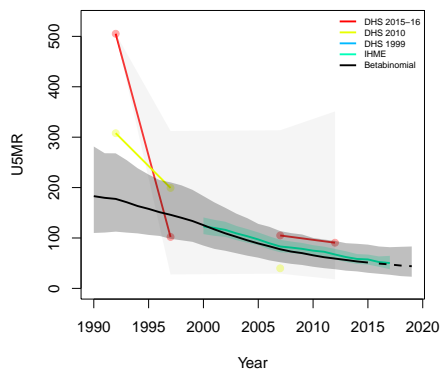
**Kasulu Township Authority, Kigoma**



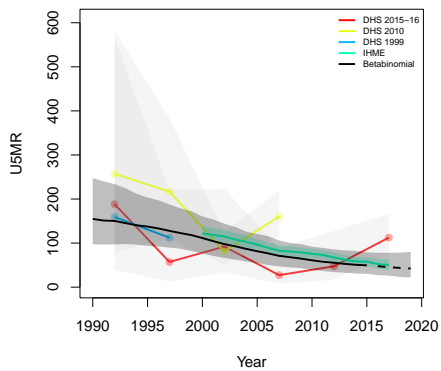
**Kibondo, Kigoma**



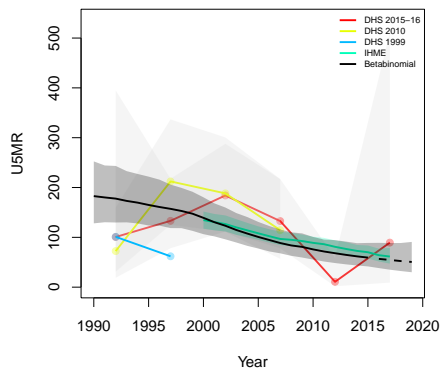
**Kigoma Rural, Kigoma**



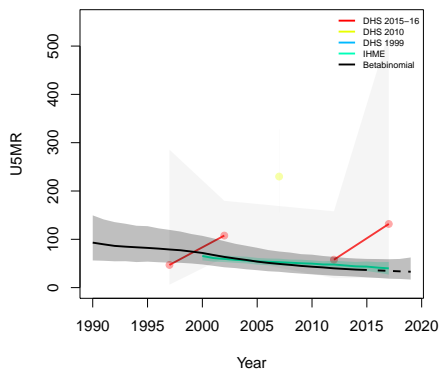
**Kigoma Urban, Kigoma**



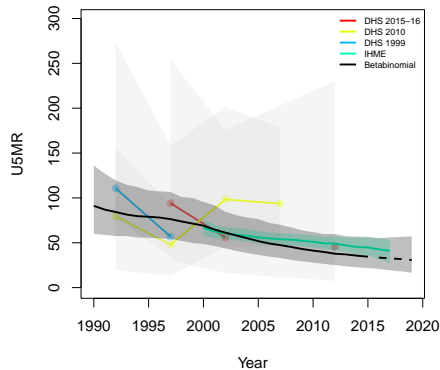
**Uvinza, Kigoma**



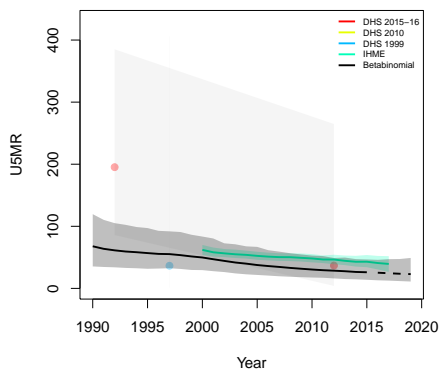
**Hai, Kilimanjaro**



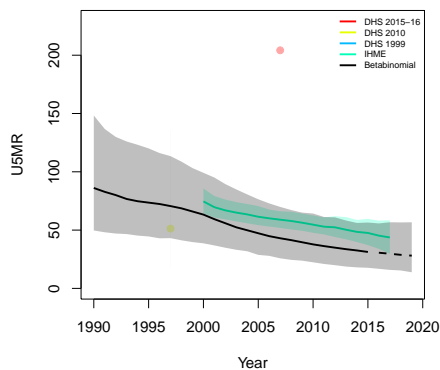
**Moshi Rural, Kilimanjaro**



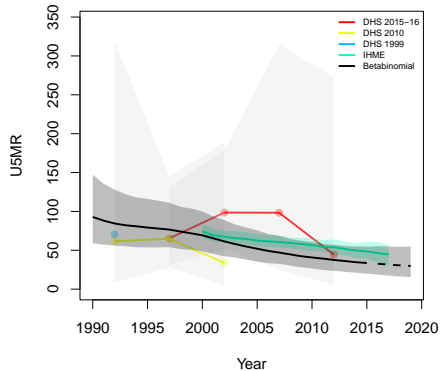
**Moshi Urban, Kilimanjaro**



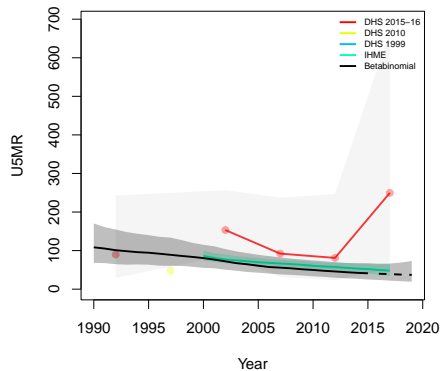
**Mwanga, Kilimanjaro**



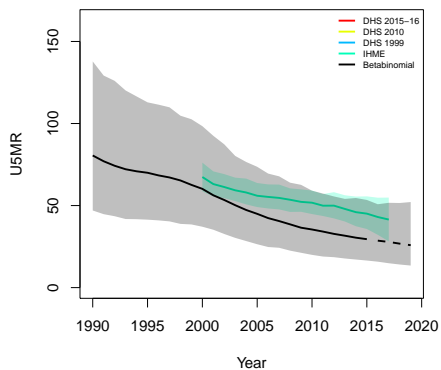
**Rombo, Kilimanjaro**



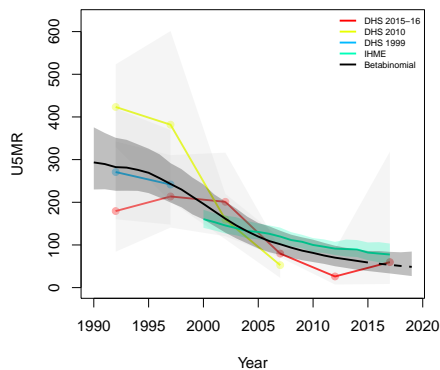
**Same, Kilimanjaro**



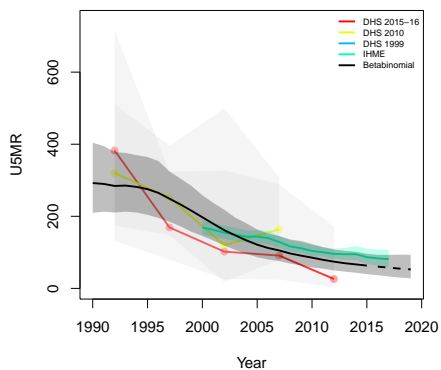
**Siha, Kilimanjaro**



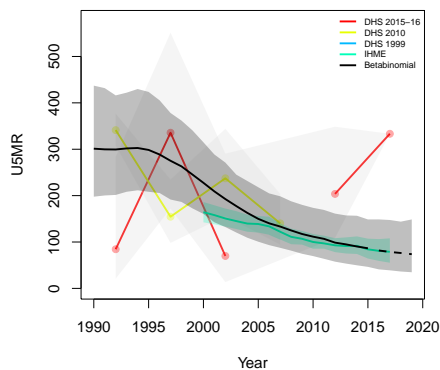
**Kilwa, Lindi**



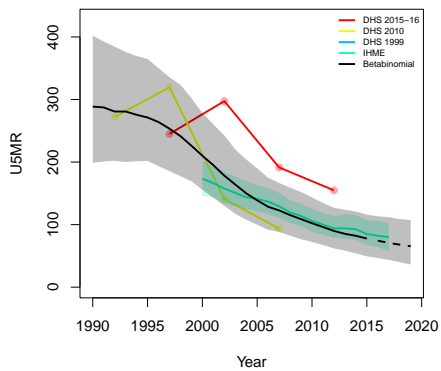
**Lindi Rural, Lindi**



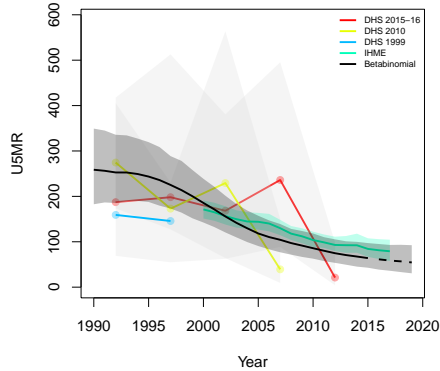
**Lindi Urban, Lindi**



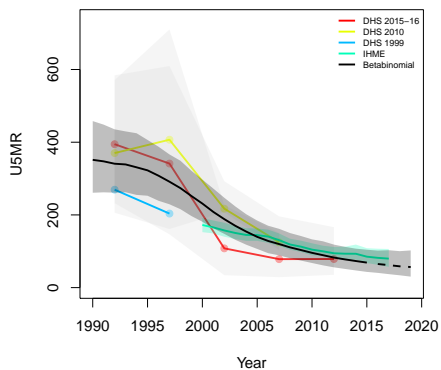
**Liwale, Lindi**



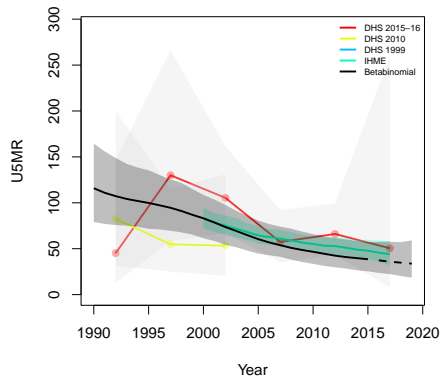
**Nachingwea, Lindi**



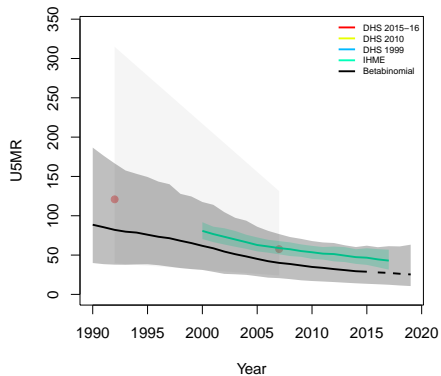
**Ruangwa, Lindi**



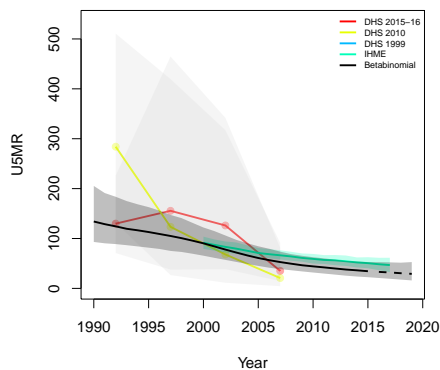
**Babati, Manyara**



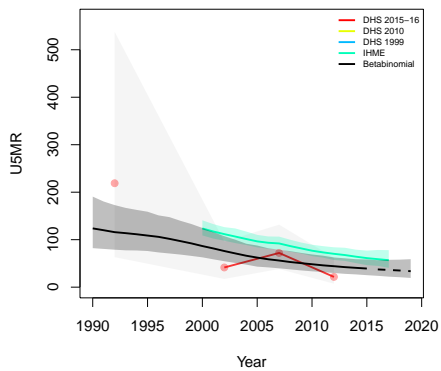
**Babati Urban, Manyara**



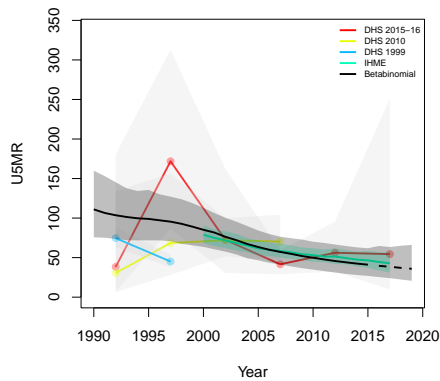
**Hanang, Manyara**



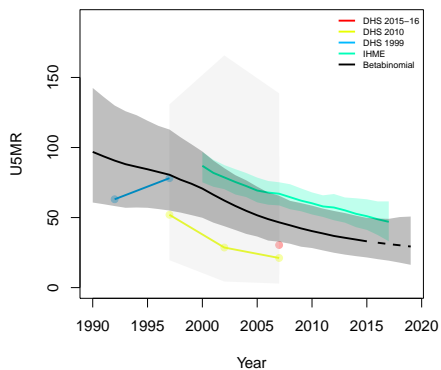
**Kiteto, Manyara**



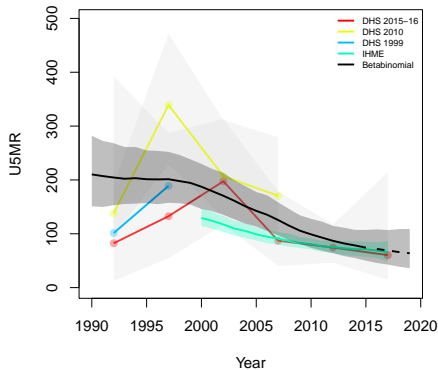
**Mbulu, Manyara**



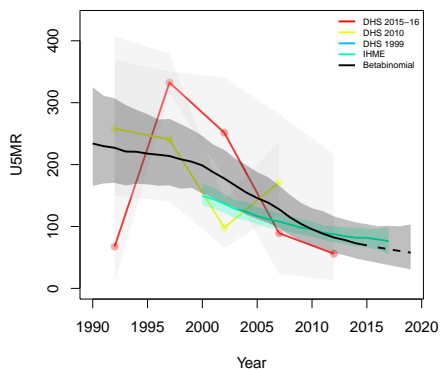
**Simanjiro, Manyara**



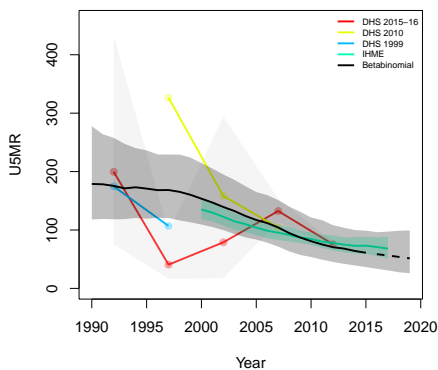
**Bunda, Mara**



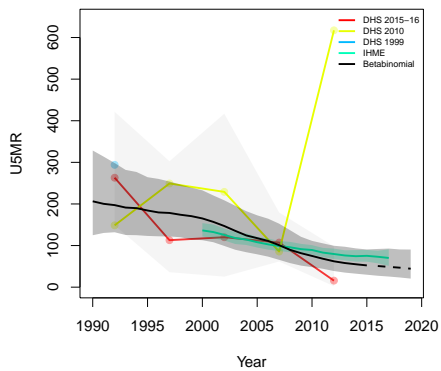
**Butiama, Mara**



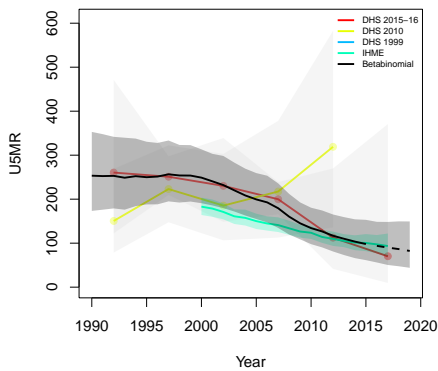
**Musoma Rural, Mara**



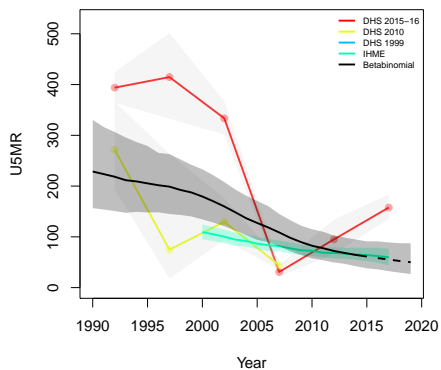
**Musoma Urban, Mara**



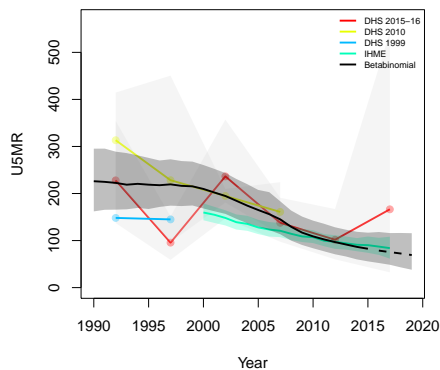
**Rorya, Mara**



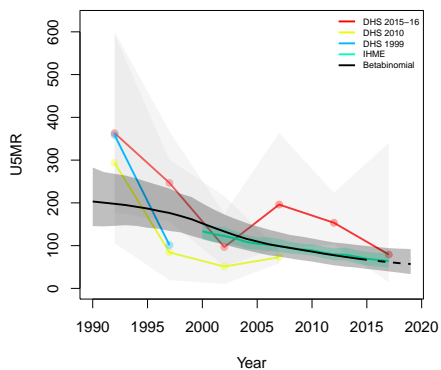
**Serengeti, Mara**



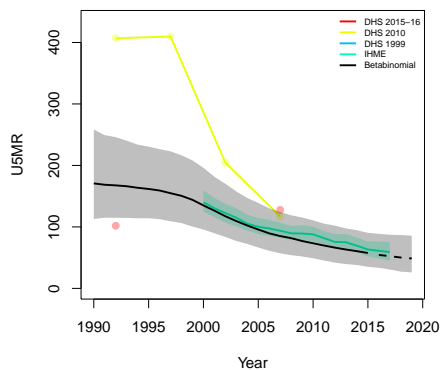
**Tarime, Mara**



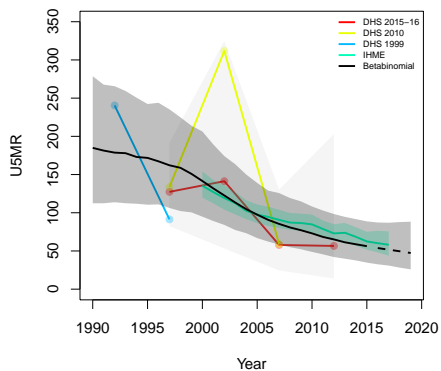
**Chunya, Mbeya**



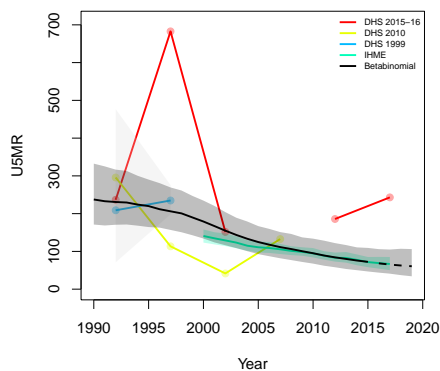
**Ileje, Mbeya**

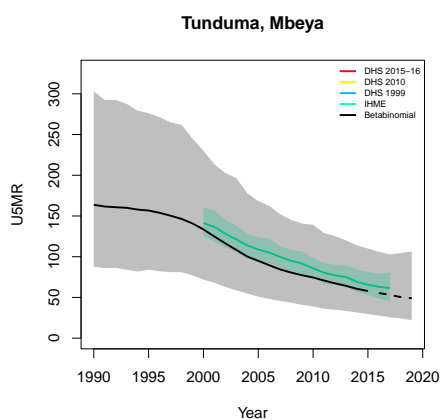
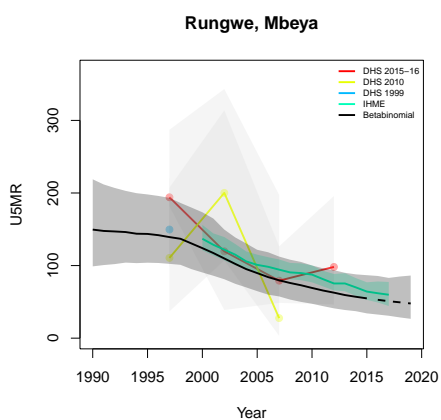
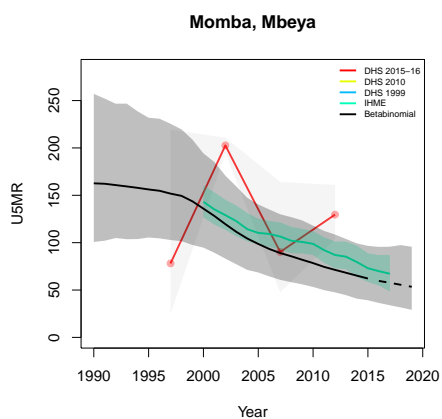
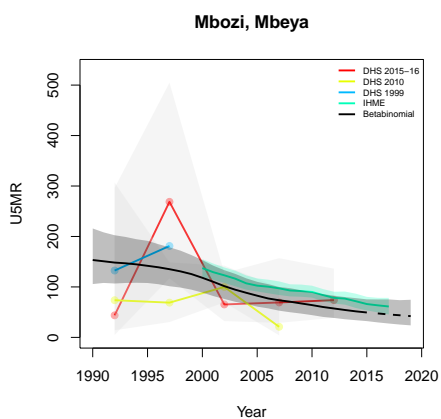
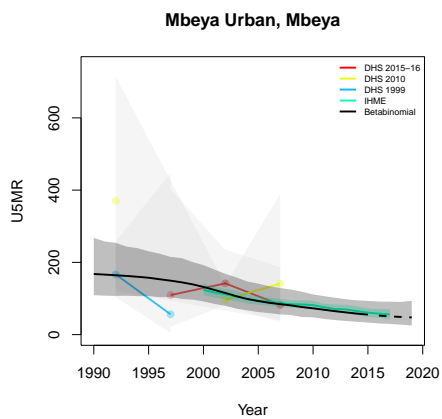
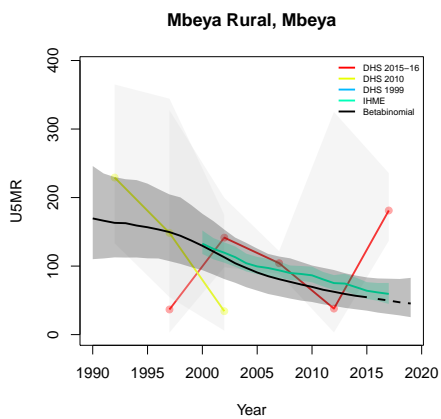


**Kyela, Mbeya**

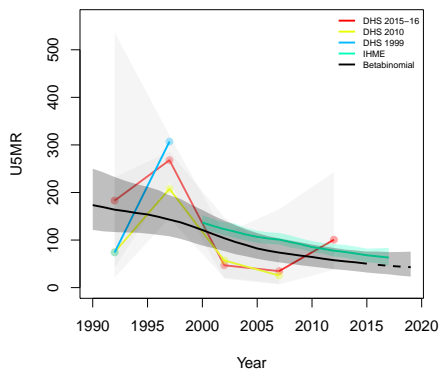


**Mbarali, Mbeya**

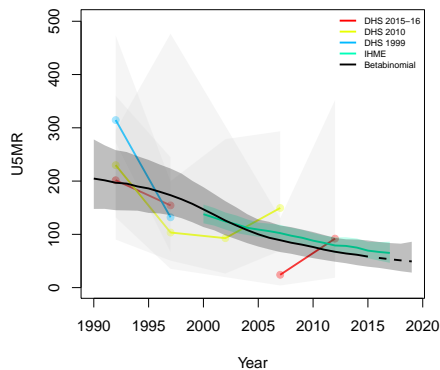




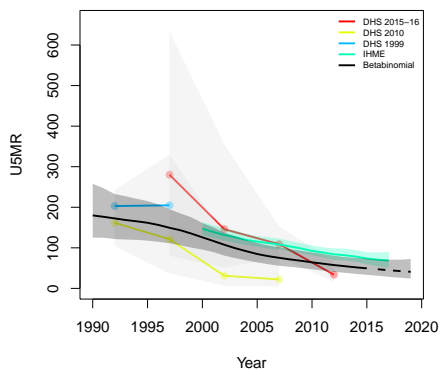
**Gairo, Morogoro**



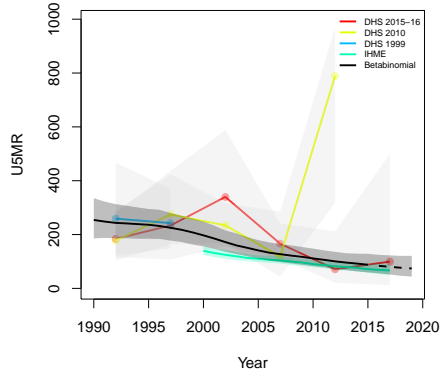
**Kilombero, Morogoro**



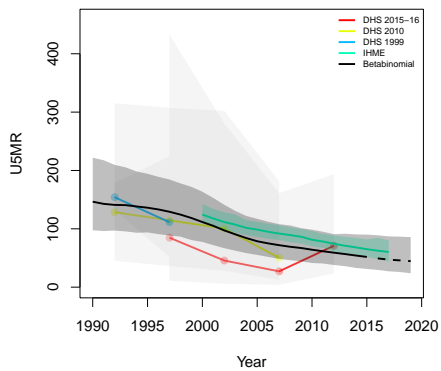
**Kilosa, Morogoro**



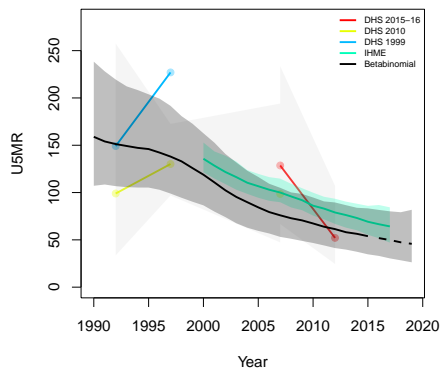
**Morogoro Rural, Morogoro**

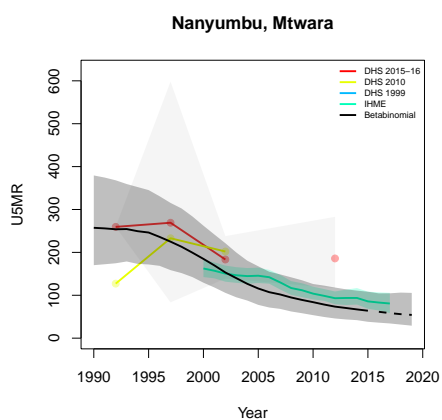
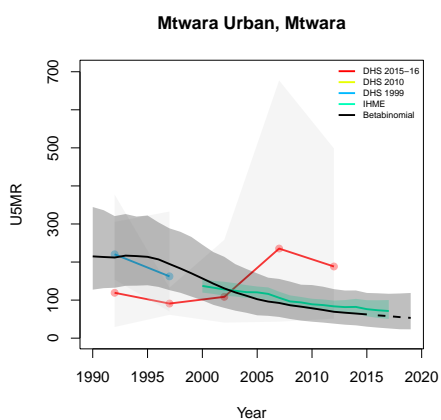
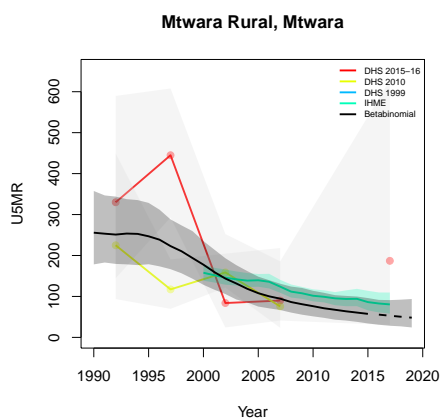
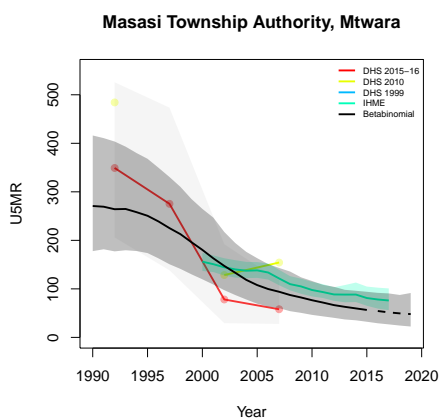
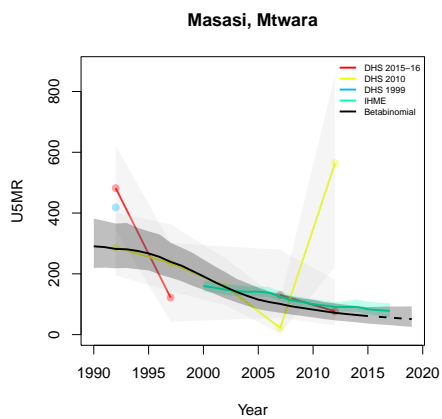
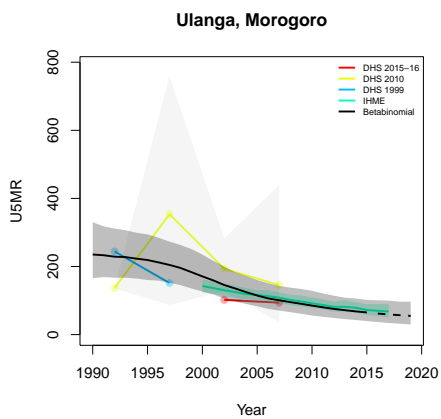


**Morogoro Urban, Morogoro**

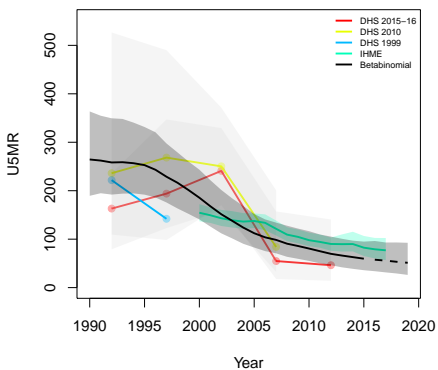


**Mvomero, Morogoro**

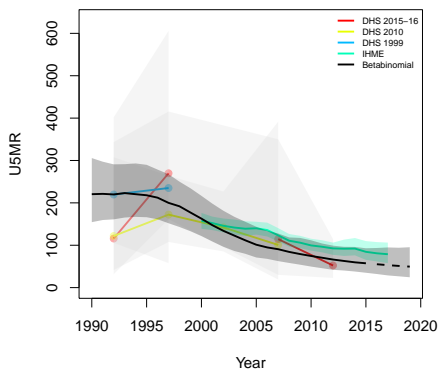




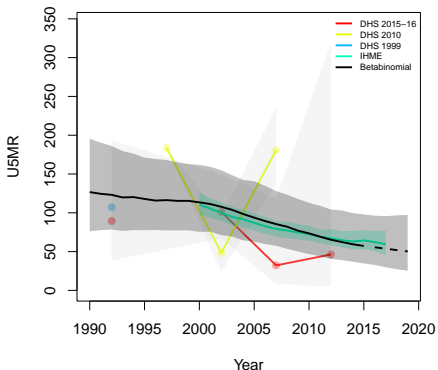
**Newala, Mtwara**



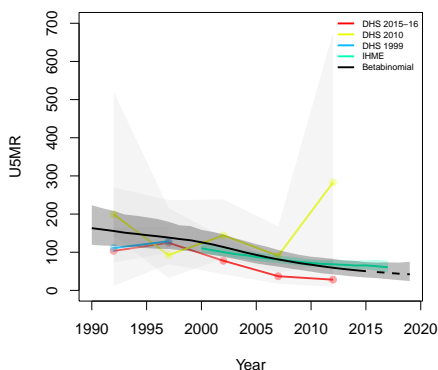
**Tandahimba, Mtwara**



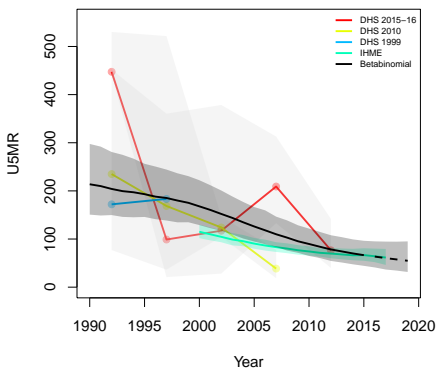
**Ilemela, Mwanza**



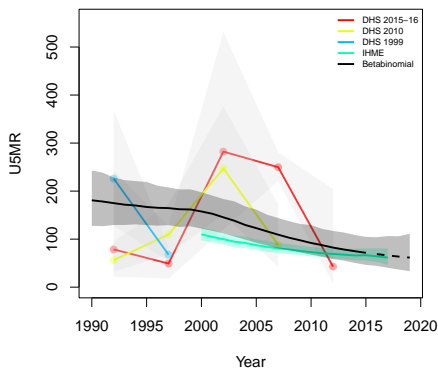
**Kwimba, Mwanza**



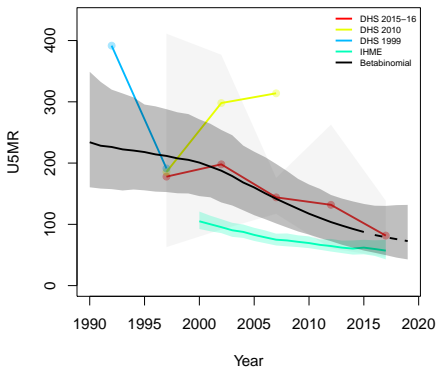
**Magu, Mwanza**



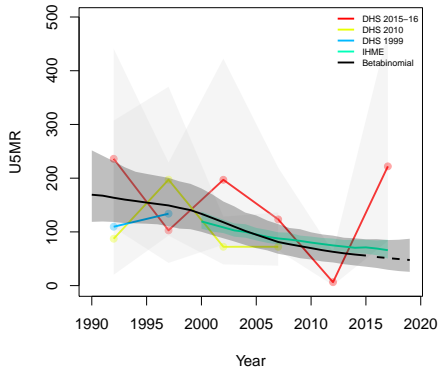
**Misungwi, Mwanza**



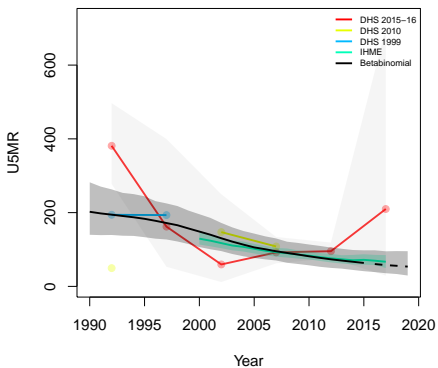
**Nyamagana, Mwanza**



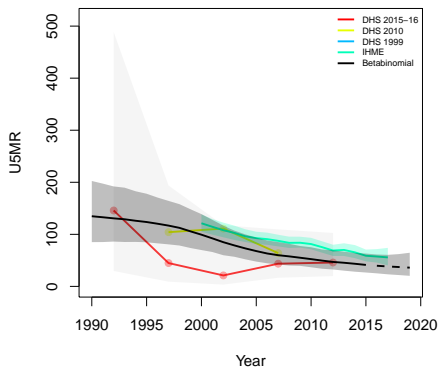
**Sengerema, Mwanza**



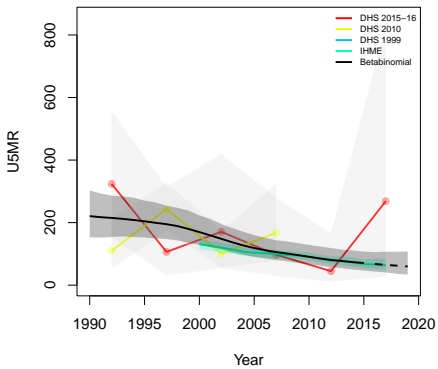
**Ukerewe, Mwanza**



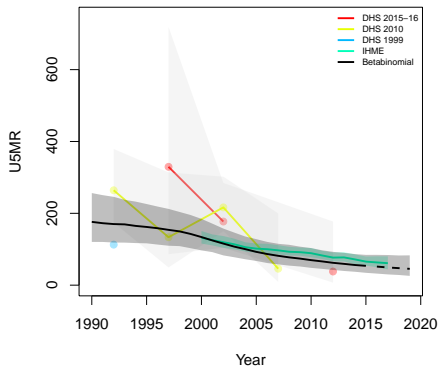
**Ludewa, Njombe**



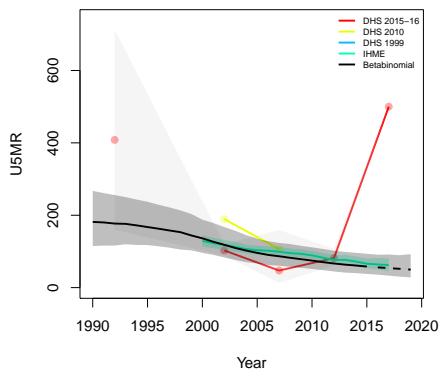
**Makambako Township Authority, Njombe**



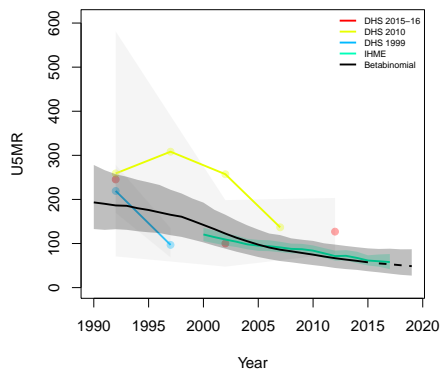
**Makete, Njombe**



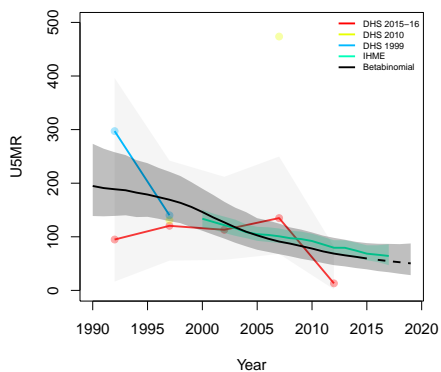
**Njombe, Njombe**



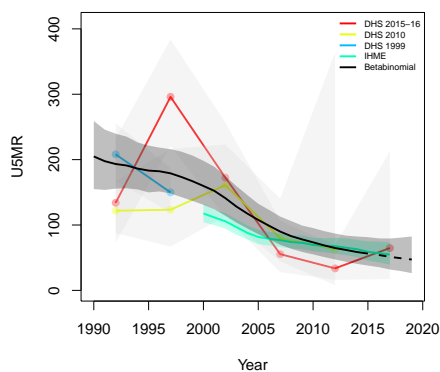
**Njombe Urban, Njombe**



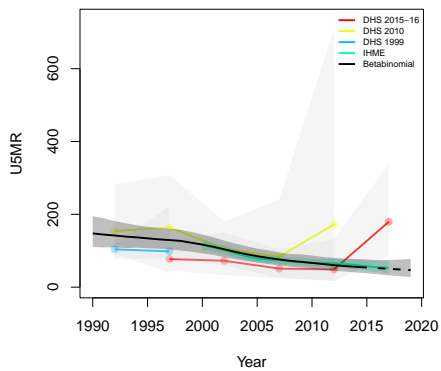
**Wanging'ombe, Njombe**



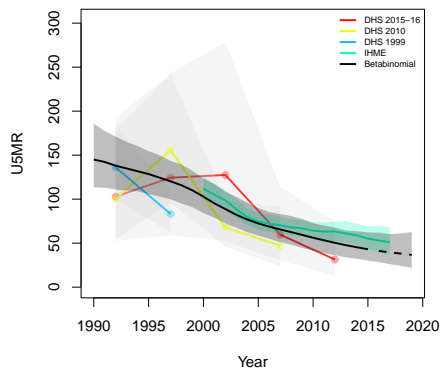
**Micheweni, Pemba North**



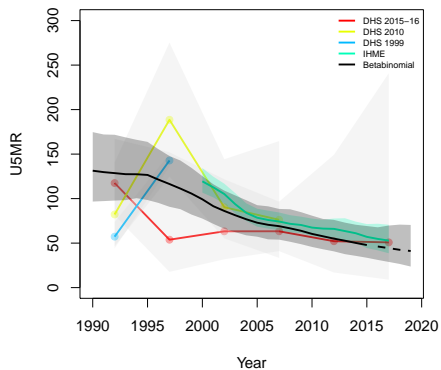
**Wete, Pemba North**



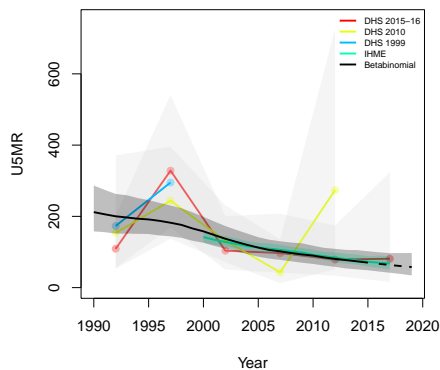
**Chake, Pemba South**



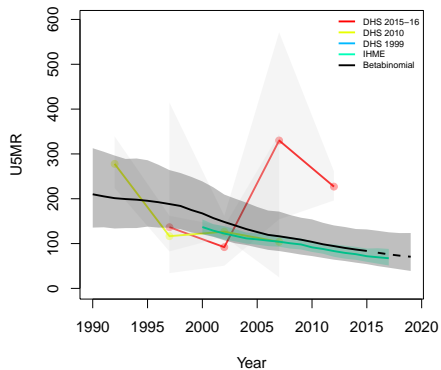
**Mkoani, Pemba South**



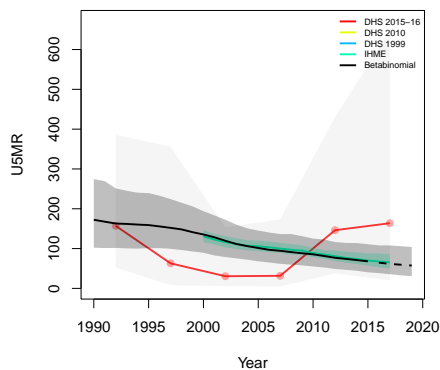
**Bagamoyo, Pwani**



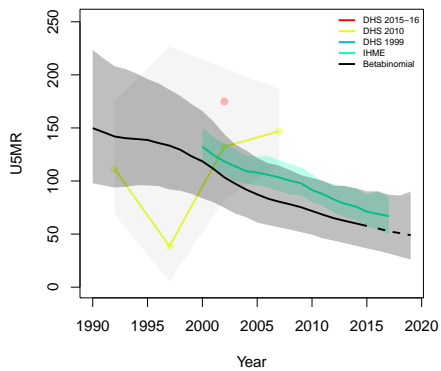
**Kibaha, Pwani**



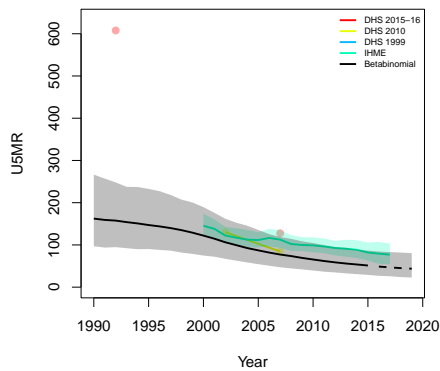
**Kibaha Urban, Pwani**

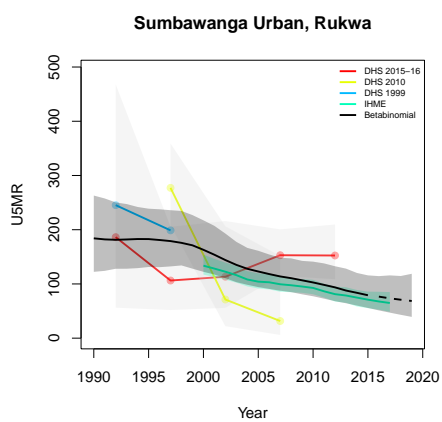
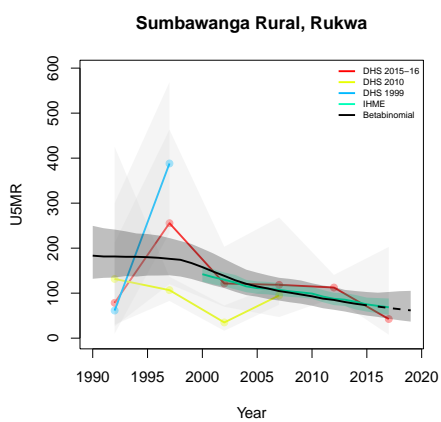
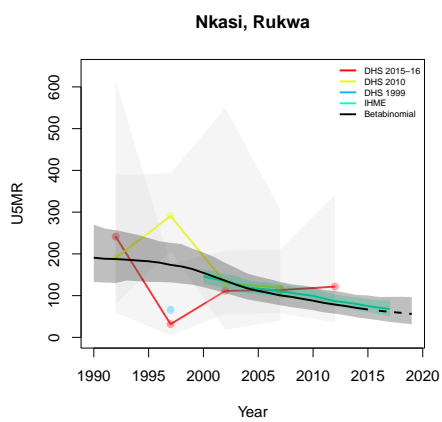
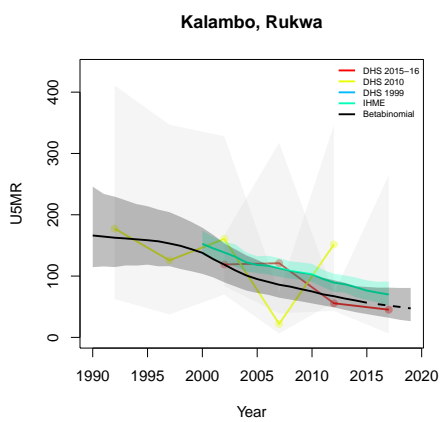
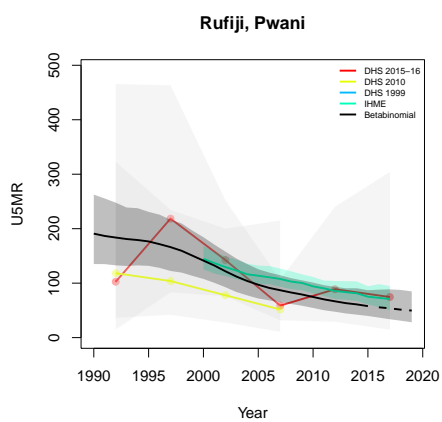
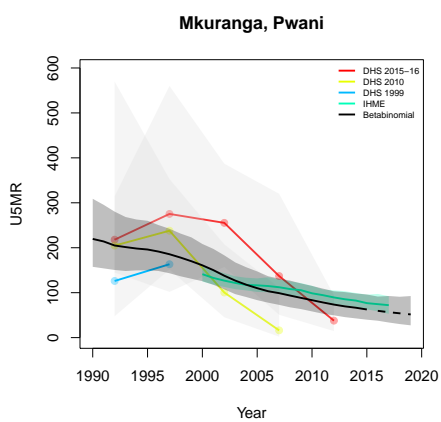


**Kisarawe, Pwani**

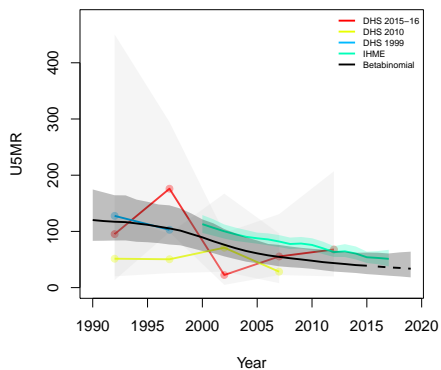


**Mafia, Pwani**

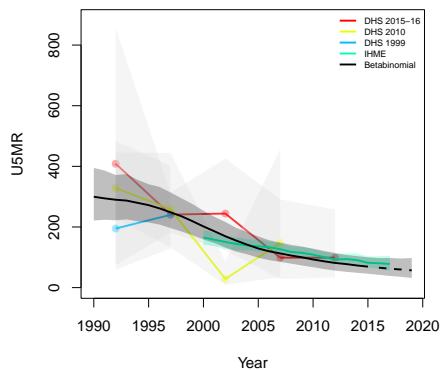




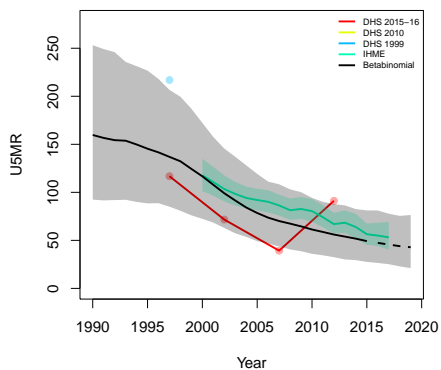
**Mbinga, Ruvuma**



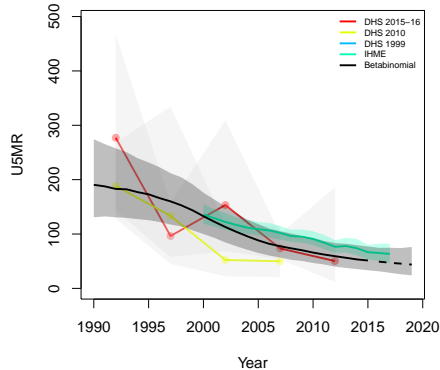
**Namtumbo, Ruvuma**



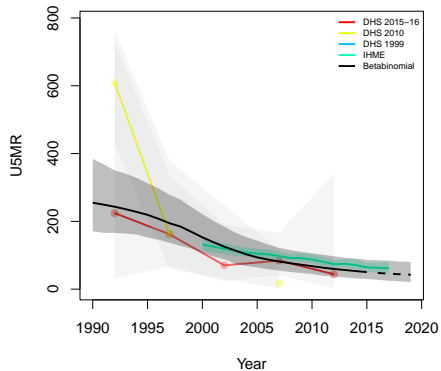
**Nyasa, Ruvuma**



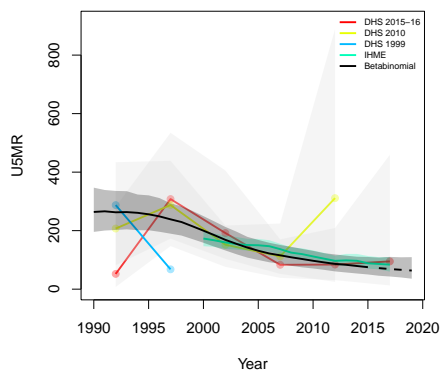
**Songea Rural, Ruvuma**

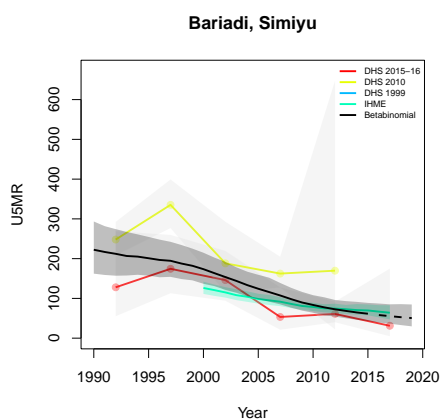
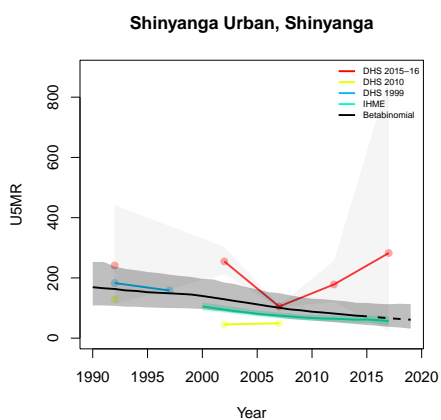
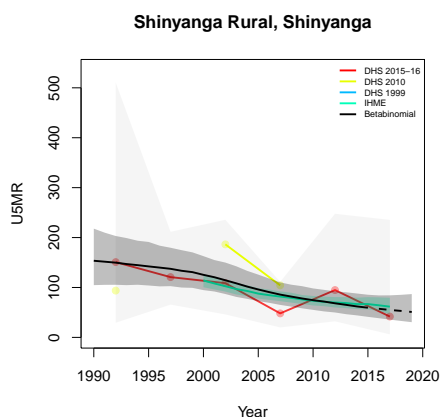
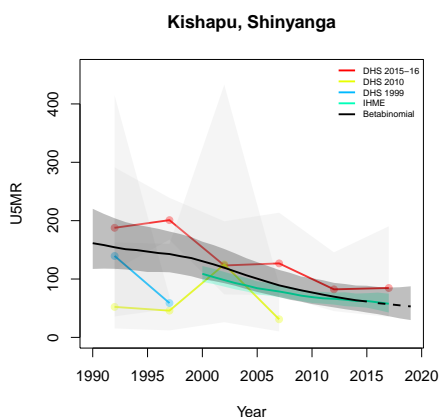
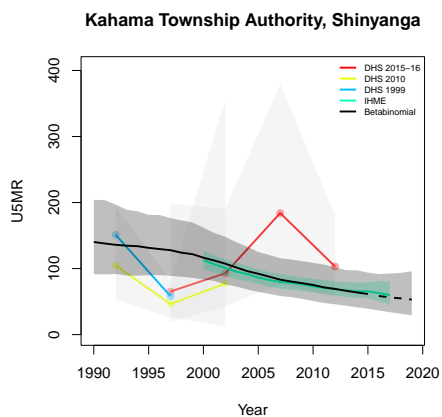
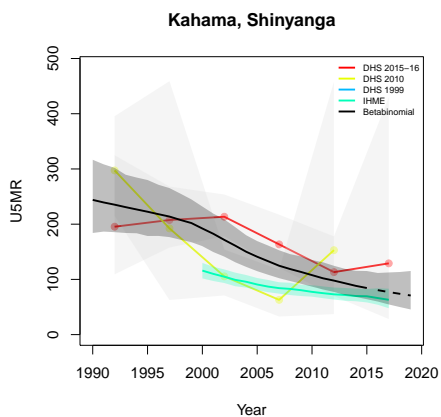


**Songea Urban, Ruvuma**

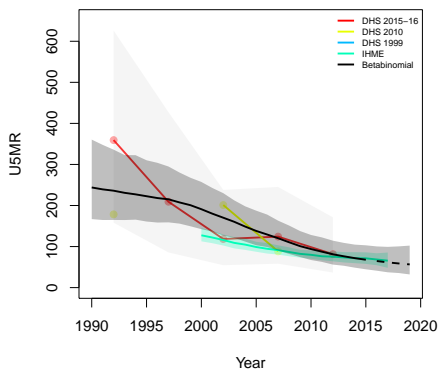


**Tunduru, Ruvuma**

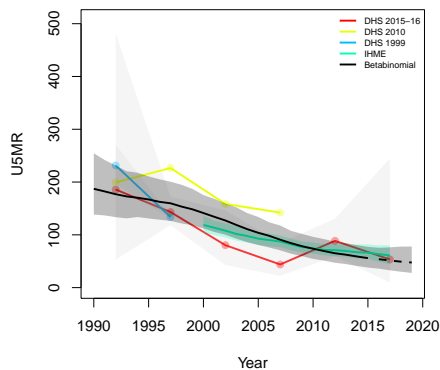




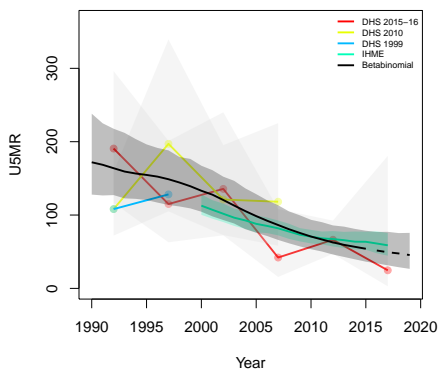
**Busega, Simiyu**



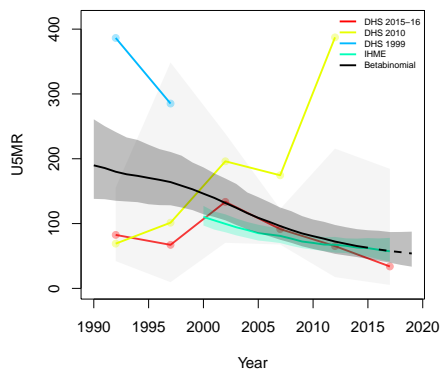
**Itilima, Simiyu**



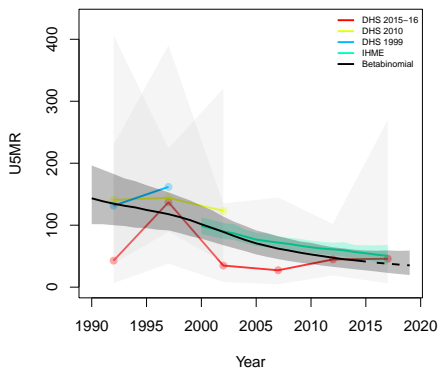
**Maswa, Simiyu**



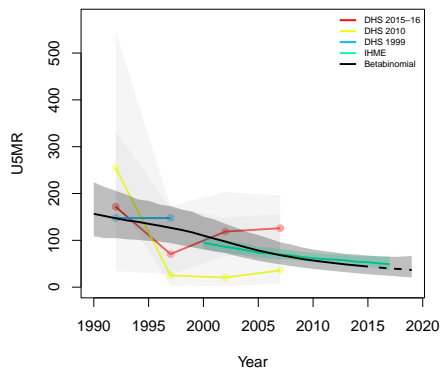
**Meatu, Simiyu**



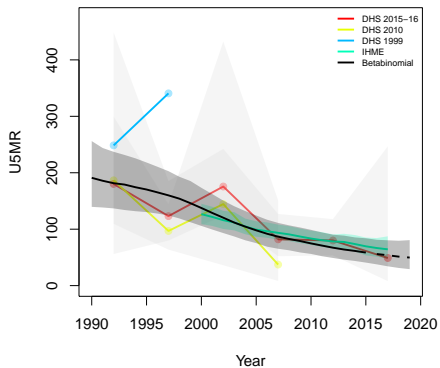
**Ikungi, Singida**



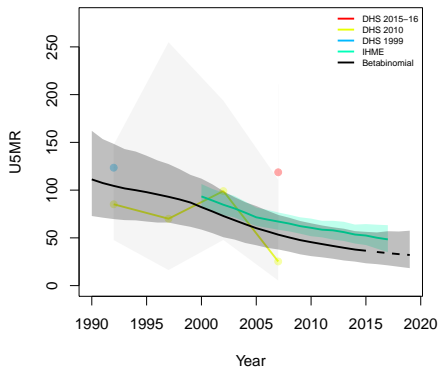
**Iramba, Singida**



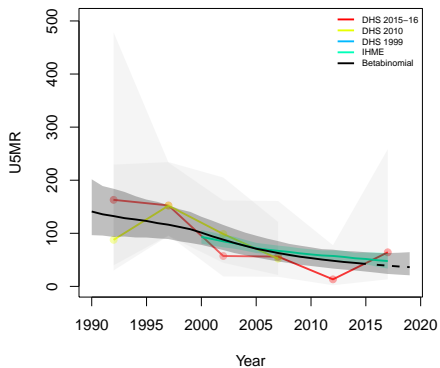
**Manyoni, Singida**



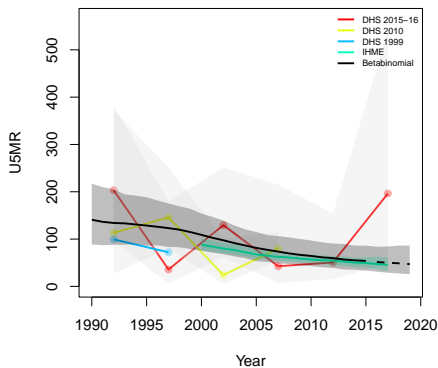
**Mkalama, Singida**



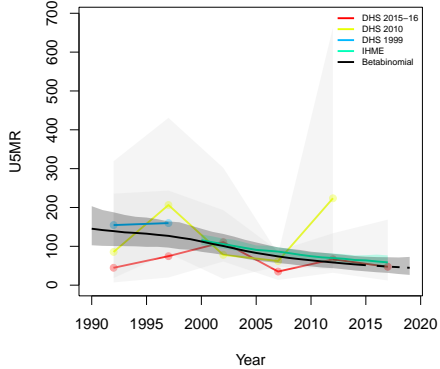
**Singida Rural, Singida**



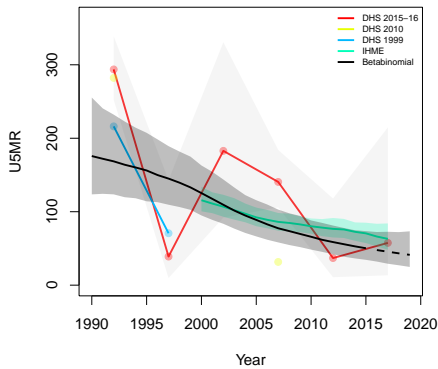
**Singida Urban, Singida**

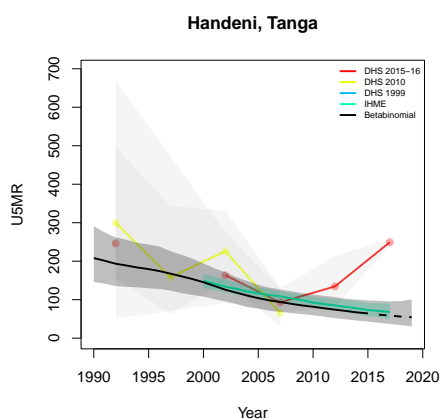
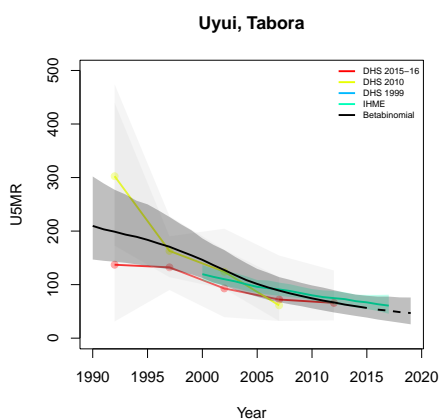
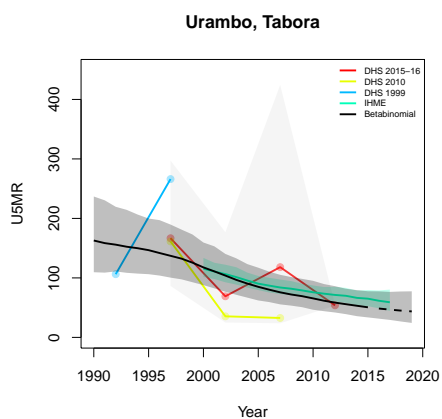
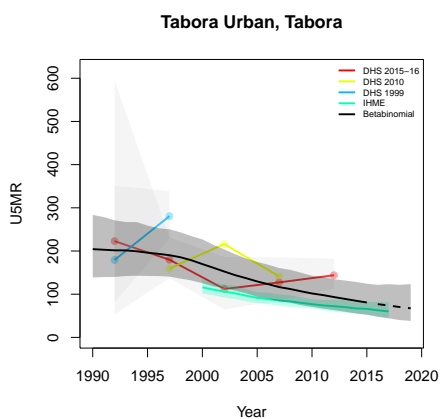
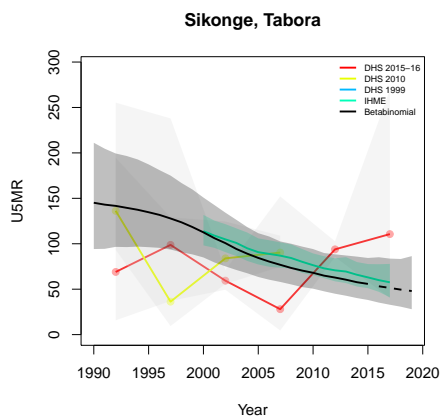
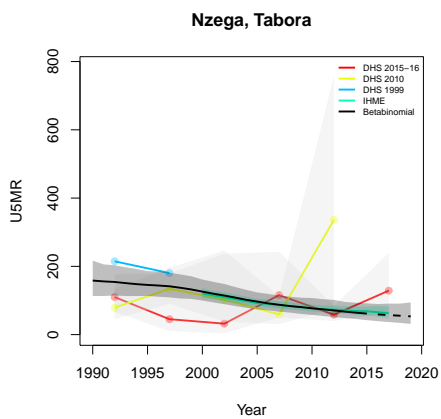


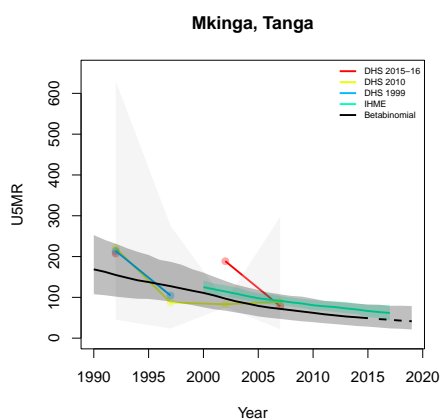
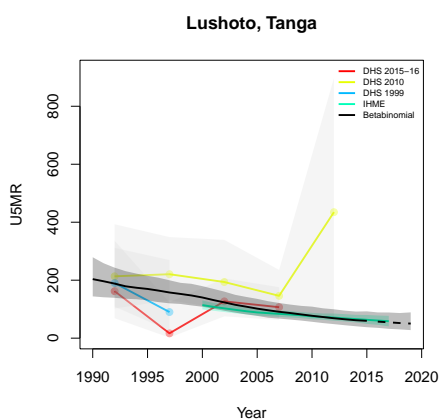
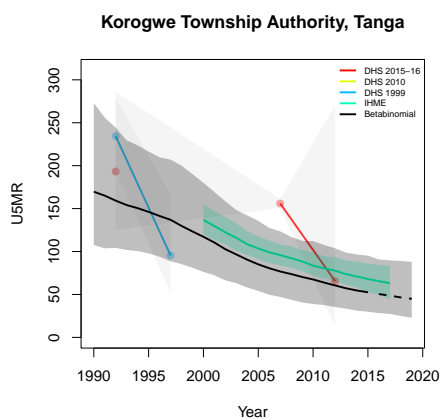
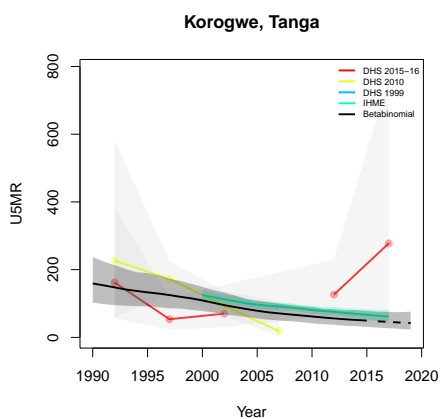
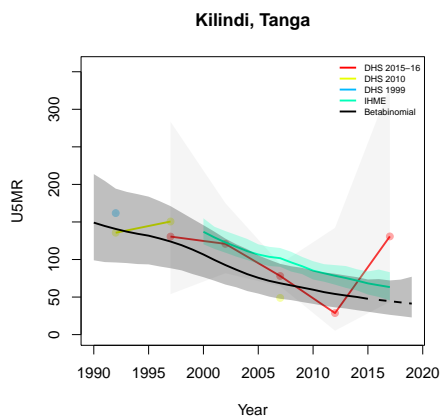
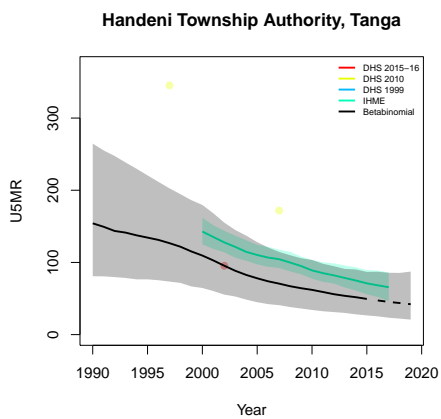
**Igunga, Tabora**



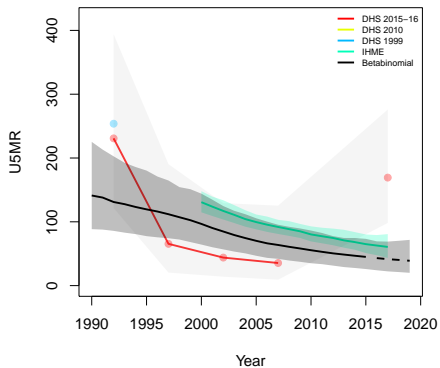
**Kaliua, Tabora**



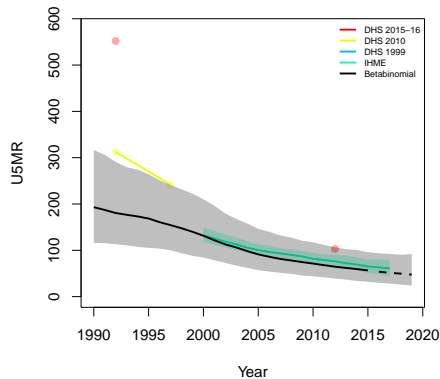




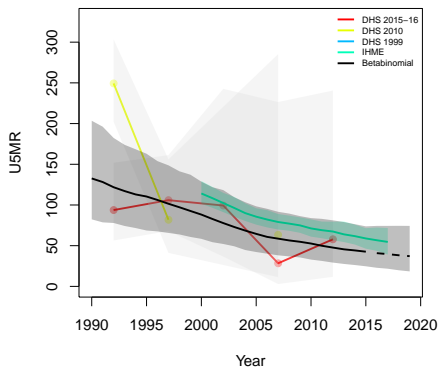
**Muheza, Tanga**



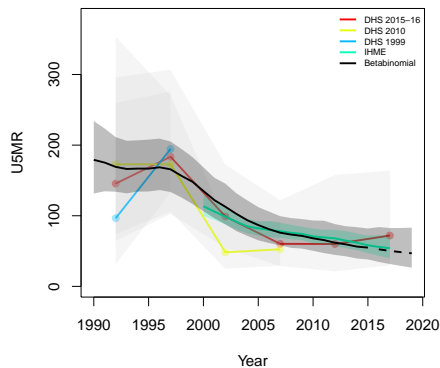
**Pangani, Tanga**



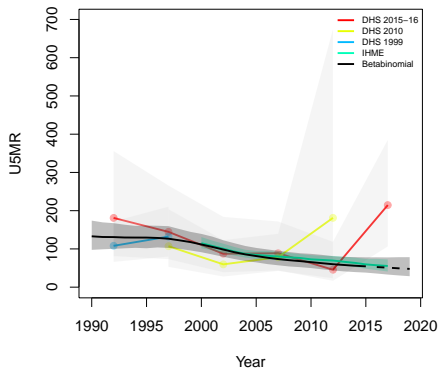
**Tanga, Tanga**



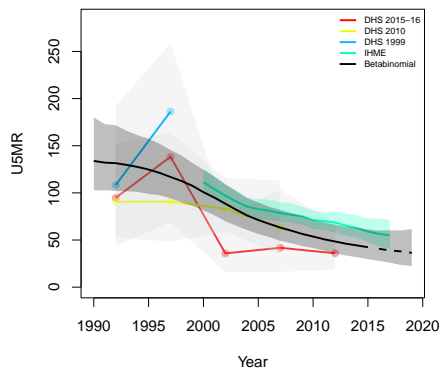
**Kaskazini 'A', Zanzibar North**

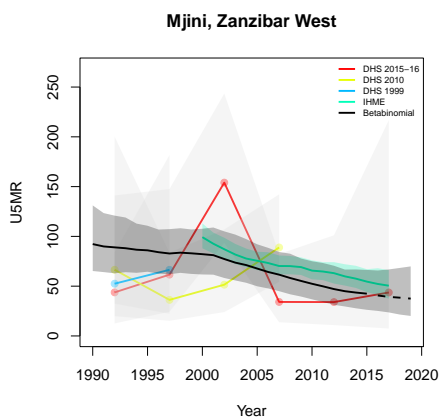
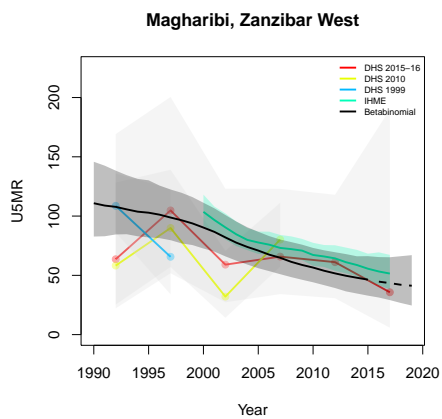
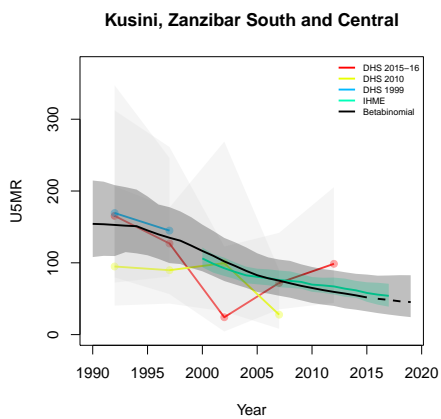


**Kaskazini 'B', Zanzibar North**



**Kati, Zanzibar South and Central**





**B.18 Togo**

Age	Survey	Clusters			Deaths			Agemonths		
		Urban	Rural	Total	Urban	Rural	Total	Urban	Rural	Total
0	1998	134	153	287	99	372	471	2583	9207	11790
	2013	128	202	330	191	622	813	6169	18482	24651
1-11	1998	134	153	287	66	409	475	27002	94938	121940
	2013	128	202	330	117	542	659	63436	189064	252500
12-23	1998	134	153	287	35	193	228	28520	97543	126063
	2013	128	202	330	51	383	434	64202	190482	254684
24-35	1998	134	153	287	24	185	209	28455	94668	123123
	2013	128	202	330	35	285	320	59829	179016	238845
36-47	1998	134	153	287	31	184	215	27856	90570	118426
	2013	128	202	330	34	206	240	55636	166685	222321
48-59	1998	134	153	287	11	96	107	27520	85479	112999
	2013	128	202	330	21	93	114	51439	155133	206572

Table B.18: **Data summary for Togo.** Total numbers of clusters (Columns 3–5) with observations in each age group by survey in urban and rural areas and combined. Numbers of deaths (Columns 6–8) and number of agemonths (Columns 9–10) observed in each age group by survey in urban and rural areas and combined.

*B.18.1 Admin-1*

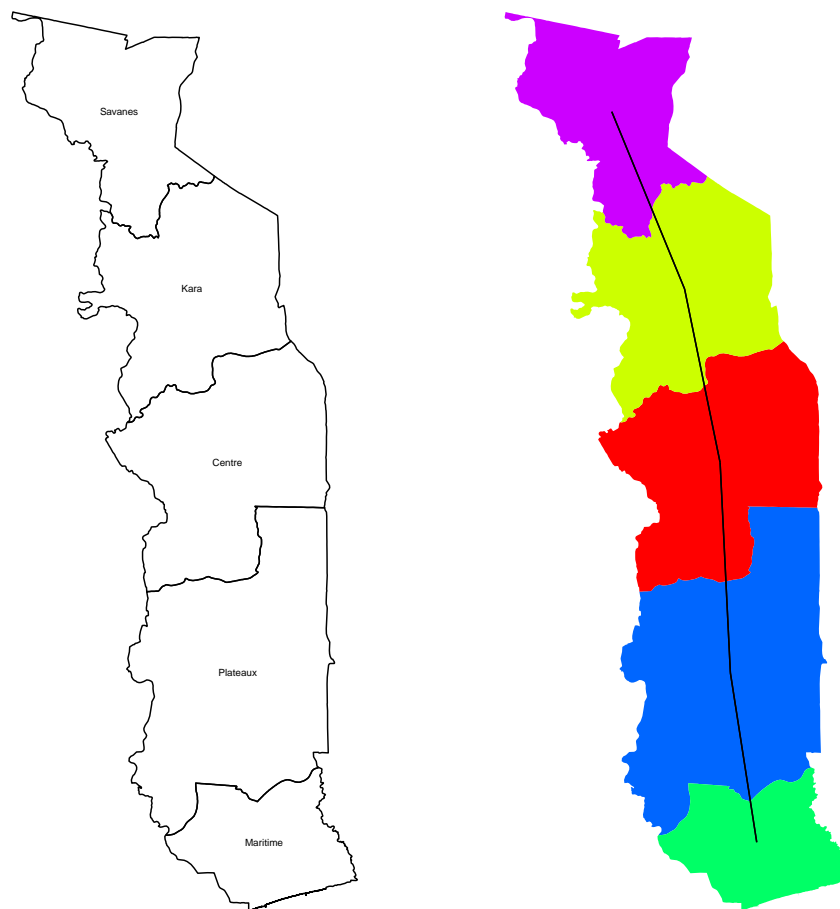


Figure B.88: **Left:** The names of the 5 Admin-1 areas of Togo . **Right:** The neighborhood structure of Admin-1 areas in Togo .

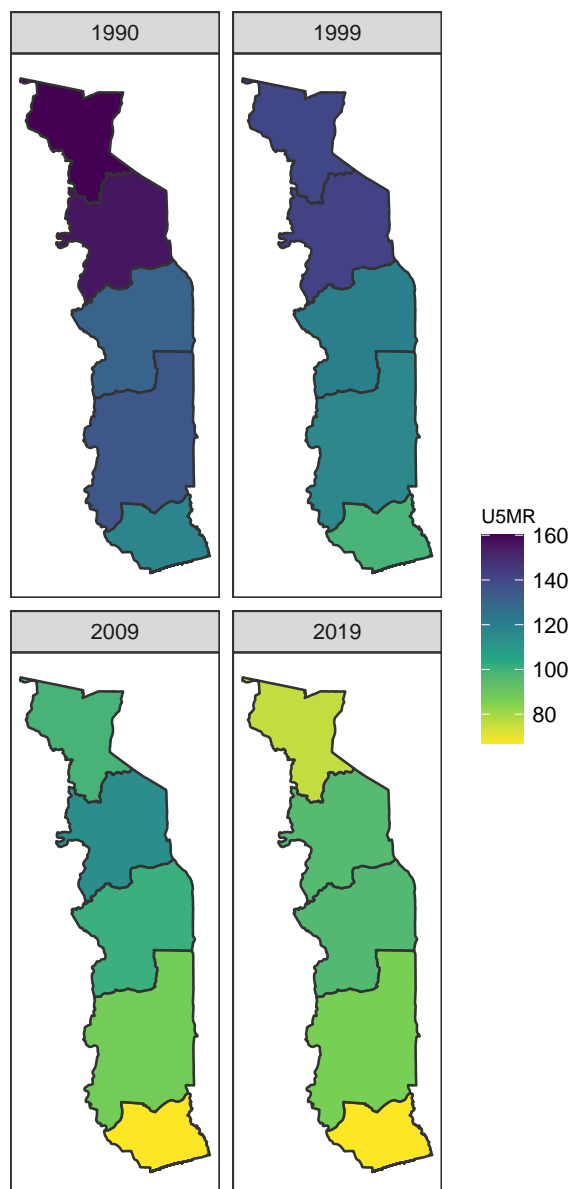


Figure B.89: Median U5MR estimates for years 1990, 1999, 2009, 2019 for Admin-1 areas in Togo .

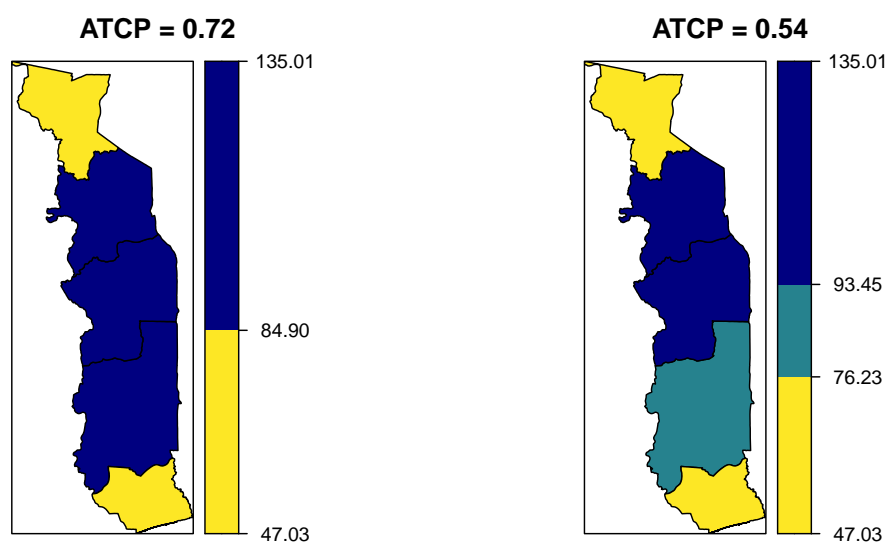
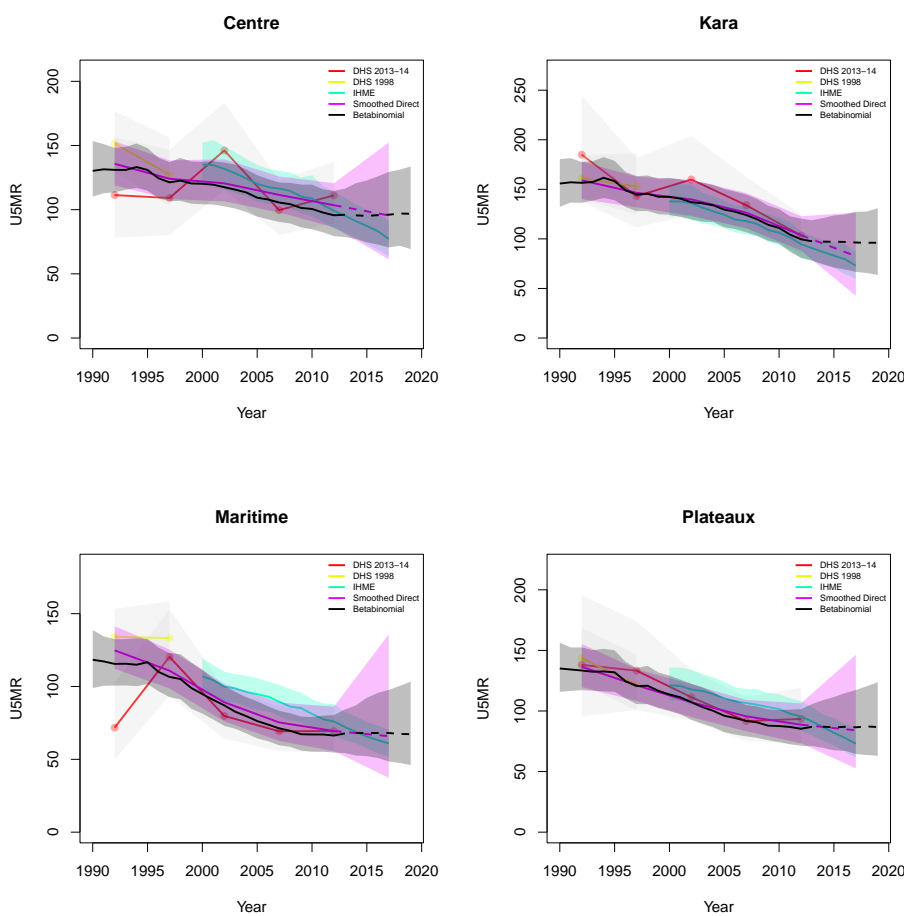
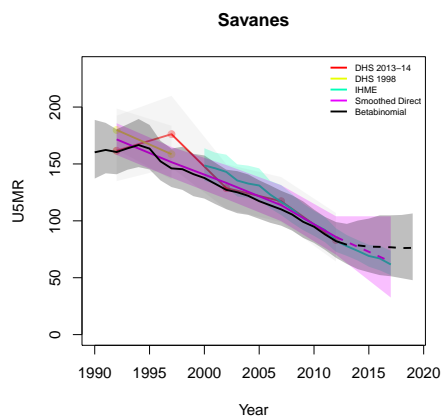


Figure B.90: Expression of uncertainty of U5MR (deaths per 1000 children) estimates for Admin-1 areas based on the average true classification probability (ATCP) in 2019 using  $K = 2, 3$  colors.

*Data and estimates over time by area*

Colored lines with circular points and light grey uncertainty bands are 5-year survey-weighted estimates of U5MR for years 1990–1994 up to 2015–2019 depending on survey timing. For a survey that ends in the middle of a 5-year period, we plot the estimates at the mid-point of the years in that interval for which the survey provides data. Black lines and corresponding intervals represent posterior medians and 95% uncertainty intervals respectively for the betabinomial model. IHME’s estimates and corresponding intervals, where we can compare, are in aquamarine.





*B.18.2 Admin-2*

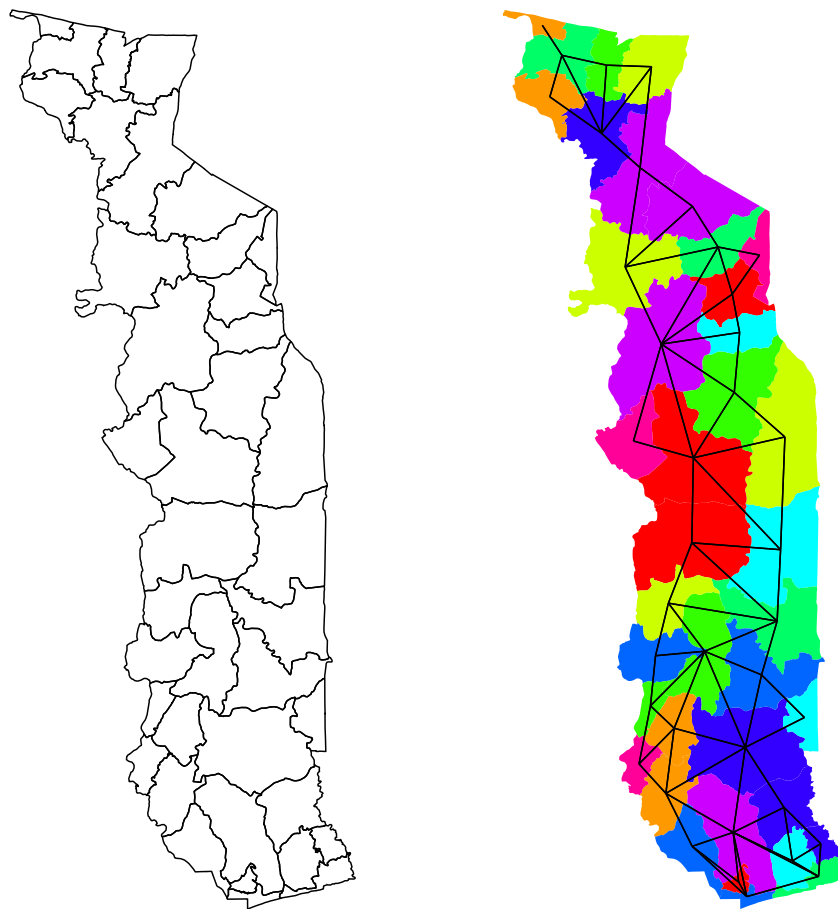


Figure B.91: **Left:** The names of the 39 Admin-2 areas of Togo . **Right:** The neighborhood structure of Admin-2 areas in Togo .

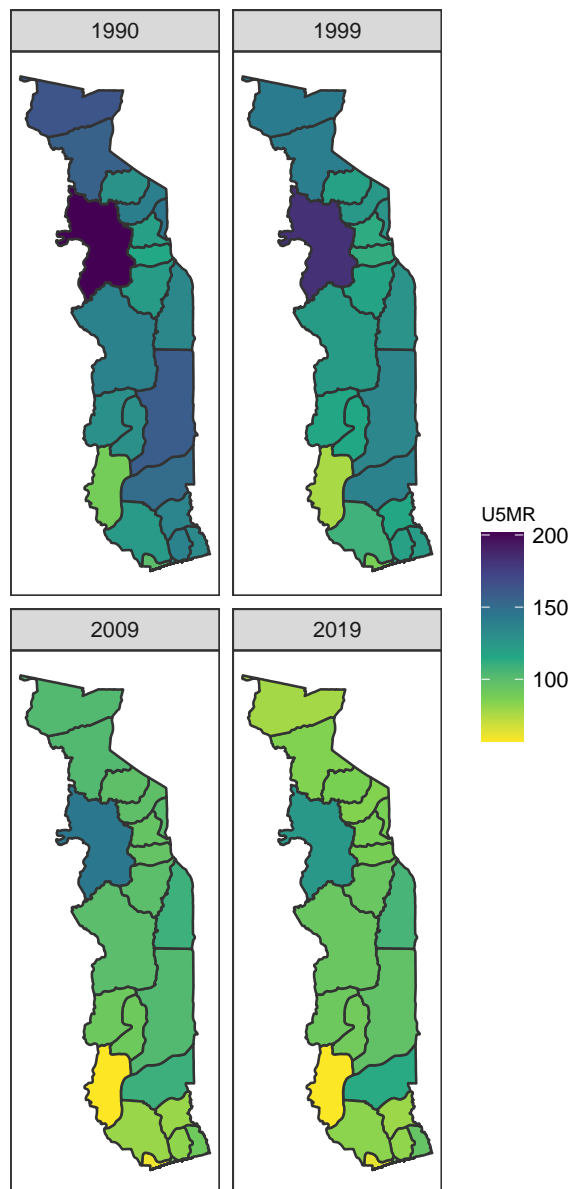


Figure B.92: Median U5MR estimates for years 1990, 1999, 2009, 2019 for Admin-2 areas in Togo .

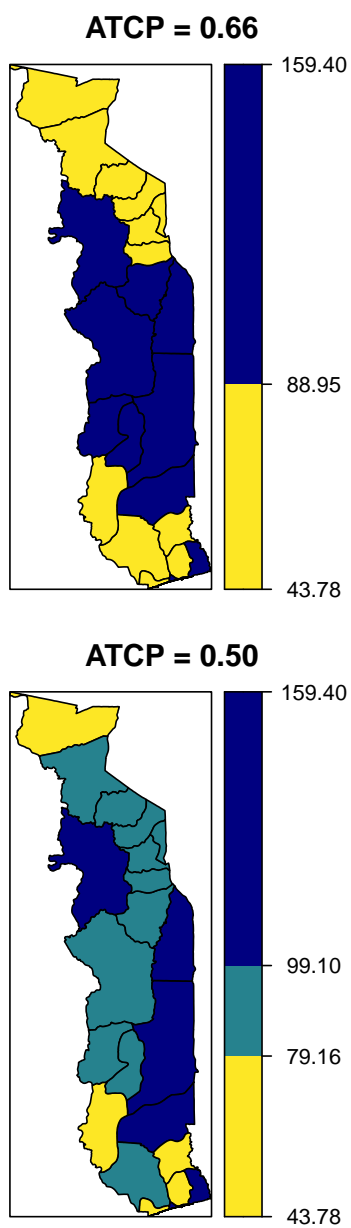
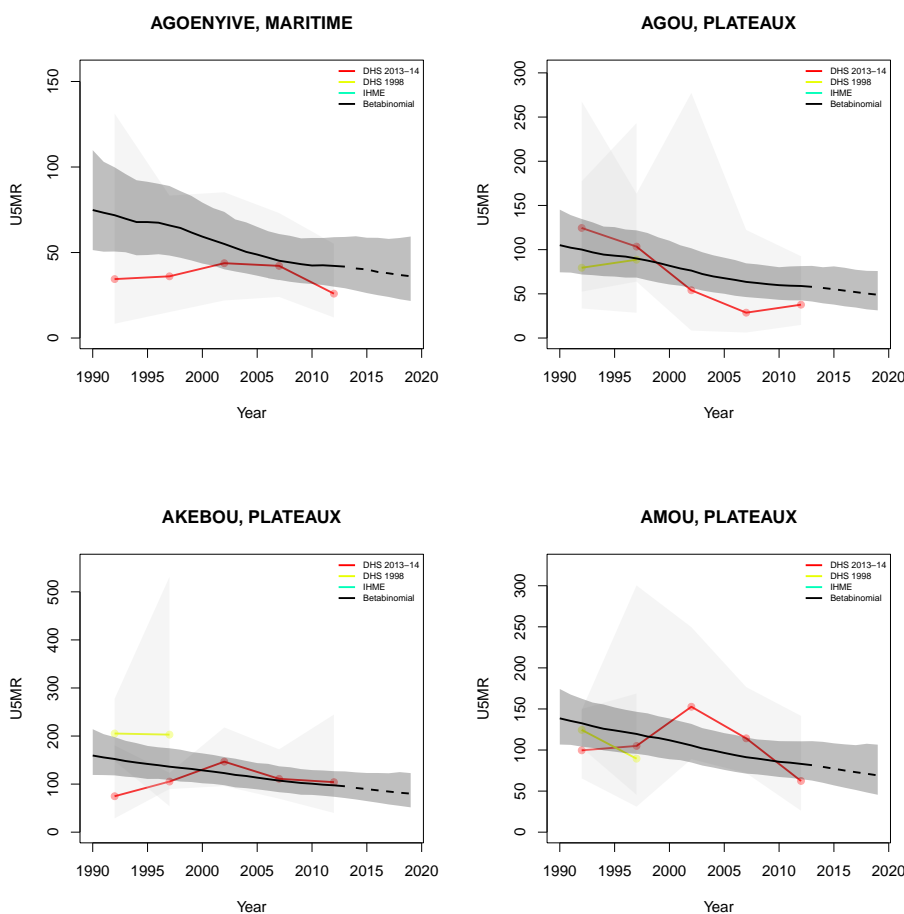


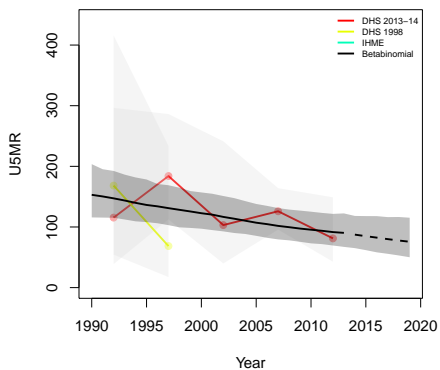
Figure B.93: Expression of uncertainty of U5MR (deaths per 1000 children) estimates for Admin-1 areas based on the average true classification probability (ATCP) in 2019 using  $K = 2, 3$  colors.

*Data and estimates over time by area*

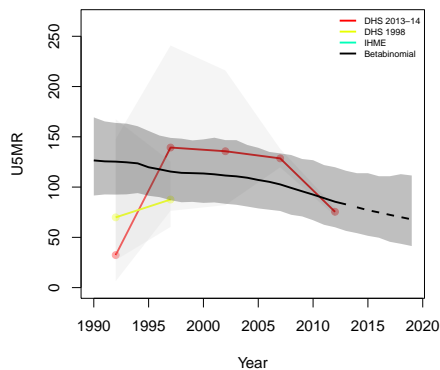
Colored lines with circular points and light grey uncertainty bands are 5-year survey-weighted estimates of U5MR for years 1990–1994 up to 2015–2019 depending on survey timing. For a survey that ends in the middle of a 5-year period, we plot the estimates at the mid-point of the years in that interval for which the survey provides data. Black lines and corresponding intervals represent posterior medians and 95% uncertainty intervals respectively for the betabinomial model. IHME’s estimates and corresponding intervals, where we can compare, are in aquamarine.



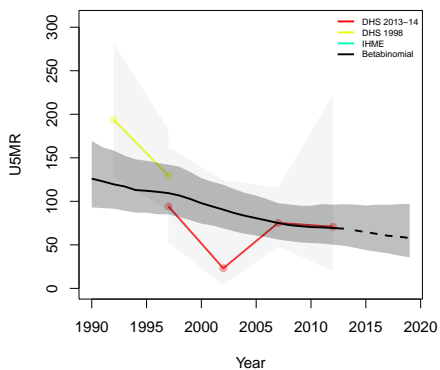
**ANIE, PLATEAUX**



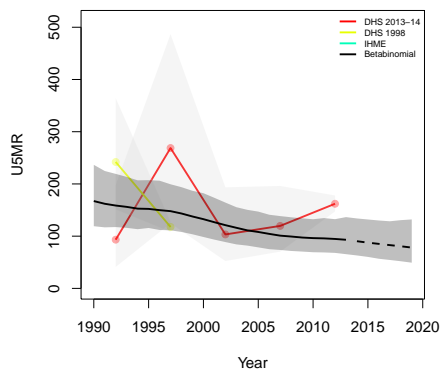
**ASSOLI, KARA**



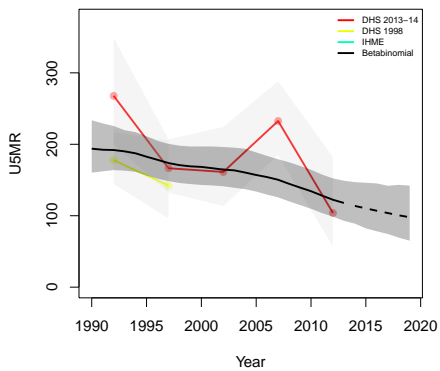
**AVE, MARITIME**



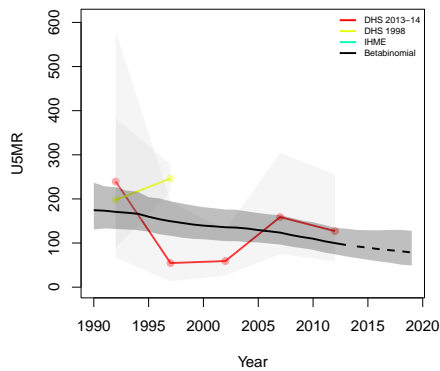
**BAS-MONO, MARITIME**



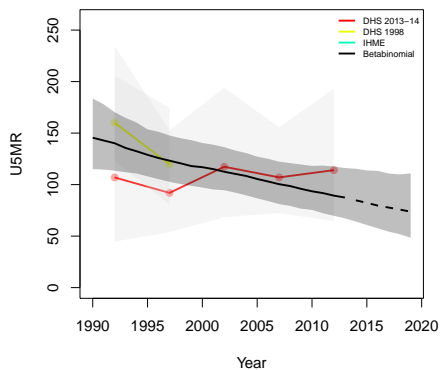
**BASSAR, KARA**



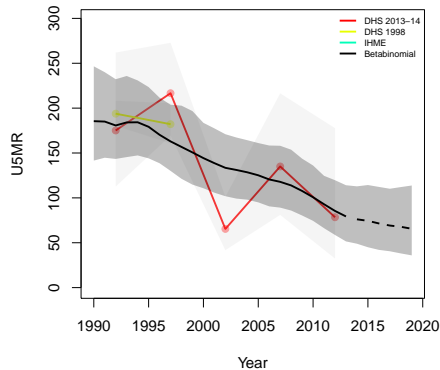
**BINAH, KARA**



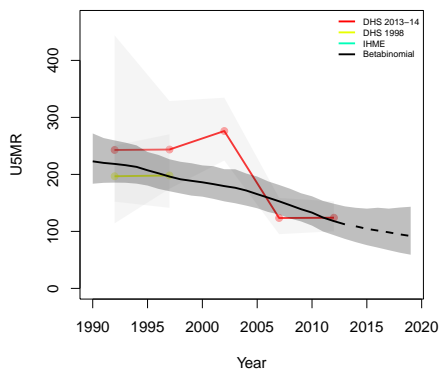
**BLITTA, CENTRALE**



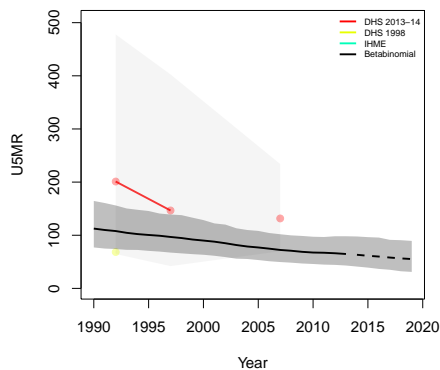
**CINKASSE, SAVANES**



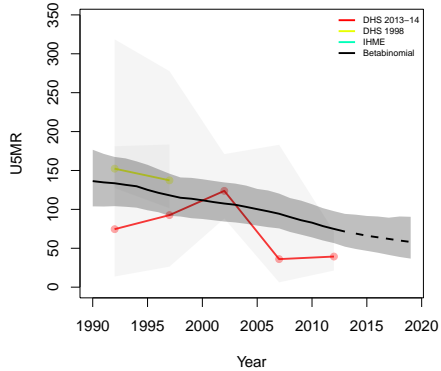
**DANKPEN, KARA**



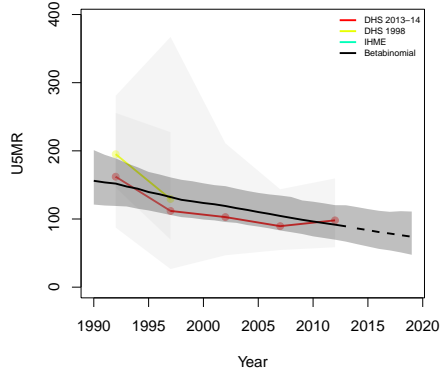
**DANYI, PLATEAUX**

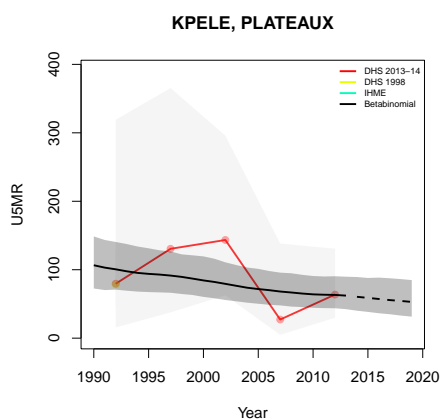
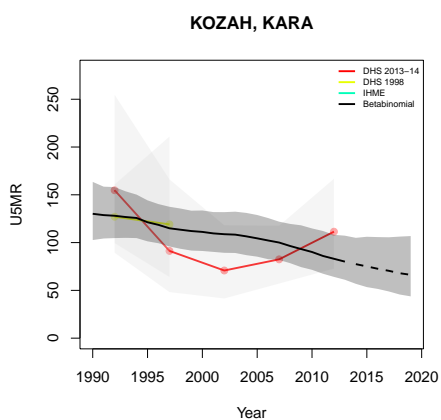
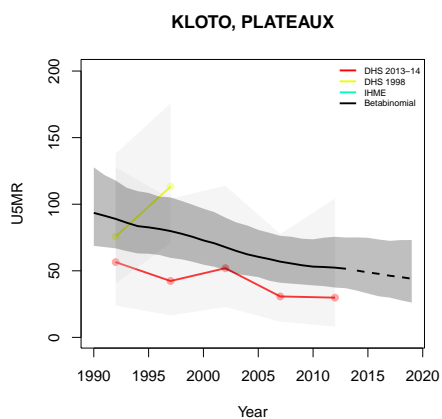
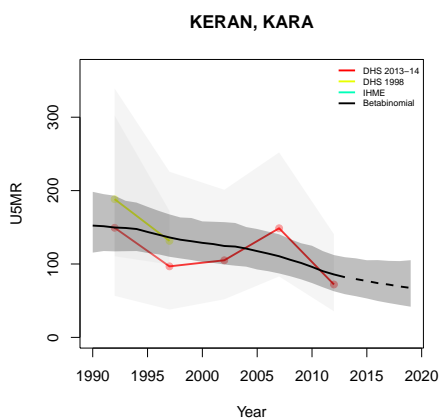
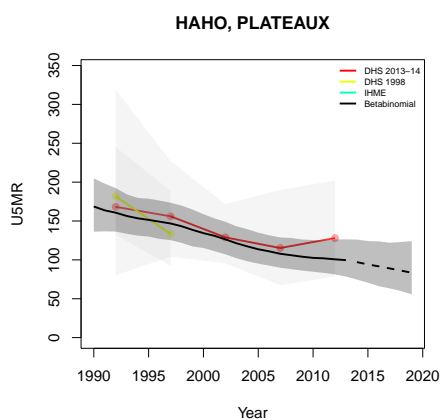
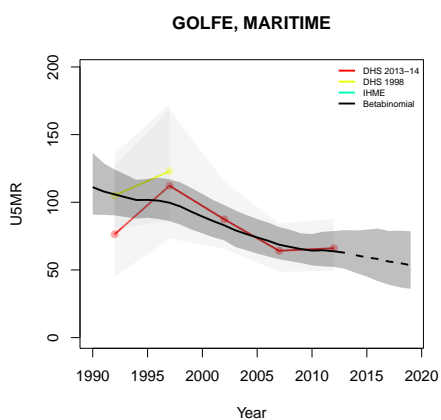


**DOUFELGOU, KARA**

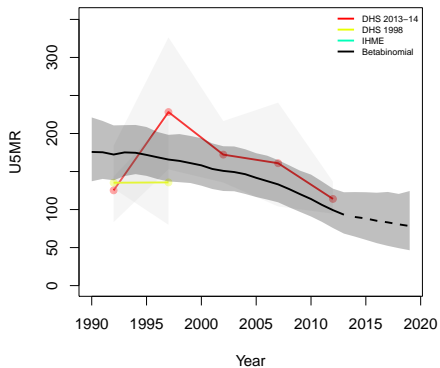


**EST-MONO, PLATEAUX**

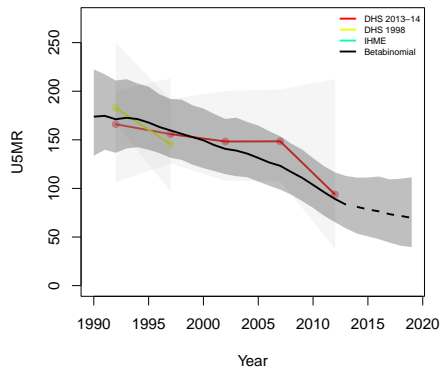




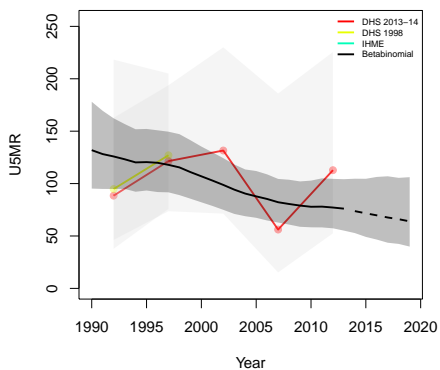
**KPENDJAL, SAVANES**



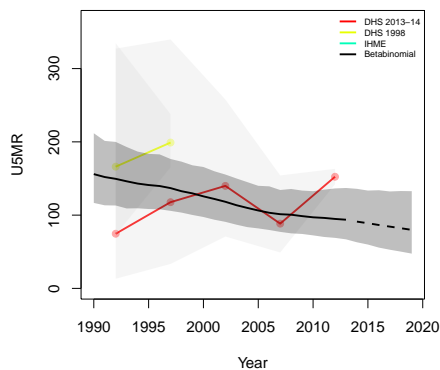
**KPENDJAL OUEST, SAVANES**



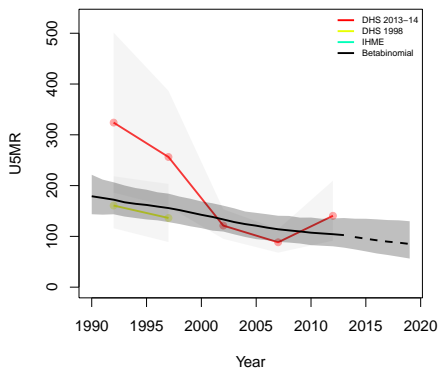
**LACS, MARITIME**



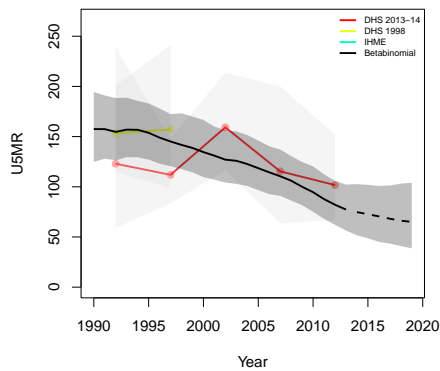
**MOYEN-MONO, PLATEAUX**



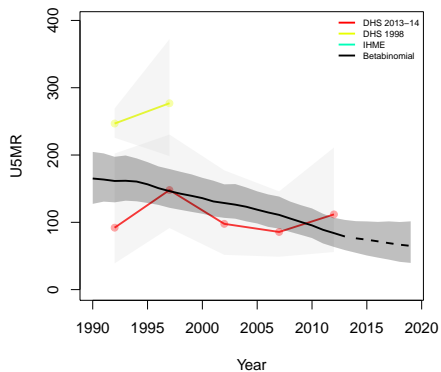
**OGOUE, PLATEAUX**



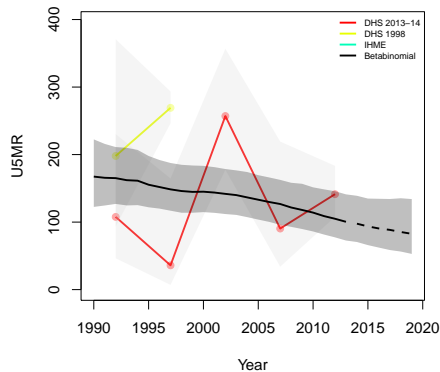
**OTI, SAVANES**



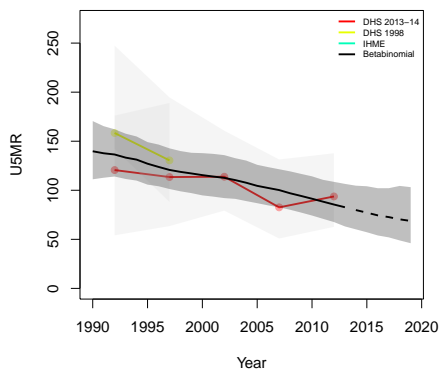
**OTI SUD, SAVANES**



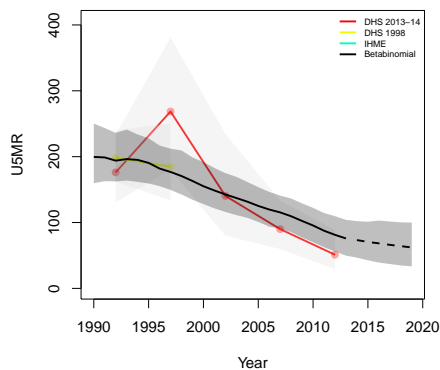
**MO, CENTRALE**



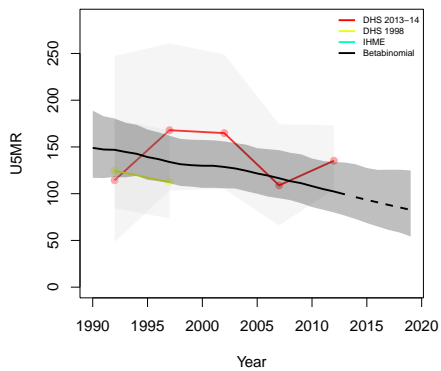
**SOTOUBOUA, CENTRALE**



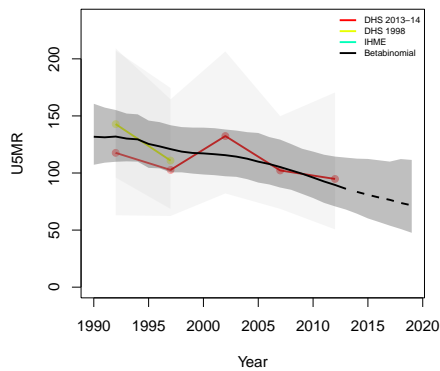
**TANDJOARE, SAVANES**

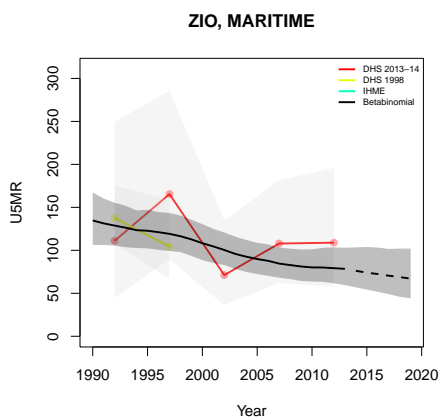
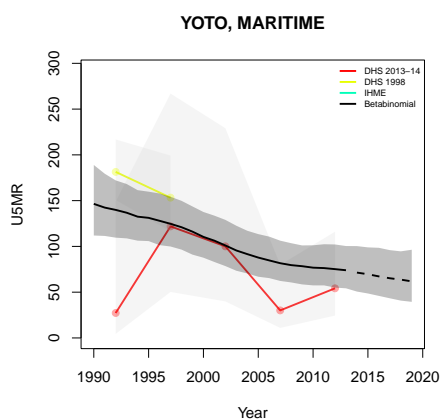
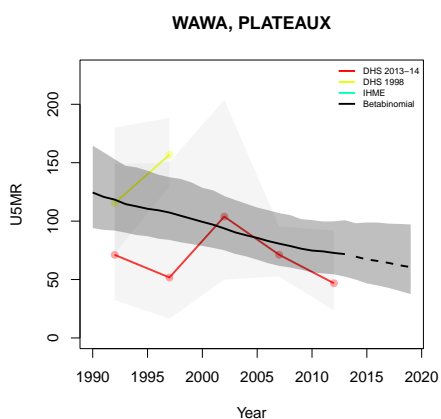
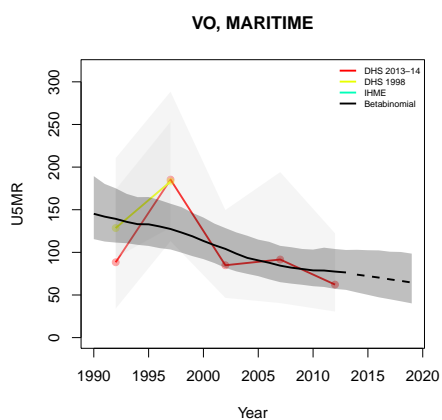
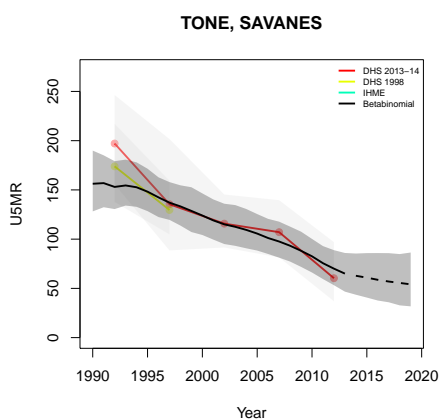


**TCHAMBA, CENTRALE**



**TCHAOUDJO, CENTRALE**





**B.19 Uganda**

Age	Survey	Clusters			Deaths			Agemonths		
		Urban	Rural	Total	Urban	Rural	Total	Urban	Rural	Total
0	2000	90	176	266	72	343	415	2990	9459	12449
	2006	56	280	336	56	644	700	2330	19377	21707
	2011	119	281	400	146	650	796	5171	20436	25607
	2016	158	527	685	267	1314	1581	9354	45850	55204
1-11	2000	90	176	266	111	528	639	30858	95517	126375
	2006	56	280	336	93	1154	1247	23969	195869	219838
	2011	119	281	400	169	883	1052	52914	208183	261097
	2016	158	527	685	252	1684	1936	96352	468877	565229
12-23	2000	90	176	266	54	278	332	31038	94080	125118
	2006	56	280	336	41	532	573	24202	193492	217694
	2011	119	281	400	88	416	504	52988	208177	261165
	2016	158	527	685	95	754	849	97289	471455	568744
24-35	2000	90	176	266	36	142	178	29416	87553	116969
	2006	56	280	336	20	317	337	22864	180906	203770
	2011	119	281	400	47	279	326	49521	194955	244476
	2016	158	527	685	59	482	541	90632	439474	530106
36-47	2000	90	176	266	23	91	114	28100	82481	110581
	2006	56	280	336	16	181	197	21485	169173	190658
	2011	119	281	400	27	171	198	46194	181266	227460
	2016	158	527	685	34	273	307	84499	408986	493485
48-59	2000	90	176	266	12	63	75	26667	77314	103981
	2006	56	280	336	6	95	101	20258	158478	178736
	2011	119	281	400	10	67	77	42847	169169	212016
	2016	158	527	685	24	139	163	78415	380313	458728

Table B.19: **Data summary for Uganda.** Total numbers of clusters (Columns 3–5) with observations in each age group by survey in urban and rural areas and combined. Numbers of deaths (Columns 6–8) and number of agemonths (Columns 9–10) observed in each age group by survey in urban and rural areas and combined.

*B.19.1 Admin-1*

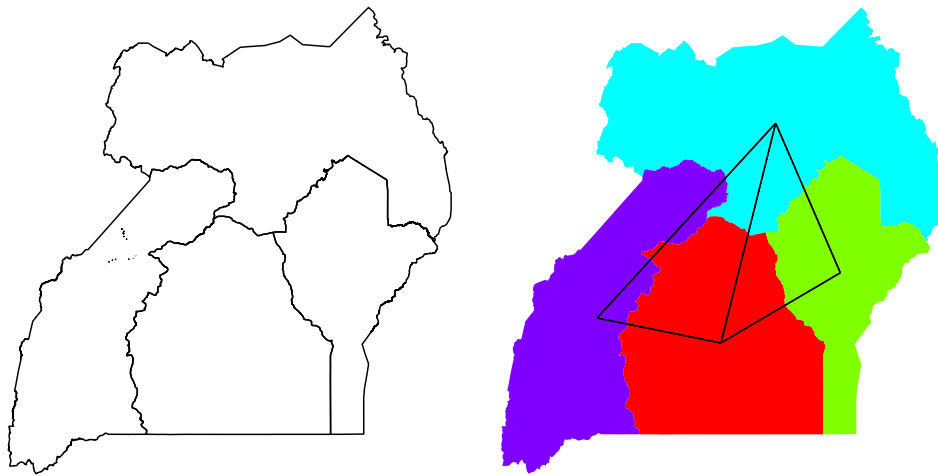


Figure B.94: **Left:** The names of the 4 Admin-1 areas of Uganda . **Right:** The neighborhood structure of Admin-1 areas in Uganda .

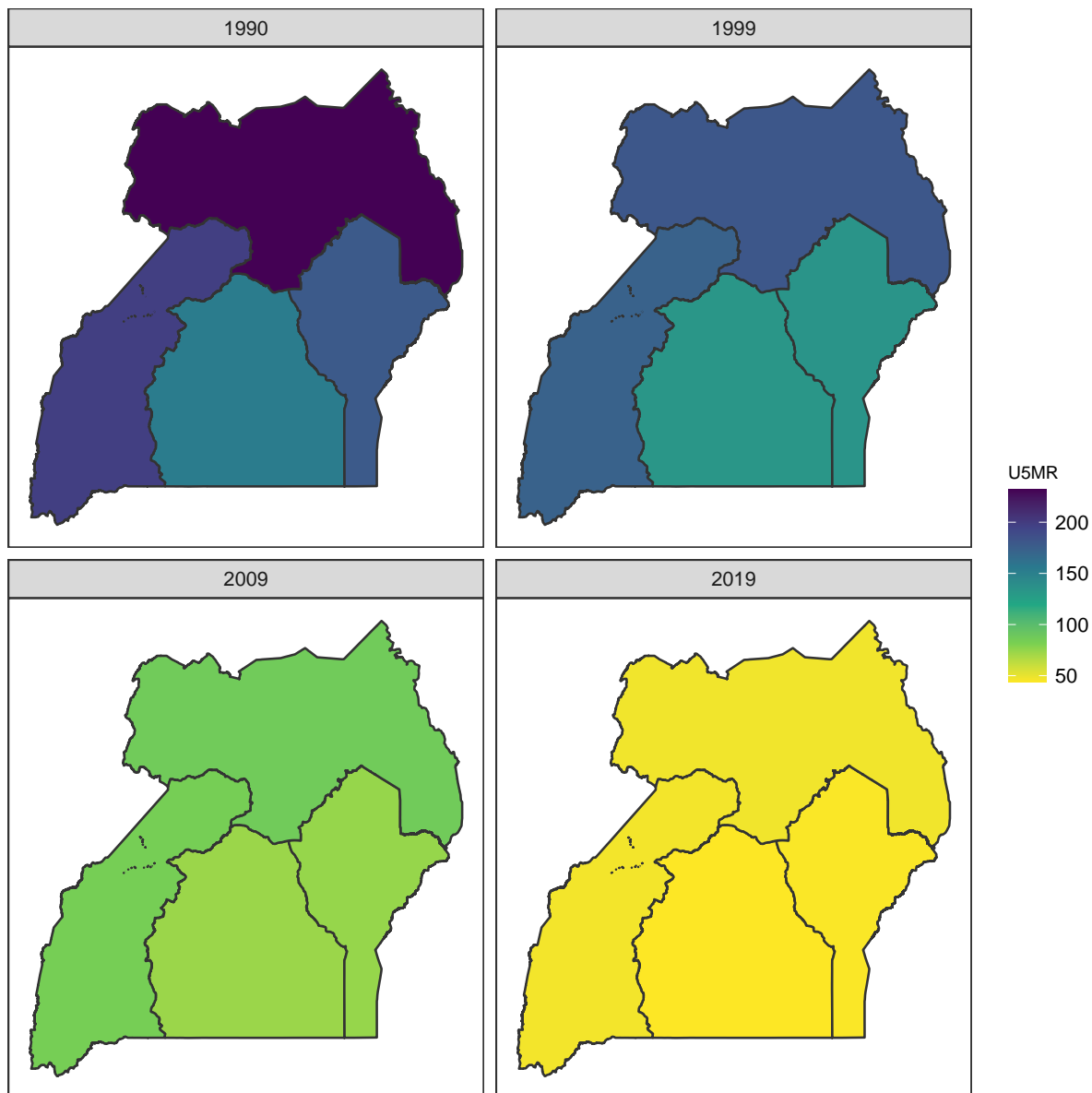


Figure B.95: Median U5MR estimates for years 1990, 1999, 2009, 2019 for Admin-1 areas in Uganda .

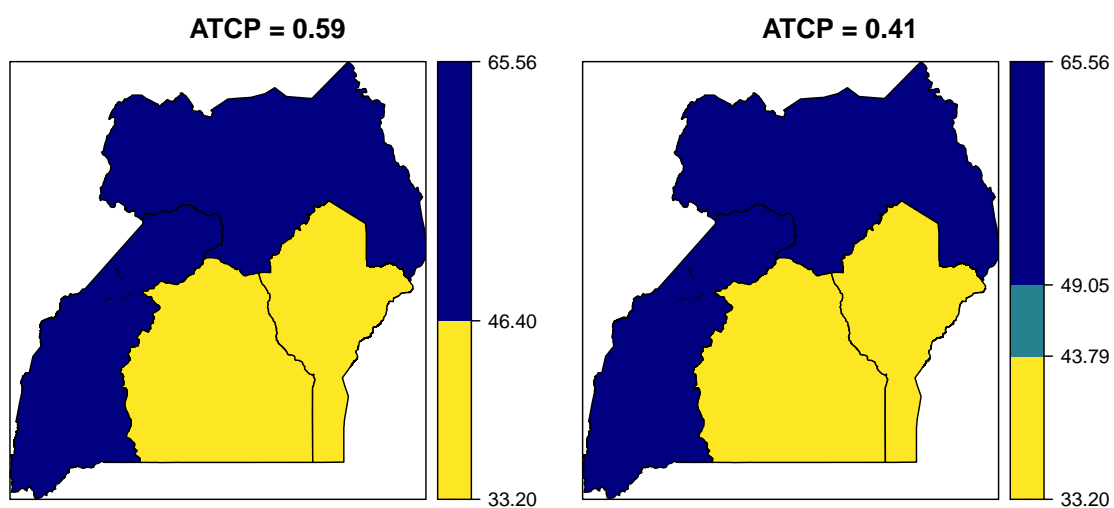
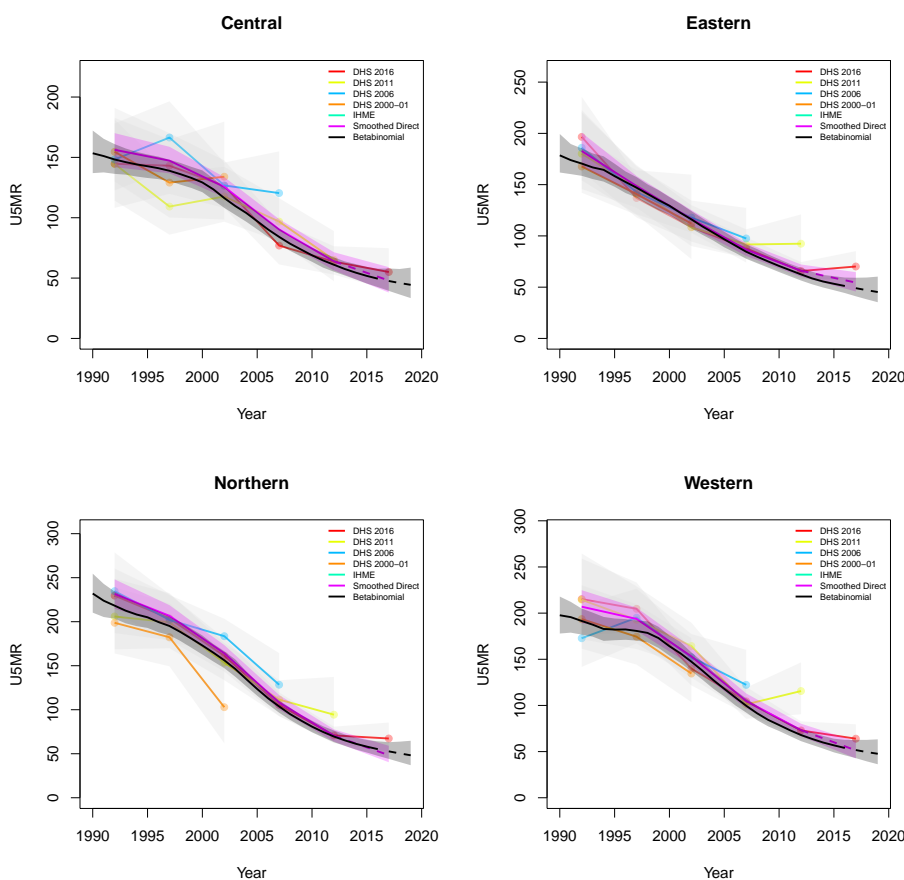


Figure B.96: Expression of uncertainty of U5MR (deaths per 1000 children) estimates for Admin-1 areas based on the average true classification probability (ATCP) in 2019 using  $K = 2, 3$  colors.

### Data and estimates over time by area

Colored lines with circular points and light grey uncertainty bands are 5-year survey-weighted estimates of U5MR for years 1990–1994 up to 2015–2019 depending on survey timing. For a survey that ends in the middle of a 5-year period, we plot the estimates at the mid-point of the years in that interval for which the survey provides data. Black lines and corresponding intervals represent posterior medians and 95% uncertainty intervals respectively for the betabinomial model. IHME's estimates and corresponding intervals, where we can compare, are in aquamarine.



*B.19.2 Admin-2*

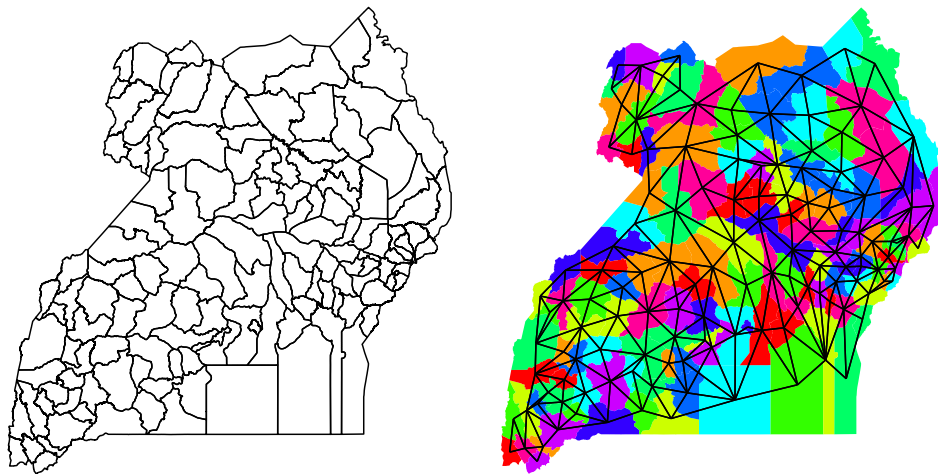


Figure B.97: **Left:** The names of the 135 Admin-2 areas of Uganda . **Right:** The neighborhood structure of Admin-2 areas in Uganda .

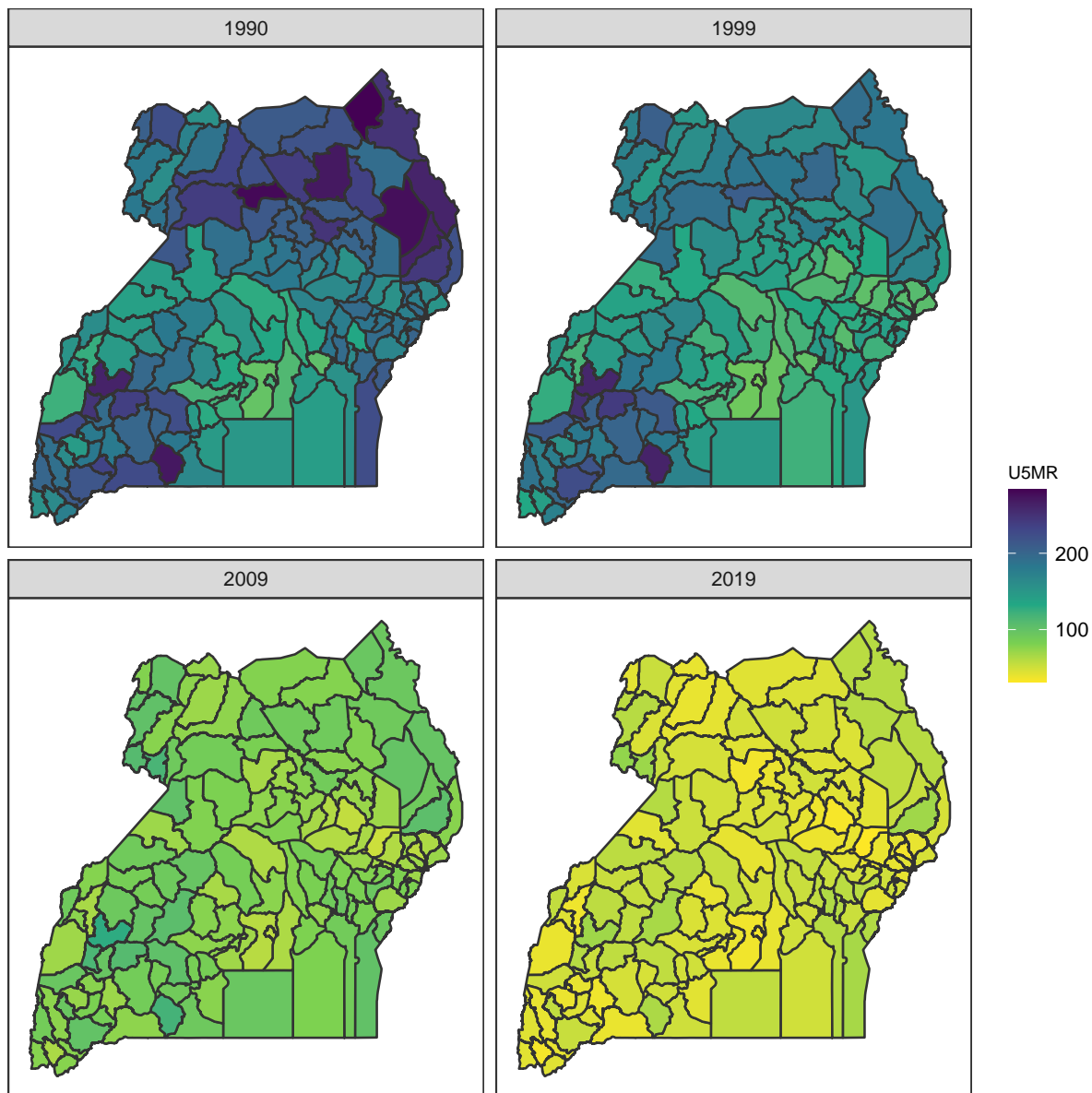


Figure B.98: Median U5MR estimates for years 1990, 1999, 2009, 2019 for Admin-2 areas in Uganda .

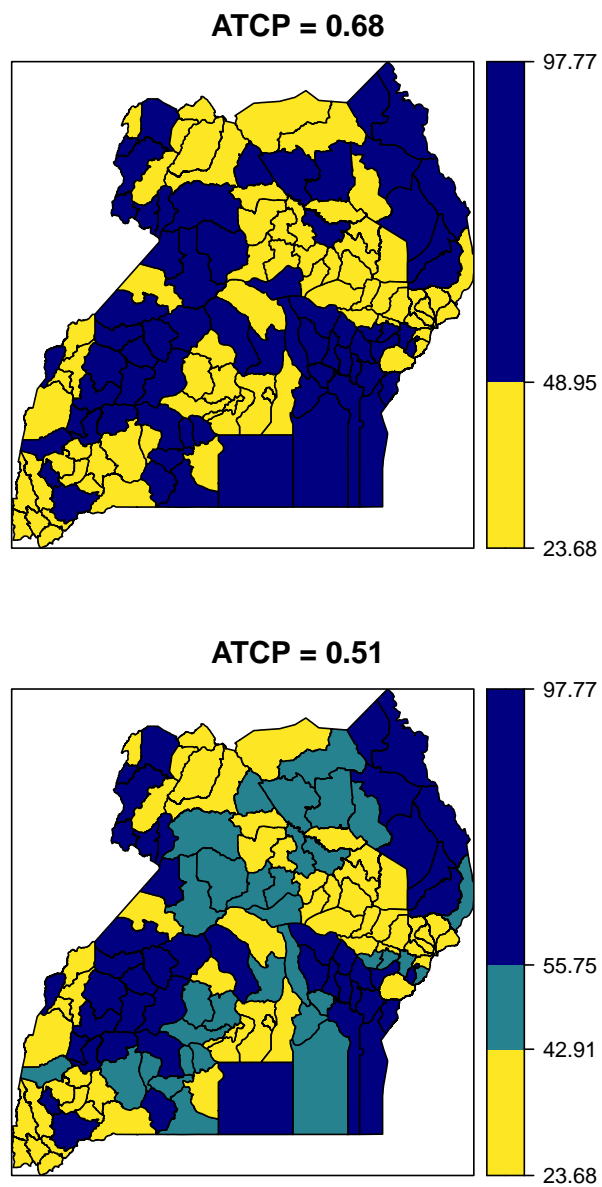
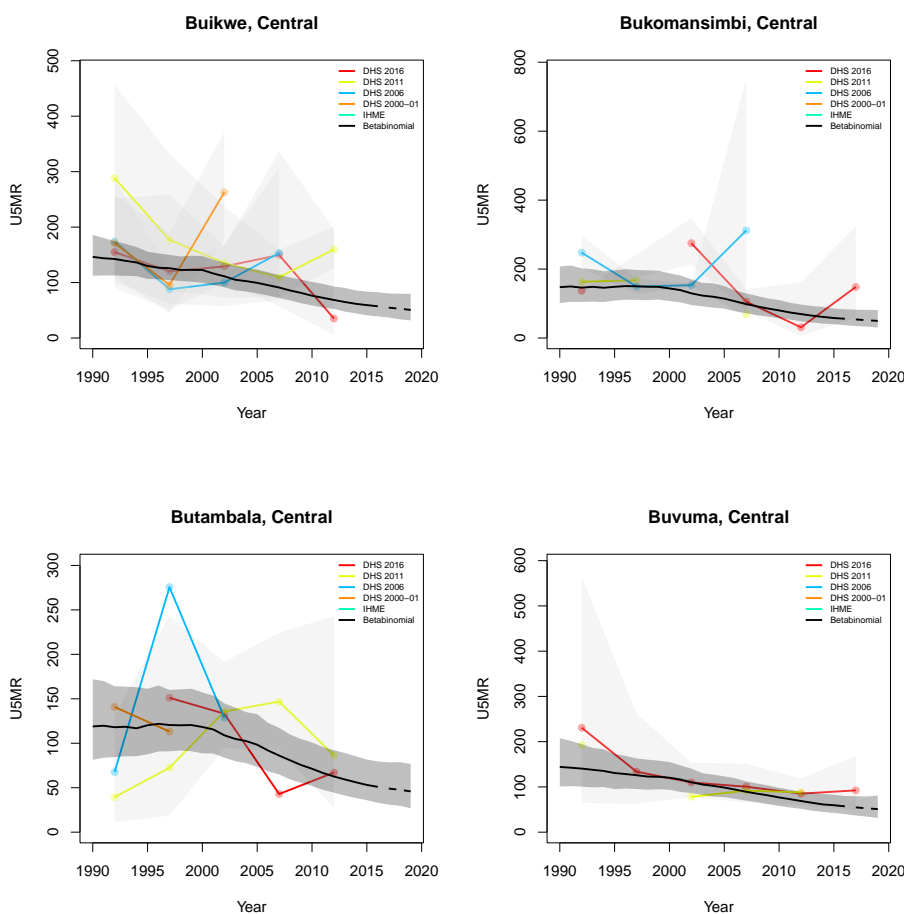


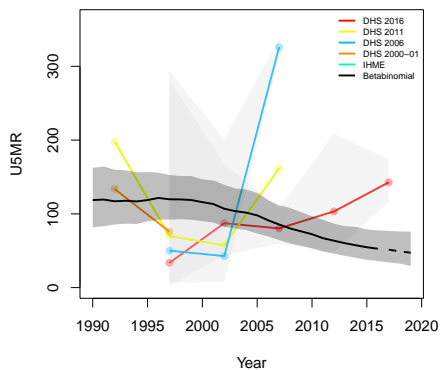
Figure B.99: Expression of uncertainty of U5MR (deaths per 1000 children) estimates for Admin-1 areas based on the average true classification probability (ATCP) in 2019 using  $K = 2, 3$  colors.

*Data and estimates over time by area*

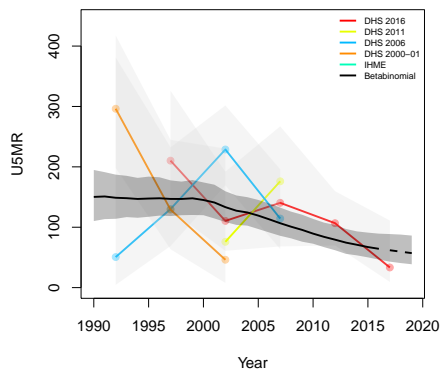
Colored lines with circular points and light grey uncertainty bands are 5-year survey-weighted estimates of U5MR for years 1990–1994 up to 2015–2019 depending on survey timing. For a survey that ends in the middle of a 5-year period, we plot the estimates at the mid-point of the years in that interval for which the survey provides data. Black lines and corresponding intervals represent posterior medians and 95% uncertainty intervals respectively for the betabinomial model. IHME’s estimates and corresponding intervals, where we can compare, are in aquamarine.



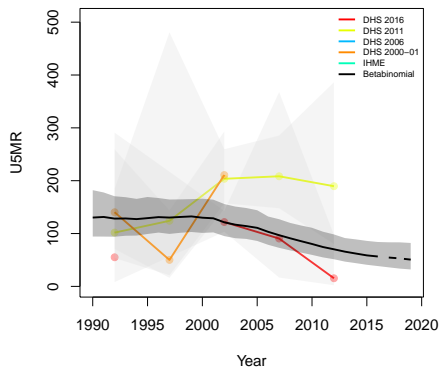
**Gomba, Central**



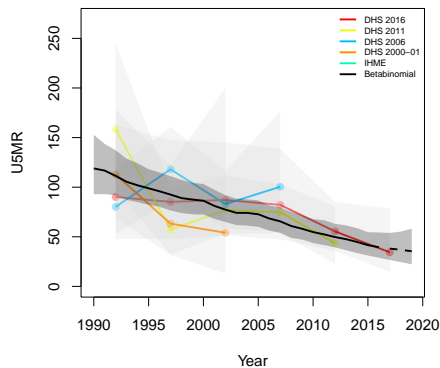
**Kalangala, Central**



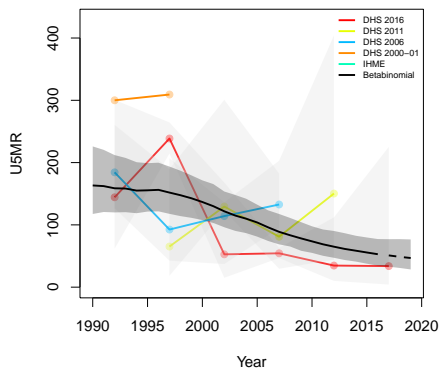
**Kalungu, Central**



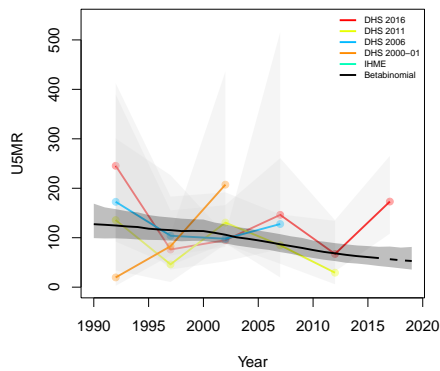
**Kampala, Central**



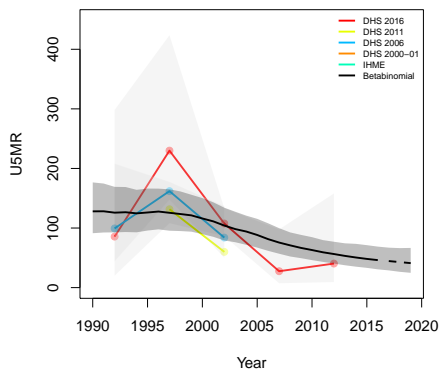
**Kassanda, Central**



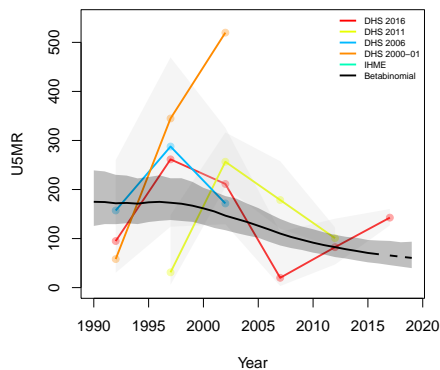
**Kayunga, Central**



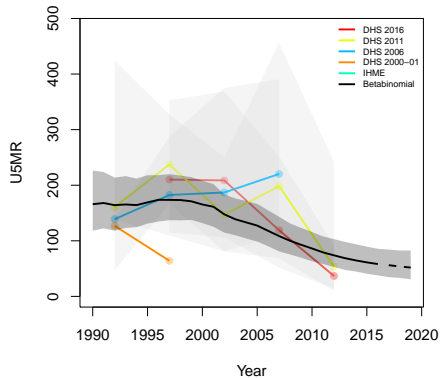
**Kiboga, Central**



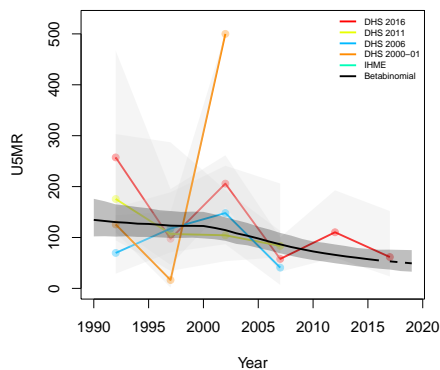
**Kyankwanzi, Central**



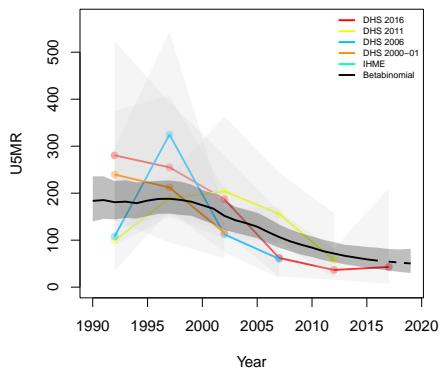
**Kyotera, Central**



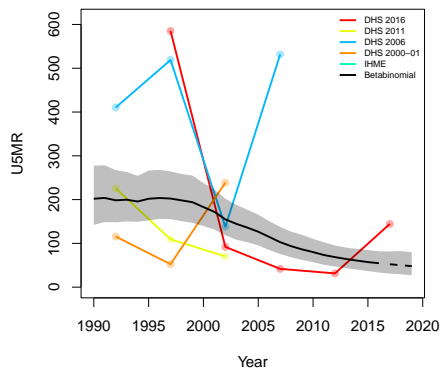
**Luwero, Central**



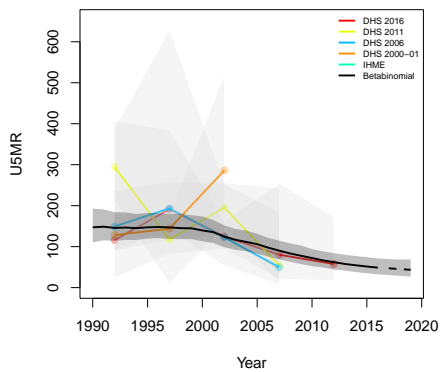
**Lwengo, Central**



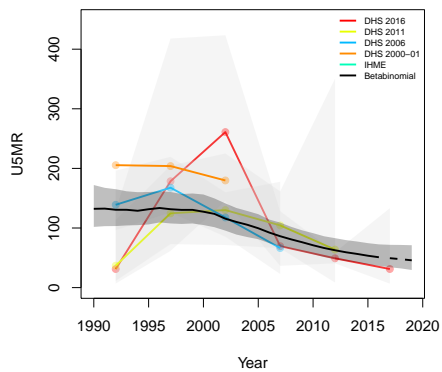
**Lyantonde, Central**



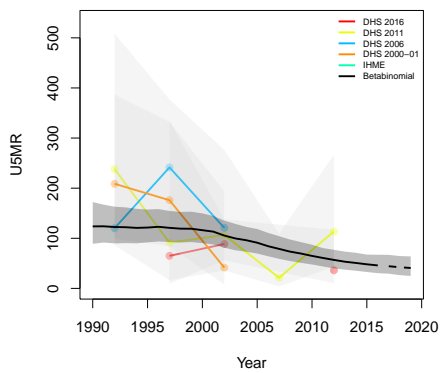
**Masaka, Central**



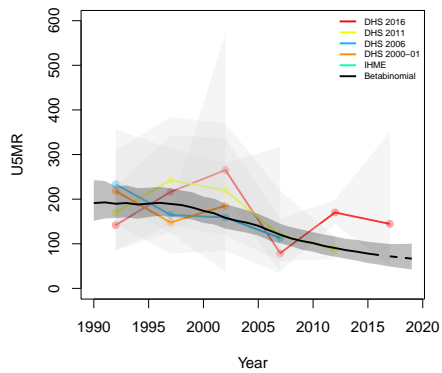
**Mityana, Central**



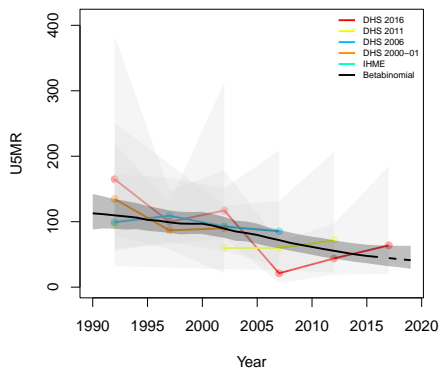
**Mpigi, Central**



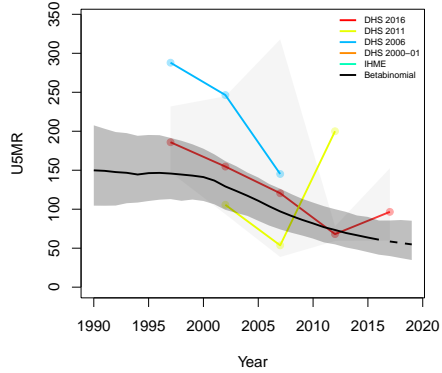
**Mubende, Central**



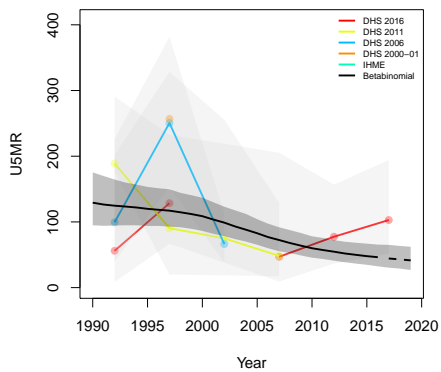
**Mukono, Central**



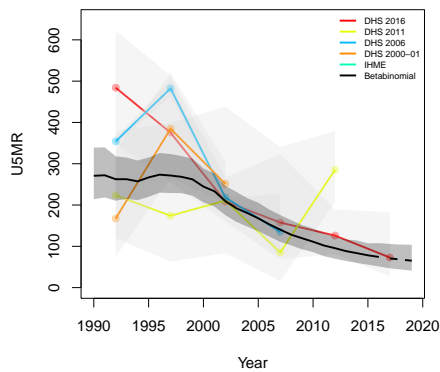
**Nakaseke, Central**



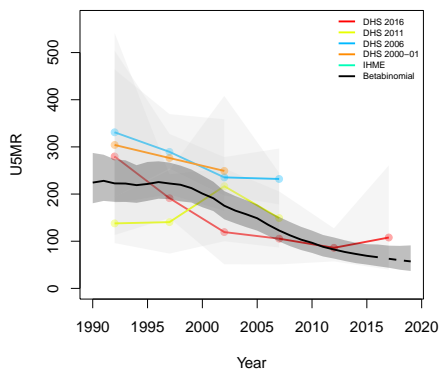
**Nakasongola, Central**



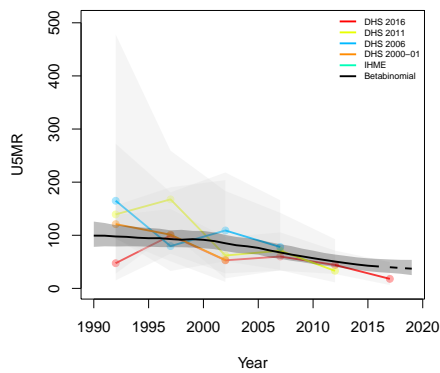
**Rakai, Central**



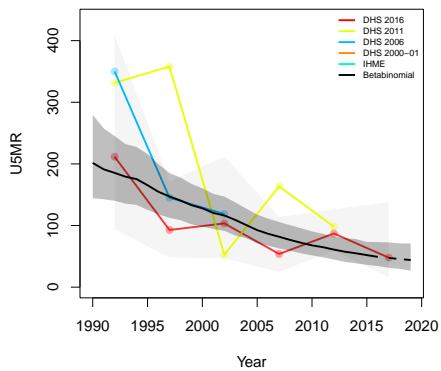
**Ssembabule, Central**



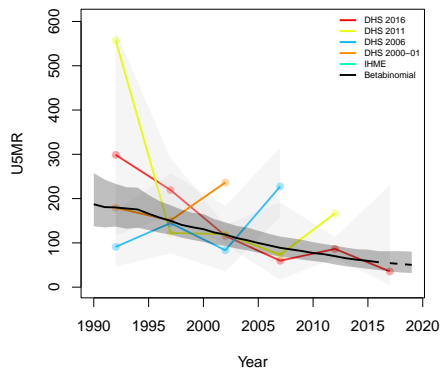
**Wakiso, Central**

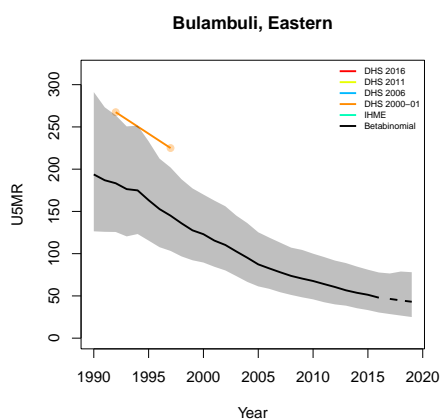
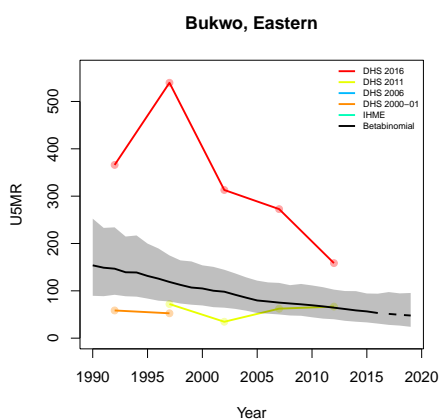
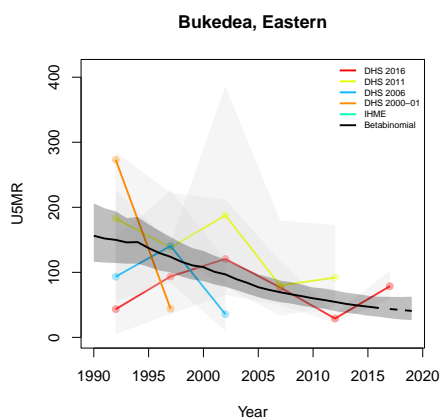
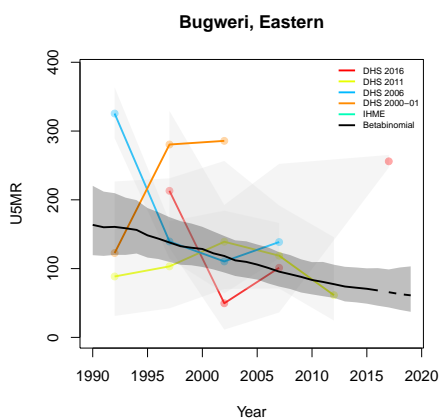
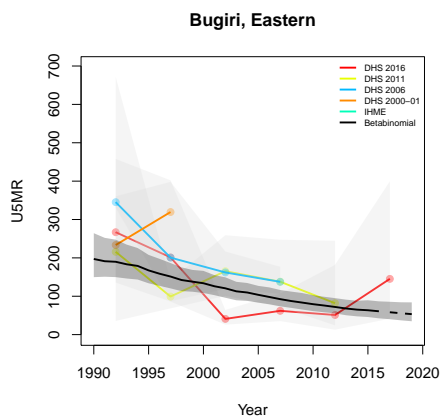
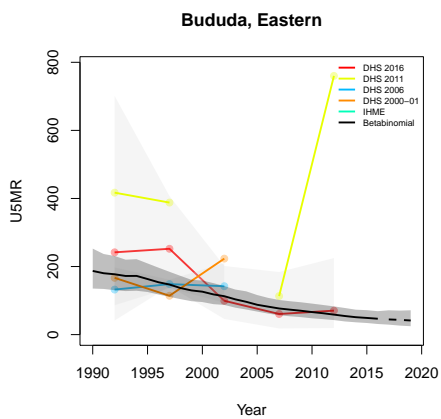


**Amuria, Eastern**

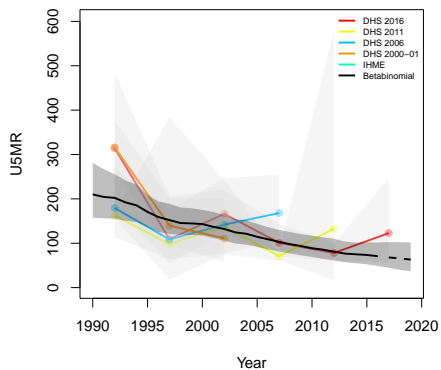


**Budaka, Eastern**

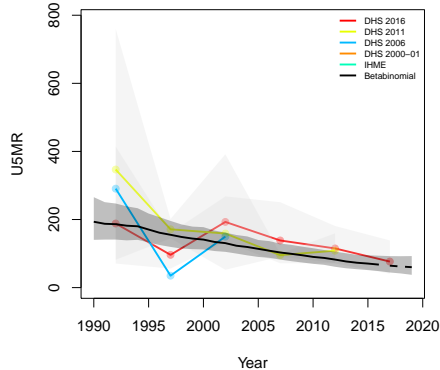




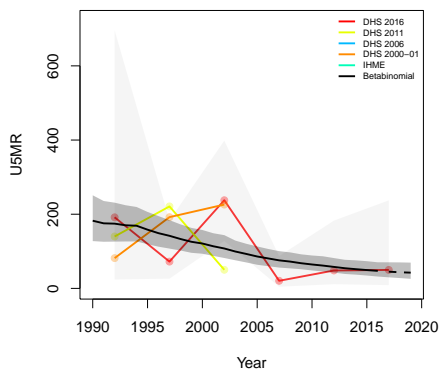
**Busia, Eastern**



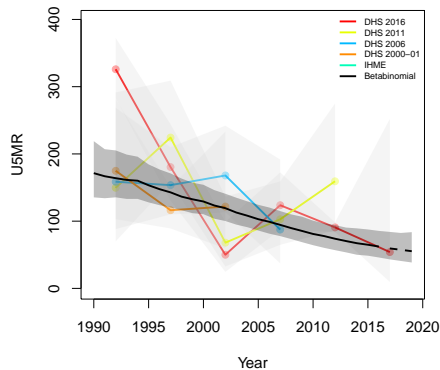
**Butaleja, Eastern**



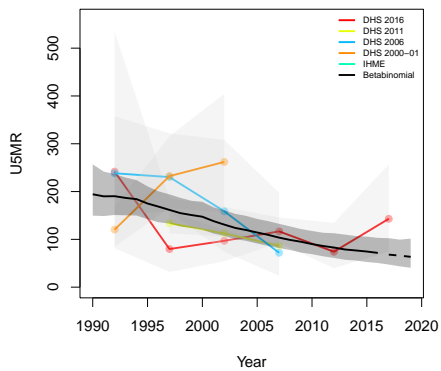
**Butebo, Eastern**



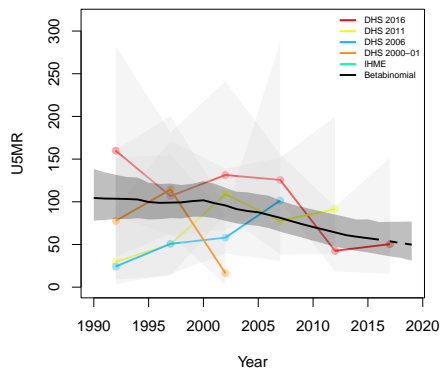
**Buyende, Eastern**



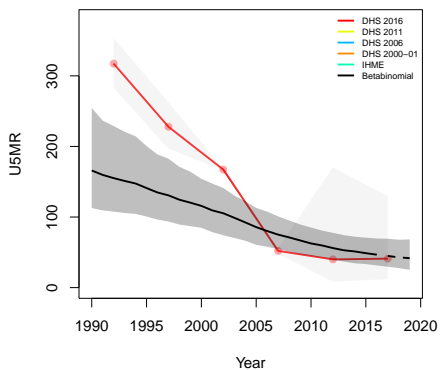
**Iganga, Eastern**



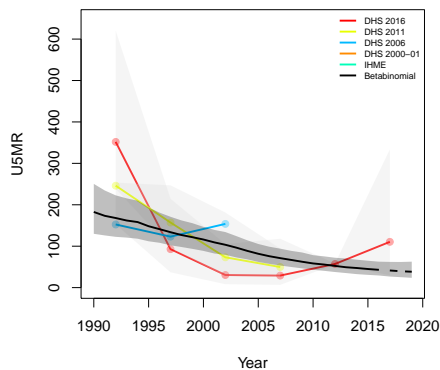
**Jinja, Eastern**



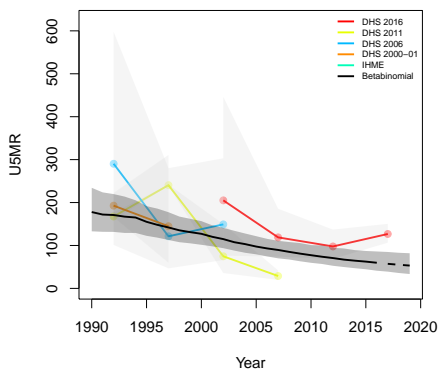
**Kaberaimaido, Eastern**



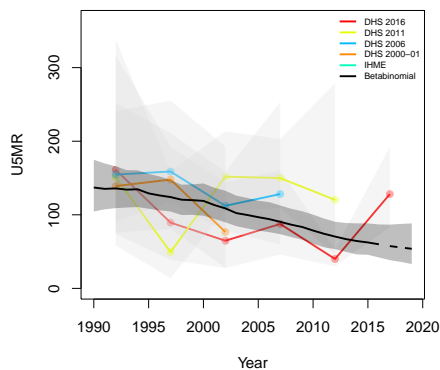
**Kalaki, Eastern**



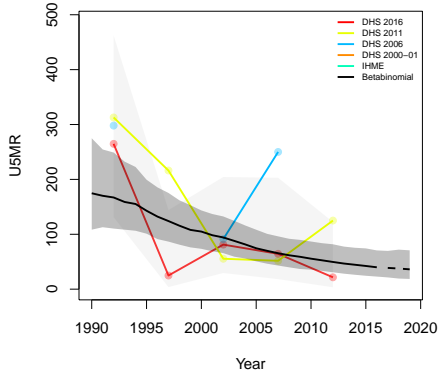
**Kaliro, Eastern**



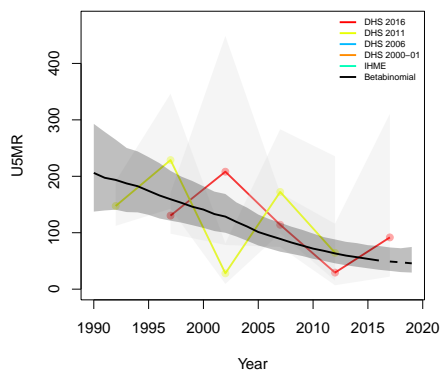
**Kamuli, Eastern**



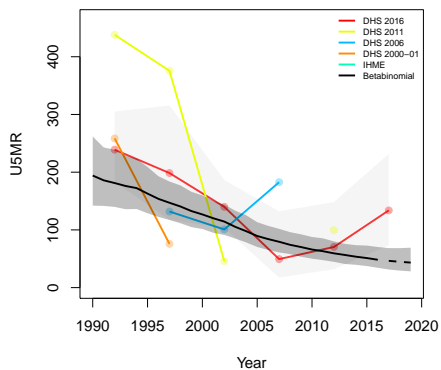
**Kapchorwa, Eastern**



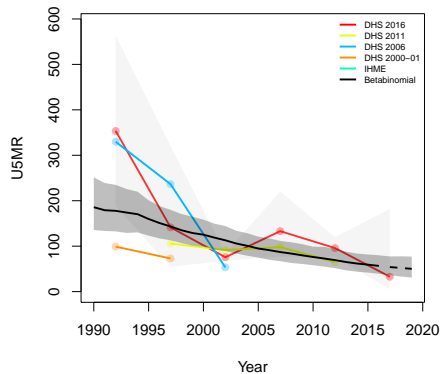
**Kapelebyong, Eastern**



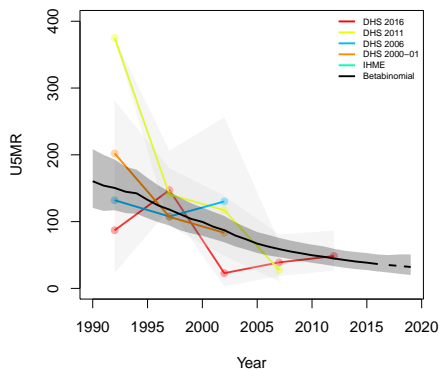
**Katakwi, Eastern**



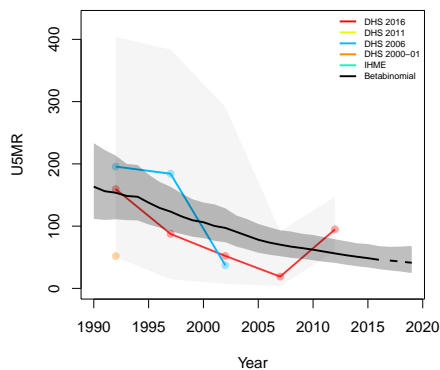
**Kibuku, Eastern**



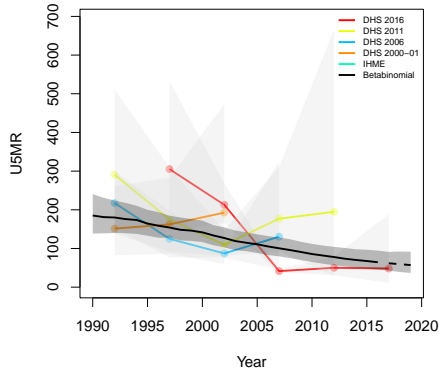
**Kumi, Eastern**



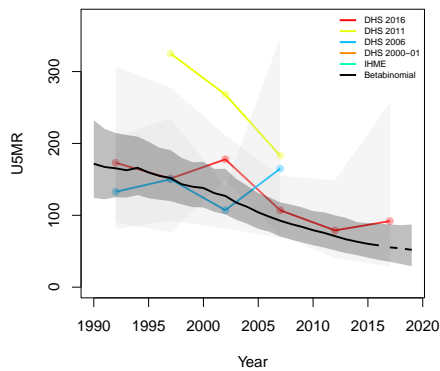
**Kween, Eastern**



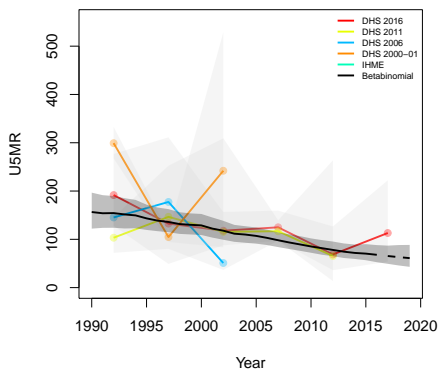
**Luuka, Eastern**



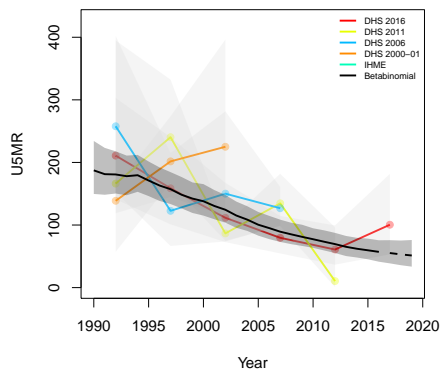
**Manafwa, Eastern**



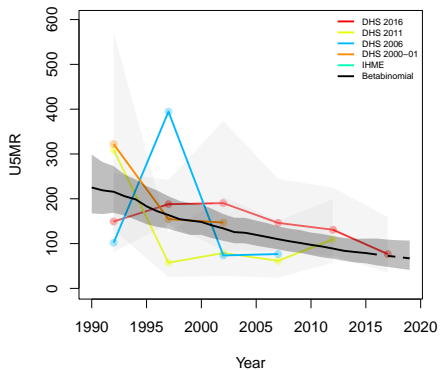
**Mayuge, Eastern**



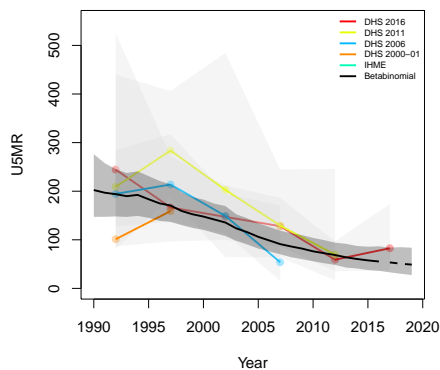
**Mbale, Eastern**



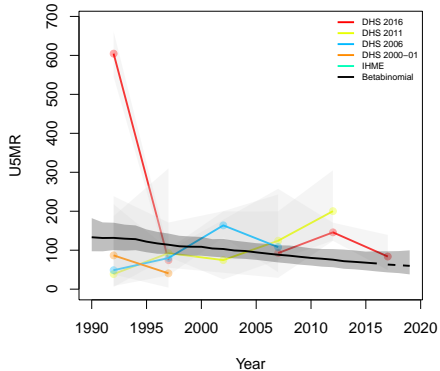
**Namayingo, Eastern**



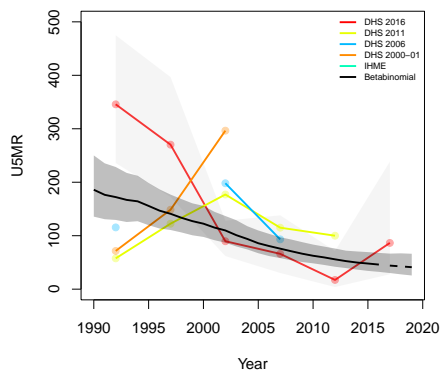
**Namisindwa, Eastern**



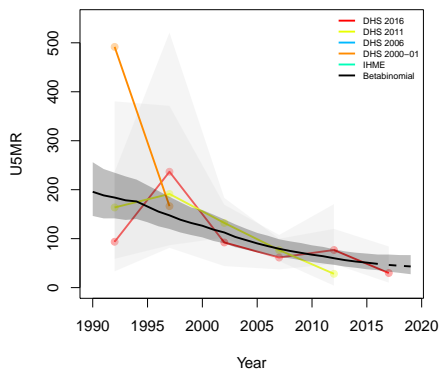
**Namutumba, Eastern**



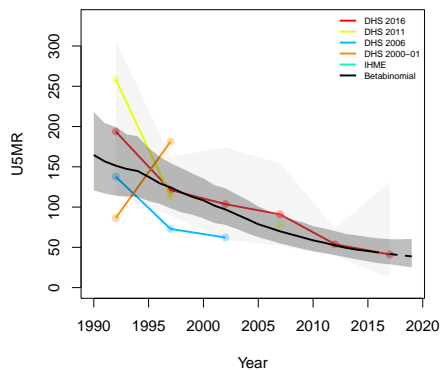
**Ngora, Eastern**



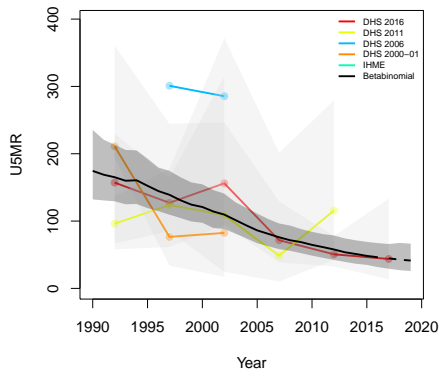
**Pallisa, Eastern**



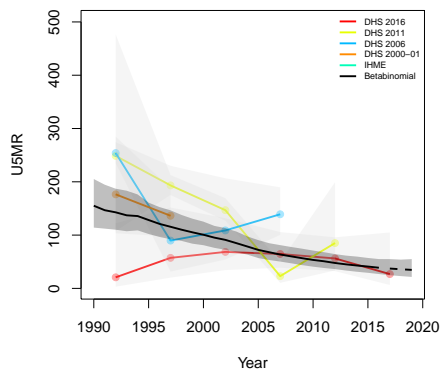
**Serere, Eastern**



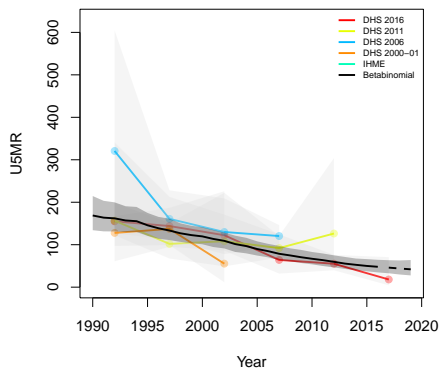
**Sironko, Eastern**



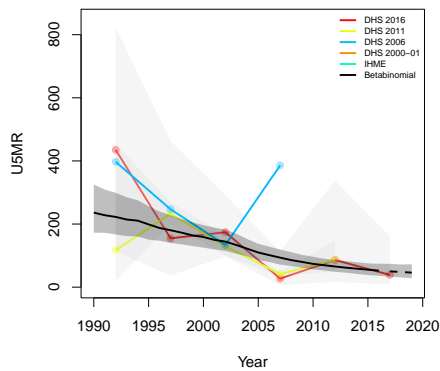
**Soroti, Eastern**

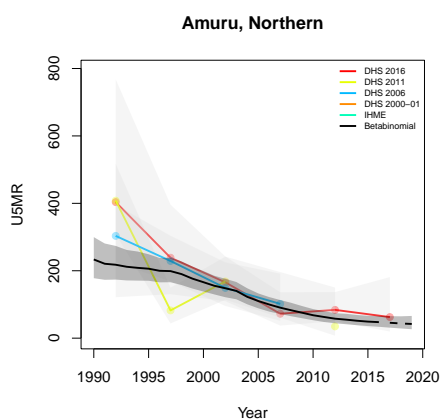
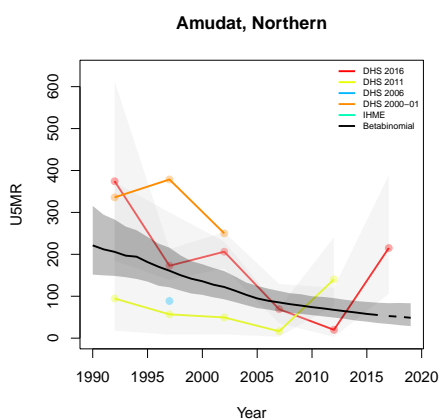
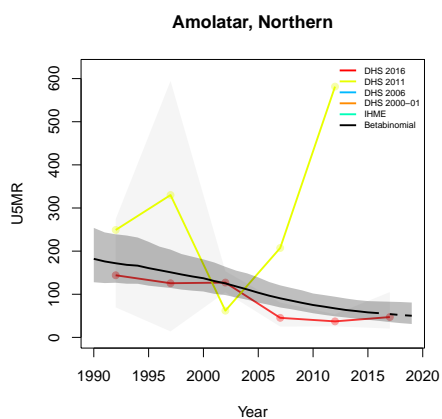
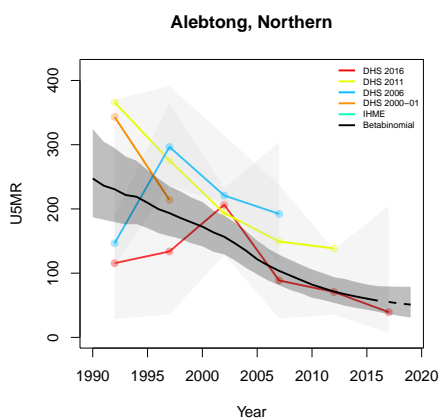
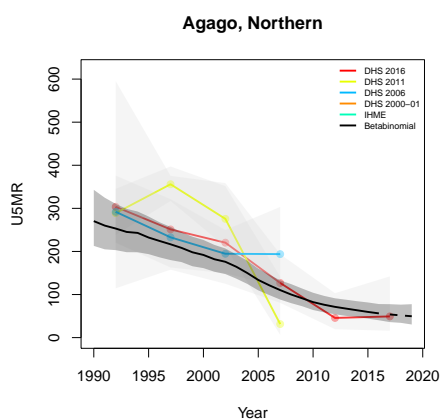
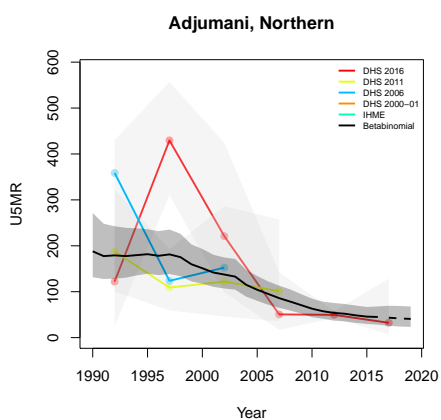


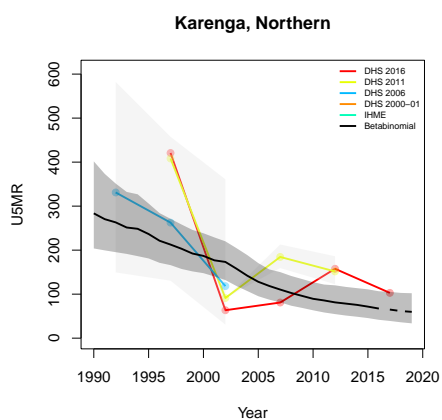
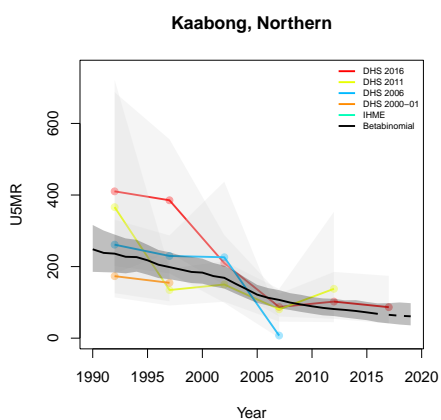
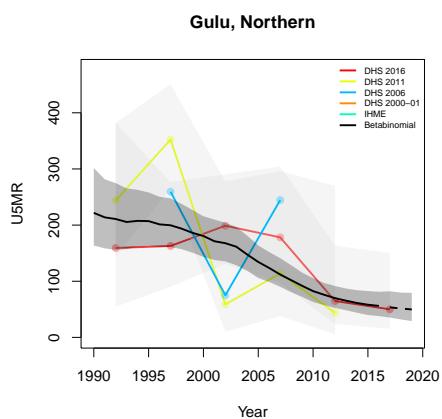
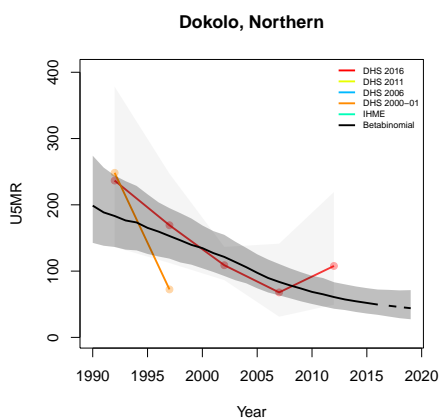
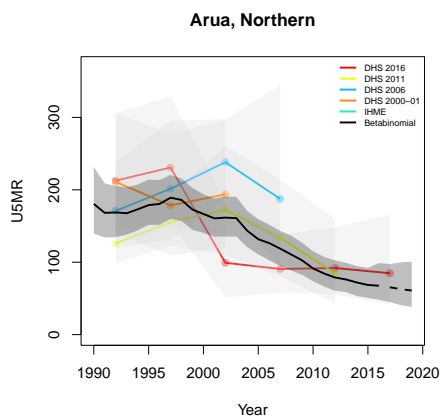
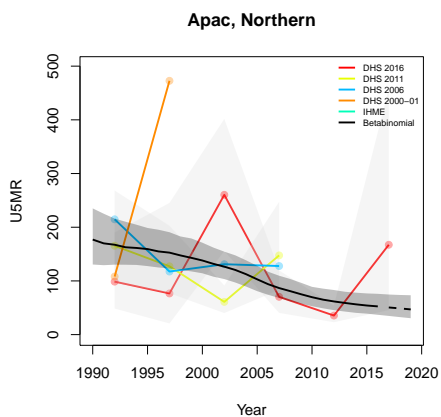
**Tororo, Eastern**



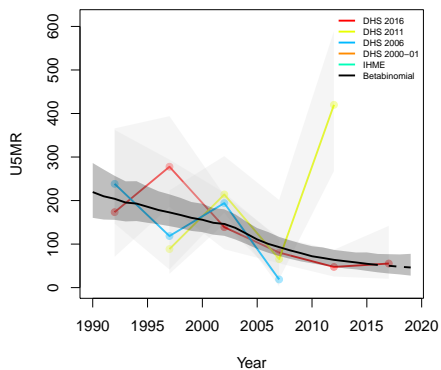
**Abim, Northern**



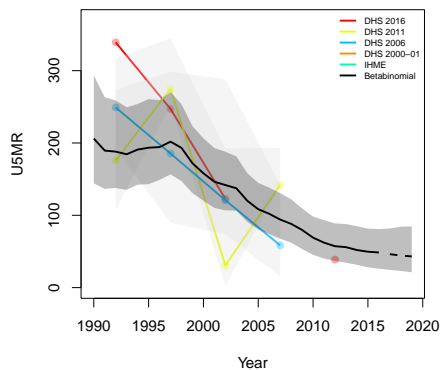




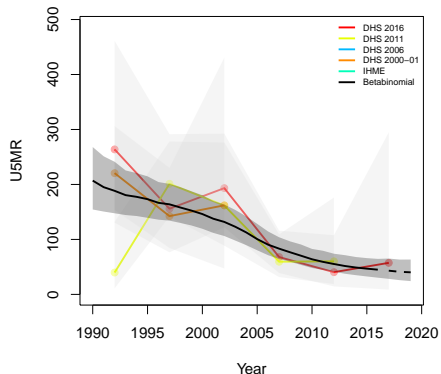
**Kitgum, Northern**



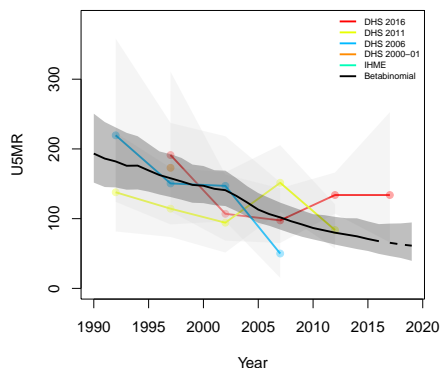
**Koboko, Northern**



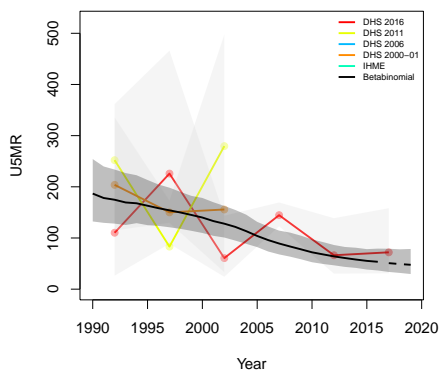
**Kole, Northern**



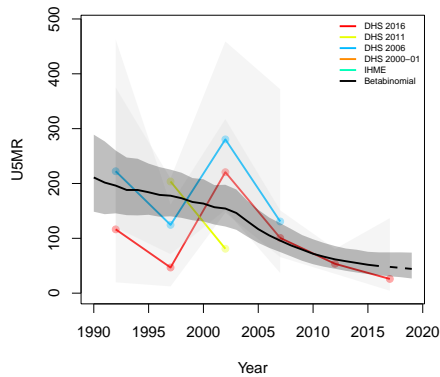
**Kotido, Northern**



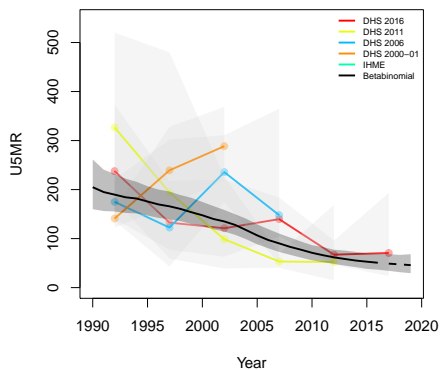
**Kwania, Northern**



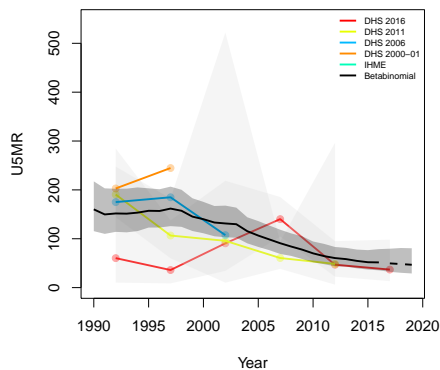
**Lamwo, Northern**



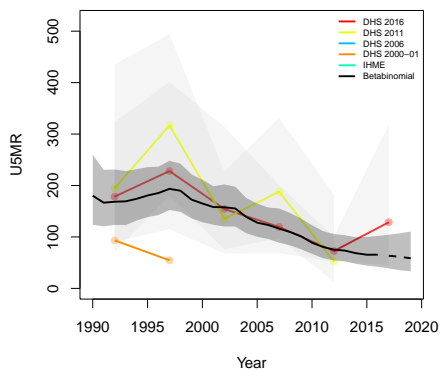
**Lira, Northern**



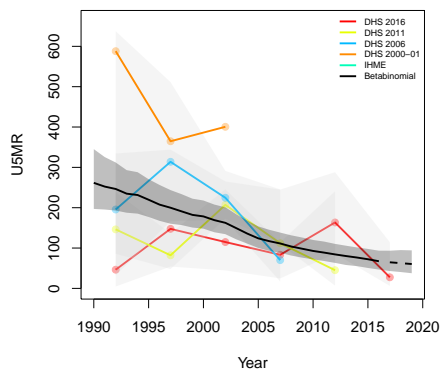
**Madi okollo, Northern**



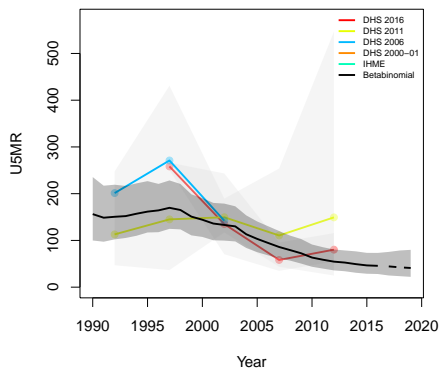
**Maracha, Northern**



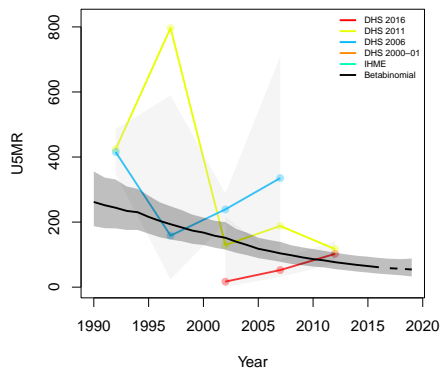
**Moroto, Northern**



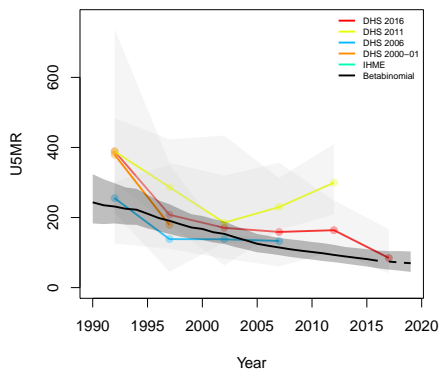
**Moyo, Northern**



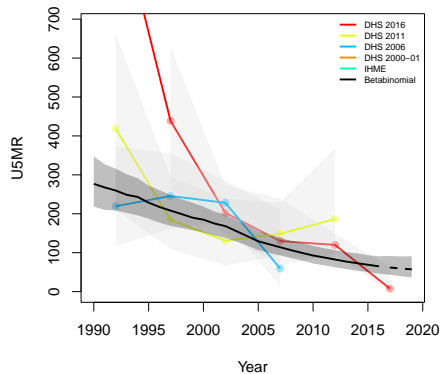
**Nabiatuk, Northern**



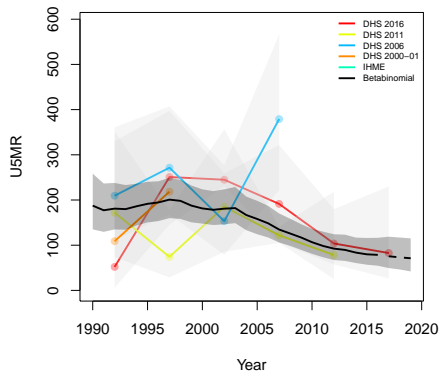
**Nakapiripit, Northern**



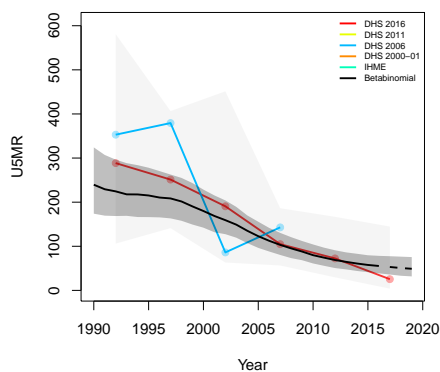
**Napak, Northern**



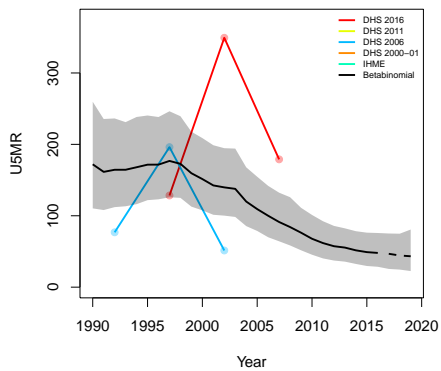
**Nebbi, Northern**



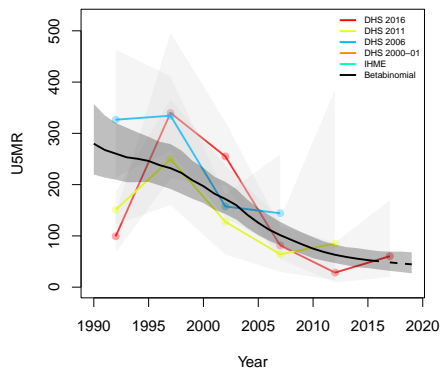
**Nwoya, Northern**



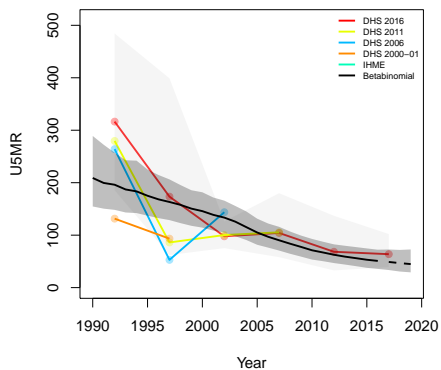
**Obongi, Northern**



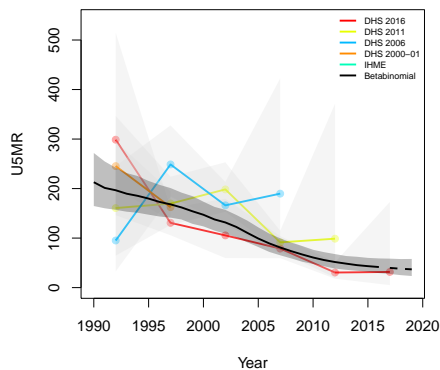
**Omoro, Northern**



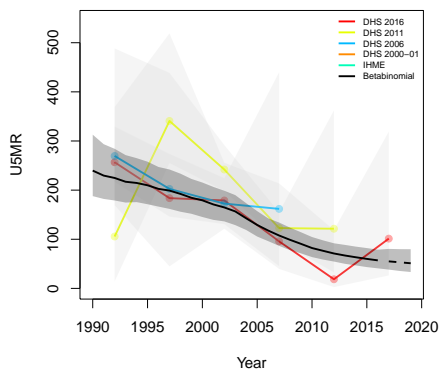
**Otuke, Northern**



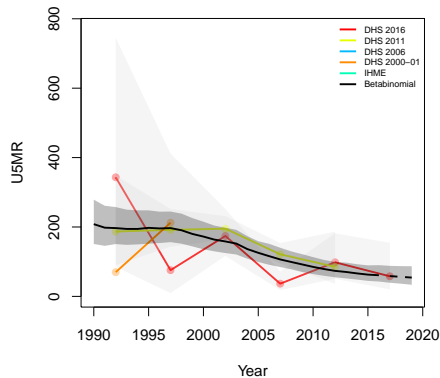
**Oyam, Northern**



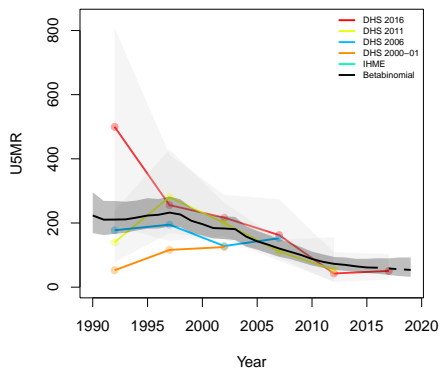
**Pader, Northern**



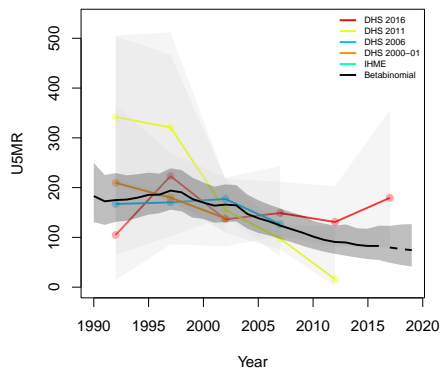
**Pakwach, Northern**



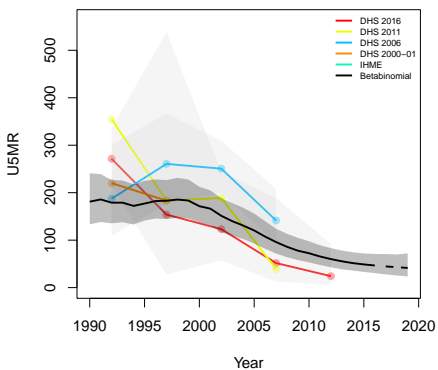
**Yumbe, Northern**



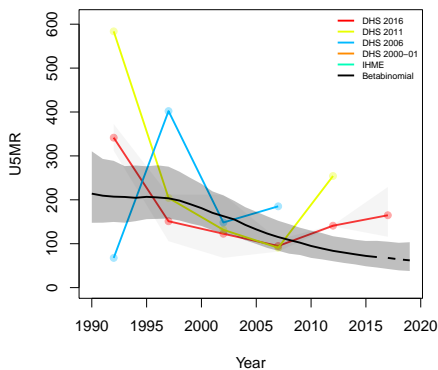
**Zombo, Northern**



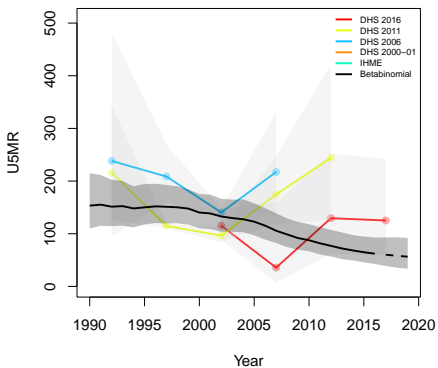
**Buhweju, Western**



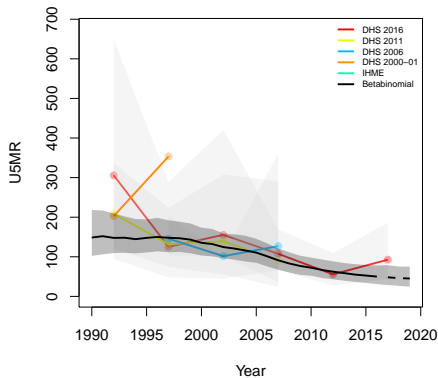
**Buliisa, Western**



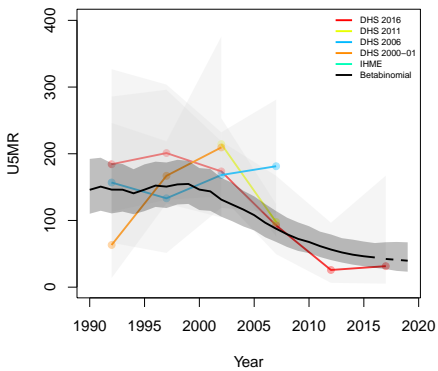
**Bundibugyo, Western**



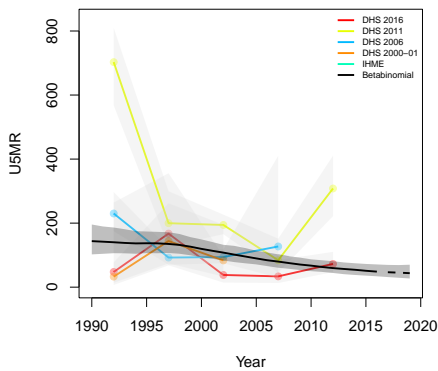
**Bunyangabu, Western**



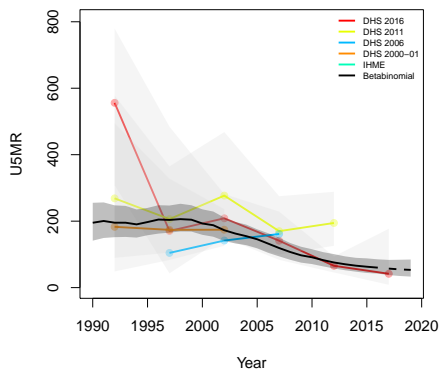
**Bushenyi, Western**



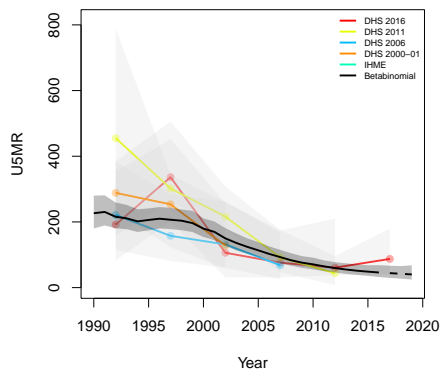
**Hoima, Western**



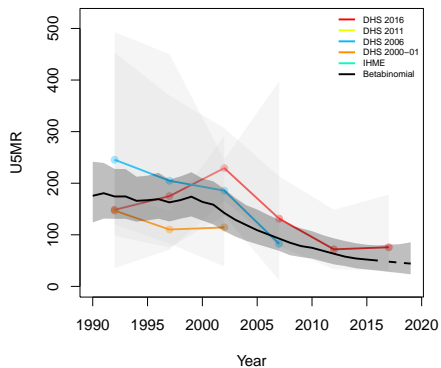
Ibanda, Western



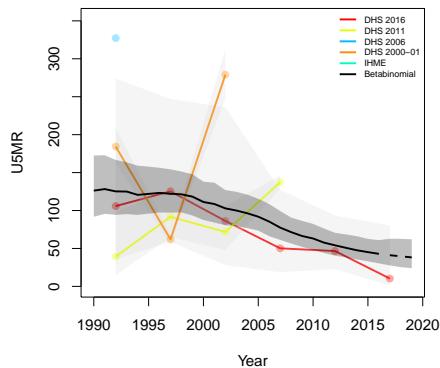
Isingiro, Western



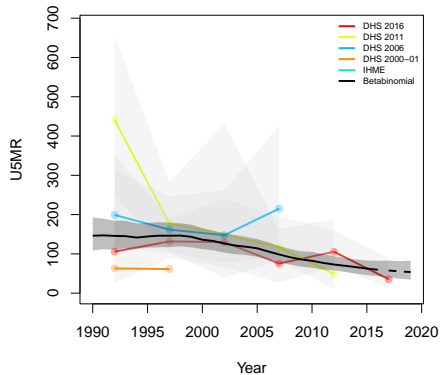
Kabale, Western



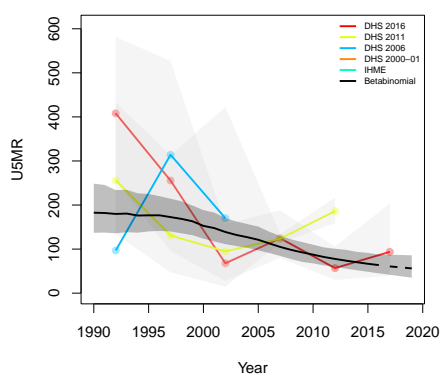
Kabarole, Western



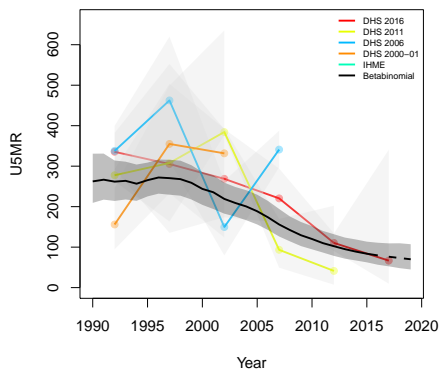
Kagadi, Western



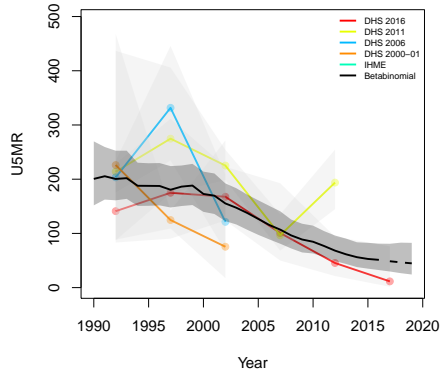
Kakumiro, Western



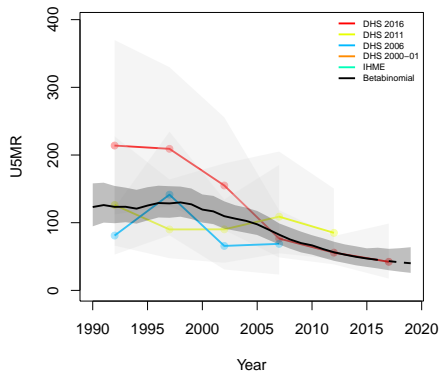
**Kamwenge, Western**



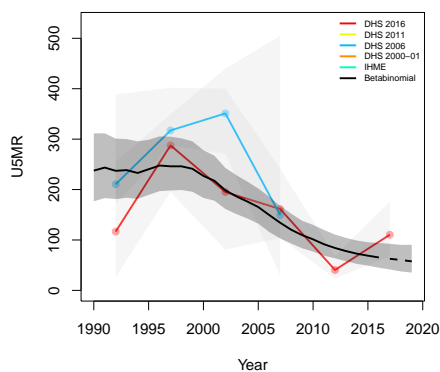
**Kanungu, Western**



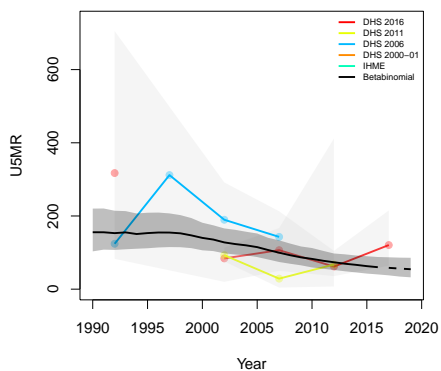
**Kasese, Western**



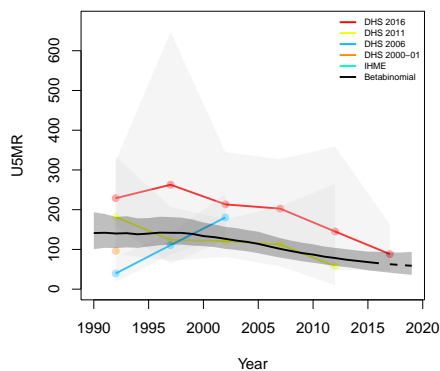
**Kazo, Western**



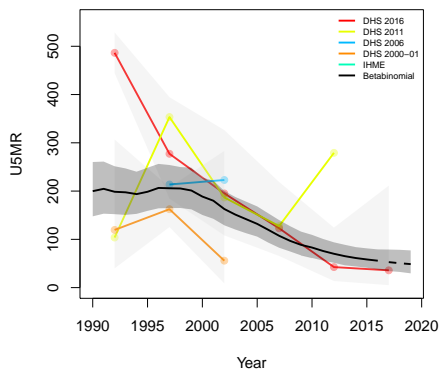
**Kibaale, Western**



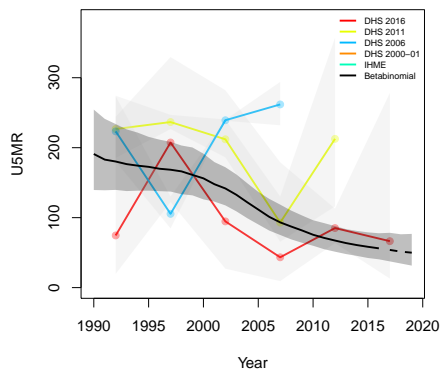
**Kikuube, Western**



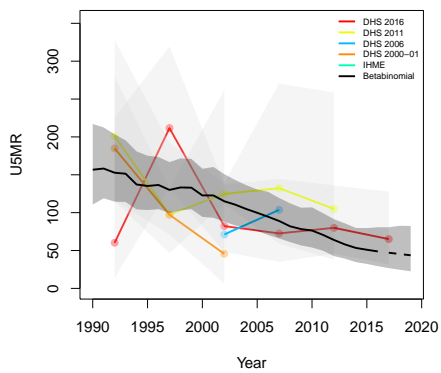
**Kiruhura, Western**



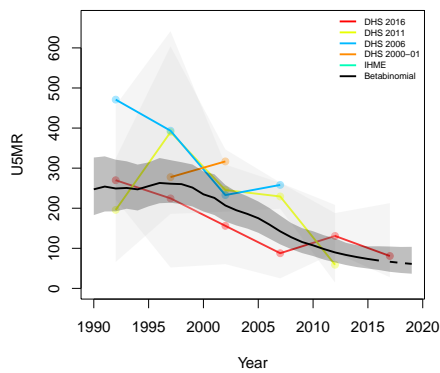
**Kiryandongo, Western**



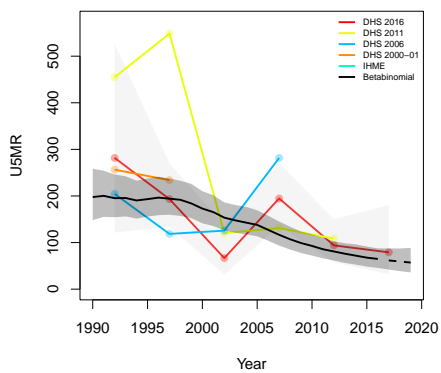
**Kisoro, Western**



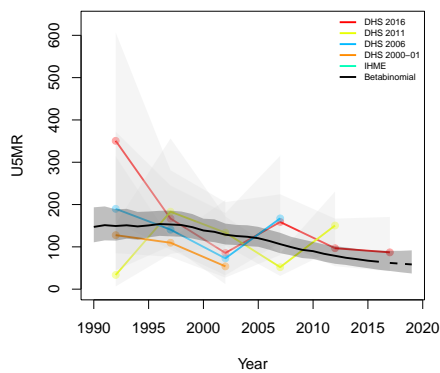
**Kitagwenda, Western**



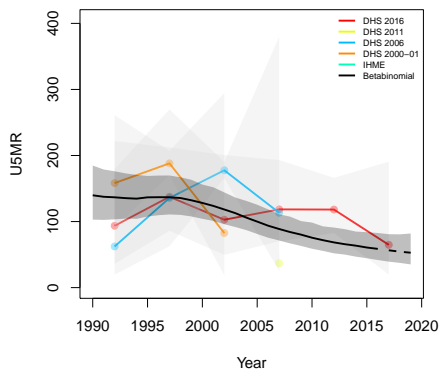
**Kyegegwa, Western**



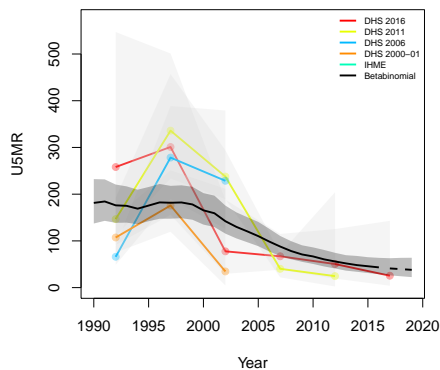
**Kyenjojo, Western**



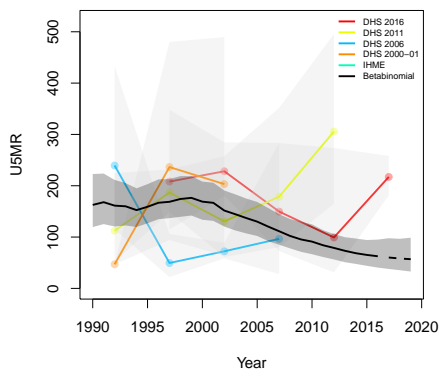
**Masindi, Western**



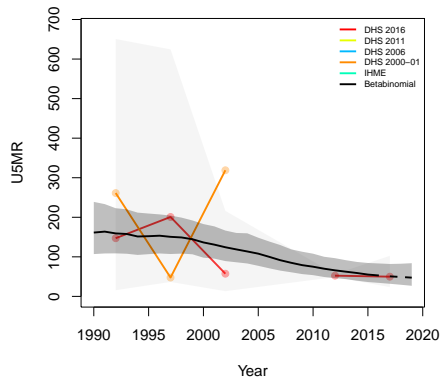
**Mbarara, Western**



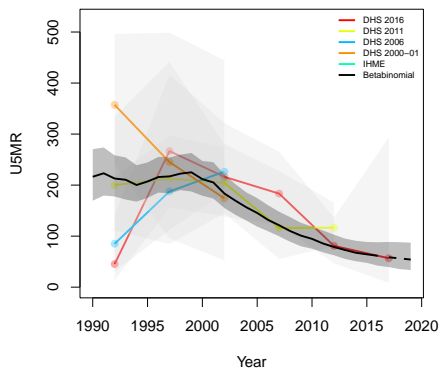
**Mitooma, Western**



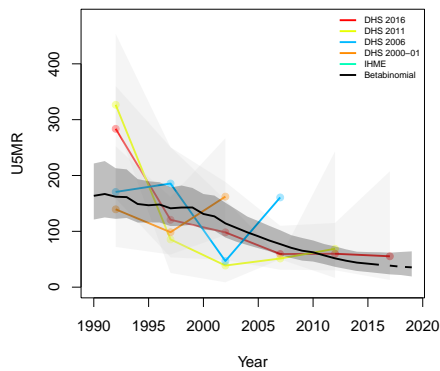
**Ntoroko, Western**

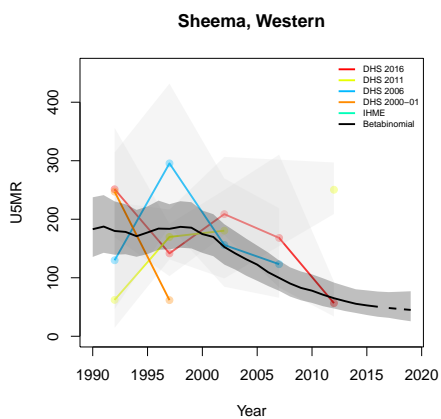
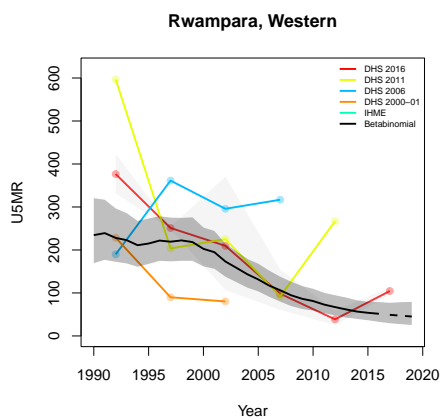
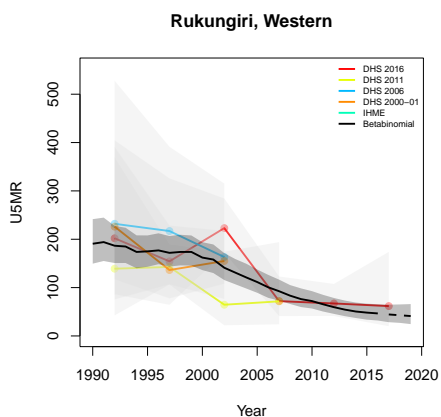
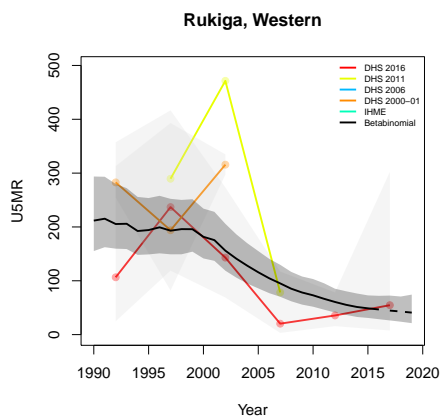
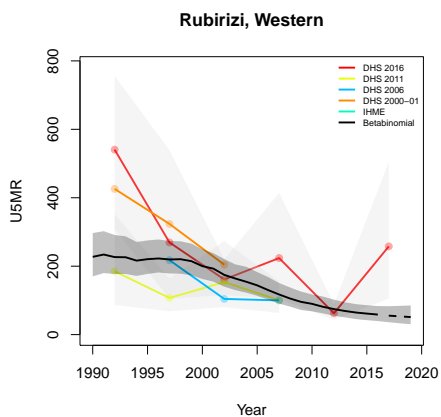


**Ntungamo, Western**



**Rubanda, Western**





**B.20 Zambia**

Age	Survey	Clusters			Deaths			Agemonths		
		Urban	Rural	Total	Urban	Rural	Total	Urban	Rural	Total
0	2007	116	202	318	176	416	592	5956	11432	17388
	2013	305	413	718	463	849	1312	17421	28501	45922
	2018	196	339	535	268	622	890	11536	25464	37000
1-11	2007	116	202	318	277	536	813	60799	115060	175859
	2013	305	413	718	585	985	1570	179776	291800	471576
	2018	196	339	535	317	684	1001	119610	262779	382389
12-23	2007	116	202	318	129	266	395	60957	113069	174026
	2013	305	413	718	290	538	828	181323	292843	474166
	2018	196	339	535	140	310	450	121433	266281	387714
24-35	2007	116	202	318	88	193	281	56958	104074	161032
	2013	305	413	718	173	355	528	169981	272752	442733
	2018	196	339	535	99	255	354	113580	248834	362414
36-47	2007	116	202	318	37	78	115	53317	95868	149185
	2013	305	413	718	77	197	274	159099	253302	412401
	2018	196	339	535	50	125	175	106033	232173	338206
48-59	2007	116	202	318	25	46	71	50169	88496	138665
	2013	305	413	718	47	97	144	148273	234208	382481
	2018	196	339	535	34	53	87	99190	216520	315710

Table B.20: **Data summary for Zambia.** Total numbers of clusters (Columns 3–5) with observations in each age group by survey in urban and rural areas and combined. Numbers of deaths (Columns 6–8) and number of agemonths (Columns 9–10) observed in each age group by survey in urban and rural areas and combined.

*B.20.1 Admin-1*

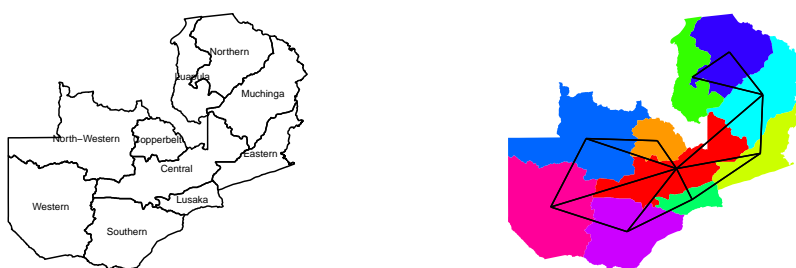


Figure B.100: **Left:** The names of the 10 Admin-1 areas of Zambia . **Right:** The neighborhood structure of Admin-1 areas in Zambia .

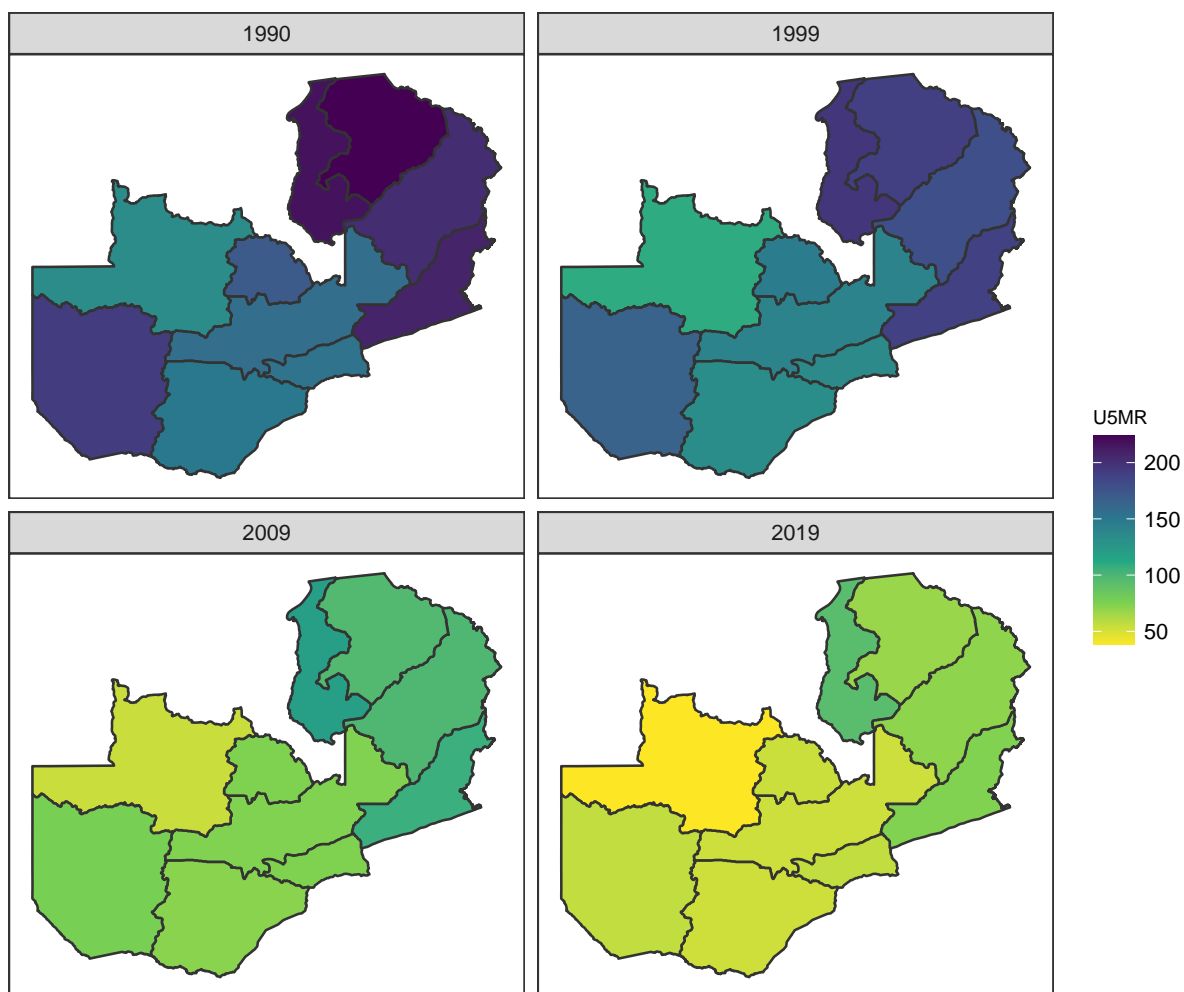


Figure B.101: Median U5MR estimates for years 1990, 1999, 2009, 2019 for Admin-1 areas in Zambia .

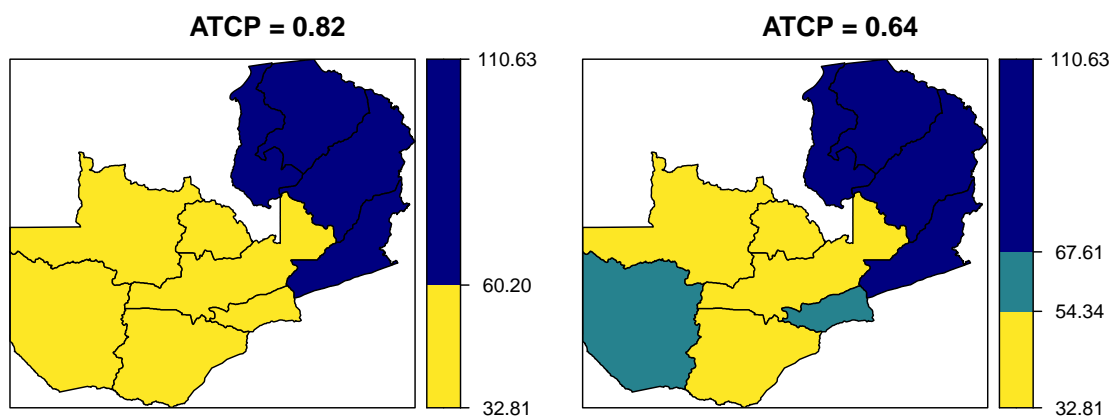
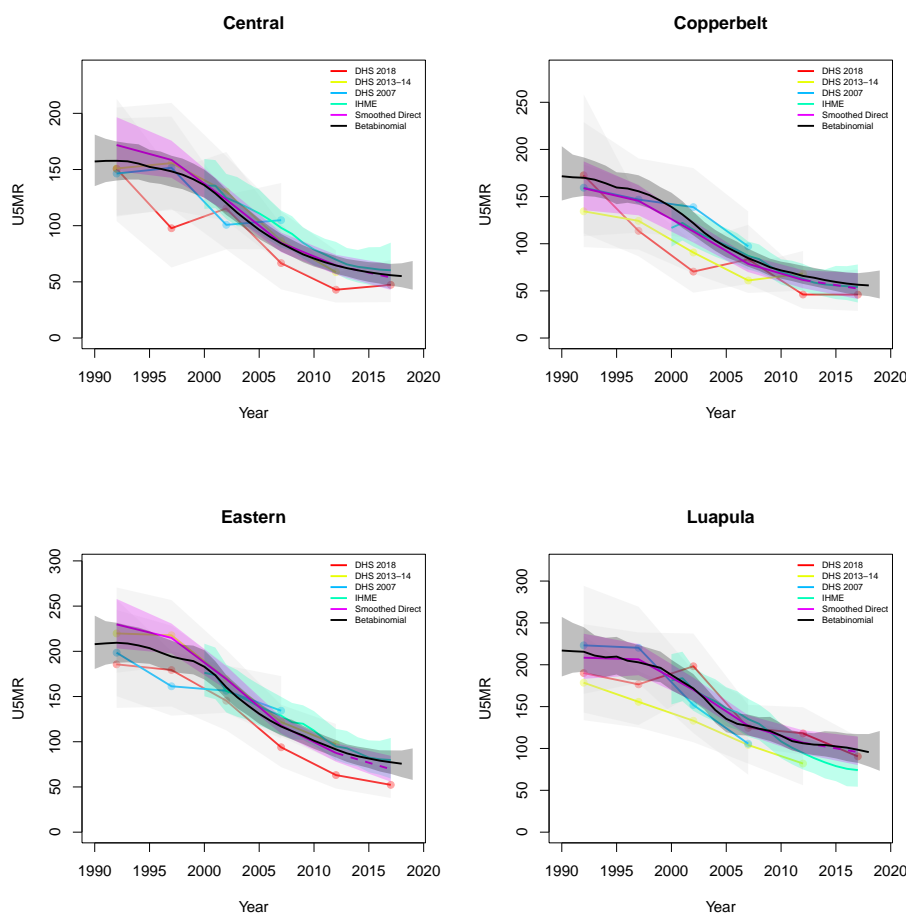


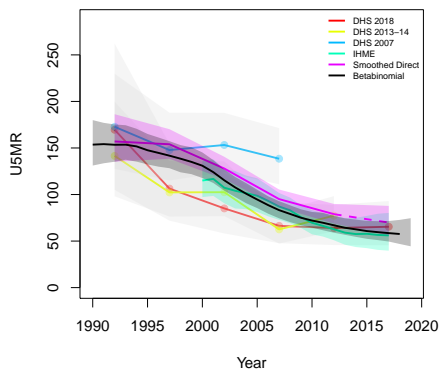
Figure B.102: Expression of uncertainty of U5MR (deaths per 1000 children) estimates for Admin-1 areas based on the average true classification probability (ATCP) in 2019 using  $K = 2, 3$  colors.

*Data and estimates over time by area*

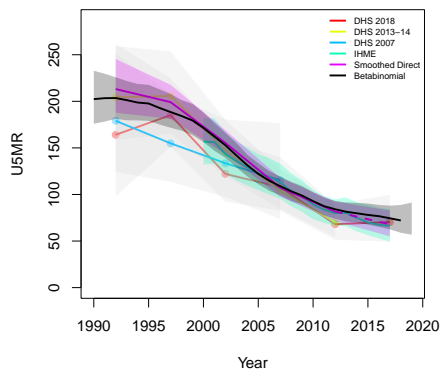
Colored lines with circular points and light grey uncertainty bands are 5-year survey-weighted estimates of U5MR for years 1990–1994 up to 2015–2019 depending on survey timing. For a survey that ends in the middle of a 5-year period, we plot the estimates at the mid-point of the years in that interval for which the survey provides data. Black lines and corresponding intervals represent posterior medians and 95% uncertainty intervals respectively for the betabinomial model. IHME’s estimates and corresponding intervals, where we can compare, are in aquamarine.



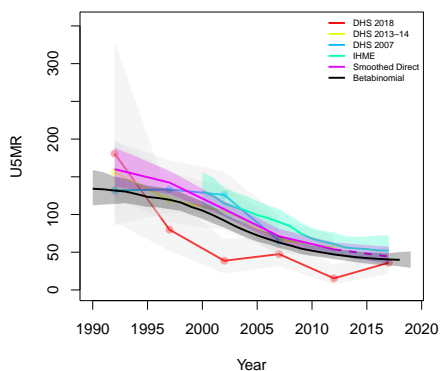
**Lusaka**



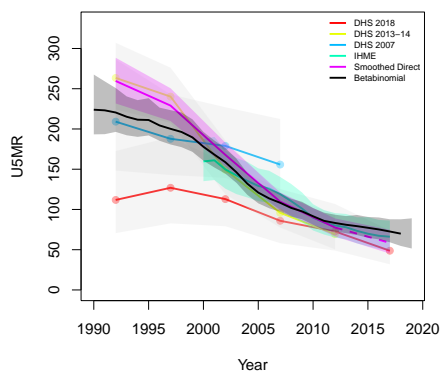
**Muchinga**



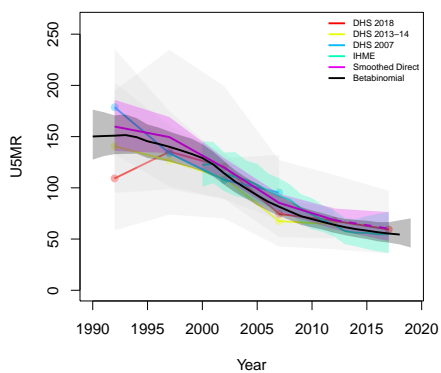
**North-Western**



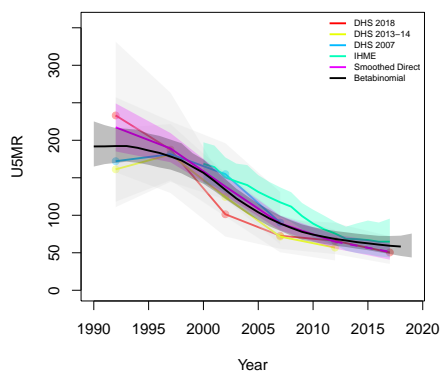
**Northern**



**Southern**



**Western**



*B.20.2 Admin-2*



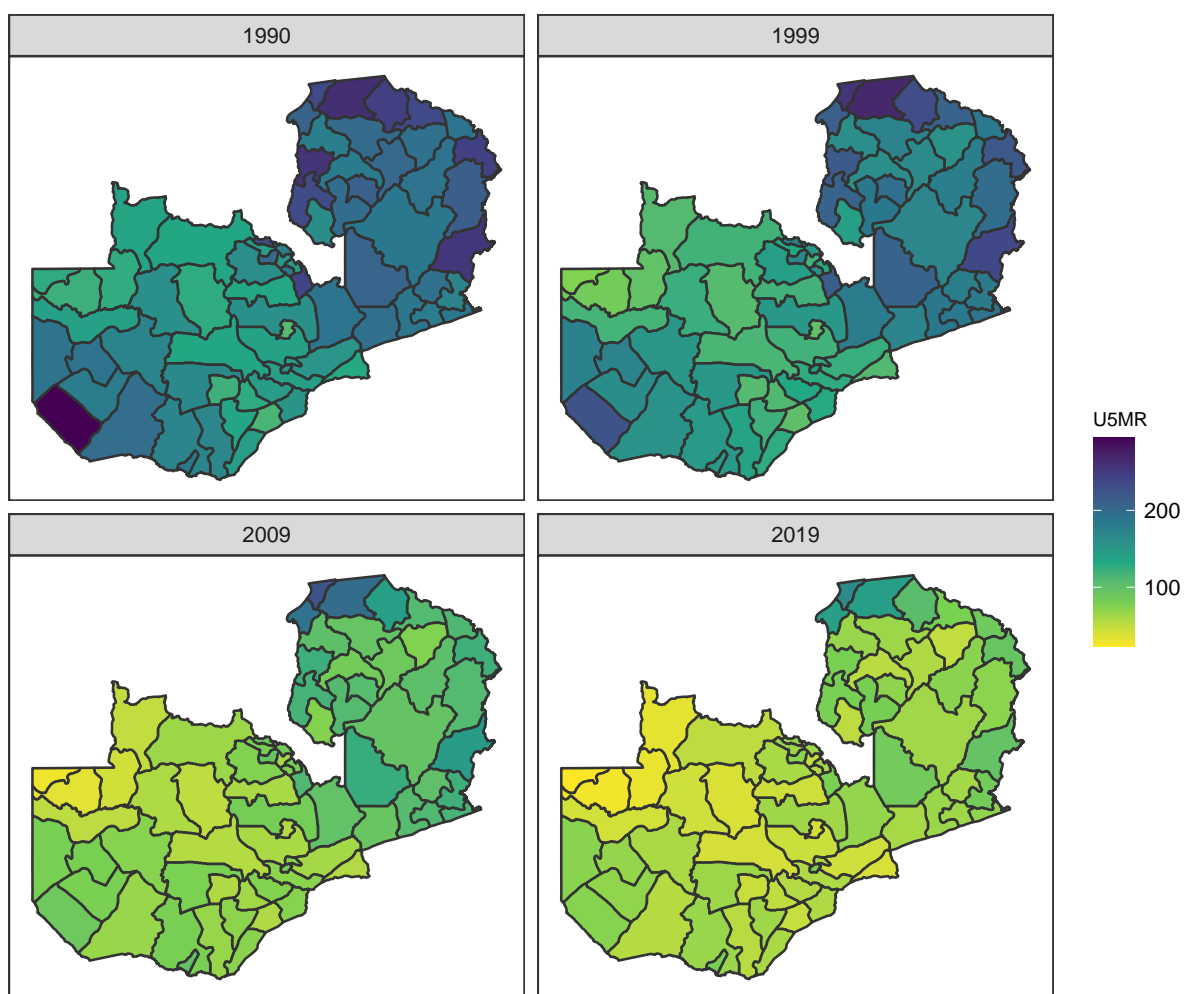


Figure B.104: Median U5MR estimates for years 1990, 1999, 2009, 2019 for Admin-2 areas in Zambia .

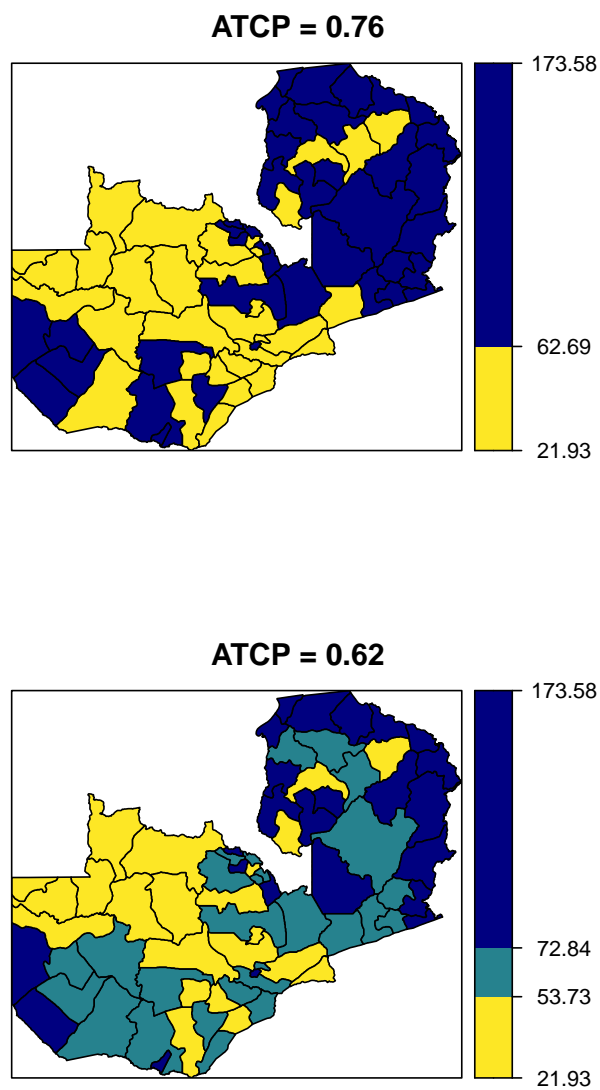
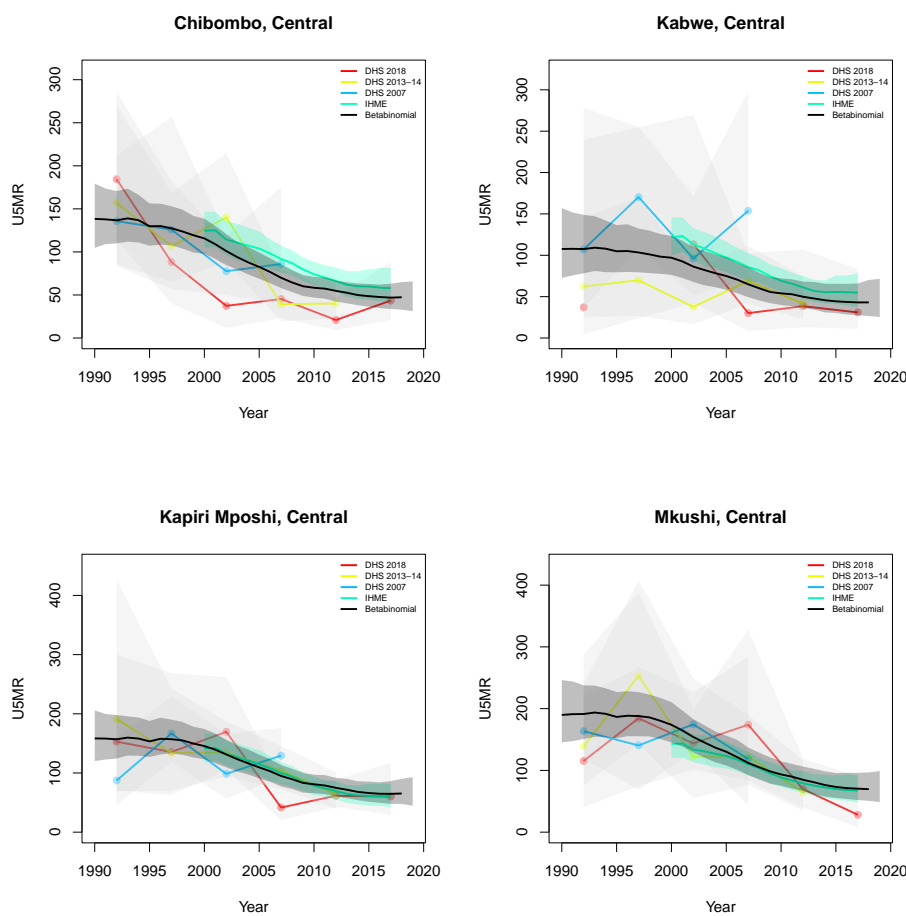


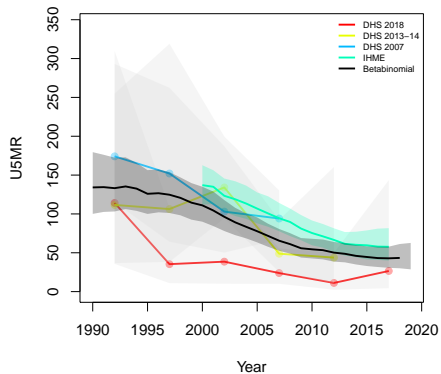
Figure B.105: Expression of uncertainty of U5MR (deaths per 1000 children) estimates for Admin-1 areas based on the average true classification probability (ATCP) in 2019 using  $K = 2, 3$  colors.

*Data and estimates over time by area*

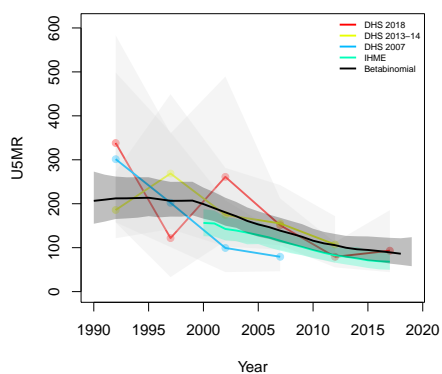
Colored lines with circular points and light grey uncertainty bands are 5-year survey-weighted estimates of U5MR for years 1990–1994 up to 2015–2019 depending on survey timing. For a survey that ends in the middle of a 5-year period, we plot the estimates at the mid-point of the years in that interval for which the survey provides data. Black lines and corresponding intervals represent posterior medians and 95% uncertainty intervals respectively for the betabinomial model. IHME’s estimates and corresponding intervals, where we can compare, are in aquamarine.



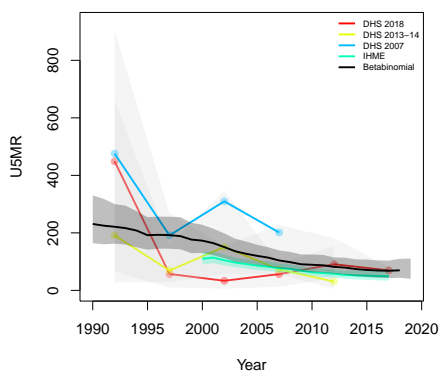
**Mumbwa, Central**



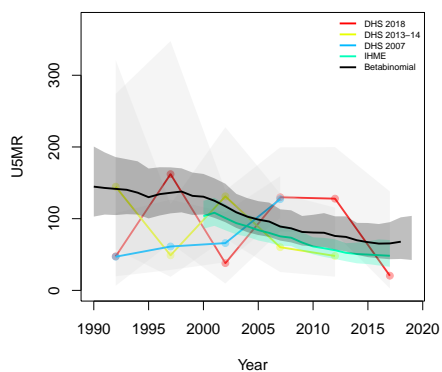
**Serenje, Central**



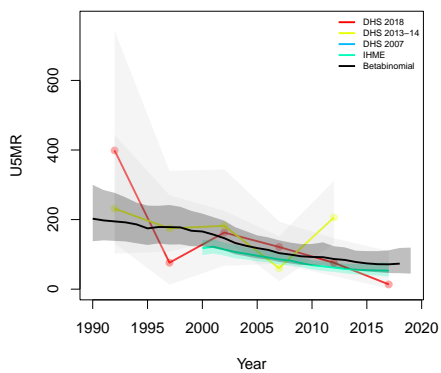
**Chililabombwe, Copperbelt**



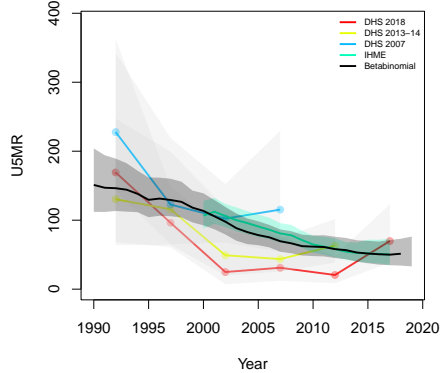
**Chingola, Copperbelt**



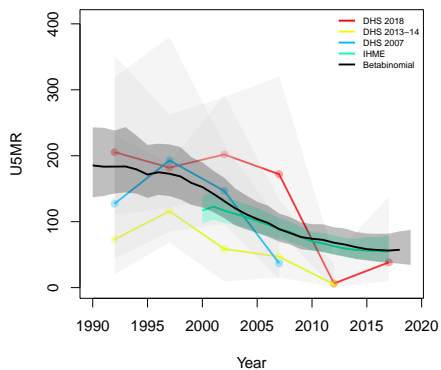
**Kalulushi, Copperbelt**



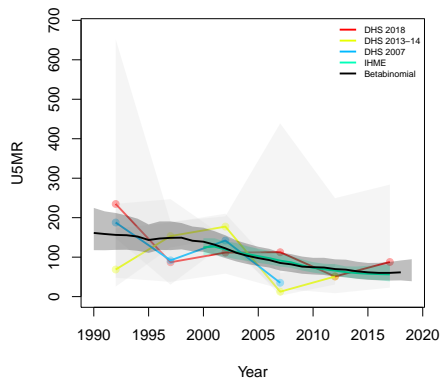
**Kitwe, Copperbelt**



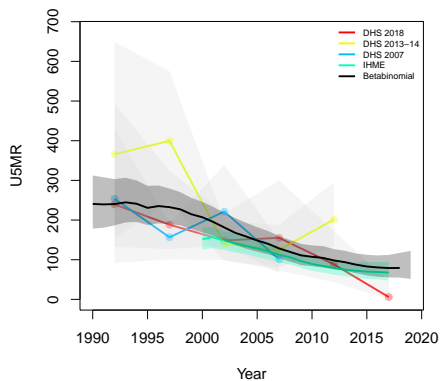
**Luanshya, Copperbelt**



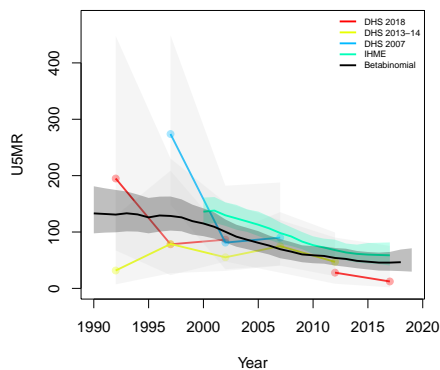
**Lufwanyama, Copperbelt**



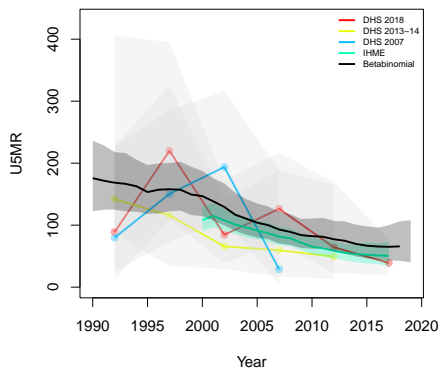
**Masaiti, Copperbelt**



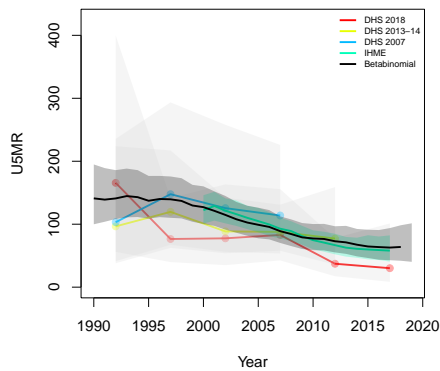
**MPongwe, Copperbelt**



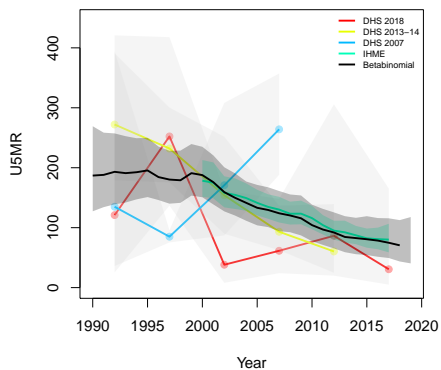
**Mufulira, Copperbelt**



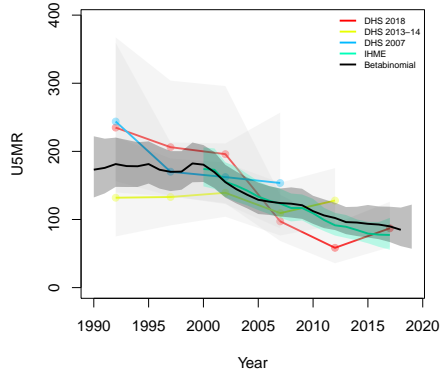
**Ndola, Copperbelt**



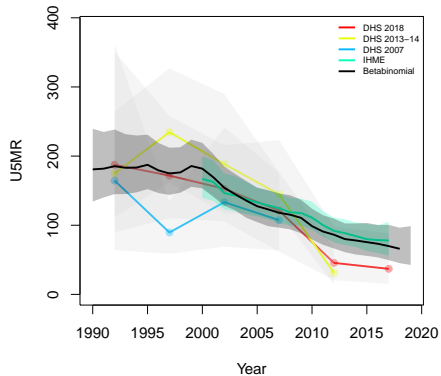
**Chadiza, Eastern**



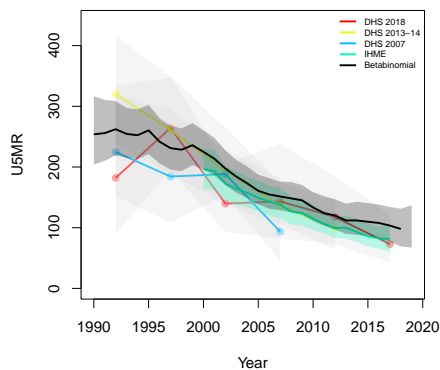
**Chipata, Eastern**



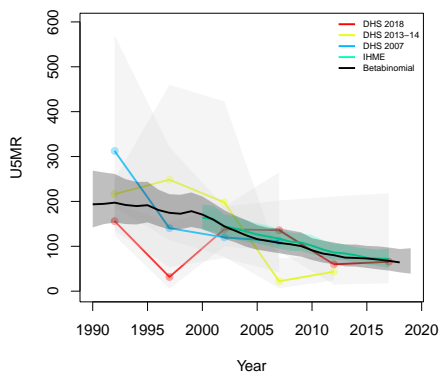
**Katete, Eastern**



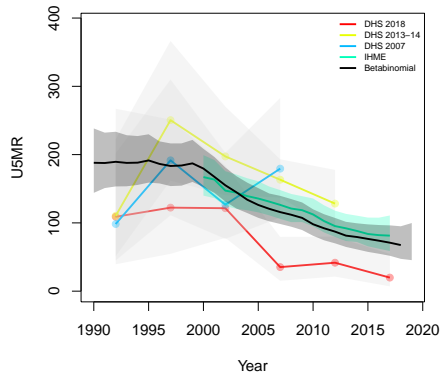
**Lundazi, Eastern**



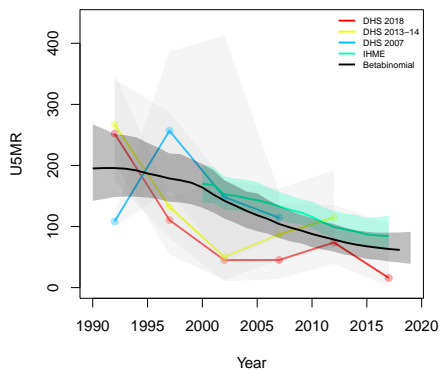
**Mambwe, Eastern**



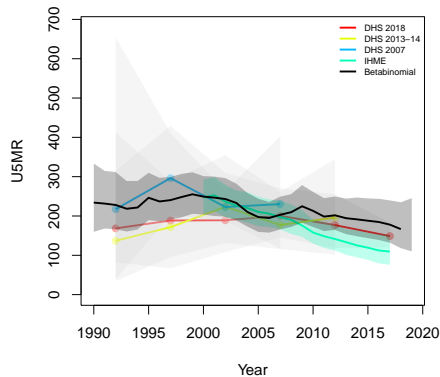
**Nyimba, Eastern**



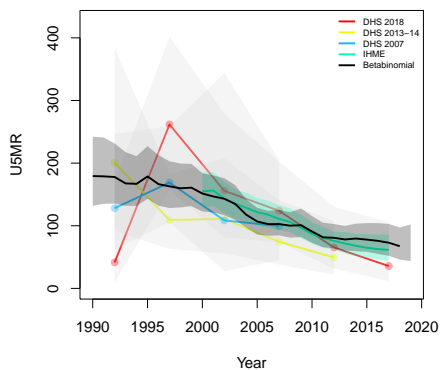
**Petauke, Eastern**



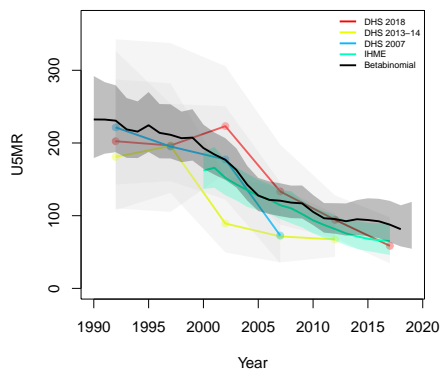
**Chiengi, Luapula**



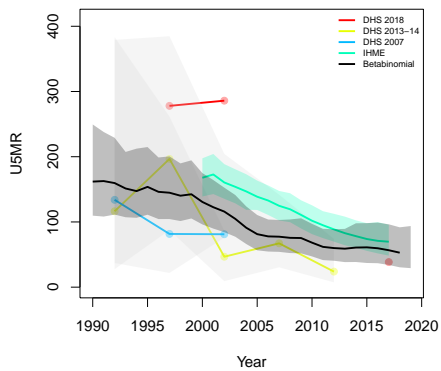
**Kawambwa, Luapula**



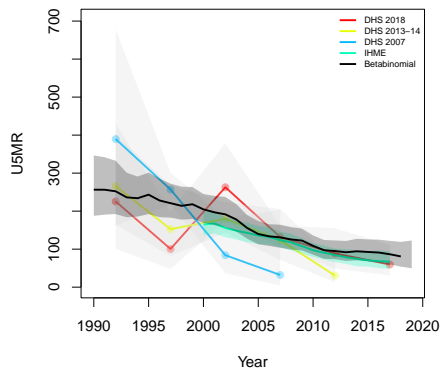
**Mansa, Luapula**



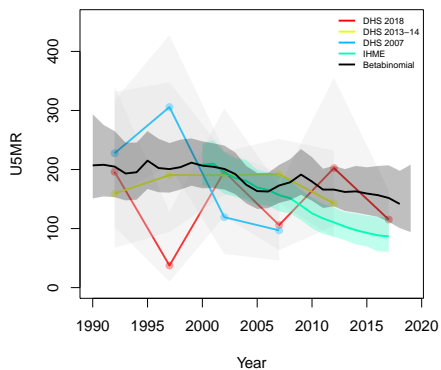
**Milenge, Luapula**



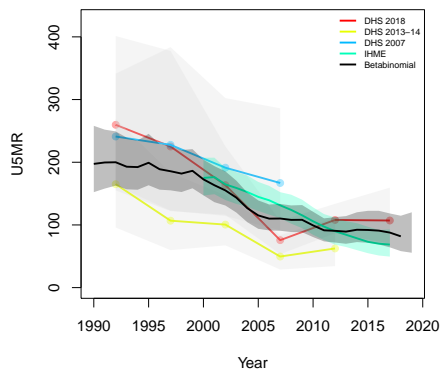
**Mwense, Luapula**



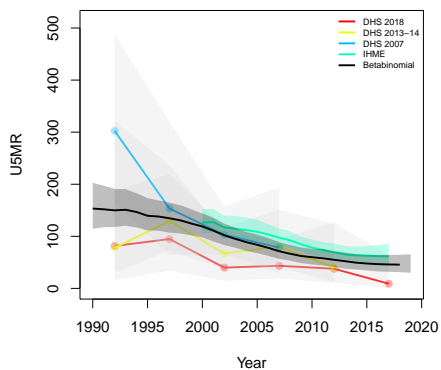
**Nchelenge, Luapula**



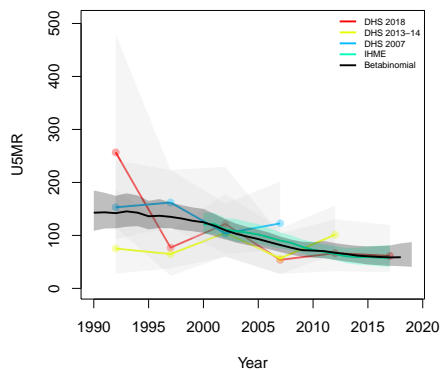
**Samfya, Luapula**



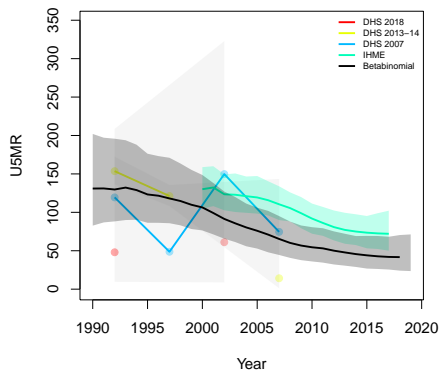
**Chongwe, Lusaka**



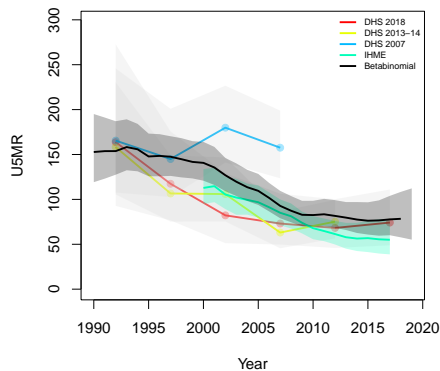
**Kafue, Lusaka**



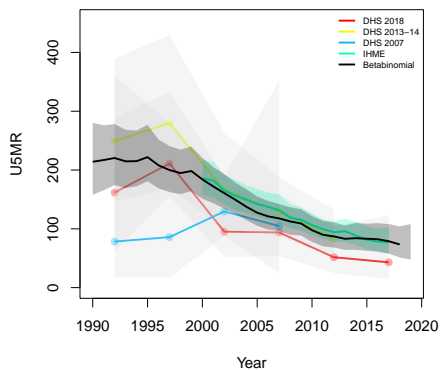
**Luangwa, Lusaka**



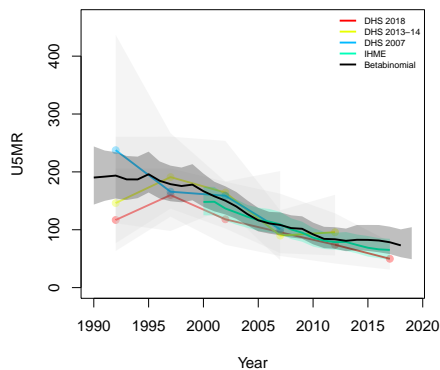
**Lusaka, Lusaka**



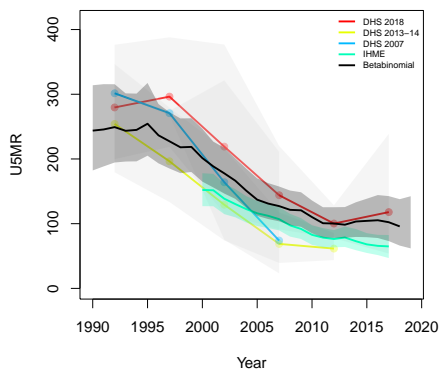
**Chama, Muchinga**



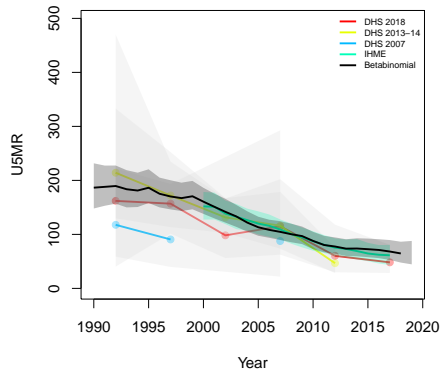
**Chinsali, Muchinga**



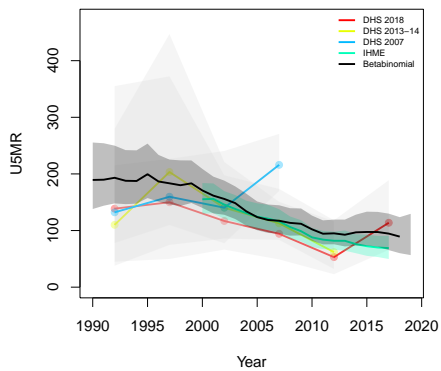
**Isoka, Muchinga**



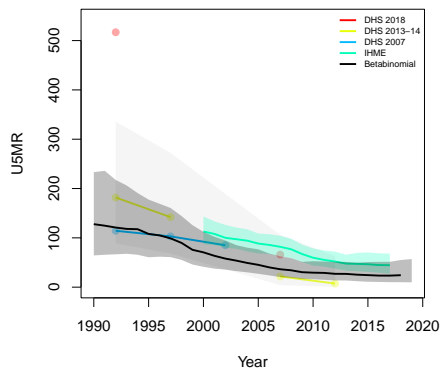
**Mpika, Muchinga**



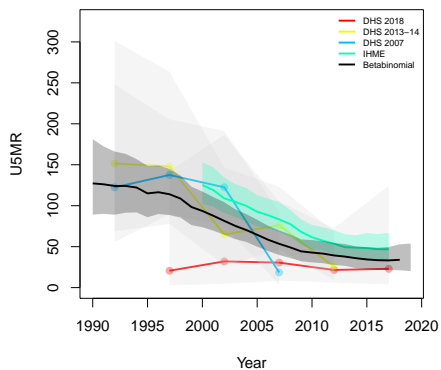
**Nakonde, Muchinga**



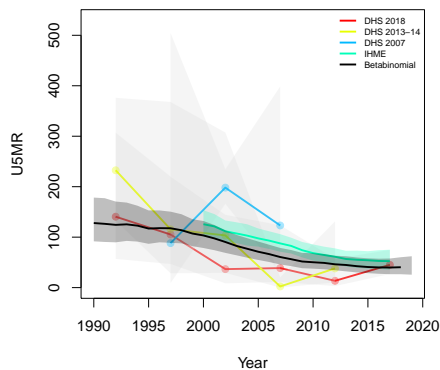
**Chavuma, North-Western**



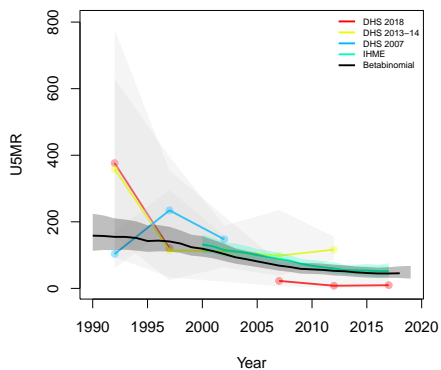
**Kabompo, North-Western**



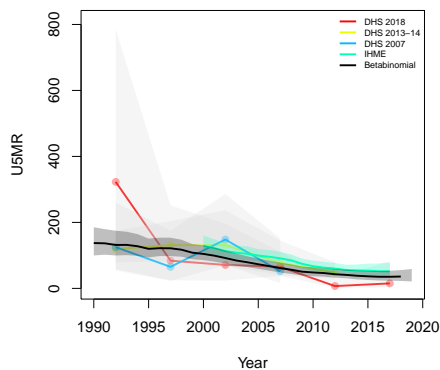
**Kasempa, North-Western**



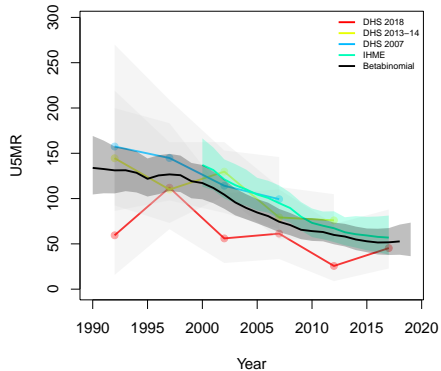
**Mufumbwe, North-Western**



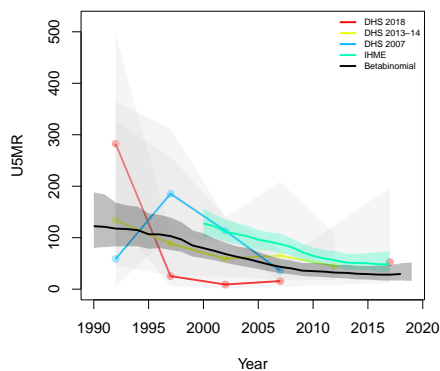
**Mwinilunga, North-Western**



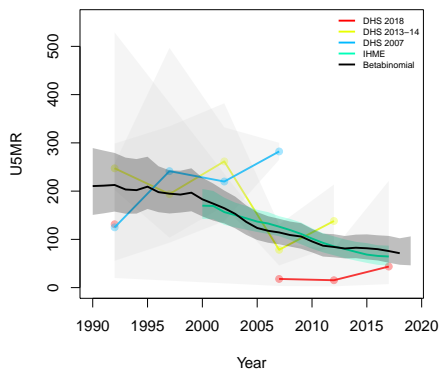
**Solwezi, North-Western**



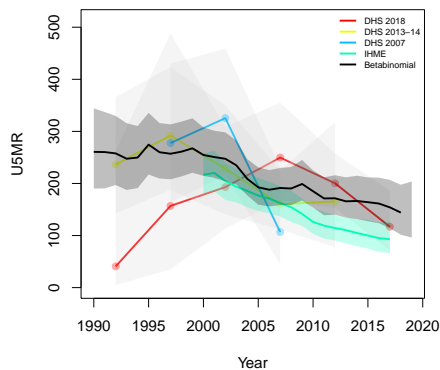
**Zambezi, North-Western**



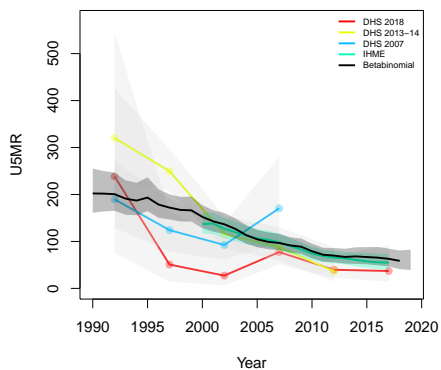
**Chilubi, Northern**



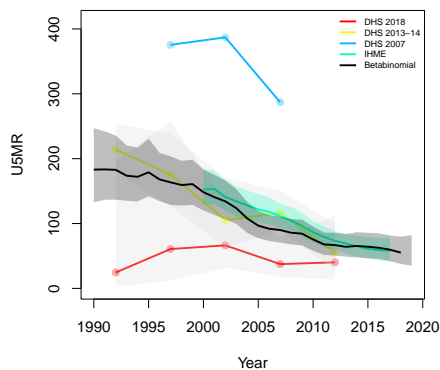
**Kaputa, Northern**



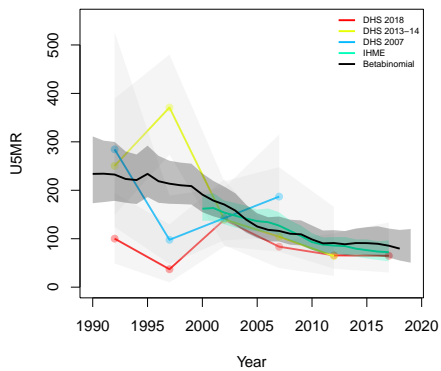
**Kasama, Northern**



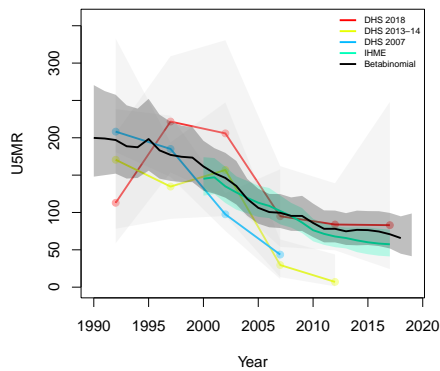
**Luwingu, Northern**



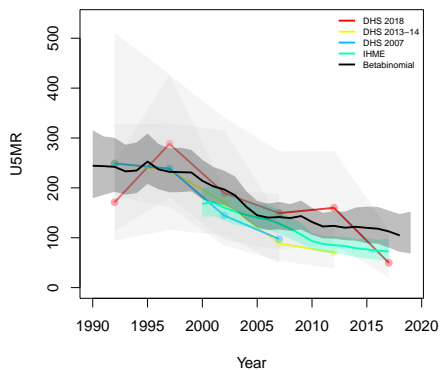
**Mbala, Northern**



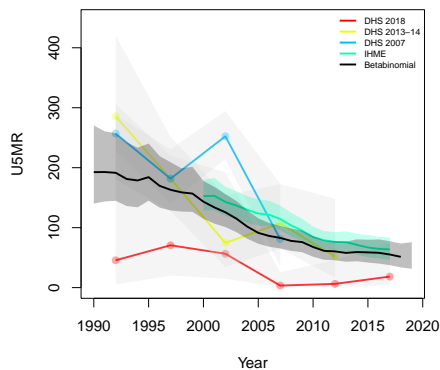
**Mporokoso, Northern**



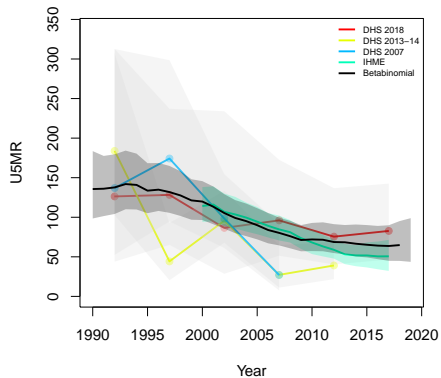
**Mpulungu, Northern**



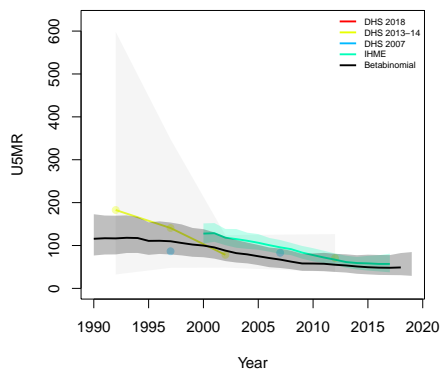
**Mungwi, Northern**



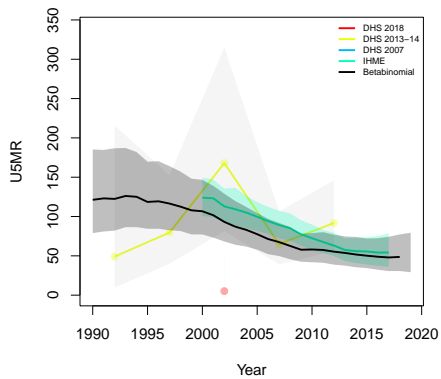
**Choma, Southern**



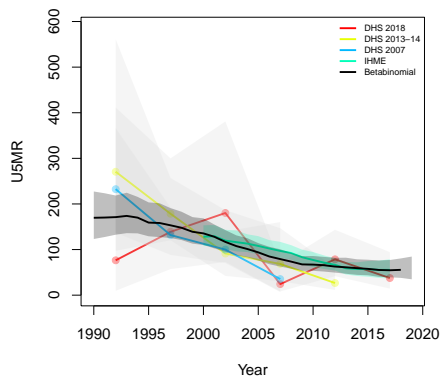
**Gwembe, Southern**



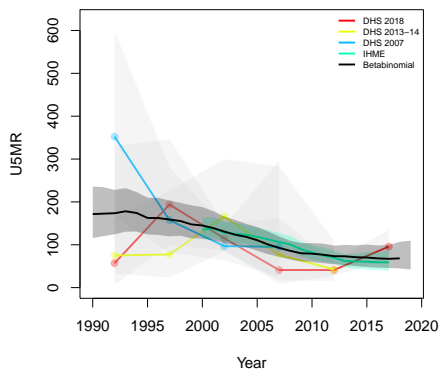
**Itzhi-Tezhi, Southern**



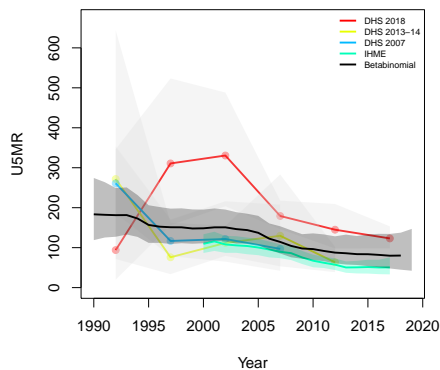
**Kalomo, Southern**



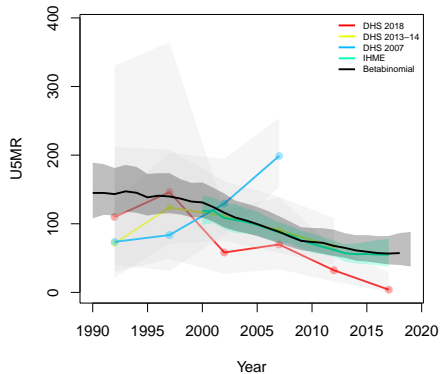
**Kazungula, Southern**



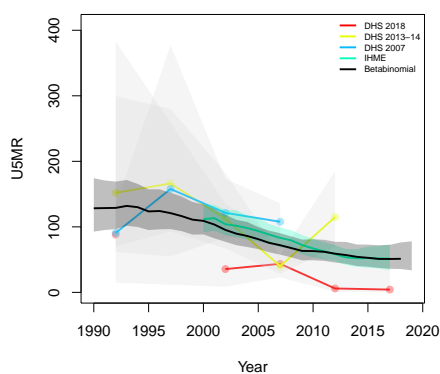
**Livingstone, Southern**



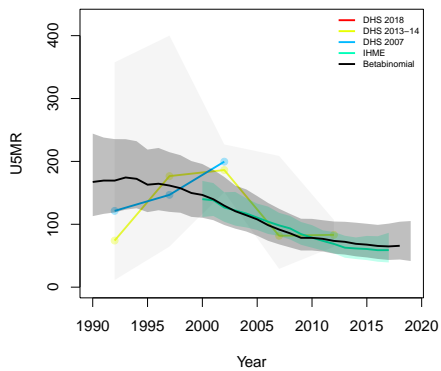
**Mazabuka, Southern**



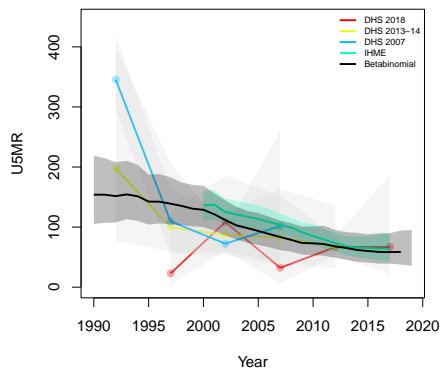
**Monze, Southern**



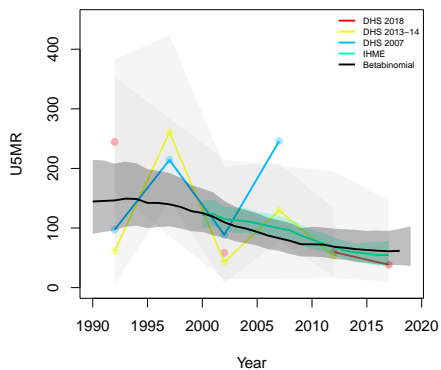
**Namwala, Southern**



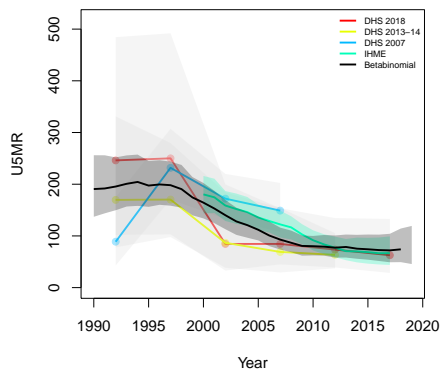
**Siavonga, Southern**



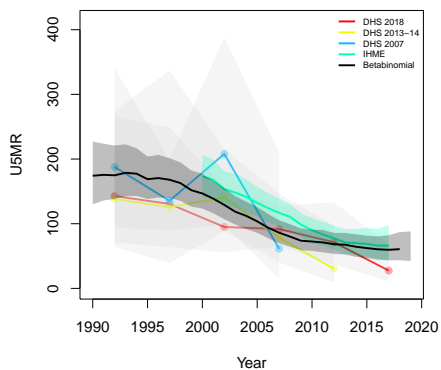
**Sinazongwe, Southern**



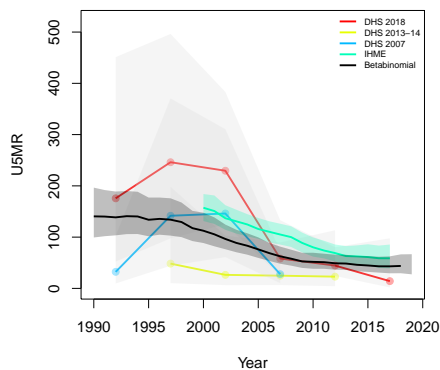
**Kalabo, Western**



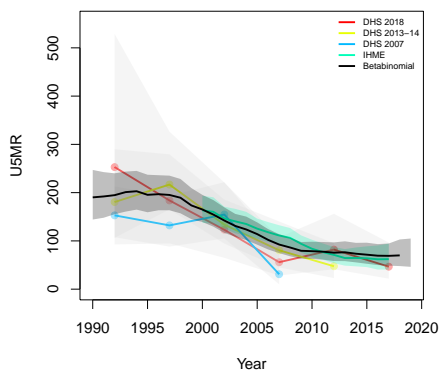
**Kaoma, Western**



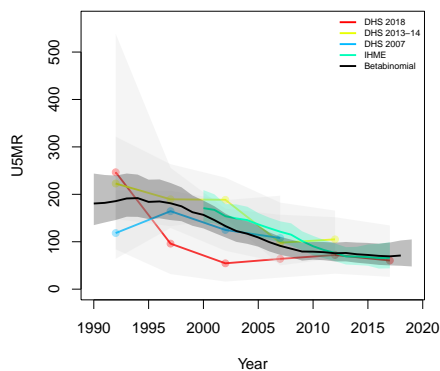
**Lukulu, Western**

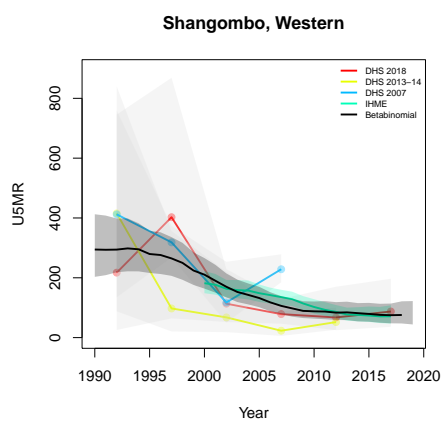
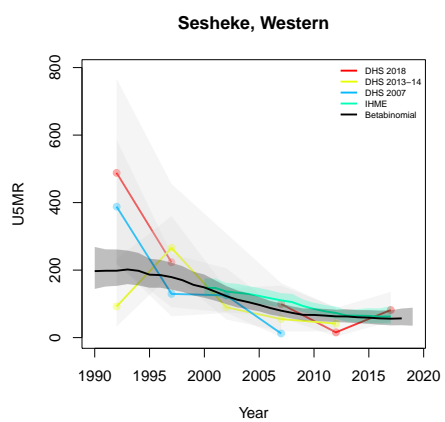


**Mongu, Western**



**Senanga, Western**





## B.21 Zimbabwe

Age	Survey	Clusters			Deaths			Agemonths		
		Urban	Rural	Total	Urban	Rural	Total	Urban	Rural	Total
0	1999	76	145	221	37	138	175	1438	5188	6626
	2010	165	228	393	92	282	374	4559	11547	16106
	2015	166	234	400	179	354	533	7337	12745	20082
1-11	1999	76	145	221	32	178	210	15104	54238	69342
	2010	165	228	393	100	257	357	46952	118782	165734
	2015	166	234	400	134	351	485	76039	130713	206752
12-23	1999	76	145	221	15	78	93	15783	56563	72346
	2010	165	228	393	40	100	140	47672	120157	167829
	2015	166	234	400	52	151	203	77844	132018	209862
24-35	1999	76	145	221	5	40	45	15010	54882	69892
	2010	165	228	393	25	67	92	44809	113782	158591
	2015	166	234	400	19	89	108	72801	123383	196184
36-47	1999	76	145	221	5	19	24	14319	53972	68291
	2010	165	228	393	13	23	36	42121	107838	149959
	2015	166	234	400	13	33	46	67461	115131	182592
48-59	1999	76	145	221	4	14	18	13805	53175	66980
	2010	165	228	393	8	8	16	39949	102413	142362
	2015	166	234	400	12	30	42	61932	107189	169121

Table B.21: **Data summary for Zimbabwe.** Total numbers of clusters (Columns 3–5) with observations in each age group by survey in urban and rural areas and combined. Numbers of deaths (Columns 6–8) and number of agemonths (Columns 9–10) observed in each age group by survey in urban and rural areas and combined.

*B.21.1 Admin-1*

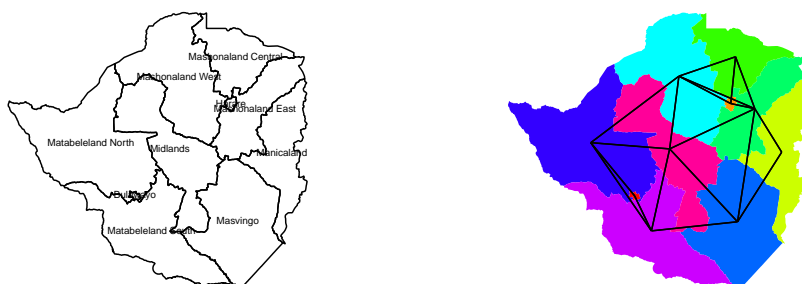


Figure B.106: **Left:** The names of the 10 Admin-1 areas of Zimbabwe . **Right:** The neighborhood structure of Admin-1 areas in Zimbabwe .

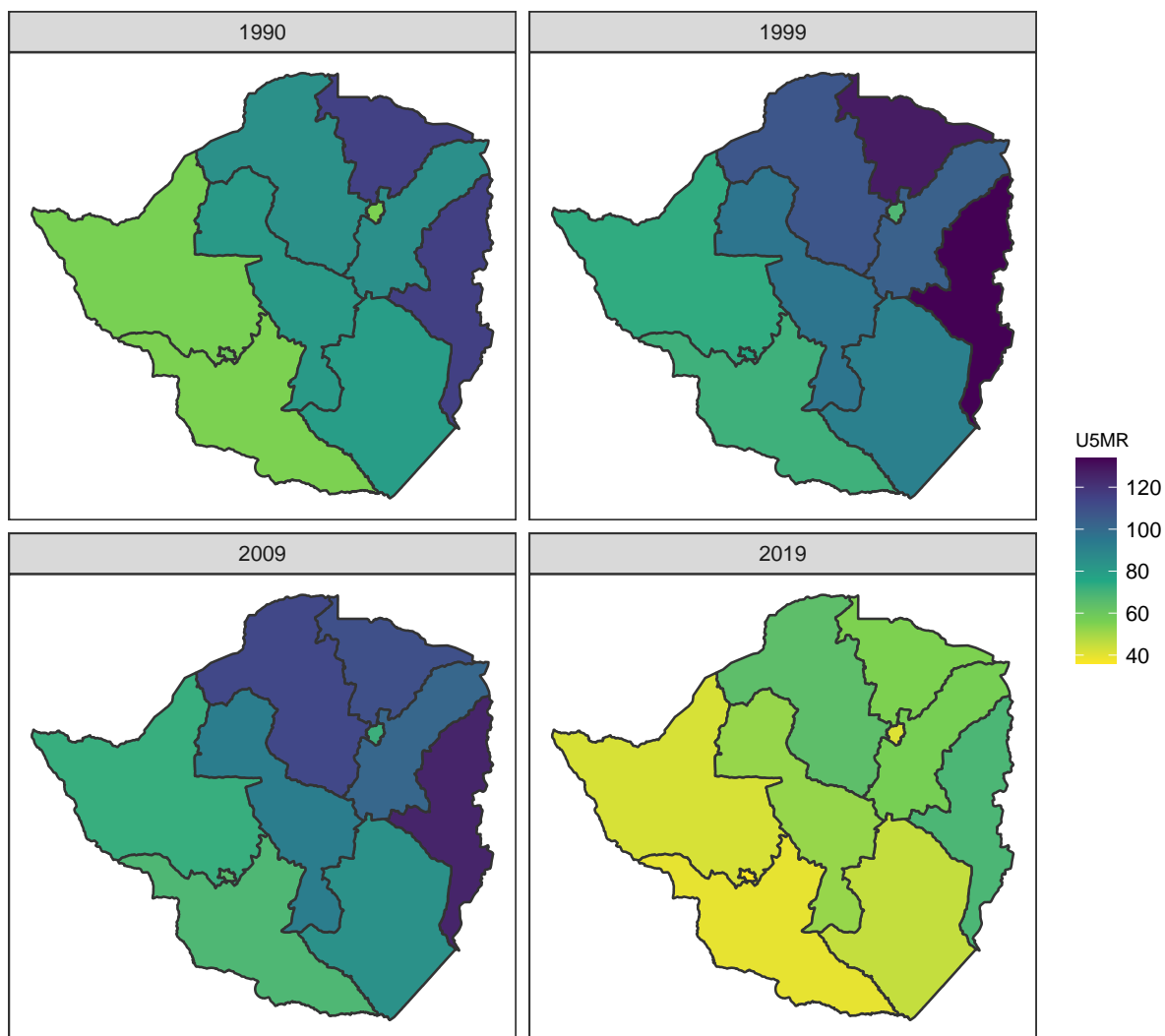


Figure B.107: Median U5MR estimates for years 1990, 1999, 2009, 2019 for Admin-1 areas in Zimbabwe .

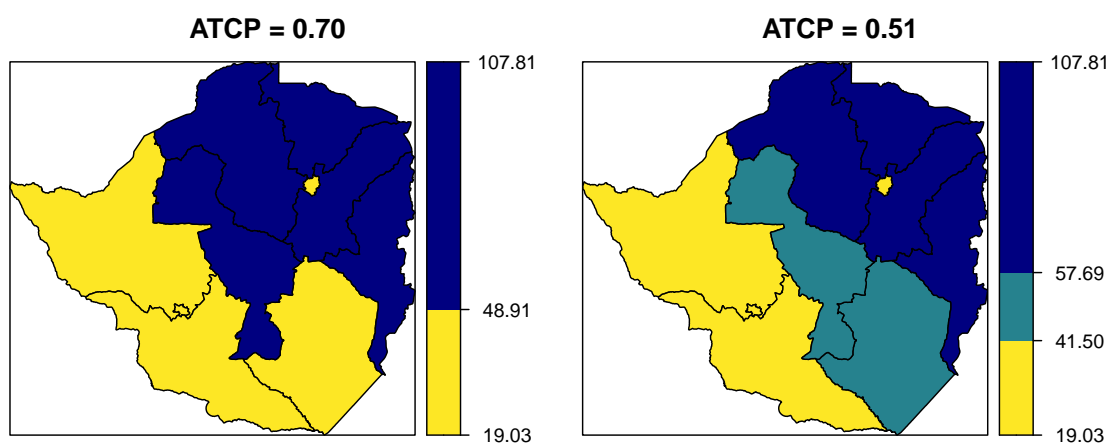
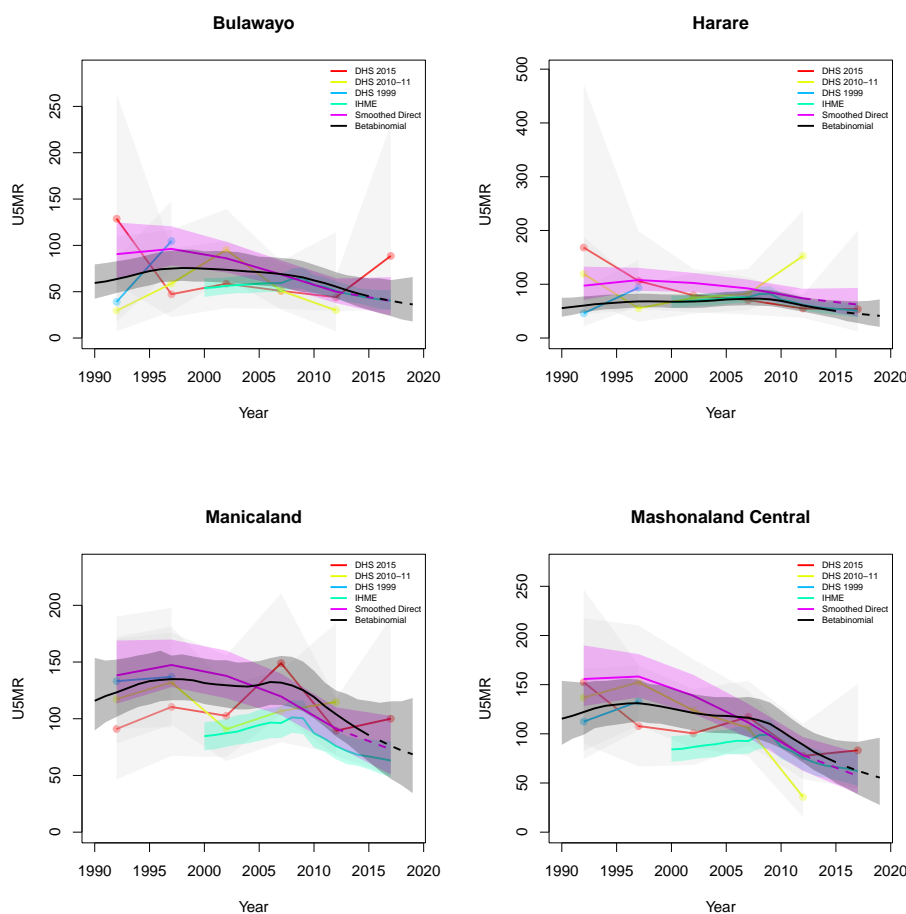


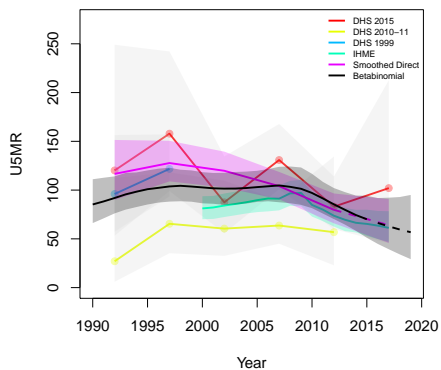
Figure B.108: Expression of uncertainty of U5MR (deaths per 1000 children) estimates for Admin-1 areas based on the average true classification probability (ATCP) in 2019 using  $K = 2, 3$  colors.

*Data and estimates over time by area*

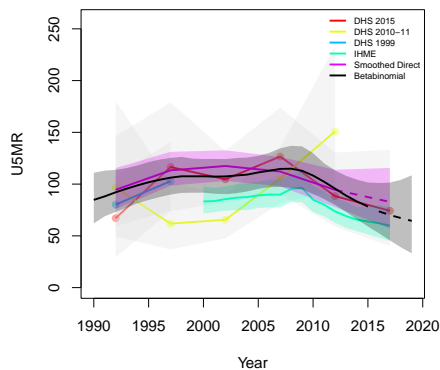
Colored lines with circular points and light grey uncertainty bands are 5-year survey-weighted estimates of U5MR for years 1990–1994 up to 2015–2019 depending on survey timing. For a survey that ends in the middle of a 5-year period, we plot the estimates at the mid-point of the years in that interval for which the survey provides data. Black lines and corresponding intervals represent posterior medians and 95% uncertainty intervals respectively for the betabinomial model. IHME’s estimates and corresponding intervals, where we can compare, are in aquamarine.



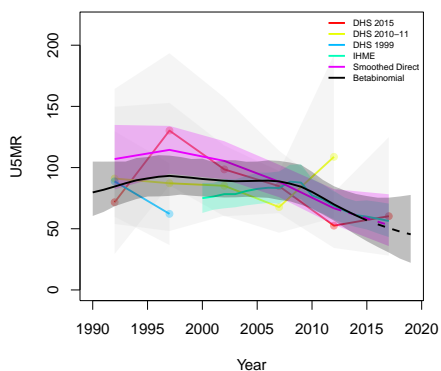
**Mashonaland East**



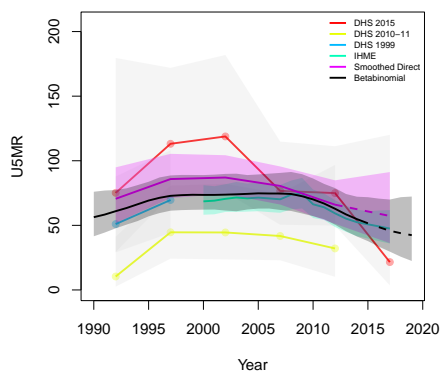
**Mashonaland West**



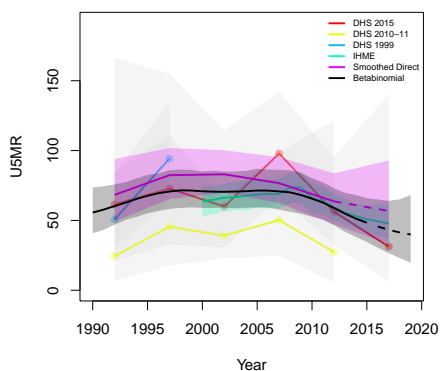
**Masvingo**



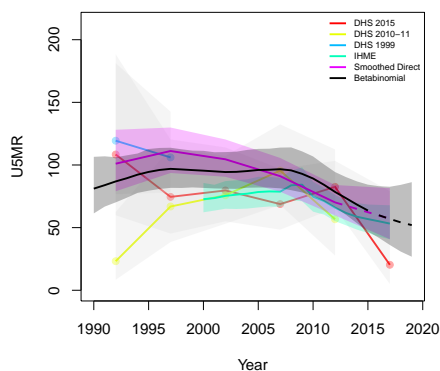
**Matabeleland North**



**Matabeleland South**



**Midlands**



*B.21.2 Admin-2*



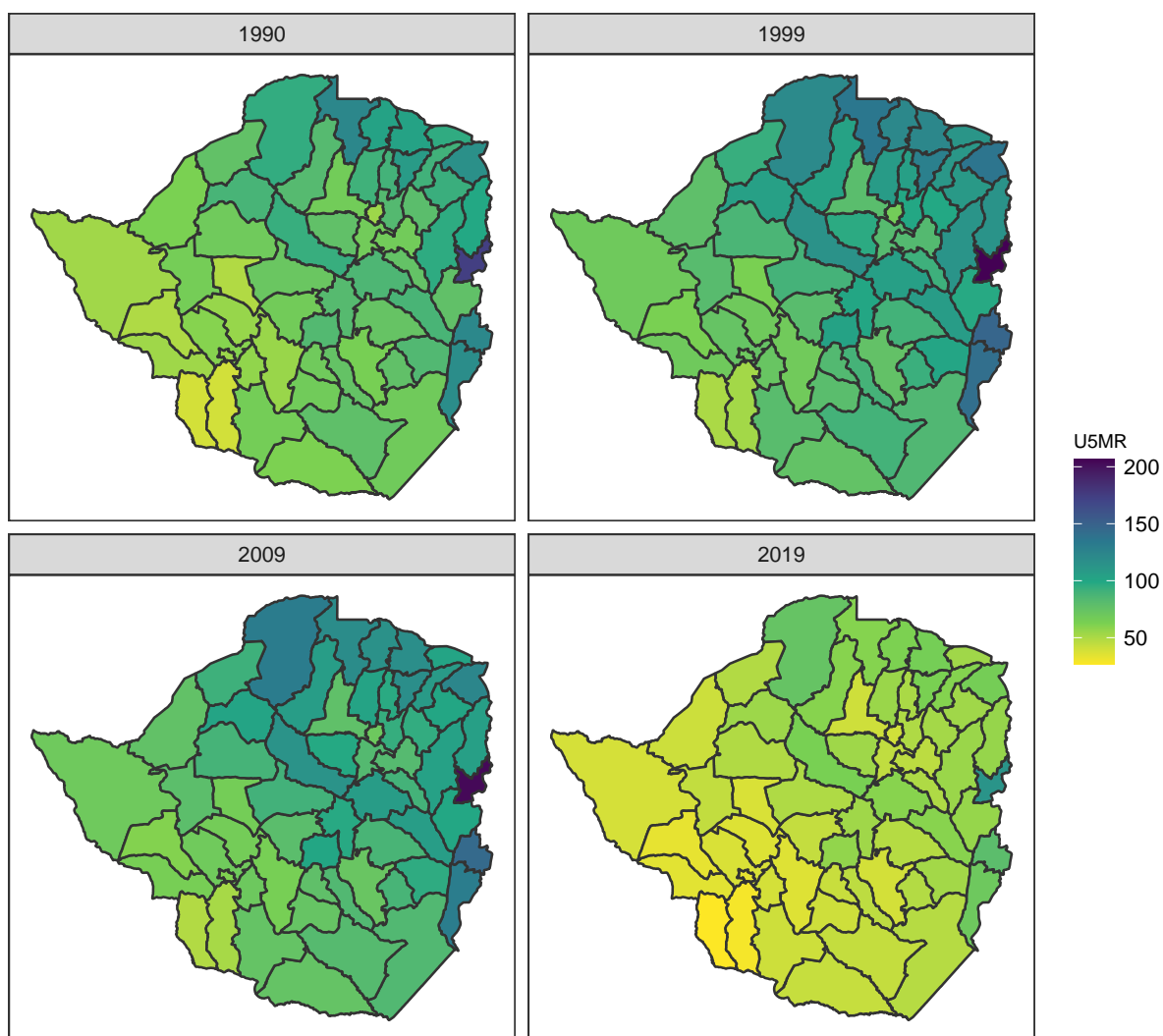


Figure B.110: Median U5MR estimates for years 1990, 1999, 2009, 2019 for Admin-2 areas in Zimbabwe .

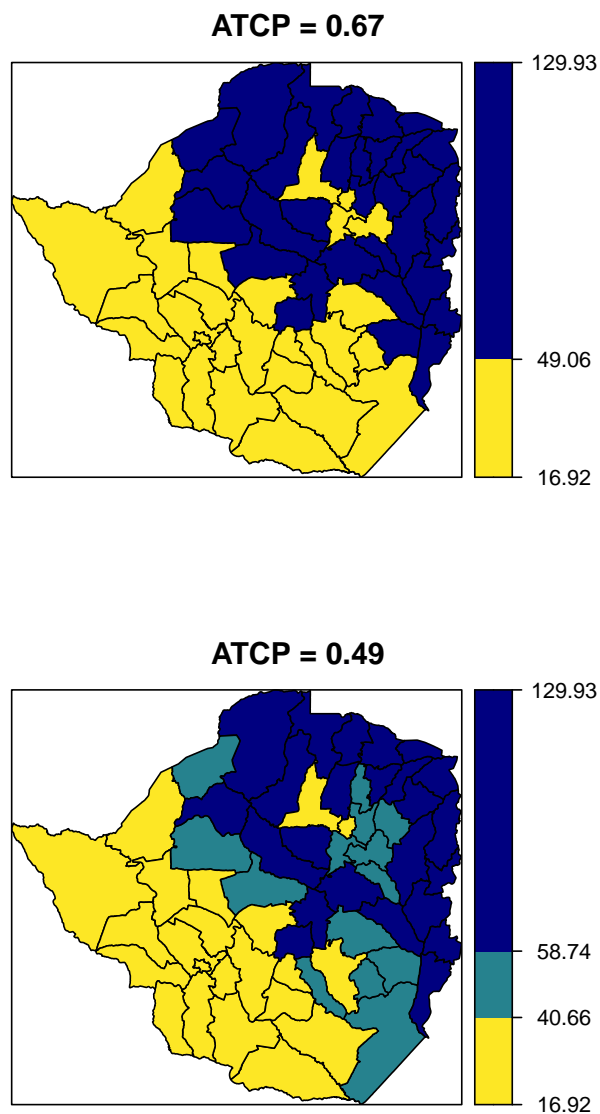
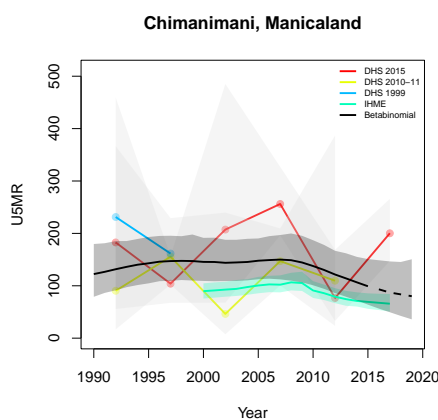
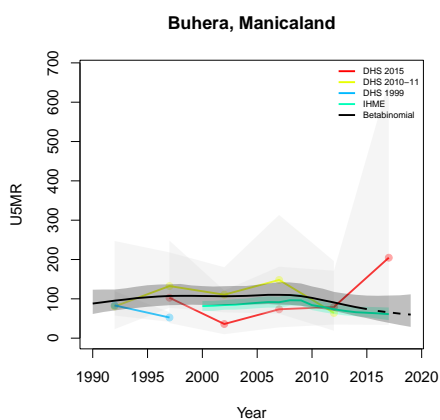
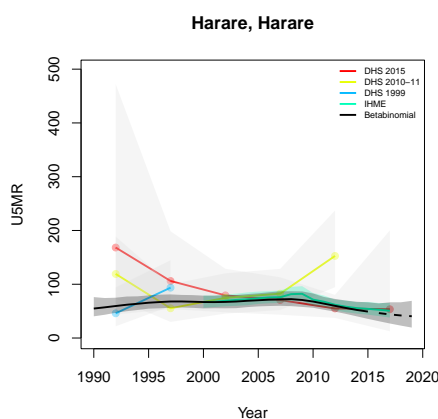
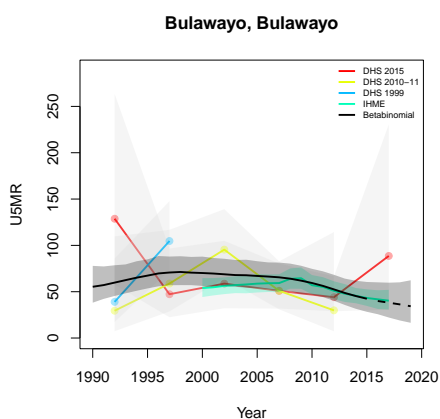


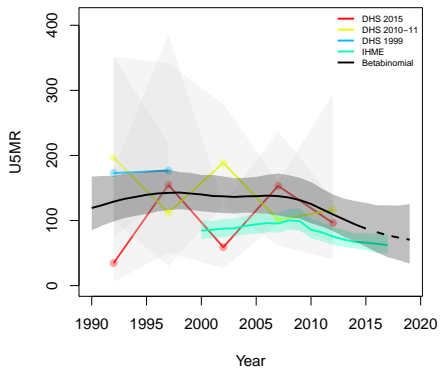
Figure B.111: Expression of uncertainty of U5MR (deaths per 1000 children) estimates for Admin-1 areas based on the average true classification probability (ATCP) in 2019 using  $K = 2, 3$  colors.

*Data and estimates over time by area*

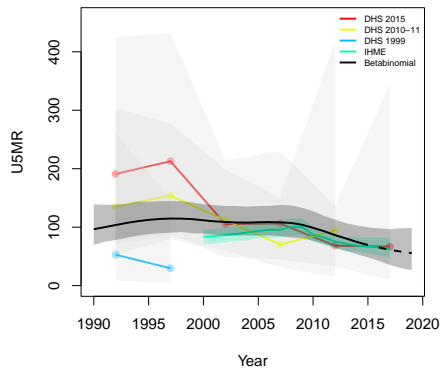
Colored lines with circular points and light grey uncertainty bands are 5-year survey-weighted estimates of U5MR for years 1990–1994 up to 2015–2019 depending on survey timing. For a survey that ends in the middle of a 5-year period, we plot the estimates at the mid-point of the years in that interval for which the survey provides data. Black lines and corresponding intervals represent posterior medians and 95% uncertainty intervals respectively for the betabinomial model. IHME’s estimates and corresponding intervals, where we can compare, are in aquamarine.



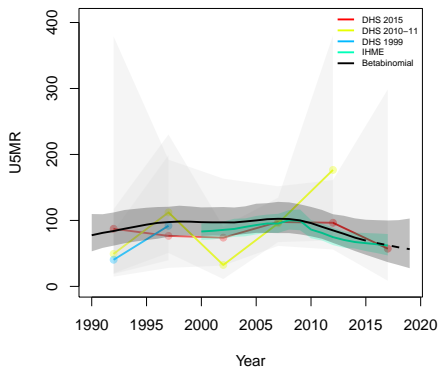
**Chipinge, Manicaland**



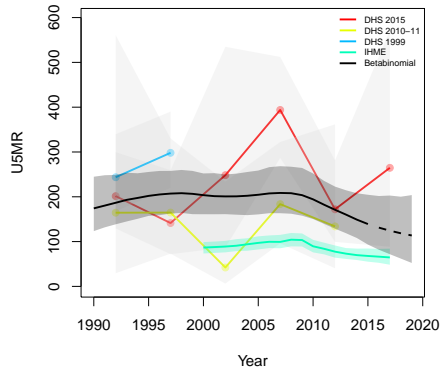
**Makoni, Manicaland**



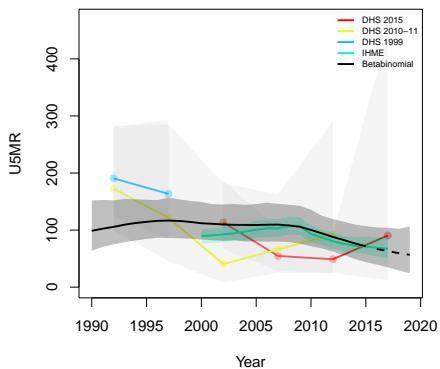
**Mutare, Manicaland**



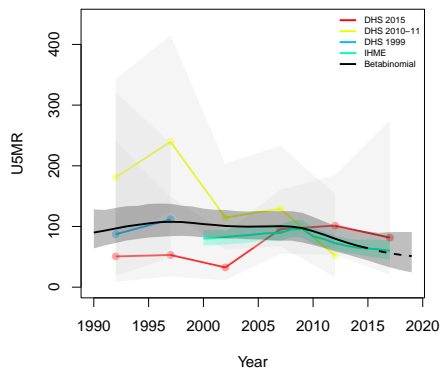
**Mutasa, Manicaland**

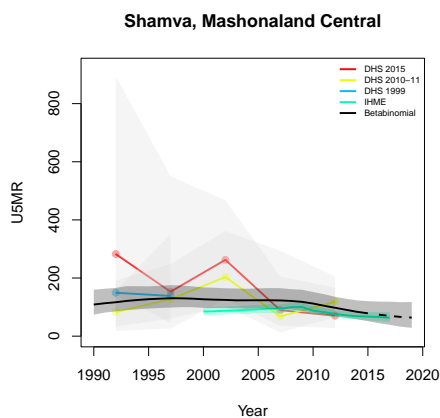
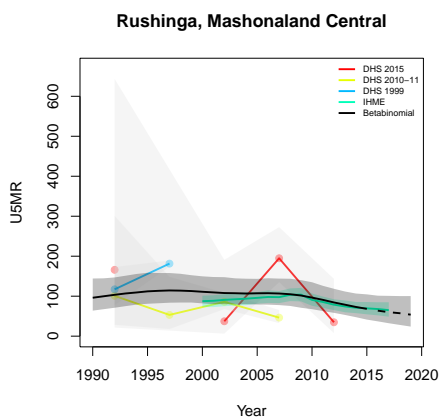
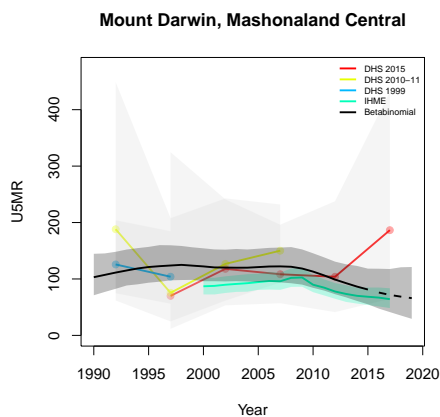
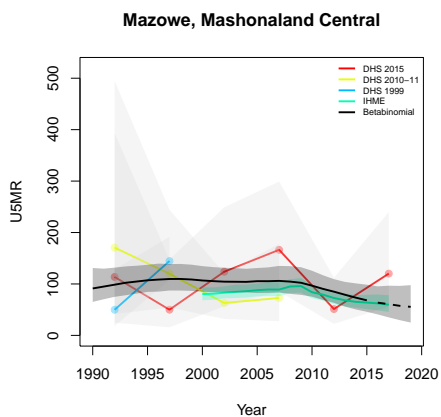
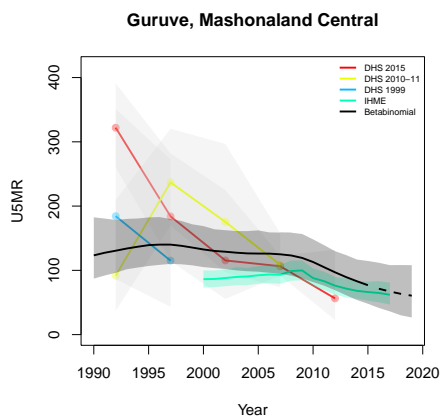
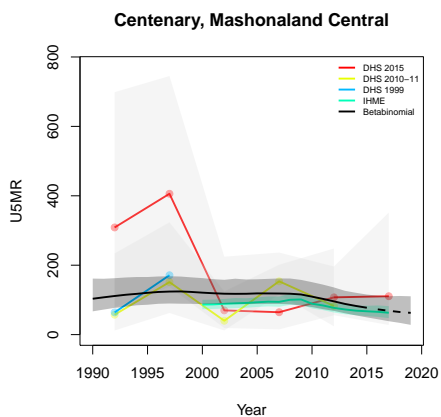


**Nyanga, Manicaland**

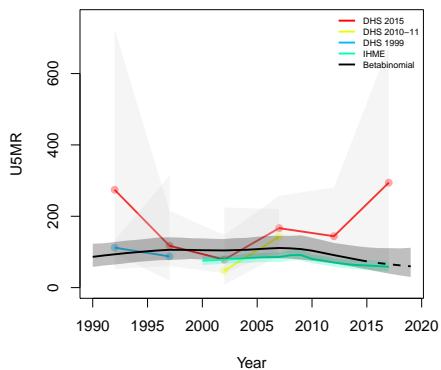


**Bindura, Mashonaland Central**

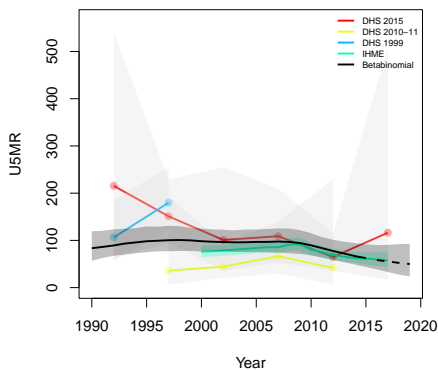




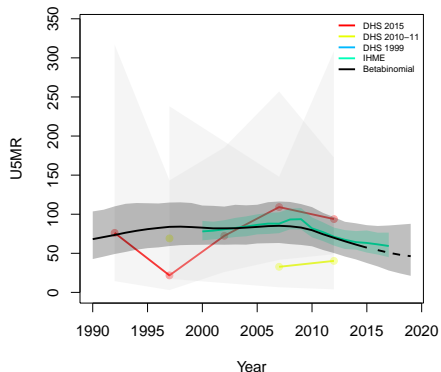
**Chikomba, Mashonaland East**



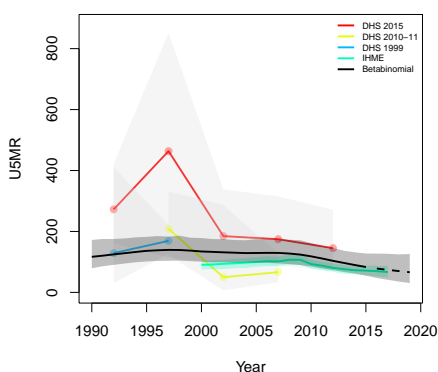
**Goromonzi, Mashonaland East**



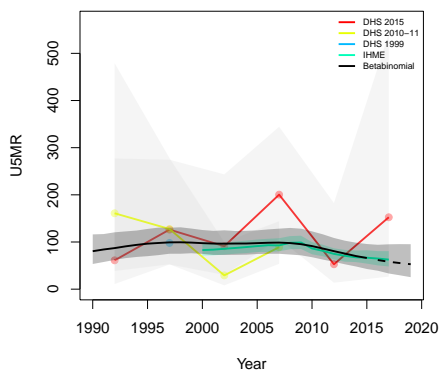
**Marondera, Mashonaland East**



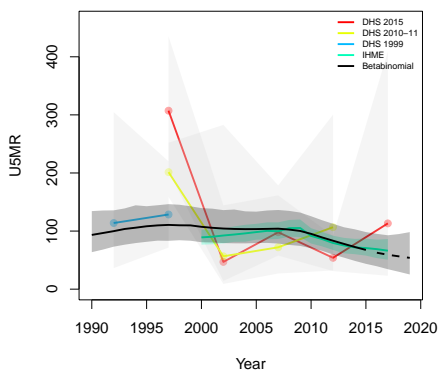
**Mudzi, Mashonaland East**



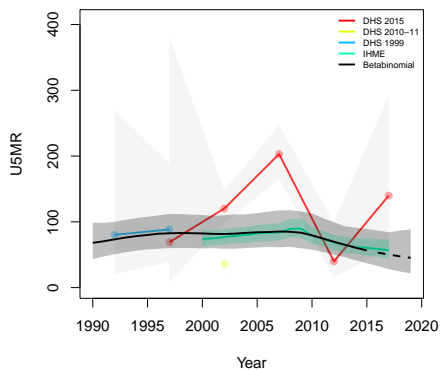
**Murehwa, Mashonaland East**



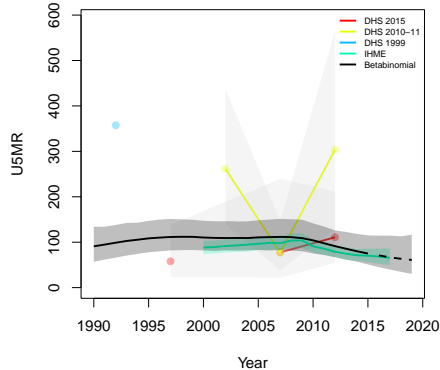
**Mutoko, Mashonaland East**



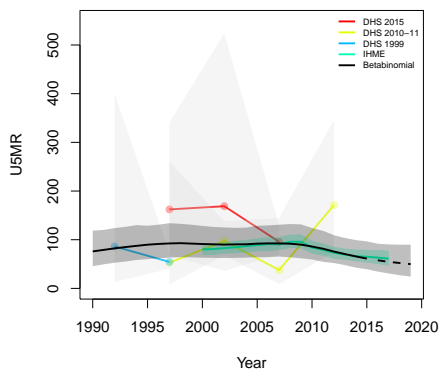
**Seke, Mashonaland East**



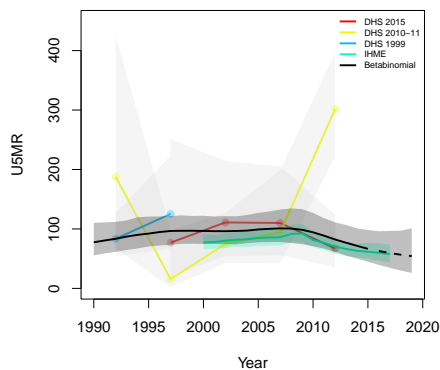
**UMP, Mashonaland East**



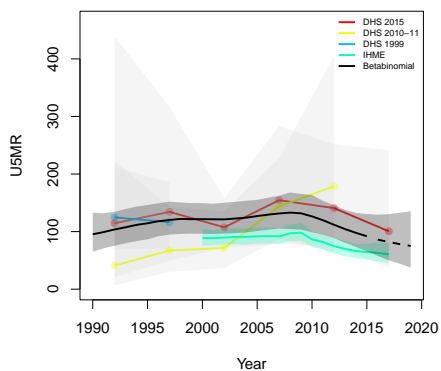
**Wedza, Mashonaland East**



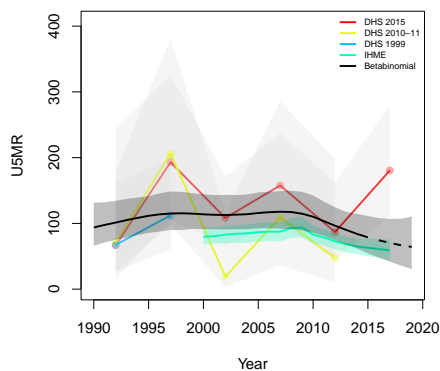
**Chegut, Mashonaland West**



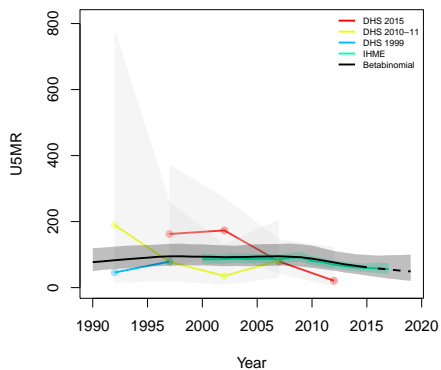
**Hurungwe, Mashonaland West**



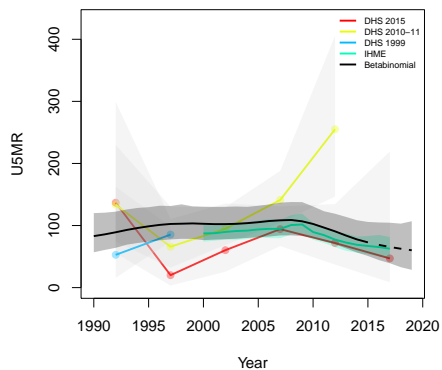
**Kadoma, Mashonaland West**



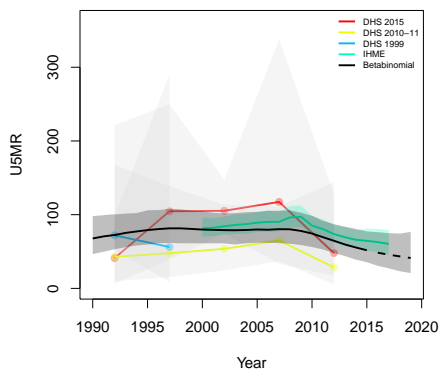
Kariba, Mashonaland West



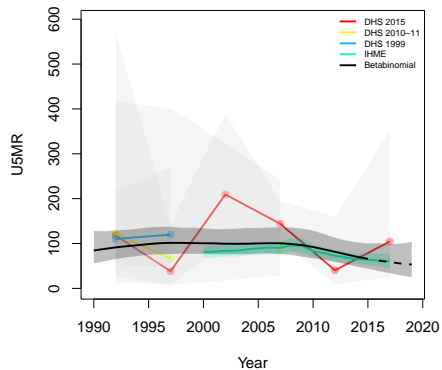
Makonde, Mashonaland West



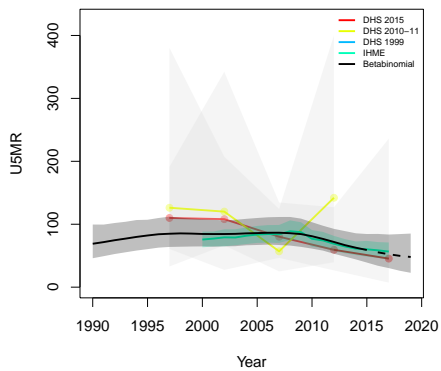
Zvimba, Mashonaland West



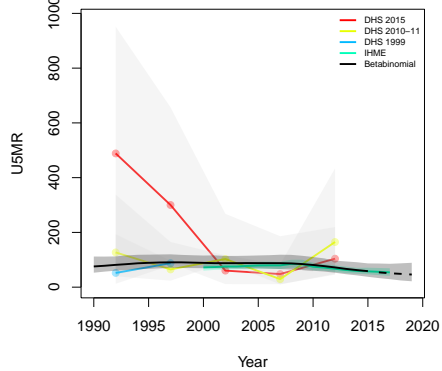
Bikita, Masvingo



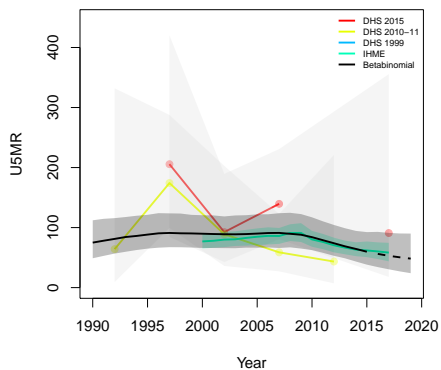
Chiredzi, Masvingo



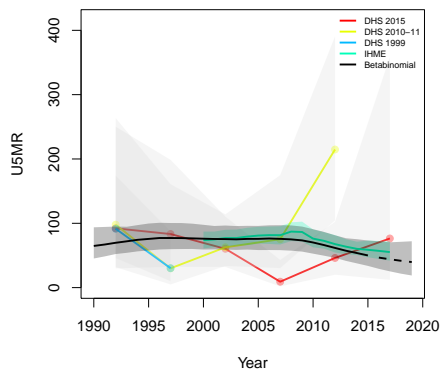
Chivi, Masvingo



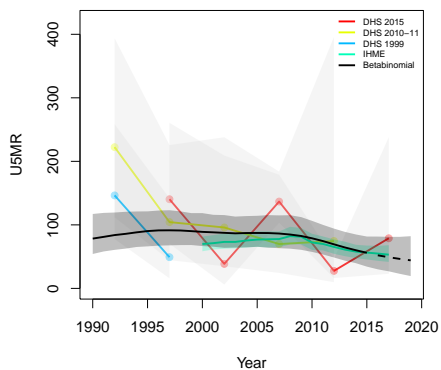
**Gutu, Masvingo**



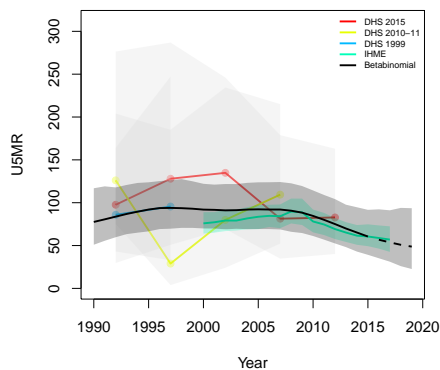
**Masvingo, Masvingo**



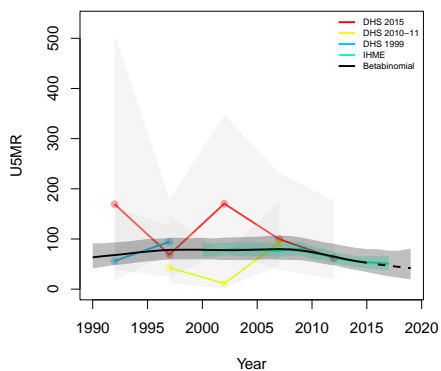
**Mwenezi, Masvingo**



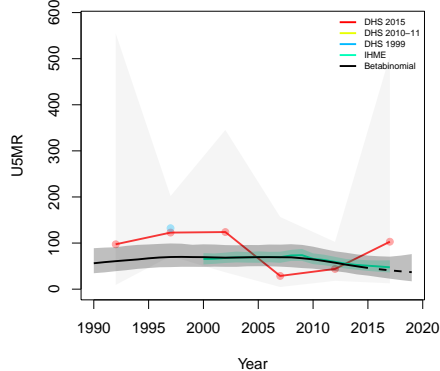
**Zaka, Masvingo**



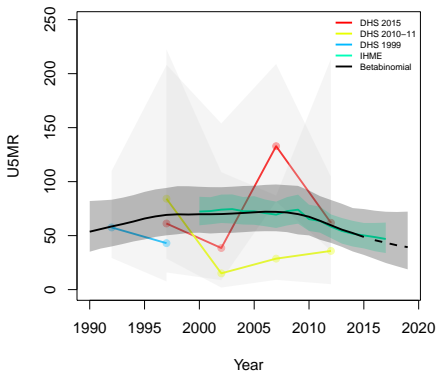
**Binga, Matabeleland North**



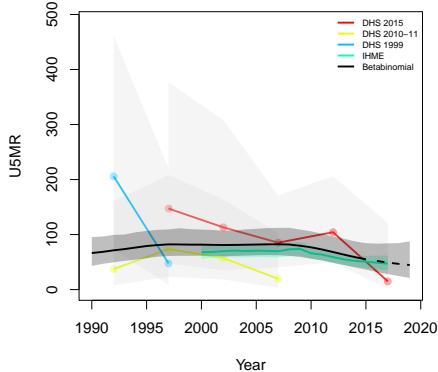
**Bubi, Matabeleland North**



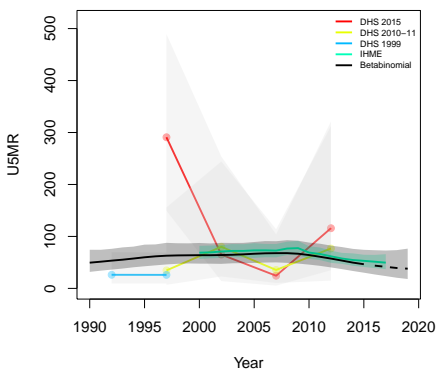
**Hwange, Matabeleland North**



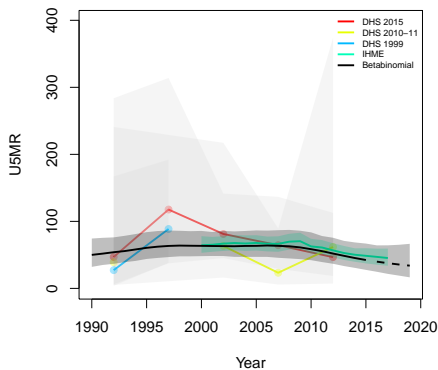
**Lupane, Matabeleland North**



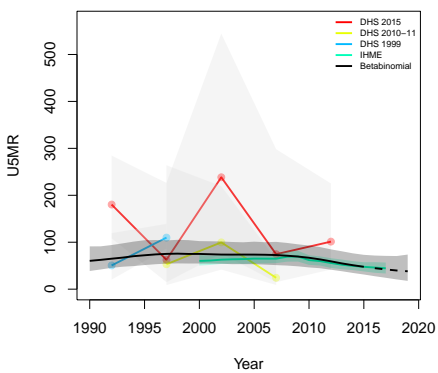
**Nkayi, Matabeleland North**



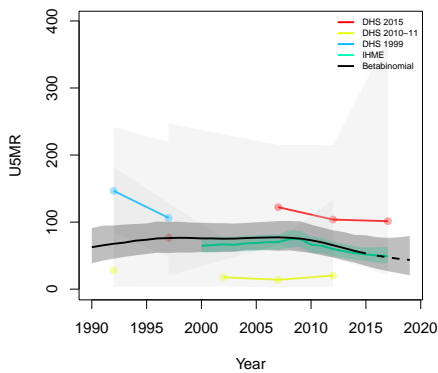
**Tsholotsho, Matabeleland North**



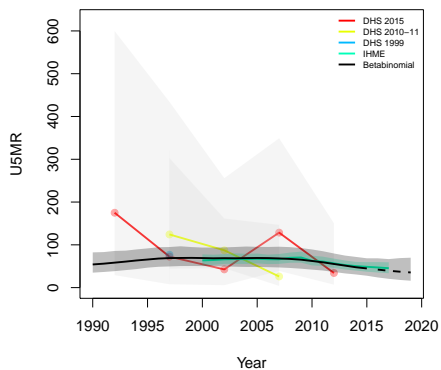
**Umguza, Matabeleland North**



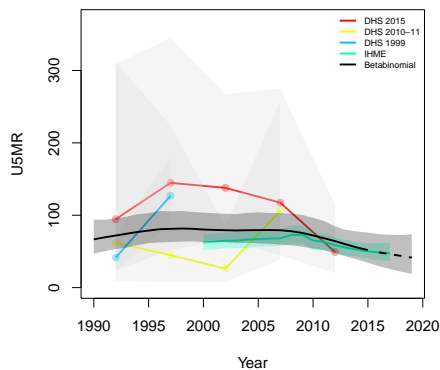
**Beitbridge, Matabeleland South**



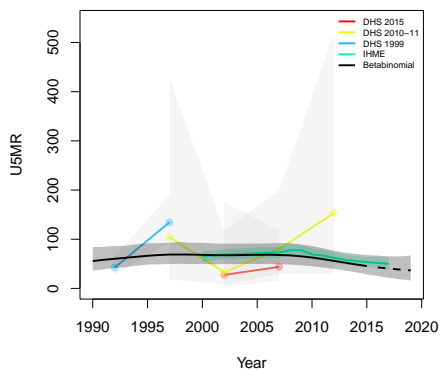
**Bulilima (North), Matabeleland South**



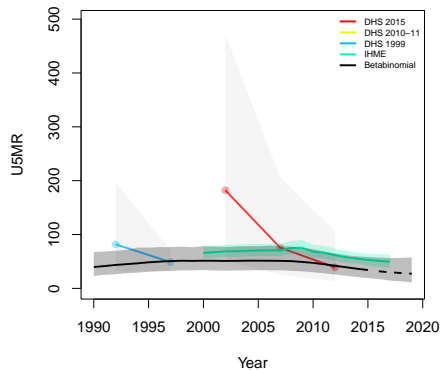
**Gwanda, Matabeleland South**



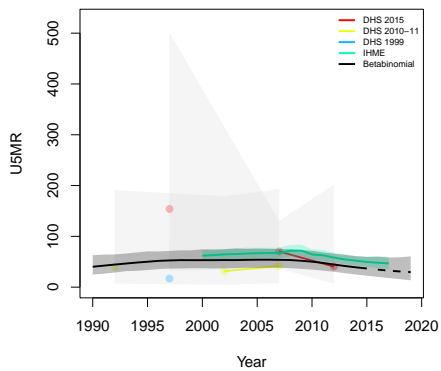
**Insiza, Matabeleland South**



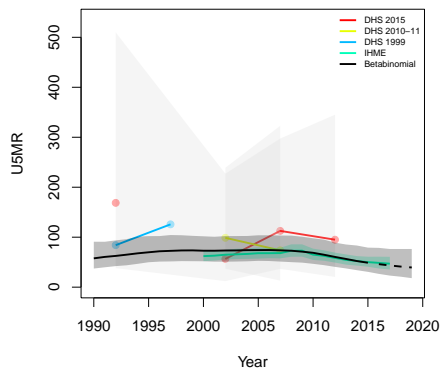
**Mangwe (South), Matabeleland South**



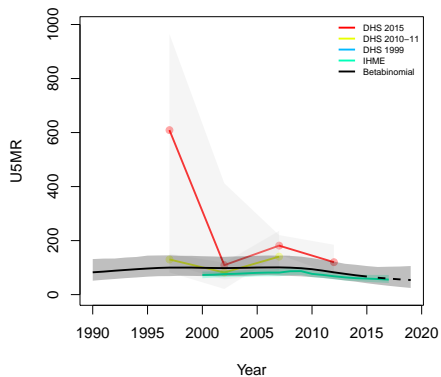
**Matobo, Matabeleland South**



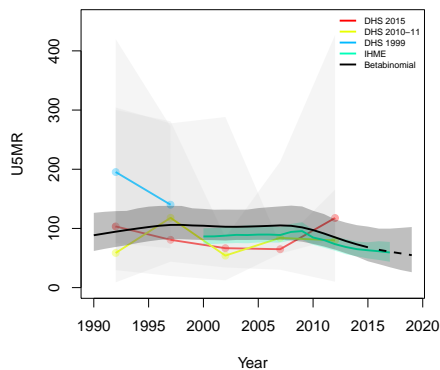
**Umzingwane, Matabeleland South**



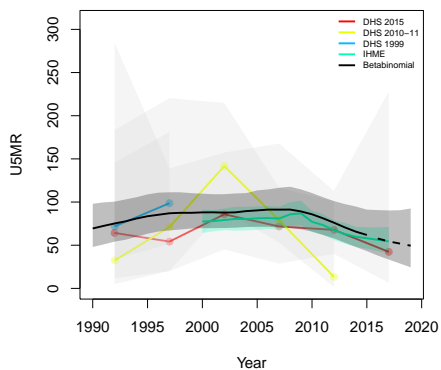
**Chirumhanzu, Midlands**



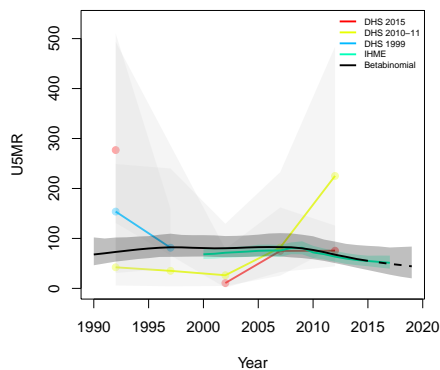
**Gokwe North, Midlands**



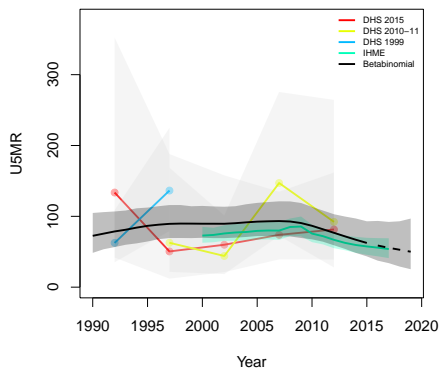
**Gokwe South, Midlands**



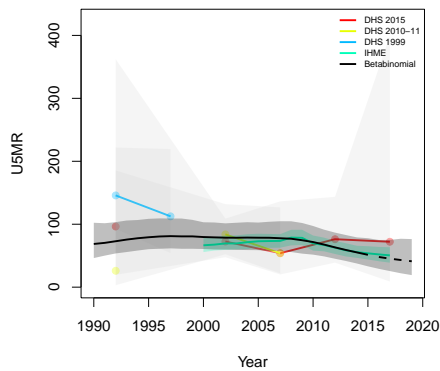
**Gweru, Midlands**



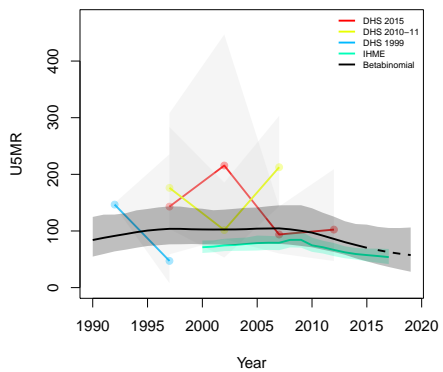
**Kwekwe, Midlands**



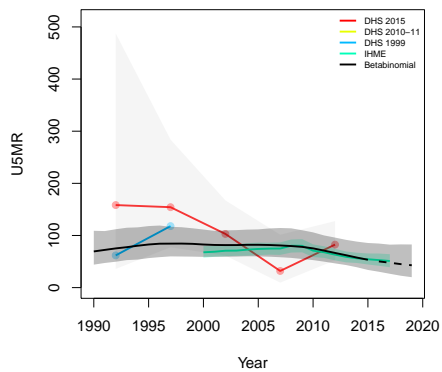
**Mberengwa, Midlands**



Shurugwi, Midlands



Zvishavane, Midlands



↳