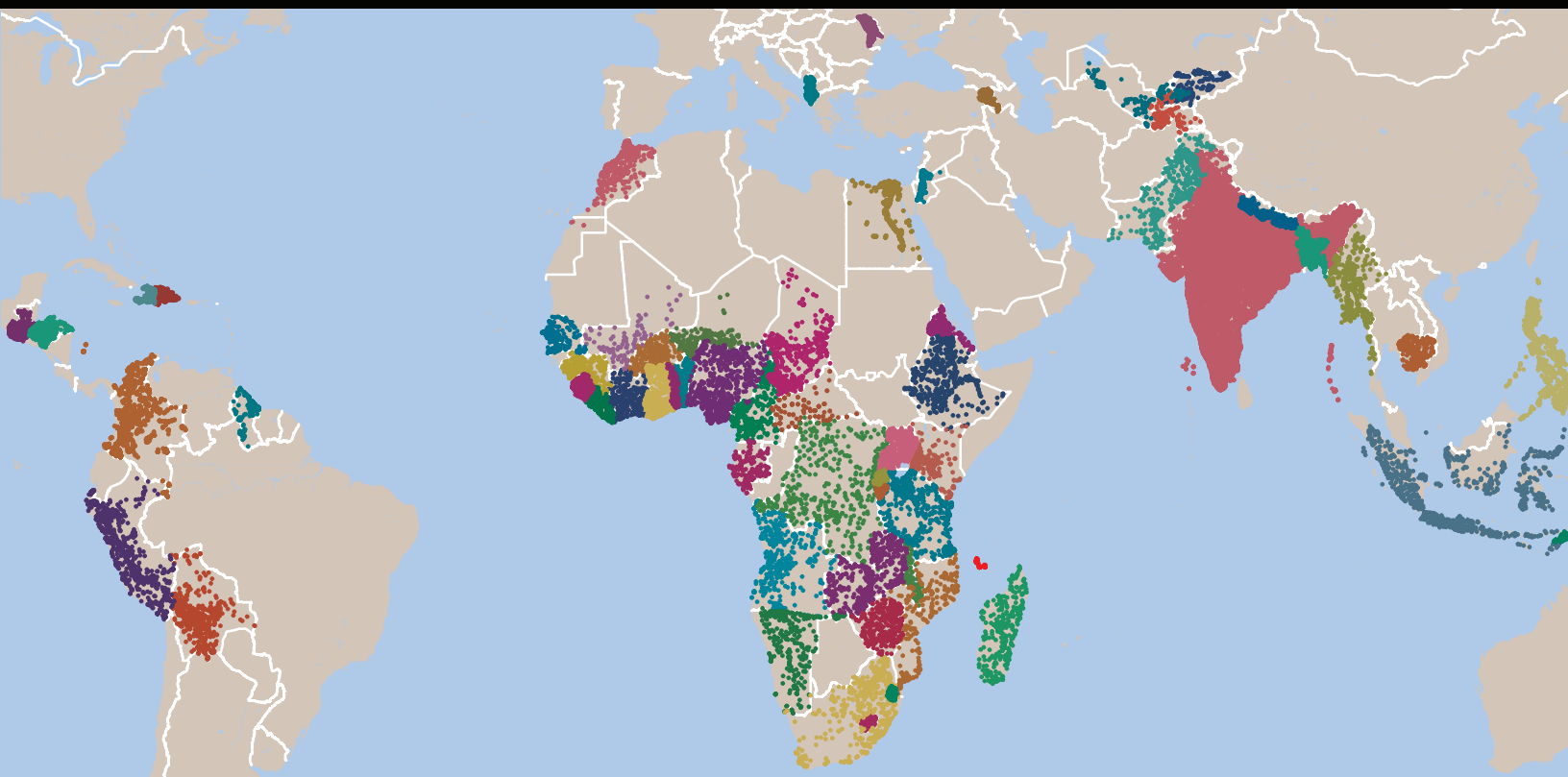




USAID
FROM THE AMERICAN PEOPLE

SPATIAL MODELING FOR SUBNATIONAL ADMINISTRATIVE LEVEL 2 SMALL-AREA ESTIMATION

DHS SPATIAL ANALYSIS REPORTS 21



SEPTEMBER 2021

This publication was produced for review by the United States Agency for International Development (USAID). The report was prepared by Yunhan Wu, Zehang Richard Li, Benjamin K. Mayala, Houjie Wang, Peter A. Gao, John Paige, Geir-Arne Fuglstad, Caitlin Moe, Jessica Godwin, Rose E. Donohue, Bradley Janocha, Trevor N. Croft, and Jon Wakefield.

DHS Spatial Analysis Reports No. 21

Spatial Modeling for Subnational Administrative Level 2 Small-Area Estimation

Yunhan Wu,¹ Zehang Richard Li,² Benjamin K. Mayala,³ Houjie Wang,¹
Peter A. Gao,¹ John Paige,⁴ Geir-Arne Fuglstad,⁴ Caitlin Moe,¹ Jessica Godwin,¹
Rose E. Donohue,³ Bradley Janocha,³ Trevor N. Croft,³ and Jon Wakefield^{1,5}

The DHS Program
ICF
Rockville, Maryland, USA

September 2021

¹ Department of Statistics, University of Washington, Seattle, WA

² Department of Statistics, University of California, Santa Cruz, CA

³ The DHS Program, ICF International, Rockville, MD

⁴ Department of Mathematical Sciences, Norwegian University of Science and Technology, Trondheim, Norway

⁵ Department of Biostatistics, University of Washington, Seattle, WA

Corresponding authors:

Jon Wakefield, Department of Statistics and Department of Biostatistics, University of Washington, Seattle, WA 98195, USA; email: jonno@uw.edu

Benjamin K. Mayala, The DHS Program, 530 Gaither Road, Suite 500, Rockville, MD 20850, USA; phone: +1 301-572-0507; fax: +1 301-407-6501; email: Ben.Mayala@icf.com

Acknowledgments: The authors wish to thank Emanuele Giorgi, Peter J. Diggle, and David Kline for their review of this report.

Editor: Diane Stoy

Document Production: Natalie Shattuck

This study was implemented with support from the United States Agency for International Development (USAID) through The DHS Program (#720-OAA-18C-00083). The views expressed are those of the authors and do not necessarily reflect the views of USAID or the United States Government.

The DHS Program assists countries worldwide in the collection and use of data to monitor and evaluate population, health, and nutrition programs. Additional information about The DHS Program can be obtained from ICF, 530 Gaither Road, Suite 500, Rockville, MD 20850 USA; telephone: +1 301-407-6500, fax: +1-301-407-6501, email: info@DHSprogram.com, internet: www.DHSprogram.com.

Recommended citation:

Wu, Yunhan, Zehang Richard Li, Benjamin K. Mayala, Houjie Wang, Peter Gao, Johnny Paige, Geir-Arne Fuglstad, Caitlin Moe, Jessica Godwin, Rose E. Donohue, Bradley Janocha, Trevor N. Croft, and Jon Wakefield, 2021. *Spatial Modeling for Subnational Administrative Level 2 Small-Area Estimation*. DHS Spatial Analysis Reports No. 21. Rockville, Maryland, USA: ICF.

CONTENTS

TABLES	v
FIGURES	vii
PREFACE	ix
ABSTRACT	xi
ACRONYMS AND ABBREVIATIONS	xiii
1 BACKGROUND AND OBJECTIVES	1
1.1 Background	1
1.2 Objectives	2
1.3 Report Outline	3
2 AREA-LEVEL MODELS	5
2.1 Spatial Modeling of Prevalence	5
2.2 Space-Time Modeling of Prevalence	6
2.3 Space-Time Modeling of U5MR	7
3 CLUSTER-LEVEL MODELS	9
3.1 Spatial Modeling of Prevalence	9
3.2 Space-Time Modeling of Prevalence	10
3.3 Space-Time Modeling of U5MR	11
3.4 Creating an Urban/Rural Stratification Surface	14
4 ADDITIONAL CONSIDERATIONS	15
4.1 Covariate Modeling	15
4.2 Continuous Spatial Models	16
4.3 Aggregation	16
4.4 Model Assessment	17
5 THE SUMMER SOFTWARE	19
6 U5MR MODELING IN EIGHT COUNTRIES	21
6.1 Eight Country Summary	21
6.2 Validation	27
6.3 Detailed Results for Zambia	28
6.4 Covariate Modeling	37
7 DISCUSSION	41
REFERENCES	42
SUPPLEMENTARY APPENDICES [link]	

TABLES

Table 1	Numbers of Admin 1 and Admin 2 areas and survey years for the eight countries whose data we analyze.....	21
Table 2	Spread of Admin 1 estimates (rate per 1,000 live births) for the year of the survey	22
Table 3	Spread of Admin 2 estimates (rate per 1,000 live births) for the year of the survey	22
Table 4	Validation results for both the smoothed direct model, and the beta-binomial Admin 1 and Admin 2 models, first averaged over Admin 1 areas within each country, and then averaged over all countries. Bias is reported in children per thousand births.	28
Table 5	Fixed effects posterior summaries for cluster-level model with Admin 2 level covariates.....	37

FIGURES

Figure 1	Hazard rates (urban/rural) over time for 6 age bands. There are 3 temporal trends for 0–1 month, 1–12 months, and for all age groups >12 months (i.e., the last 4 age groups). Similarly, there are 3 urban/rural hazard rate adjustments.....	13
Figure 2	U5MR estimates and uncertainty at year of survey for Bangladesh.....	22
Figure 3	U5MR estimates and uncertainty at year of survey for Cameroon.....	23
Figure 4	U5MR estimates and uncertainty at year of survey for Ethiopia.....	23
Figure 5	U5MR estimates and uncertainty at year of survey for Kenya	24
Figure 6	U5MR estimates and uncertainty at year of survey for Malawi	25
Figure 7	U5MR estimates and uncertainty at year of survey for Nepal	25
Figure 8	U5MR estimates and uncertainty at year of survey for Nigeria	26
Figure 9	U5MR estimates and uncertainty at year of survey for Zambia.....	26
Figure 10	Comparison of direct and smoothed direct estimates, Zambia Admin 1, 2016-2018	29
Figure 11	Aggregated beta-binomial national estimates versus direct national estimates, over time, and with 95% error bands, Zambia	30
Figure 12	Probability for Admin 2 estimates exceeding national direct estimates (Zambia 2018).....	31
Figure 13	Odds ratios (urban/rural) over time for the age bands 0–1 month, 1–12 months, >12 months	32
Figure 14	Trends of U5MR for Admin 1 areas in Zambia over time	33
Figure 15	Trends of U5MR for Admin 2 areas in Zambia over time	33
Figure 16	Ridgeplot representation of posterior distribution of U5MR for Admin 2 areas in Luapula (Zambia 2018)	34
Figure 17	Admin 2 U5MR ranking in Zambia, 2018. On the left we show the regions with the highest U5MR, and on the right those with the lowest U5MR.	35
Figure 18	Average True Classification Probabilities (ATCP) for (a) $K = 2$, (b) $K = 3$, and (c) $K = 4$ for Zambia in 2018	36
Figure 19	Comparison of yearly Admin 2 U5MR estimates for the beta-binomial cluster level model with Admin 2 level covariates and base (no covariate) model	38
Figure 20	Comparison of 2018 admin 2 U5MR (a) point and (b) uncertainty estimates for Zambia using cluster-level model with Admin 2 level covariates and base model..	39

PREFACE

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

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

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

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

Sunita Kishor
Director, The DHS Program

ABSTRACT

Subnational estimates of the health and demographic indicators recorded in the Demographic and Health Surveys Program are of great importance for prioritizing resources and assessing if target levels for indicators are being attained. In this report, we examine subnational variation in the under-5 mortality rate by using small area estimation models with the goal of estimating at the Admin 2 level. We describe spatio-temporal modeling and consider discrete spatial models in detail, the possibility of including covariates, and accounting for urban/rural stratification, model assessment, and visualization of results. We offer recommendations for subnational modeling, describe an analysis pipeline, and include R code to perform the various steps, using the SUMMER package. We illustrate methods and provide results for Bangladesh, Cameroon, Ethiopia, Kenya, Nepal, Nigeria, Malawi, and Zambia.

A supplementary appendix is provided at <https://dhsprogram.com/publications/publication-SASAR21-Spatial-Analysis-Reports.cfm>.

Keywords: Bayesian inference, complex survey designs, integrated nested Laplace approximation, prevalence mapping, small area estimation, spatial smoothing.

ACRONYMS AND ABBREVIATIONS

Admin 1	first subnational administrative level
Admin 2	second subnational administrative level
BYM	Besag, York, Mollié model
CAR	conditional autoregressive
DHS	Demographic and Health Survey
GP	Gaussian process
HMIS	health management information system
ICAR	intrinsic conditional autoregressive
iid	independent and identically distributed
INLA	integrated nested Laplace approximation
LMICs	low and middle-income countries
MBG	model-based geostatistics
MIS	Malaria Indicator Survey
MRF	Markov random field
RMSE	root mean squared error
RW	random walk
SAE	small area estimation
SDG	Sustainable Development Goal
SPDE	stochastic partial differential equations
SUMMER	Small-Area-Estimation Unit/Area Models and Methods for Estimation in R
TCP	True Classification Probability
UN IGME	United Nations Inter-agency Group for Mortality Estimation
U5MR	under-5 mortality rate

1 BACKGROUND AND OBJECTIVES

1.1 Background

Most household surveys such as the Demographic Health Surveys (DHS) and Malaria Indicator Surveys (MIS) provide reliable estimates of survey indicators primarily at the national level, as well as the first subnational administrative level – Admin 1 (states, provinces, or regions). Since national-level estimates are useful for comparing nations and aggregating data across large world regions, their natural audience includes international policy makers and donors (Li et al. 2019). The analysis at Admin 1 is useful in understanding the distribution of health and demographic phenomena, but it does not provide comprehensive estimates at lower levels such as the second subnational administrative level (Admin 2), where health programs are designed and implemented (Li et al. 2019, Mayala et al. 2019). Admin 2 areas are often referred to as districts, counties, cercles, and communes.

Local officials and in-country partners have long expressed a desire for more localized DHS estimates. Countries now have an even greater need for these data because health program planning and implementation are increasingly decentralized to the Admin 2 level. Decision-makers at this level are often constrained by a lack of routinely available local data for key indicators that would allow for data-driven policymaking (Wickremasinghe et al. 2016). A need exists for local data that is routinely produced, encompasses a variety of demographic and health subject areas, and is easily accessible and interpretable. As local needs demand, these data can be used for priority setting at the Admin 2 level, identification of poorly performing localities, and more equitable resource allocation.

In addition to local needs, international development goals also help drive the demand for Admin 2 estimates (Utazi et al. 2021). During the last several years and within the framework of the Sustainable Development Goals (SDGs), there has been an expressed need to improve the measurement and understanding of local geographic patterns to support more decentralized decision-making and more efficient program implementation (United Nations General Assembly 2015). In an effort to improve health outcomes for all, the SDGs prioritize reducing within-country inequalities because within-country heterogeneity was often overlooked when progress was monitored with national averages (Hosseinpour, Bergen, and Magar 2015).

To better address the need for fine spatial and lower-level estimates, there are three existing options: (i) scaling up the nationally representative survey data collection process by increasing the sample size, (ii) using data derived from routine health management information systems (HMIS) from health facilities, or (iii) using a spatial modeling approach.

Increasing sample size is costly and time-consuming. The HMIS data quality is not always reliable, and the data are not easily accessed. Instead, spatial modeling techniques that can leverage existing survey data and spatial similarity between survey clusters have become increasingly popular in mapping key development indicators (Mayala et al. 2019; Utazi et al. 2018).

Statistically, the estimation of area-level characteristics at the Admin 2 level falls under the category of small area estimation (SAE). There are two philosophically distinct ways of approaching SAE: design-

based and model-based inference (Skinner and Wakefield, 2017). Design-based methods assess the frequentist properties of estimators, and average over all possible samples that could have been drawn, under the specified sampling design. In this paradigm, the values of the responses in the population are viewed as fixed rather than random. So-called direct estimates, which in the simplest case are just weighted estimates, can be used. These adjust for the design and provide an appropriate variance, but are unstable or cannot be calculated for most Admin 2 areas because of data sparsity. The DHS surveys are typically powered to the Admin 1 level. In this report, we use national and Admin 1 weighted estimates as comparison measures for the model-based approaches that are our focus. The classic text on SAE is Rao and Molina (2015).

Model-based approaches can be either frequentist or Bayesian. Frequentist refers to the conventional version in which estimators are judged by repeated sampling of the outcome, in contrast to frequentist design-based inference in which the data are viewed as fixed, and the indices of the sampled units are viewed as random. If a hypothetical infinite population model-based approach is used, a probabilistic model is specified for the responses, which are now viewed as random variables. In the context of the DHS Program, one must consider the stratification and the cluster sampling to obtain valid inference. Random effects modeling and the inclusion of covariates are popular within the model-based approach to SAE, because they allow smoothing across space and covariates to provide more reliable estimates.

Prevalence mapping (Wakefield 2020) is the production of maps that display the prevalence of health and demographic outcomes. This approach clearly has large overlap with SAE. The SAE smoothing methods often use area-level models, while prevalence mapping uses model-based geostatistics (MBG) methods, which specify a continuous spatial model. Examples of prevalence mapping that use area-level SAE techniques include HIV prevalence (Gutreuter et al. 2019) and the under-5 mortality rate (U5MR) (Dwyer-Lindgren et al. 2014; Li et al. 2019; Mercer et al. 2015). Examples of prevalence mapping with MBG include HIV prevalence (Dwyer-Lindgren et al. 2019), malaria (Gething et al. 2016), U5MR (Burstein et al. 2019), and vaccination coverage (Mosser et al. 2019; Utazi et al. 2018).

In this report, we describe a range of models that have been proposed for SAE, and more specifically, a class of discrete spatial models, and discuss whether to use covariate information. We provide a computational pipeline, along with R scripts that use the SUMMER package (Li et al. 2020), to produce Admin 2 level estimates of U5MR.

1.2 Objectives

A model-based approach is required to obtain reliable Admin 2 estimates. In this report, we review approaches to SAE and focus on a particular model, with a fast computational implementation that we believe is a good choice for routine work. The sampling model (assumed distribution for the death indicators at the child level) is beta-binomial and is specified at the level of the cluster (enumeration area), uses discrete spatial random effects that are specified at the Admin 2 level, and includes urban/rural strata as an explanatory variable. Aggregation from the cluster-level mean produces the required Admin 2 estimates. The overall objective is to provide practical guidelines for SAE, via a concrete set of models, along with an accessible implementation in the SUMMER package.

1.3 Report Outline

In the SAE literature, there are two distinct models: area-level models take a weighted estimate in each area as the response, and then smooth, while cluster-level models model the cluster level totals individually. Even when only interested in the most recent year, spatio-temporal modeling can be advantageous because it leverages temporal smoothing, and can use more data from previous years in order to alleviate sparse data in the most recent year. In Section 2 we describe area-level models and in Section 3 we consider cluster-level models. In each section, we first consider spatial models for a generic binary indicator, before extending to space-time models and then estimating U5MR. Section 4 describes model extensions and alternative models. In Section 5 we describe the R package SUMMER, which may be used to fit all the models that we fit in the report. Section 6 presents summary results for eight countries, and detailed results for Zambia. The detailed analyses for the remaining seven countries can be found in the Appendices. Section 7 provides a concluding discussion.

2 AREA-LEVEL MODELS

2.1 Spatial Modeling of Prevalence

We will let i index the areal units for which estimation is required, with n areas in total. Assume there are N_i individuals in the population with responses $y_{ik}, k = 1, \dots, N_i$, in area $i, i = 1, \dots, n$. Let R_i denote the set of indices of the individuals sampled by the survey in area i . A *direct estimate* in a specific area only uses response data on the variable of interest from that area. For simplicity, we assume that the weights are simply the reciprocal of the inclusion probabilities. We will also not explicitly index the clusters in this section since the discussion is relevant for general designs. Let d_{ik} be the design weight associated with individual k in area i , whose response is y_{ik} . Within area i , the design-based weighted (direct) estimator (Hájek 1971; Horvitz and Thompson 1952) is

$$\hat{m}_i^{\text{HT}} = \frac{\sum_{k \in R_i} d_{ik} y_{ik}}{\sum_{k \in R_i} d_{ik}}. \quad (1)$$

The variance of this estimator, V_i^* , does not in general have an exact formula and is usually computed with either linearization or replication methods (Wolter 2007). In a child mortality context, jackknife methods are commonly used, as in Pedersen and Liu (2012). A starting point for SAE analysis is mapping the weighted estimates, if there are sufficient data in each area for those to be computed. These weighted estimates have excellent properties (if the weights are reliable and stable), such as design consistency, which means that as we sample an increasing proportion of the complete population in the area, we approach the true proportion of the condition of interest.

When the data are relatively sparse in an area, the direct estimates will have unacceptably large associated uncertainty. In a major advance, Fay and Herriot (1979) introduced a very clever approach that models a transformation of the weighted estimate to gain precision by using a random effects model. For binary outcomes, one choice of transformation is $Z_i = \text{logit}(\hat{m}_i^{\text{HT}})$. We denote the associated estimated design-based variance by V_i (which can be derived from V_i^* using the delta method).

An *area-level model* is,

$$Z_i | \theta_i \sim N(\theta_i, V_i) \quad (2)$$

$$\theta_i = \alpha + x_i^T \beta + e_i + S_i, \quad (3)$$

where θ_i is the logit of the true proportion in area i . Area-specific deviations from the regression model are modeled by using a pair of random effects. The independent and identically distributed (iid) terms are $e_i \sim \text{iid}N(0, \sigma_e^2)$, while the collection $\mathbf{S} = [S_1, \dots, S_n]$ is assigned a spatial distribution where, recall, n is the number of areas. The original Fay-Herriot model did not include the spatial random effects \mathbf{S} , but the iid random effects only. Choices for the spatial distribution are described in Banerjee et al. (2015), with common forms being the conditional autoregressive (CAR) and intrinsic CAR (ICAR) models. Both of these choices capture the concept that, in general, outcomes are likely to be similar in nearby locations. Such similarity can be due to unmeasured covariates that are associated with the outcome of interest. The ICAR model has the form

$$S_i | S_j, j \in \mathcal{N}_i \sim N\left(\frac{1}{m_i} \sum_{j \in \mathcal{N}_i} S_j, \frac{\sigma_s^2}{m_i}\right),$$

where \mathcal{N}_i is the set of neighbors of area i , and m_i is the number of such neighbors. Hence, the random effect S_i is smoothed to the mean of the neighbors of area i spatial contributions, where neighbors are commonly defined as sharing a boundary. This choice is somewhat arbitrary but has been shown to be useful for defining a smoothing model in many applications. The inclusion of iid and ICAR random effects corresponds to the popular Besag, York, Mollié (BYM) model introduced by Besag et al. (1991). The ICAR choice is a Markov random field (MRF) model, which offers computational advantages (Rue and Held 2005). The MRF models may be fit with either a frequentist or Bayesian approach in which priors are placed on the fixed effects (the β 's) and on the variances/spatial dependence parameters. Design-based consistency is achieved (if the priors do not assign zero probability density to the true θ_i), since the V_i term will tend to zero and the bias due to the random effects smoothing disappears asymptotically.

There are two practical difficulties with this approach. The direct estimates may be on the boundary for a summary parameter that is not on the whole real line. For example, in the binary case we may have \hat{m}_i^{HT} equal to 0 or 1. In this case, Z_i will be undefined. Further, a transform of the weighted estimator may not share the same design-based properties as the untransformed estimator, such as being design unbiased. These problems may be alleviated by using an unmatched sampling and linking model (You and Rao 2002). A second difficulty is that reliable variance estimates V_i may be unavailable, particularly for areas with few samples. In this case, variance smoothing can be used (Rao and Molina 2015, Section 6.4.1).

2.2 Space-Time Modeling of Prevalence

We now extend the model of the previous section to the situation in which spatially indexed data are available over time. In this case, we can exploit the tendency of demographic indicators to be smooth over time, in the absence of shocks such as natural disasters or abrupt outbreaks of conflict. Now let $\hat{m}_{i,t}^{HT}$ represent the weighted estimate in area i and (say) year t with $Z_{i,t} = \text{logit}(\hat{m}_{i,t}^{HT})$. An *area-level temporal model* is,

$$Z_{i,t} | \theta_{i,t} \sim N(\theta_{i,t}, V_{i,t}) \quad (4)$$

$$\theta_{i,t} = \alpha + x_{i,t}^T \beta + \underbrace{e_i + S_i}_{\substack{\text{Spatial} \\ \text{Main Effects}}} + \underbrace{\tau_t + \lambda_t}_{\substack{\text{Temporal} \\ \text{Main Effects}}} + \underbrace{\delta_{i,t}}_{\substack{\text{Space-Time} \\ \text{Interaction}}}, \quad (5)$$

where $\theta_{i,t}$ is the logit of the true proportion in area i and year t , $V_{i,t}$ is the design-based variance estimate, e_i and S_i are (as before) random effects indexed by space that are iid and have spatial structure, τ_t are iid temporal random effects, λ_t are temporal random effects with structure in time, and $\delta_{i,t}$ are random effects that model the space-time interaction (deviations from the main effects of space and time). The temporal smoothing models that we heavily utilize are random walk models that are local smoothers. Random walk 1 (RW1) models have,

$$\lambda_t - \lambda_{t-1} \sim \text{iid} N(0, \sigma_\lambda^2),$$

so that the contribution at time t , λ_t , are pulled toward the previous value. Random walk 2 (RW2) models have,

$$(\lambda_t - \lambda_{t-1}) - (\lambda_{t-1} - \lambda_{t-2}) \sim iid N(0, \sigma_\lambda^2),$$

so that adjacent slopes are encouraged to be similar. The interaction terms $\delta_{i,t}$ can be one of the type I to IV interactions, as described in Knorr-Held (2000). The type I model assumes the $\delta_{i,t}$ are iid, the type II model has smooth functions of time (RW2, say) crossed with iid space, the type III has smooth functions of space (ICAR) crossed with iid time, and the type IV model has smooth functions of space crossed with a smooth function of time (RW2 crossed with ICAR). More details of the smoothers are provided in the next section. For the model to be identifiable, the necessary sum-to-zero constraints are imposed on each group of random effects. These constraints are complex (Knorr-Held 2000) but are the default choice in the SUMMER package.

2.3 Space-Time Modeling of U5MR

For composite indicators such as U5MR, the direct estimates require additional modeling. The SUMMER package implements the discrete hazards model described in Mercer et al. (2015). In this section, we focus on the estimation of U5MR. Following previous work by Mercer et al. (2015), we use discrete time survival analysis to estimate age-specific monthly probabilities of dying in user-defined age groups. We assume constant hazards within the age bands. The default choice uses the monthly age bands

$$[0,1), [1,12), [12,24), [24,36), [36,48), [48,60).$$

In standard demographic notation, we let ${}_a q_{x_a}^{it}$ be the probability of death in age group $[x_a, x_a + n_a)$ in area i and year t , where x_a and n_a are the start and end of the a -th age group. Using a synthetic cohort approach (in which a hypothetical child passes through the six age bands in a single year), the U5MR for area i and year t is calculated as

$$\hat{m}_{i,t}^{HT} = {}_{60} \hat{q}_0^{it} = 1 - \prod_{a=1}^6 (1 - n_a \hat{q}_{x_a}^{it}) \quad (6)$$

The constant one-month hazards in each age band can be estimated by fitting a weighted logistic regression model (Binder 1983):

$$\text{logit}({}_1 q_m^{it}) = \beta_{a[m]}^{it} \quad (7)$$

where

$$a[m] = \begin{cases} 1 & \text{if } m = 0, \\ 2 & \text{if } m = 1, \dots, 11, \\ 3 & \text{if } m = 12, \dots, 23, \\ 4 & \text{if } m = 24, \dots, 35, \\ 5 & \text{if } m = 36, \dots, 47, \\ 6 & \text{if } m = 48, \dots, 59. \end{cases} \quad (8)$$

The design-based variance of $\text{logit}(\widehat{m}_{i,t}^{\text{HT}})$ may then be estimated using the delta method, although resampling methods such as the jackknife can also be used (Pedersen and Liu 2012). The smoothing of the direct estimates can then proceed by using the space-time model described in equations (4) and (5).

When multiple surveys exist, one may choose to either model the survey-specific effects as fixed or random (for example, Mercer et al. (2015) describe a random effects model), or first aggregate the direct estimates from multiple surveys to obtain a ‘meta-analysis’ estimate in each area and time period (Li et al. 2019). To mitigate the sparsity of available data in each year, Li et al. (2019) also consider a yearly temporal model, while the direct estimates are calculated over multiyear periods. All these variations can be fit with the SUMMER package, which we describe in Section 5. Mercer et al. (2015) and Li et al. (2019) used these *smoothed direct models* in a space-time context, with spatial ICAR and temporal random walk components, along with a space-time interaction term. This implicitly implies that the spatial, temporal, and spatio-temporal random effects are shared across all age bands.

Unfortunately, when moving to Admin 2, the data become very sparse, and it is not generally possible to form (6) for all areas and years. If there are only a small number of area-year combinations in which an estimate and variance cannot be reliably formed, then it is possible to “fill-in” these observations by using the approach described by Godwin and Wakefield (2021). However, for most situations in which Admin 2 yearly estimates are required, a cluster-level model-based approach is required, as we describe in the next section.

3 CLUSTER-LEVEL MODELS

Battese et al. (1988) describe a nested error regression model at the level of the sampling unit (thus, the name *unit-level* models). For DHS data, the units correspond to clusters (that is, enumeration areas that are the basic sampling units of the DHS), and this contrasts with the *area-level* models considered in the previous section, in which the data from all clusters in an area are combined. We begin with a model that is defined over space only for a binary response, before extending to space-time for a prevalence and then space-time for U5MR. Cluster-level models directly model the observed responses in a conventional model-based approach. Since the weights do not appear in the formulation, we must adjust for the sampling design using terms within the model. In general, model terms may be included to acknowledge stratification, while random effects account for the dependence of responses in the same cluster aspect of the design. A further complication for SAE is that we must aggregate from the cluster level to the area level. This step was not necessary in the area-level models since they directly model at the required scale.

3.1 Spatial Modeling of Prevalence

Let Y_c denote the number of events occurred in cluster c , and n_c denote the number of individuals at risk $c = 1, \dots, C$. An important distinction is that the random variables are now the responses Y_c , whereas in design-based inference, it is the units that are selected which are treated as random. Consider an overdispersed binomial, cluster-level model,

$$Y_c | p_c, d \sim \text{BetaBinomial}(n_c, p_c, d) \quad (9)$$

$$p_c = \text{expit}(\alpha + x_c^T \beta + e_{i[S_c]} + S_{i[S_c]}) \quad (10)$$

where p_c is the risk associated with cluster c , α is the intercept, x_c are cluster-specific covariates, with β the accompanying log odds ratios, and d the overdispersion parameter. Overdispersion (excess-binomial variation) is commonly seen when modeling health and demographic variables over space and time. Here, we attribute the overdispersion to the cluster sampling (dependence within clusters). The marginal variance is $\text{var}(Y_c | p_c, d) = n_c p_c (1 - p_c) [1 + (n_c - 1)d]$ so that small values of d correspond to little overdispersion.

With respect to the model specified by (9) and (10), we have two area-level random effects: an independent contribution $e_i \sim_{iid} N(0, \sigma_e^2)$ and an intrinsic conditional autoregressive (ICAR) spatial component $[S_1, \dots, S_n] \sim \text{ICAR}(\sigma_s^2, \phi)$. The notation here $i[S_c]$ reads as, “the area i within which the cluster s_c resides.” Hence, we use the BYM2 parameterization (Riebler et al. 2016), which consists of an overall variance parameter σ^2 for the random effects e_i and S_i , and a parameter, ϕ , that represents the proportion of this variance that is ICAR. In all analyses we use penalized complexity (PC) priors (Simpson et al. 2017) on variance and correlation components such as σ^2 and ϕ .

In the absence of cluster-level covariates (so $x_c = x_{i[S_c]}$ for all clusters c in area i) and stratification, we can aggregate in the above setting to the area level in a straightforward fashion to give the area-level risk:

$$p_i = \text{expit}(\alpha + x_i^T \beta + e_i + S_i), \quad i = 1, \dots, n.$$

This calculation is straightforward since we have assumed that risk is constant within the area.

Now suppose we have a binary stratification variable, which for concreteness we will label as urban/rural. In DHS surveys, it is typical for there to be oversampling of urban clusters relative to rural clusters. The extent of bias is often difficult to determine from the extent of the oversampling of clusters alone, since urban and rural clusters typically contain, on average, different numbers of households, and have differing numbers of children. The oversampling can potentially lead to bias, since rural risk is often different from urban risk. In this case, if one ignores the oversampling, we will distort the area risk because we have a nonrepresentative sample. For distortion to occur we need there to be both: (i) an association between urban/rural and the response, and (ii) oversampling of urban or rural clusters.

To make the stratification adjustment clear, we write the model as:

$$p_c = \text{expit}(\alpha + \gamma \times I(s_c \in \text{urban}) + x_{i[s_c]}^T \beta + e_{i[s_c]} + S_{i[s_c]}), \quad (11)$$

with $I(s_c \in \text{urban}) = 1$ if cluster c is urban, so that α is the intercept for rural clusters and $\alpha + \gamma$ for urban clusters. To obtain the area-level risk, we need to mix the rural and urban risks for the *aggregated risk*:

$$p_i = q_i \times \text{expit}(\alpha + x_i^T \beta + e_i + S_i) + (1 - q_i) \times \text{expit}(\alpha + \gamma + x_i^T \beta + e_i + S_i). \quad (12)$$

where q_i , the proportion of the relevant population in area i that is rural and $1 - q_i$ the proportion that is urban. The definition of urban/rural is from the time at which the sampling frame was created, for example, based on the most recent census. The DHS does not change the urban/rural classification at the time of the survey. It does not matter if a rural cluster has become urban, because the original classification is what is relevant (since this is what defines the geographical partition by which sampling is carried out). We discuss the urban and rural modeling in more detail in Section 3.3.

3.2 Space-Time Modeling of Prevalence

The space-time version of model (11) is

$$\begin{aligned} p_{c,t} &= \text{expit}(\eta_{c,t} + \tau_t) \\ \eta_{c,t} &= \alpha + \alpha_t \times I(s_c \in \text{rural}) \\ &\quad + (\gamma + \gamma_t) \times I(s_c \in \text{urban}) \\ &\quad + x_{i[s_c],t}^T \beta + e_{i[s_c]} + S_{i[s_c]} + \delta_{i[s_c],t}, \end{aligned} \quad (13)$$

where τ_t is an iid normal temporal term that we discuss in more detail shortly, α is the intercept for rural clusters, and $\alpha + \gamma$ is the intercept for urban clusters (both are with respect to time $t = 0$), and α_t and γ_t are the time-varying parameters for the rural and urban clusters, respectively. These parameters are both assigned RW2 smoothing priors to acknowledge that we expect similarity over time. For the iid shocks τ_t it is a contextual choice as to whether these terms are included in the fitted curves. If one believes they correspond to the ‘true’ signal, then we would include, while if we think they reflect local biases (recall bias, for example), then we should not include. The default in the SUMMER package is to exclude these terms. Further discussion is provided in the next section.

To obtain the area-level risk over time, we need to mix the rural and urban risks to provide the *aggregated risk*:

$$p_{i,t} = q_{i,t} \times \text{expit}(\alpha + \alpha_t + x_{i,t}^T \beta + e_i + S_i + \delta_{i,t}) + (1 - q_{i,t}) \times \text{expit}(\alpha + \gamma + \gamma_t + x_{i,t}^T \beta + e_i + S_i + \delta_{i,t}) \quad (14)$$

where $q_{i,t}$ is the proportion of the relevant population in the area that is rural in year t , and $1 - q_{i,t}$ is the proportion of the relevant population in the area that is urban.

3.3 Space-Time Modeling of U5MR

We assume a discrete hazards model as before in Section 2.3. We consider a beta-binomial model for the probability (hazard) of death from month m to $m + 1$ at cluster location s_c in year t , because we expect overdispersion. Assuming constant hazards within age bands, we assume that the number of deaths occurring within age band $a[m]$, in cluster c and at time t , follow the beta-binomial distribution,

$$Y_{a[m],c,t} \mid h_{a[m],c,t}, d \sim \text{BetaBinomial}(n_{a[m],c,t}, h_{a[m],c,t}, d), \quad (15)$$

where $h_{a[m],c,t}$ is the monthly hazard within age band $a[m]$, in cluster c , at time t , and d is the overdispersion parameter.

In general, we want to model the hazards as a function of space, time, child's age, and urban/rural strata. Since the sparsity of data do not support using most of the possible interactions, parsimonious choices must be made.

- **Age Terms:** We consider $a = 1, \dots, 6$ age bands (as in (8)), in the sense of having 6 intercepts, and allow $a^* = 1, 2, 3$ time trends – one each for the first two age bands, and a final common trend for the remaining 4 age bands (the default choice in the SUMMER package). We write this as,

$$a^*[m] = \begin{cases} 1 & \text{if } m = 0, \\ 2 & \text{if } m = 1, \dots, 11, \\ 3 & \text{if } m = 12, \dots, 59. \end{cases} \quad (16)$$

We could include 6 time trends, one for each age band, but for the sake of parsimony, we focus modeling efforts on the earlier ages, where deaths are more prevalent.

- **Spatial and Spatio-Temporal Terms:** The spatial jumps e_i and S_i , and the space-time interactions $\delta_{i,t}$ are assumed to act equally on all child age groups. We also include an iid temporal ‘shock’ τ_t . As discussed in the last section, it is a contextual choice as to whether they are used in predictions. The default in the SUMMER package is not to include, which results in smoother temporal trajectories. If we do not include, we can think of the temporal jumps as localized bias terms, or short scale variation that we do not trust as ‘real,’ and that takes attention away from the more persistent temporal trends. If there were conflicts and natural disasters that are localized in time to specific years, then we may wish to include the τ_t terms. The concern is that we would oversmooth in such situations.

- **Urban/Rural Adjustments:** We include separate urban and rural temporal terms to acknowledge the sampling design. Since urban clusters are often oversampled and have different risk from rural clusters, it is important to acknowledge this in the model (Paige et al. 2020). We assume urban/rural associations are the same in all areas and so do not depend on i . We also need to decide in which age groups the urban/rural associations are constant across (we choose the 3 age bands (16)), so that we have three odds ratio parameters. Because the true urban/rural classification is not constant over time and the association is likely to change over time, we allow different odds ratios over time, although we encourage smoothness through the use of RW2s on the log odds ratio parameters. We emphasize that we are including these terms to acknowledge the design, and not to model ‘true’ urban/rural differences in risk.

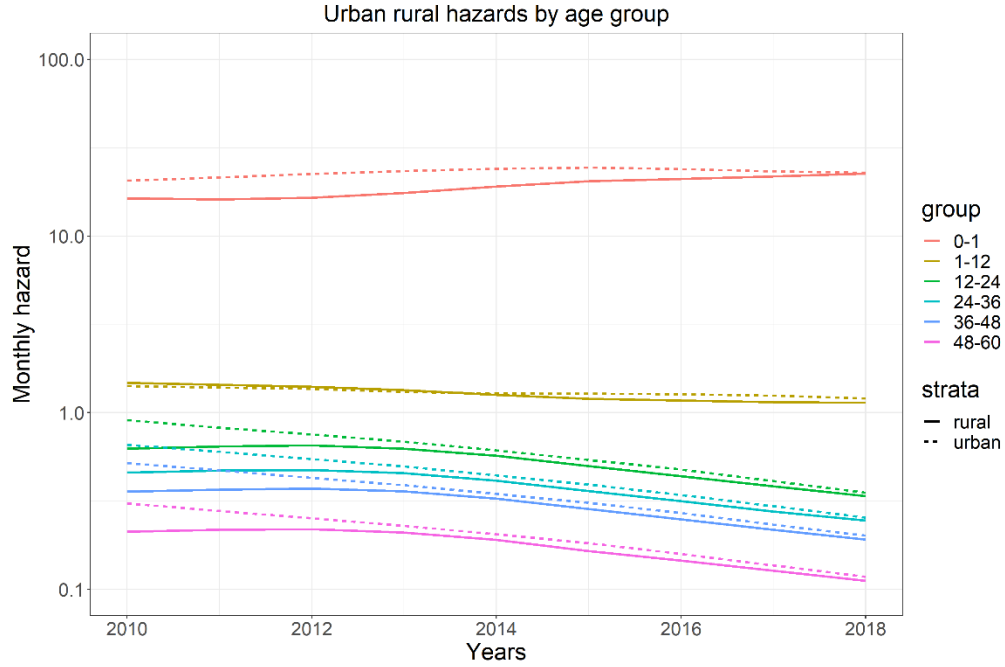
The latent logistic model we use for the 8 country analyses in Section 6 is,

$$h_{m,c,t} = \text{expit}(\eta_{m,c,t} + \tau_t) \quad (17)$$

$$\begin{aligned} \eta_{m,c,t} &= \alpha_{a[m]} + \alpha_{a^*[m],t}^* \times I(s_c \in \text{rural}) \\ &+ (\gamma_{a^*[m]} + \gamma_{a^*[m],t}^*) \times I(s_c \in \text{urban}) \\ &+ x_{i,t}^T \beta + S_{i[s_c]} + e_{i[s_c]} + \delta_{i[s_c],t}. \end{aligned} \quad (18)$$

This form includes a collection of terms that are used for prediction and a number that are not. The τ_t are unstructured (iid) temporal effects that allow for perturbations over time (in the aggregation process below we exclude these terms in our area-level estimates, see the previous discussion). We include spatial main effects S_i and e_i and the space-time interactions $\delta_{i,t}$, and the covariate model $x_{i,t}^T \beta$ (with area-level covariates) as before. As discussed, we have 6 age intercepts for rural (α_a) (the baseline values), and then 3 urban adjustments (γ_{a^*}). Over time, we have 3 temporal smoothers for each of the rural, $\alpha_{a^*[m],t}$, and urban clusters, $\gamma_{a^*[m],t}$. Figure 1 shows the estimates of the hazard rates for Zambia over time, color coded by age band, and with different line types corresponding to urban and rural. Because of the log scale, we can see the parallel curves for the 4 older age bands, and the constant association with urban/rural for these 4 age groups. For the neonatal group, there is a slight increase in hazard over time, while for the other age bands there are decreases, which are more pronounced for the last 4, which are forced to have the same temporal trend. Although U5MR is typically lower in urban than rural areas (Yaya et al. 2019), we observe the opposite in Zambia. In general, the urban/rural association becomes less pronounced over time, which is consistent with urbanization that is occurring in reality, although the urban/rural cluster labels remain constant, so that we have a type of measurement error in the label. See Section 3.4 for more discussion.

Figure 1 Hazard rates (urban/rural) over time for 6 age bands. There are 3 temporal trends for 0–1 month, 1–12 months, and for all age groups >12 months (i.e., the last 4 age groups). Similarly, there are 3 urban/rural hazard rate adjustments.



The estimated U5MRs in rural and urban regions of area i and at time t are,

$$U5MR_{i,t,R} = 1 - \prod_{m=0}^{59} \left[\frac{1}{1 + \exp(\alpha_{a[m]} + \alpha_{a[m],t}^* + S_i + e_i + \delta_{i,t})} \right] \quad (19)$$

$$U5MR_{i,t,U} = 1 - \prod_{m=0}^{59} \left[\frac{1}{1 + \exp(\alpha_{a[m]} + \gamma_{a^*}^* + \gamma_{a^*}^* + S_i + e_i + \delta_{i,t})} \right], \quad (20)$$

for the default choice of age bands. The aggregate risk in area i and in year t is

$$p_{i,t} = q_{i,t} \times U5MR_{i,t,R} + (1 - q_{i,t}) \times U5MR_{i,t,U} \quad (21)$$

where $q_{i,t}$ and $1 - q_{i,t}$ are the proportions of the under-5 population in area i that are rural and urban in year t . The process by which we estimate $q_{i,t}$ is described in the next section.

Beyond point estimates, we obtain the full posterior of $U5MR_{i,t}$, so that various summaries can be reported or mapped. The SUMMER package uses samples from an approximation to the posterior to allow inference for $U5MR_{i,t}$. The estimate constructed for U5MR is not relevant to any child, because that child would have to experience the hazards for each age group simultaneously in time period t , rather than moving through age groups over multiple time periods. Nevertheless, the resulting U5MR is a useful summary and the conventional measure used to inform on child mortality.

3.4 Creating an Urban/Rural Stratification Surface

We now discuss how the $q_{i,t}$ can be estimated. From the DHS reports, or census summary tables, we can obtain the fractions of the complete population that are rural at the Admin 1 level, at the time of the census from which the DHS sampling frame was constructed. Using WorldPop population density surfaces (Stevens et al. 2015) from the year of census sampling frame construction, we can define an area-specific population density threshold that will provide the correct proportion at the Admin 1 level. The WorldPop population density estimates are available on a 1×1 km grid, and given the thresholding procedure just described, we are able to define each point of the grid as urban or rural.

We can then obtain the 0–5 population at each grid point. Other populations can be used as the response dictates. For example, if we are examining female educational attainment for a certain age group, we can use population estimates for that age group. We assume that the same Admin 1 threshold applies in each of the constituent Admin 2 areas. Summing up the grid points within an Admin 2 geographical area therefore provides the fraction of the 0–5 population that is urban/rural.

The sampled clusters are not included in this classification process. Thus, one can use the urban/rural grid to assign the sampled clusters to urban/rural, and then can compare with the classification used in the survey. This is for the year of the census. To go forward in time (to obtain $q_{i,t}$), we can use the same classification, in most countries, since the sampling frame is not updated. Hence, the urban/rural grid labels are constant over time, although the fractions will change because the WorldPop population density grid values change across years.

It might appear strange that the urban/rural classification of grid points is constant over time, because urbanization is typically occurring. However, the original sampling frame for urban/rural classification is what is relevant, because this defines the stratification that is then used for sampling. In the same way, the design weights are also constant over time. To account for the stratification, one needs to model the association between the design variable and the hazard, and since the relationship may change over time, we include a time-varying association parameter.

4 ADDITIONAL CONSIDERATIONS

4.1 Covariate Modeling

Incorporating auxiliary information as covariates in area-level or cluster-level models can improve the accuracy of the resulting small area estimates (Rao and Molina 2015). Auxiliary information for a given covariate that is available at the cluster-level can be incorporated in the cluster-level linear predictor models in an obvious way. With covariates that vary within areas, the difficulty with a nonlinear model is the aggregation step to obtain the Admin 2 estimate of the area-level risk. The difficulty is that the aggregation requires a population density surface and a covariate surface. The population density surface is estimated, and the covariate surface may also be estimated. In practice, it is difficult to assess how closely these correspond to the true population density and the true covariate surface. For illustration, assume we have a risk model

$$p(s) = \text{expit}(\alpha + x(s)^T \beta + S(s)),$$

where α is the intercept, $x(s)$ are covariates measured at cluster spatial location s with associated log odds ratios β , and $S(s)$ is the latent residual (logit) risk at location s . Then the aggregated risk in area i is

$$p_i = \int_{A_i} p(s) \times q(s) ds \approx \sum_{l=1}^{M_i} p(s_l) \times q(s_l) \quad (22)$$

where $q(s)$ is the population density at s (normalized to integrate to 1 over A_i) and we have approximated the integral in area i using a grid indexed by $l = 1, \dots, M_i$ grid points. We emphasize that the covariate surface $x(s_l)$ is required at all of the grid points s_l .

For continuous outcomes with an identity link function, prediction is straightforward when we have area-level mean covariate values. However, when using a binary response model for which cluster-level probabilities undergo nonlinear transformation, covariate information is required for all individuals. This means that we need cluster-specific covariate information for all clusters in the sampling frame, as well as estimates of the population for each cluster. Thus, generating predictions from a model with cluster-level covariates can be difficult in a setting where we have incomplete covariate information for our entire population, which is often the case for low- and middle-income countries (LMICs).

With the cluster-level linear predictor (18), the simplest approach is to add area-level covariates, which use values that are constant for each cluster in a given area. Hence, we add $x_{i,t}^T \beta$, where β are the associated log odds ratios. For example, the area-level covariates can be population-weighted averages of pixel-level surfaces. For both binary and continuous responses, prediction is straightforward because the effect of a covariate is identical for all individuals belonging to an area and there is no within-area variation in the risk, unless we model the urban/rural strata.

Cluster-level covariate models are more appealing than area-level covariate alternatives since they are closer to the mechanism of action and reduce the possibility of ecological bias (Wakefield 2008), which arises when one tries to interpret area-level association as relevant to the individuals within the areas (though for prediction this is less important, although the loss of information in using an area-level summary

is still important). Using satellite-derived fine-scale maps of covariates such as vegetation or nighttime lights has become popular for creating maps of poverty indicators. However, when Admin 2 estimates for LMICs are the target, maps of urban/rural classifications, population density, and covariates are typically estimates and may exhibit biases and large uncertainties. These biases are not well understood, and data sources that provide uncertainty estimates often only provide marginal uncertainty at each location. More work is needed to understand the effects of uncertainty in the covariate map as well as uncertainty in population density. Jittering, or adding random noise to the data, in the cluster locations, which is commonly done for confidentiality purposes, would introduce further error in estimating cluster-level covariate information. This means that continuous spatial models are appealing and necessary for fine-scale maps of indicators, although this does not imply that they should be the default choice for producing reliable areal estimates.

4.2 Continuous Spatial Models

The modeling strategy suggested in this report for producing Admin 2 estimates is to use discrete spatial models. We encourage the use of these models due to their ease of implementation and benefits in terms of aggregation. However, a key component of recent works on subnational estimation of U5MR (Burstein et al. 2019; Golding et al. 2017; Wakefield et al. 2019) has been to produce fine-scale spatio-temporal maps. In the Appendix we provide more details on these models, and here content ourselves with references to the literature on continuous models. Stein (1999) provides extensive discussion of the theoretical aspects of continuous spatial models, while Diggle and Giorgi (2019) offer a practical guide to their use in the context of public health applications. Heaton et al. (2018) provide an extensive simulation study that compares various implementations of continuous spatial models. Paige et al. (2020) include detailed simulation studies that compare discrete and continuous spatial models and provide support for our use of the former. We favor discrete spatial models because we aim for admin estimates. A summary of our position is that we do not need to capture continuous spatial variation and aggregation gives an additional source of inaccuracy due to imprecise fine-scale population information. We acknowledge, however, the potential for traditional geostatistical models to pick up more local spatial variation that does not fit with administrative borders.

4.3 Aggregation

In this report we focus on Admin 2 estimation, although it can be of interest to aggregate these estimates to the Admin 1 or national level for model checking. We focus on aggregating Admin 2 to Admin 1. In this section, let i be the Admin 1 area index and j be the Admin 2 area index. We use T_i to denote the set of indices of Admin 2 areas within Admin 1 area i , such that $j \in \{T_i\}$ indicates that Admin 2 area j is within Admin 1 area i .

First, we calculate the proportion of the population in Admin 2 area j with respect to its upper area-level Admin 1 area i . Let f_{ij} be this fraction and let q_{ij} denote population density within Admin 2 area j . We obtain population density from WorldPop surfaces for the under-5 population and we treat them as known without uncertainty. The calculation follows as

$$f_{ij} = \frac{q_{ij}}{\sum_{j \in T_i} q_{ij}} \quad (23)$$

Aggregation to the upper area-level is carried out to give the estimand,

$$p_i = \sum_{j \in \mathcal{T}_i} p_{ij} \times f_{ij} \quad (24)$$

where p_i is the U5MR at Admin 1 area i and p_{ij} is U5MR at Admin 2 area j , which is within Admin 1 area i . The aggregated Admin 1 U5MR is a convex combination of U5MR at Admin 2 level with $\sum_{j \in \mathcal{T}_i} f_{ij} = 1$. For estimation, we obtain the posterior distribution of $\pi(p_i|\mathbf{y})$ by sampling

$$p_i^{(s)} = \sum_{j \in \mathcal{T}_i} p_{ij}^{(s)} \times f_{ij},$$

where $p_{ij}^{(s)}$ are posterior draws from $\pi(p_{ij}|\mathbf{y})$, $s = 1, \dots, S$, which we obtain from the Admin 2-level beta-binomial model.

It is important to note that if one is primarily interested in estimates at a particular level, it is more often preferable to model at that level since one does not have to perform aggregation, which can add error.

4.4 Model Assessment

We validate the smoothed direct, stratified Admin 1 level beta-binomial, and stratified Admin 2 level beta-binomial models by removing data from one Admin 1 area at a time in the final time period. Predictions for the omitted area are compared against the direct estimate for that area, and the predictions of the Admin 2 level beta-binomial model are aggregated to the Admin 1 level with the associated population weights of the constituent Admin 2 areas. This validation method allows us to evaluate the models' accuracy. We exclude data in the time period with the latest survey due to its importance for the purpose of estimation in the survey year.

We consider a number of scoring rules and metrics to help evaluate model performance: relative bias, (absolute) bias, root mean square error (RMSE), and 80% and 95% CI width. Given central predictions from the model, $\hat{\mathbf{p}} = (\hat{p}_1, \dots, \hat{p}_m)^T$ and direct estimates $\mathbf{p} = (p_1, \dots, p_m)$ for the m Admin 1 areas, we calculate the overall scoring rules and metrics for a country as:

$$\text{RelativeBias}(\hat{\mathbf{p}}, \mathbf{p}) = \frac{1}{m} \sum_{i=1}^m 100\% \times \frac{\hat{p}_i - p_i}{p_i} \quad (25)$$

$$\text{Bias}(\hat{\mathbf{p}}, \mathbf{p}) = \frac{1}{m} \sum_{i=1}^m 1000 \times (\hat{p}_i - p_i) \quad (26)$$

$$\text{RMSE}(\hat{\mathbf{p}}, \mathbf{p}) = \sqrt{\frac{1}{m} \sum_{i=1}^m (\hat{p}_i - p_i)^2}. \quad (27)$$

We use percentage units for relative bias, and children per thousand births units for absolute bias. Given the average scoring rules for each country, we then average with equal weight over all countries to obtain the overall scoring rules and metrics for each model.

Relative bias can be used for determining how much larger the predictions are relative to the direct estimates as a percentage, while bias measures the mean absolute differences between the direct estimates and predictions. RMSE can be approximately viewed as the typical error in the predictions.

A useful property of relative bias is that it acknowledges inherent differences in errors that depend on the scale of what is being predicted. A disadvantage is that if some direct estimates are very close to zero due to noise in the data, predictions only slightly higher on an absolute scale can potentially result in relative biases that are very large. Such biases are due entirely to noise in the direct estimates rather than any substantial fault in the predictions. Significant relative bias is therefore not unexpected for unbiased models.

5 THE SUMMER SOFTWARE

The models discussed in this report can be fitted with the R package SUMMER (Li et al. 2021, 2020). The SUMMER package provides a general framework for smoothing and mapping prevalence with complex survey data. For the estimation of child mortality, the SUMMER package provides the pipeline to process DHS birth records and fit both area-level and cluster-level models, and specifically, direct, smoothed direct, and model-based estimates with different model specifications. Fully Bayesian estimation techniques are used that utilize the integrated nested Laplace approximation (INLA) computational machinery as described by Rue et al. (2017). This technique has been extensively investigated in the context of spatial modeling, such as in Osgood-Zimmerman and Wakefield (2020), and the work of Blangiardo and Cameletti (2015), Krainski et al. (2018), Moraga (2019), and Gómez-Rubio (2020).

The SUMMER package can also be used for SAE with other binary and continuous outcomes. The package has been successfully utilized for a range of data including the estimation of mortality rates (Li et al. 2019; Mercer et al. 2015; Schlüter and Masquelier 2021), HIV prevalence (Wakefield et al. 2020) and vaccination coverage (Dong and Wakefield 2021). In addition to model fitting, the package provides a collection of visualization tools that summarize and present estimated prevalences.

Recently, the SUMMER package was used to obtain the official United Nations Inter-agency Group for Mortality Estimation (UN IGME) yearly estimates (1990–2018) of U5MR at Admin 2 level for 22 countries in Africa and Asia (UN IGME 2021). A variant of the beta-binomial sampling model described in Section 3.3 was used with: 1. cluster-level modeling, 2. space-time smoothing, 3. country-specific models, 4. Bayesian inference, 5. overdispersion, 6. benchmarking to UN IGME national U5MR estimates, 7. HIV adjustment, and 8. informative visualization. These eight characteristics, along with the beta-binomial likelihood, lead to the model label of BB8. In the work of this report, we do not incorporate benchmarking or HIV adjustment, although this is possible with the SUMMER package.

A standardized description of the materials and code for reproducing the work are available at <https://github.com/wu-thomas/SUMMER-DHS>

An example piece of SUMMER code is given below. The documentation describes the syntax in detail, but many of the arguments are self-explanatory. The arguments beginning with `pc` concern the PC hyperpriors that we use.

```
### Stratified beta-binomial model at Admin 2
fit.strat.admin2 <- smoothCluster(data = mod.dat,
  family = "betabinomial",
  Amat = admin2.mat,
  year_label = 2010:2018,
  strata.time.effect = TRUE,
  time.model = "rw2", st.time.model = "ar1",
  type.st = 4, pc.st.slope.u = 1, pc.st.slope.alpha = 0.01,
  age.groups = c("0", "1-11", "12-23", "24-35", "36-47", "48-59"),
  age.n = c(1, 11, 12, 12, 12, 12),
  age.rw.group = c(1, 2, 3, 3, 3, 3),
  age.strata.fixed.group = c(1, 2, 3, 3, 3, 3))
```


6 U5MR MODELING IN EIGHT COUNTRIES

6.1 Eight Country Summary

In this section, we present results obtained from eight DHS surveys (we use the last survey available in each country), which used the cluster-level model described in Section 3.3 and were implemented in the SUMMER package. We selected countries with a variety of numbers of Admin 1 and Admin 2 areas – the list of surveys and the number of admin areas are summarized in Table 1. Although our principal aim is to utilize estimates for the most recent period, we use the retrospective nature of full birth history data to fit a model to 9 years of data. The extra years of data aid in estimation, since we can leverage temporal similarity of rates.

Table 1 Numbers of Admin 1 and Admin 2 areas and survey years for the eight countries in the analysis

Country	DHS survey year	Admin 1 areas	Admin 2 areas
Bangladesh	2018	7	64
Cameroon	2018	10	58
Ethiopia	2016	11	79
Kenya	2014	47	301
Malawi	2015	3	28
Nepal	2016	5	13
Nigeria	2018	37	774
Zambia	2018	10	72

Figures 2–9 show the estimated U5MR and width of the 95% credible intervals. The Admin 2 level estimates are mapped for each country at the year of the survey. The estimates are produced from the beta-binomial model, given by equations (15), (17), and (18). Because the scales are different in the different figures and the surveys were conducted in different years, direct comparison is not possible. In general, we see spatial structure in the U5MR estimates, in that two areas that are neighbors have more similar risk than two areas that are far apart. We also see substantial variation across Admin 2 regions in each country.

The right-hand plots show that there is significant uncertainty in many of the Admin 2 areas, with the width of the 95% credible intervals routinely greater than 60 deaths per 1,000 live births. Comparison of the left and right figures reveals that the areas with higher U5MR estimates have relatively higher uncertainty, which stems from the mean-variance relationship of the beta-binomial. These plots are included to indicate the inherent uncertainty, although the Bayesian machinery allows one also to examine many other summaries such as the probabilities of exceedance of certain values, which may be more suitable for answering public health questions. We include such plots later in the report.

Tables 2 and 3 provide three number summaries of U5MR in each country at the Admin 1 and Admin 2 level, respectively. The 5th percentiles and 95th percentiles show that there is considerable subnational variation in each country. The median indicates the estimated U5MR in a typical area. Cameroon and Nigeria have relatively high levels of U5MR at the time of the survey, while Bangladesh and Nepal have relatively low levels of U5MR. In general, the range of subnational variation is greater at the Admin 2 level. For Nigeria, the ratio of the 95th to 5th percentile of the distribution of posterior medians is 3.7 for Admin 1 and 4.7 for Admin 2, which is considerable. Due to the small number of Admin 1 areas in most countries, the quantiles are approximated, and can only serve as a rough guide for the subnational variation. In general,

there is large variation across areas. Understanding this variation and equalizing the burden of U5MR is clearly a public health priority. The aim is to have equal and low risk in all subnational regions.

Table 2 Spread of Admin 1 estimates (rate per 1,000 live births) for the year of the survey

	Country	DHS Survey year	5 th Percentile	Median	95 th Percentile
1	Bangladesh	2018	35.1	38.7	42.8
2	Cameroon	2018	46.5	80.2	115.1
3	Ethiopia	2016	40.0	62.9	85.9
4	Kenya	2014	28.2	42.8	68.4
5	Malawi	2015	47.5	55.7	59.8
6	Nepal	2016	31.4	34.1	38.8
7	Nigeria	2018	49.5	98.7	181.3
8	Zambia	2018	36.4	52.7	74.0

Table 3 Spread of Admin 2 estimates (rate per 1,000 live births) for the year of the survey

	Country	DHS Survey year	5 th Percentile	Median	95 th Percentile
1	Bangladesh	2018	31.3	35.8	43.1
2	Cameroon	2018	51.6	79.4	113.2
3	Ethiopia	2016	44.5	65.1	87.5
4	Kenya	2014	27.9	42.0	75.8
5	Malawi	2015	44.4	53.2	71.6
6	Nepal	2016	30.0	34.0	38.5
7	Nigeria	2018	44.9	92.2	209.0
8	Zambia	2018	25.8	45.7	86.0

Figure 2 U5MR estimates and uncertainty at year of survey for Bangladesh

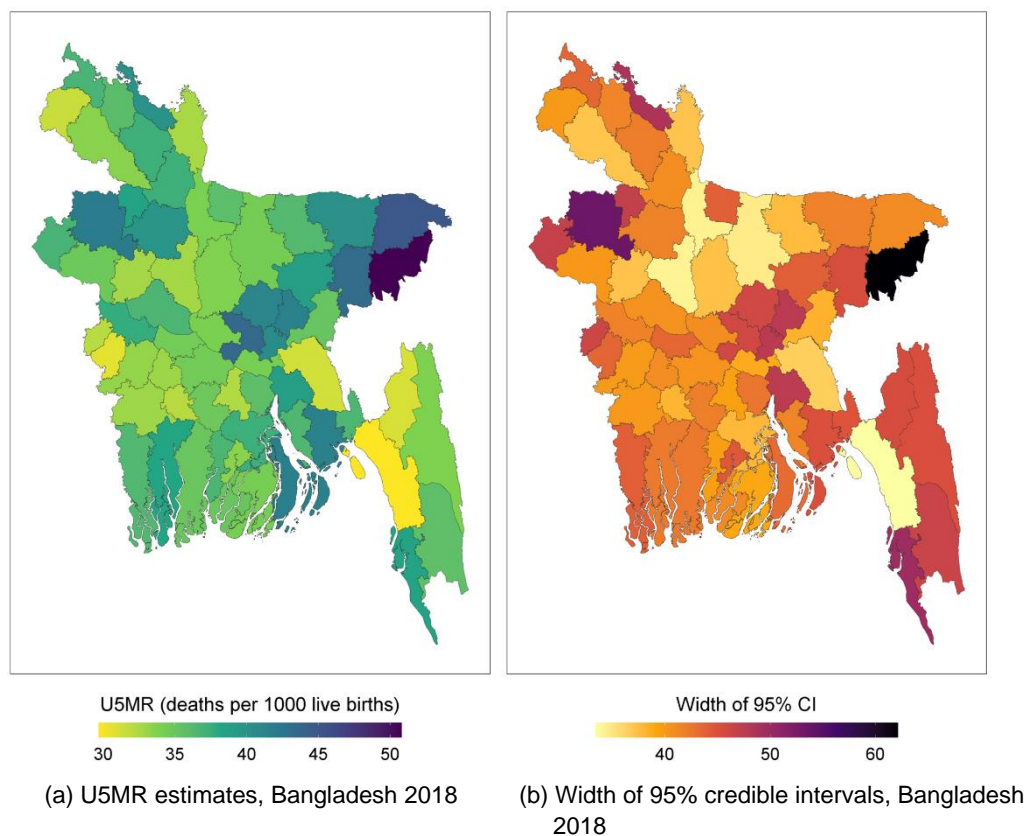


Figure 3 U5MR estimates and uncertainty at year of survey for Cameroon

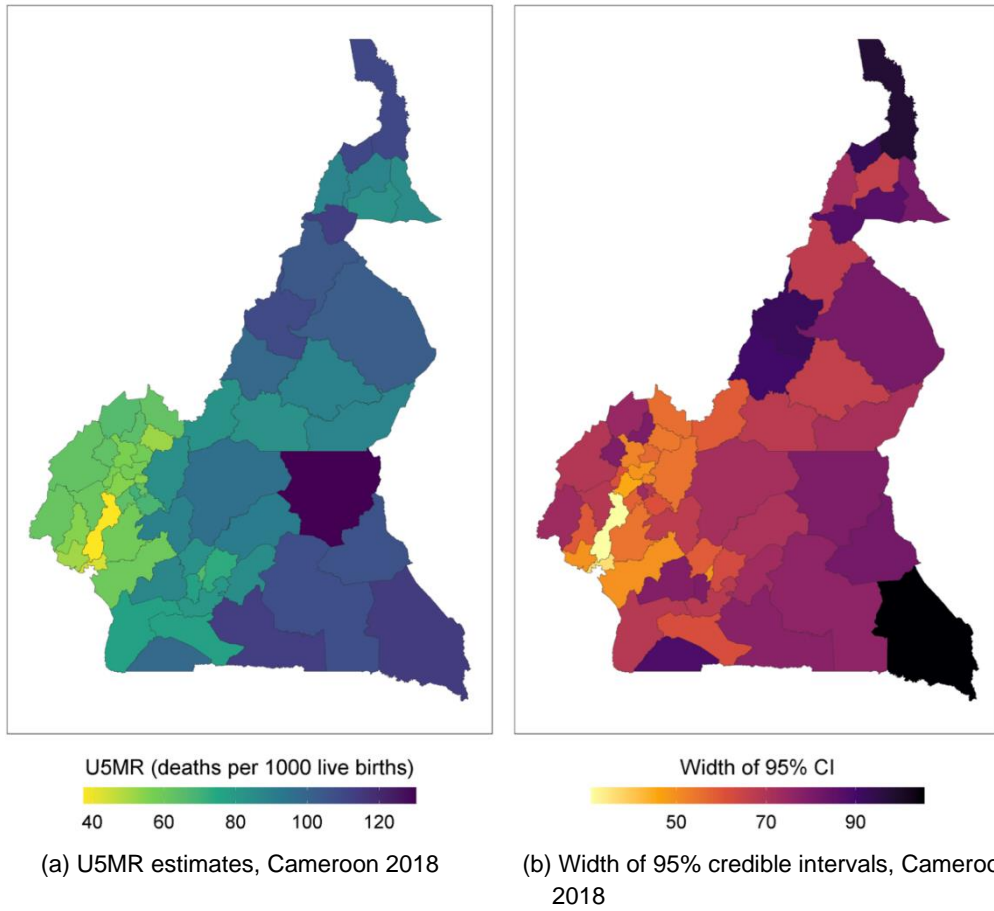


Figure 4 U5MR estimates and uncertainty at year of survey for Ethiopia

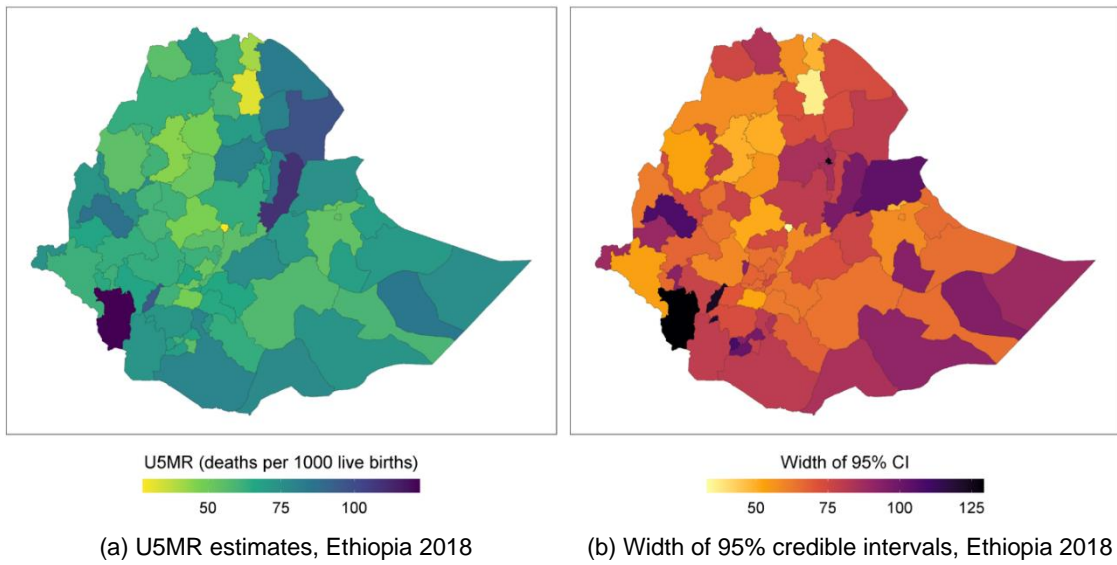


Figure 5 U5MR estimates and uncertainty at year of survey for Kenya

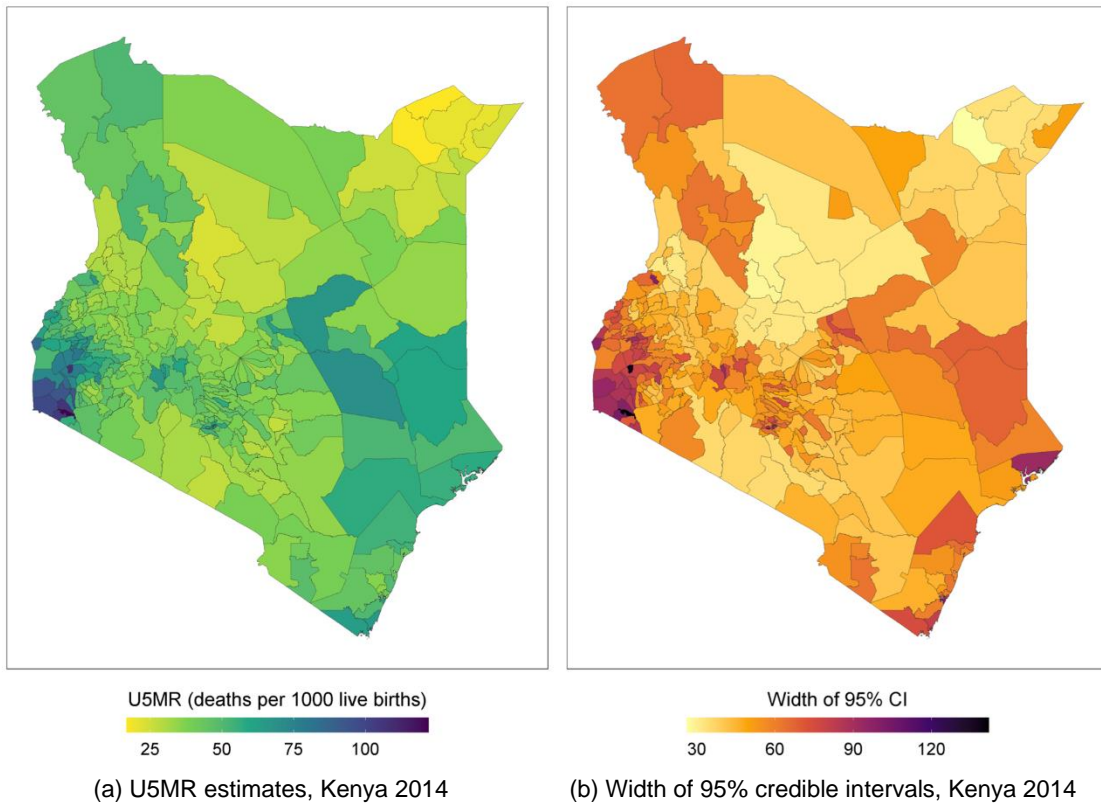


Figure 6 U5MR estimates and uncertainty at year of survey for Malawi

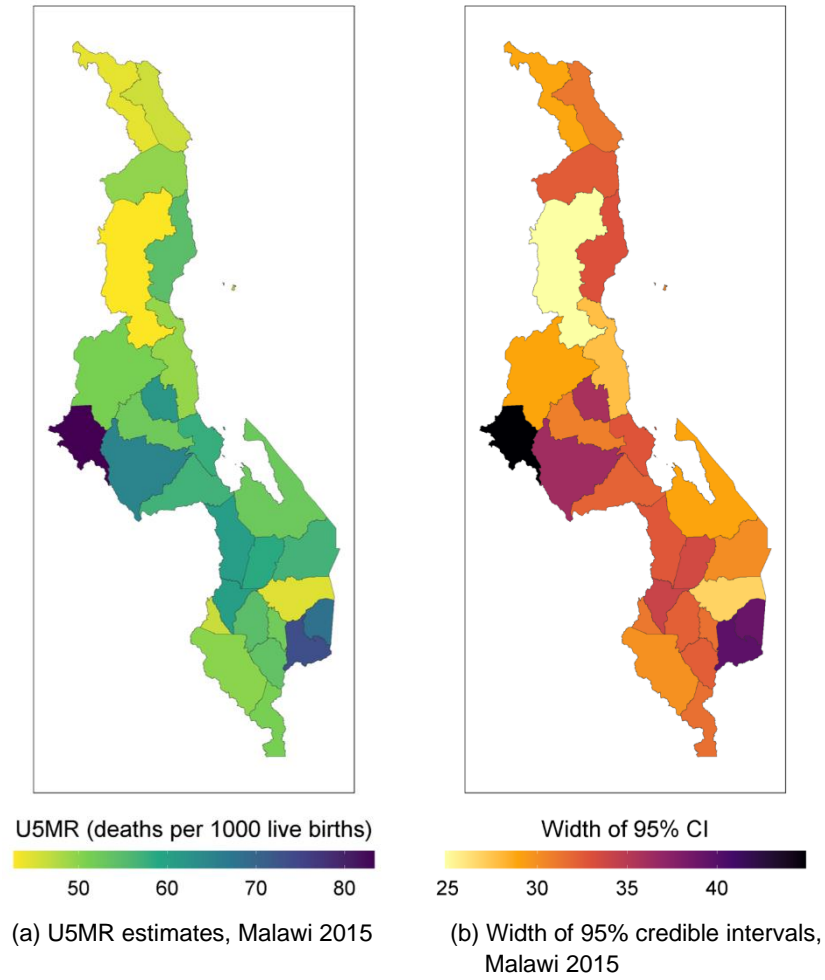


Figure 7 U5MR estimates and uncertainty at year of survey for Nepal

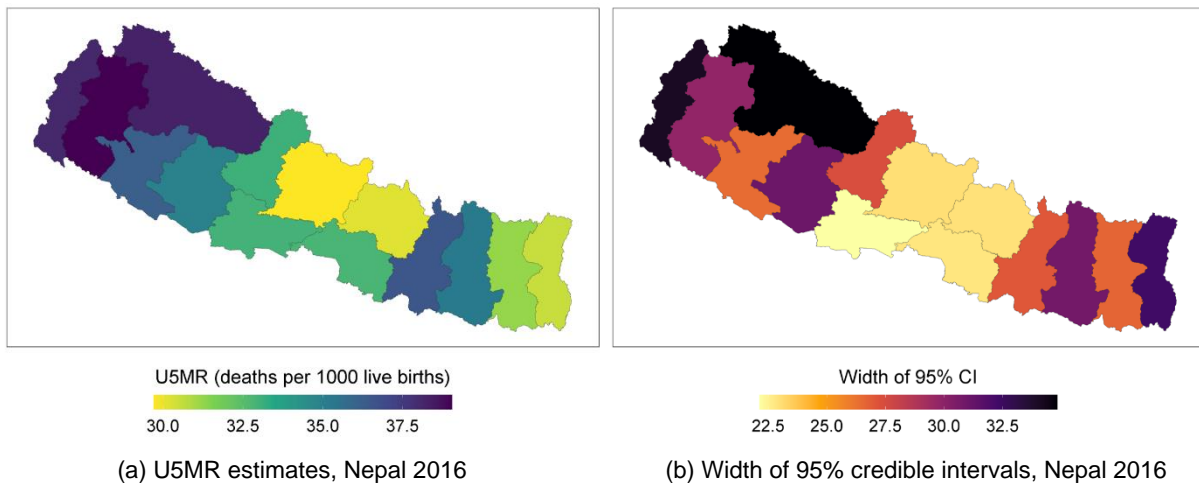


Figure 8 U5MR estimates and uncertainty at year of survey for Nigeria

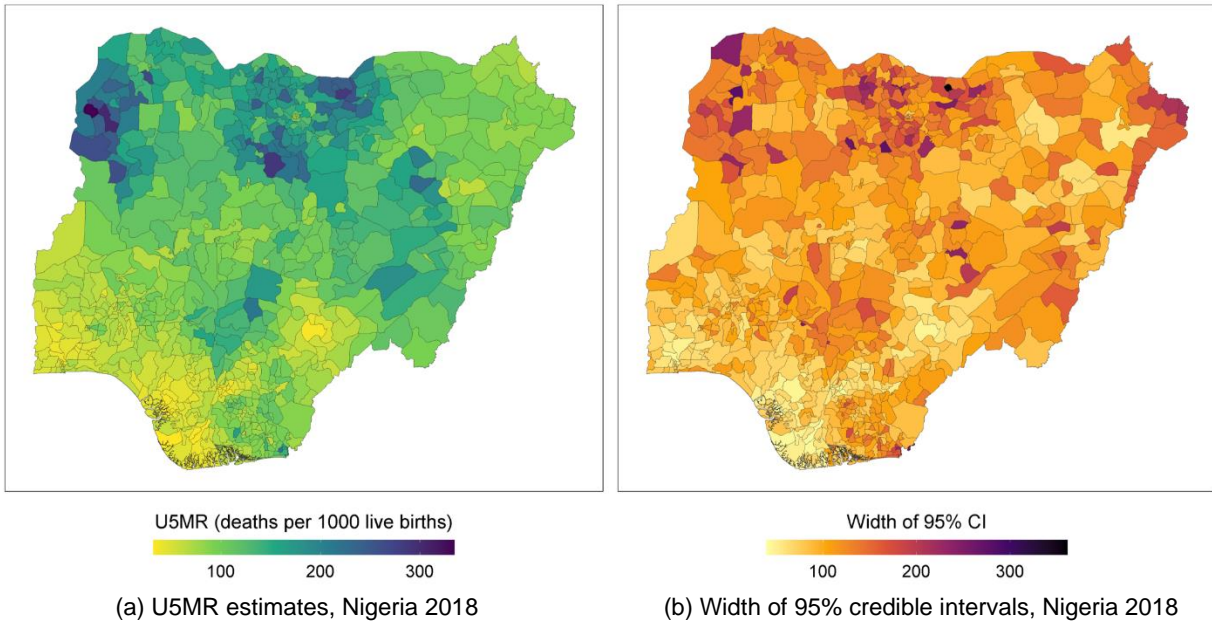
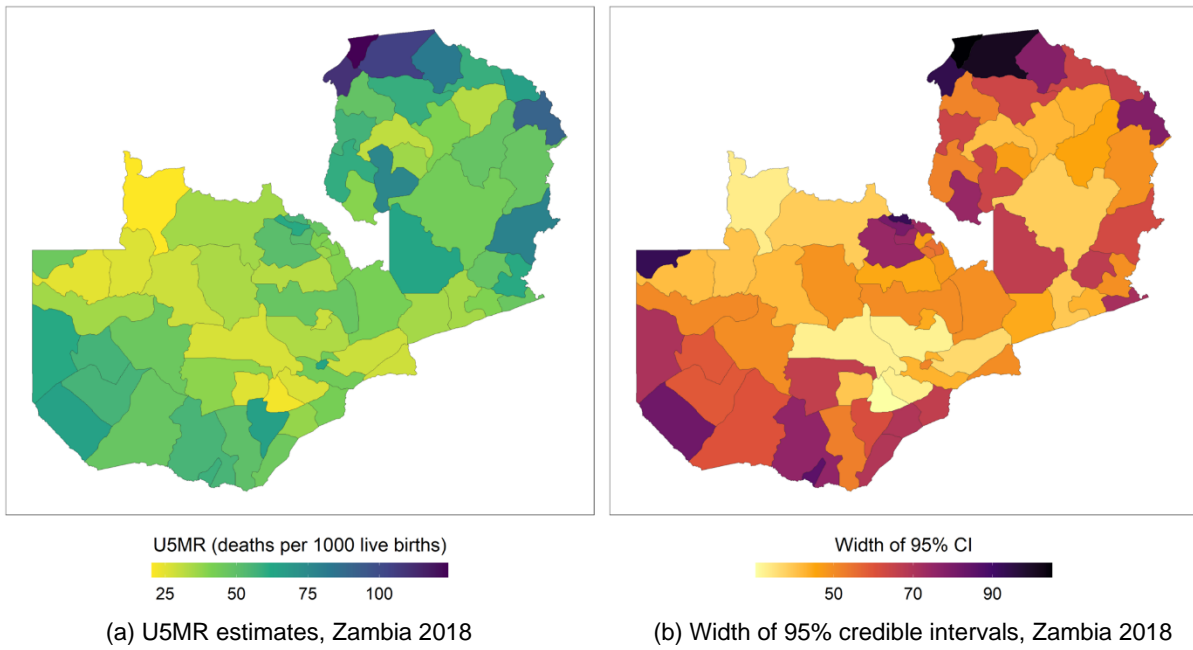


Figure 9 U5MR estimates and uncertainty at year of survey for Zambia



6.2 Validation

Table 4 summarizes the cross-validation metrics for all countries. The cross-validation experiment was devised with the following steps. For the final 3-year period, all data were removed from a single Admin 1 area, and this was repeated for each Admin 1 area. The metrics given by (25), (26), and (27) were calculated for each country, and also summed, with equal weight, across countries. Two beta-binomial cluster-level models were fitted, one with the spatial and spatio-temporal model specified at the Admin 1 level, and the other at the Admin 2 level; the cluster-level models we used were described in Section 3.3. We also fit an area-level smoothed direct model at Admin 1 (recall that the direct and smoothed direct models are not tenable at Admin 2, due to data sparsity). For the Admin 2 level model, we aggregate the beta-binomial model with spatial and spatio-temporal terms specified at the Admin 2 level, using the method described in Section 4.3. To compare estimates from yearly beta-binomial model with direct estimates for a 3-year window, we also aggregated draws from 3 years, as described in Section 4.3, to form a single posterior distribution of estimates.

Overall, the Admin 1 level beta-binomial model performed the best in terms of relative bias, bias, and RMSE. The Admin 1 level beta-binomial model RMSE was 0.015 compared to the smoothed direct and Admin 2 level beta-binomial model RMSEs of 0.017 and 0.027, respectively. This indicated better overall central predictions for the Admin 1 level beta-binomial model. The Admin 1 level beta-binomial model relative bias was also the closest to zero at 10.9%, compared to the relative bias of the smoothed direct and Admin 2 level beta-binomial models, which were 12.9% and 17.3%, respectively. The bias of the Admin 1 level beta-binomial model was 0.6 children per thousand births compared to 1.3 for the smoothed direct model and -0.2 for the Admin 2 level beta-binomial model.

Although the Admin 2 level beta-binomial model had the smallest absolute bias, it also had the largest RMSE. This suggests that modeling at a finer spatial scale than that of the predictions does not necessarily improve predictive performance, and can in fact, make the predictions worse. The relatively poor performance of the Admin 2 level beta-binomial model (when predicting at the Admin 1 level) may be due in part to variability induced by the process of aggregating predictions to coarser spatial scales. This latter additional variability is a result of uncertainty in the aggregation fractions.

We observed fairly large relative biases (over 10%) for Kenya, Nepal, Cameroon, and Bangladesh. In all cases, this is due to outlier Admin 1 areas, whose individual relative biases were sometimes in the hundreds of percent due to the direct estimates being very close to zero. In Kenya, for example, we observed relative biases in Laikipia and Tharaka-Nithi of 426% and 493%, respectively for the smoothed direct model, with direct estimates of just 8.4 and 7.2 deaths per thousand births, respectively. These were the two smallest direct estimates of all 47 Admin 1 areas in Kenya. It is therefore possible that the large relative biases of all models for Kenya, Nepal, Cameroon, and Bangladesh are due to noise in the direct estimates for a small number of Admin 1 areas.

Table 4 Validation results for both the smoothed direct model, and the beta-binomial Admin 1 and Admin 2 models, first averaged over Admin 1 areas within each country, and then averaged over all countries. Bias is reported in children per thousand births.

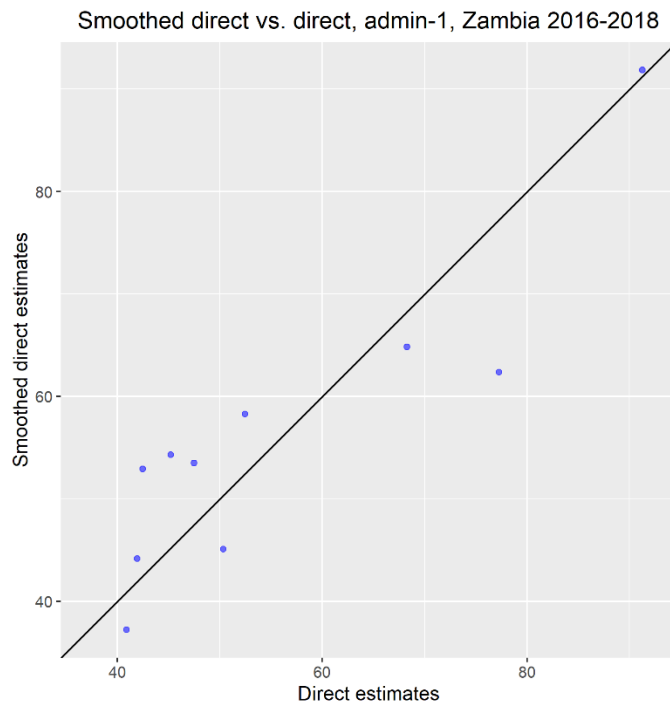
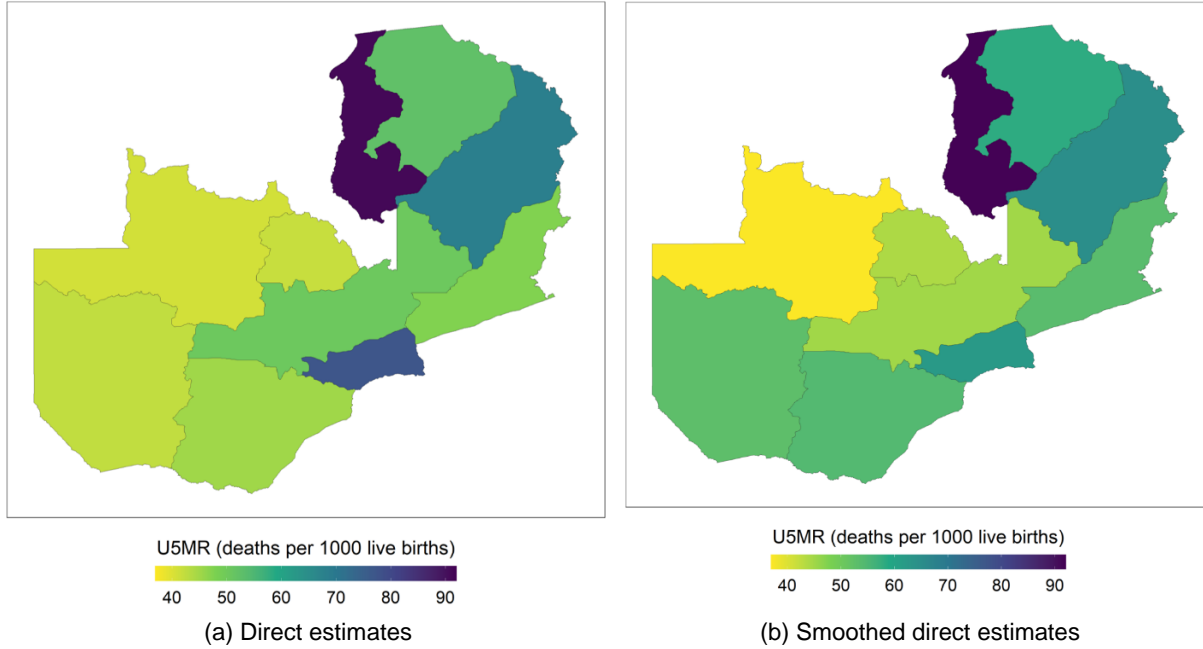
Country	Model	Relative Bias	Bias	RMSE
Bangladesh	Smoothed Direct	13.6	1.7	0.012
	Beta-Binomial Admin 1	10.8	0.8	0.012
	Beta-Binomial Admin 2	10.4	0.6	0.012
Cameroon	Smoothed Direct	14.8	2.7	0.023
	Beta-Binomial Admin 1	17.5	5.3	0.019
	Beta-Binomial Admin 2	32.6	4.7	0.042
Ethiopia	Smoothed Direct	5.1	0.9	0.014
	Beta-Binomial Admin 1	-1.9	-2.4	0.014
	Beta-Binomial Admin 2	-1.5	-2.3	0.017
Kenya	Smoothed Direct	38.5	1.0	0.022
	Beta-Binomial Admin 1	31.2	-0.9	0.021
	Beta-Binomial Admin 2	44.9	-1.0	0.029
Malawi	Smoothed Direct	2.2	0.5	0.004
	Beta-Binomial Admin 1	-0.7	-1.1	0.004
	Beta-Binomial Admin 2	-1.0	-1.1	0.003
Nepal	Smoothed Direct	15.9	2.5	0.013
	Beta-Binomial Admin 1	25.4	5.5	0.015
	Beta-Binomial Admin 2	23.7	5.0	0.014
Nigeria	Smoothed Direct	5.9	-0.9	0.032
	Beta-Binomial Admin 1	2.2	-2.0	0.020
	Beta-Binomial Admin 2	27.2	-5.4	0.077
Zambia	Smoothed Direct	7.4	2.1	0.014
	Beta-Binomial Admin 1	2.8	-0.2	0.015
	Beta-Binomial Admin 2	2.4	-1.8	0.018
Overall	Smoothed Direct	12.9	1.3	0.017
	Beta-Binomial Admin 1	10.9	0.6	0.015
	Beta-Binomial Admin 2	17.3	-0.2	0.027

6.3 Detailed Results for Zambia

The Admin 2 estimates of U5MR can have substantial uncertainties even after space-time smoothing. Here we further explore the uncertainties of the estimates by using Zambia as an example. The same set of figures for the other countries are included in the Appendices. As we noted in Section 2, direct (weighted) estimates provide a reliable summary when there are sufficient clusters in each area. Unfortunately, it will rarely be possible to obtain yearly Admin 2 direct estimates due to data sparsity. But we may often be able to produce yearly Admin 1 estimates, or Admin 2 estimates using data summed over years.

Figure 10 panel (a) shows a map of direct estimates at Admin 1 in the period 2016-2018, while panel (b) displays the smoothed direct (Fay-Herriot) estimates. We plot the estimates against each other in panel (c), and we see some smoothing, although the highest and lowest points are not attenuated.

Figure 10 Comparison of direct and smoothed direct estimates, Zambia Admin 1, 2016-2018



(c) Direct versus smoothed direct estimates

The direct and smoothed direct estimates adjust for the complex design via the weighting and the use of an appropriate variance calculation. The model-based approaches that are required for yearly Admin 2 estimation must contain appropriate terms to acknowledge the design. One check is to form Admin 1 estimates and compare them with the direct estimates. One would expect some shrinkage, but not systematic differences, either shifted higher or lower. We can also further aggregate to the national level. In Figure 11, we plot the aggregated beta-binomial estimates versus the yearly direct estimates. We aggregate by using the method described in Section 4.3. We see that the aggregated beta-binomial estimates are much smoother over time and have reduced uncertainty when compared to the direct estimates. We emphasize that we would not use Admin 2 modeling to produce Admin 1 level or national estimates. Our model is designed for Admin 2 estimates, and for Admin 1 level estimates, it is better to use an Admin 1 level model, and similarly, for a national estimate, one should just pool all the data and form a national weighted estimate. A primary reason for this advice is the uncertainty in the aggregation process.

Figure 11 Aggregated beta-binomial national estimates versus direct national estimates, over time, and with 95% error bands, Zambia

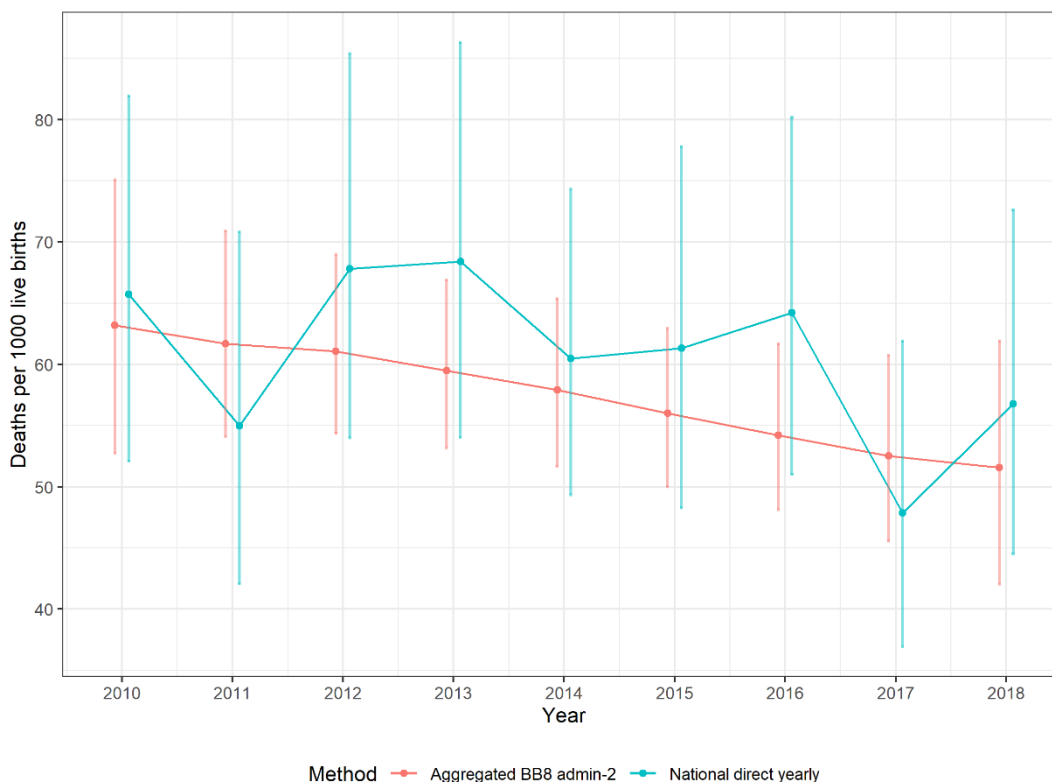
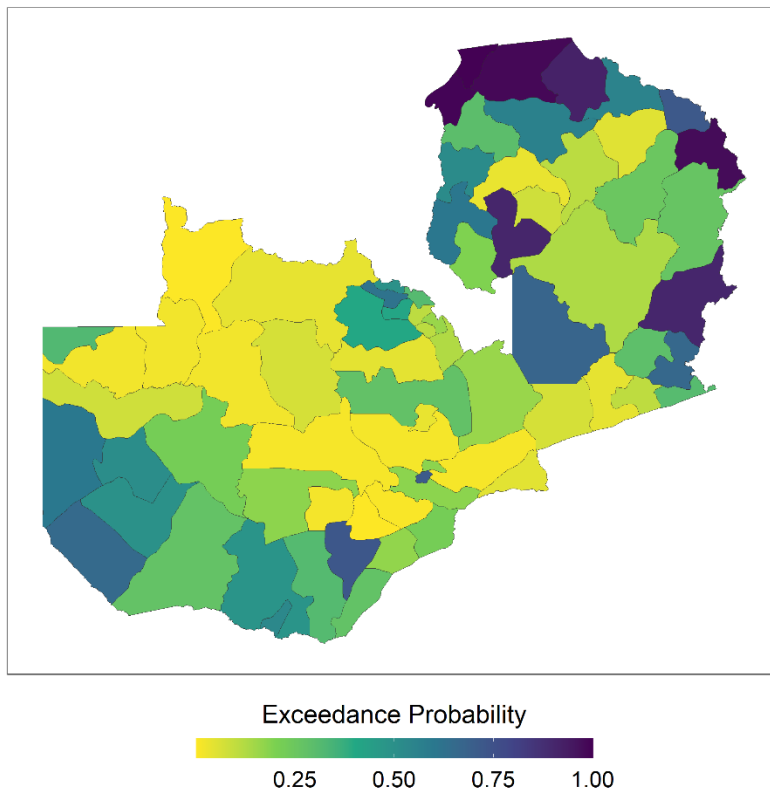


Figure 12 shows the posterior probability of the U5MR in each Admin 2 region exceeding the national U5MR at 2018. For a more stable estimate, we use the national rate during the period 2016–2018 (and its associated uncertainty), which is 56.2 per 1,000 live births. The national weighted estimate has an associated asymptotic distribution and if we use this distribution as a likelihood, and assume a flat prior on the national estimate, we can take the asymptotic distribution as an approximate posterior and then sample from this distribution. With samples from both the posterior distribution for the U5MR in the area, and samples from the posterior for the national average, it is then straightforward to calculate the exceedance probability. We are assuming that the two posterior distributions are independent, which is not quite true, although the dependence will be inconsequential, unless there are few areas that contribute to the national average. The pattern in this map is obviously similar to the left panel of Figure 9, although this map may be more useful for public health officials. For example, two of the most northerly Admin 2 regions have probabilities close to 1 of having U5MR greater than the national level.

Figure 12 Probability for Admin 2 estimates exceeding national direct estimates (Zambia 2018)



In Figure 13 we plot the urban to rural odds ratios for three age bands against year. Although there is a larger amount of uncertainty, we see stronger associations in the neonatal and > 12 months age groups, with the odds ratio becoming closer to 1 in the later years.

Figure 13 Odds ratios (urban/rural) over time for the age bands 0–1 month, 1–12 months, >12 months

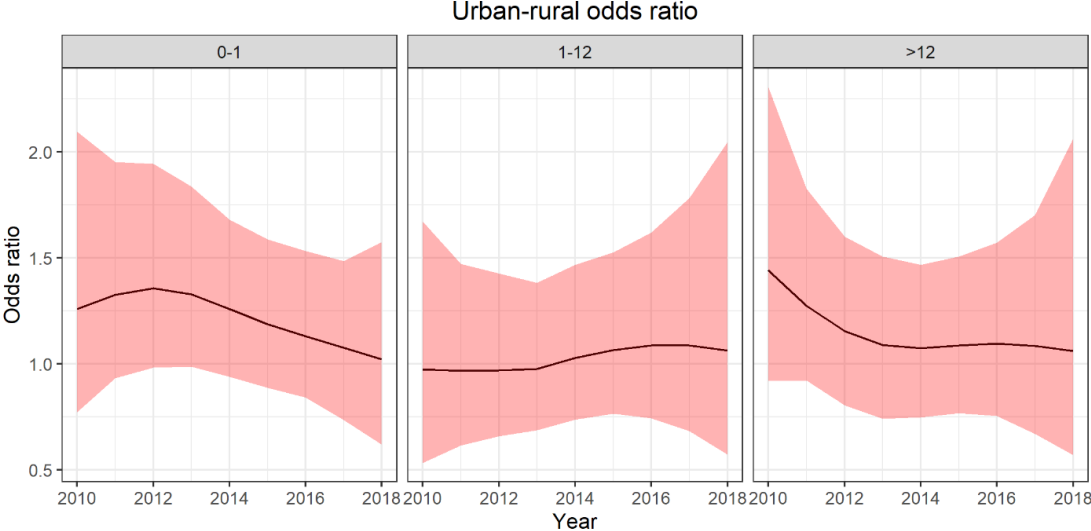


Figure 14 shows the trends across time of U5MR for Admin 1 regions in Zambia. Figure 15 shows the equivalent plot for Admin 2 regions in Zambia. The overall trend is downward, although there is strong subnational variation across areas, as was shown in Tables 2 and 3.

Figure 14 Trends of U5MR for Admin 1 areas in Zambia over time

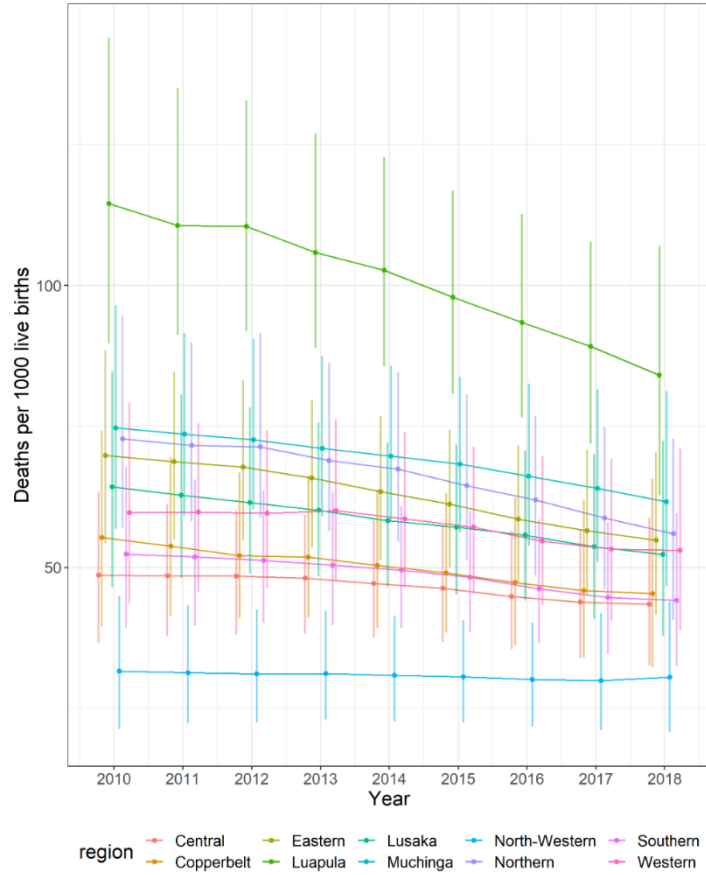
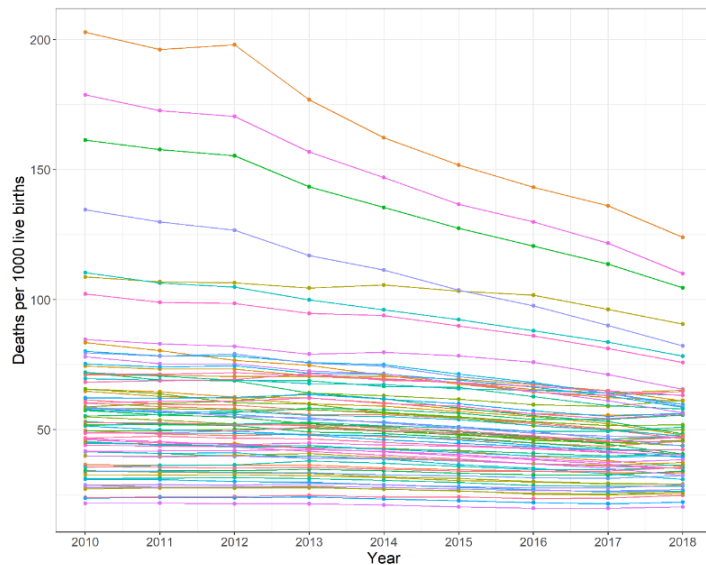


Figure 15 Trends of U5MR for Admin 2 areas in Zambia over time



Ridgeplots provide further insights into the uncertainty of estimation in each area. Showing all 72 of the Admin 2 posterior distributions in one plot would be too cluttered. As an illustration, therefore, Figure 16 shows the posterior distributions for U5MR in the Admin 2 regions that are contained in the Admin 1 region of Luapula. The large uncertainty is clear, particularly for Nchelenge and Chiengi.

Figure 16 Ridgeplot representation of posterior distribution of U5MR for Admin 2 areas in Luapula (Zambia 2018)

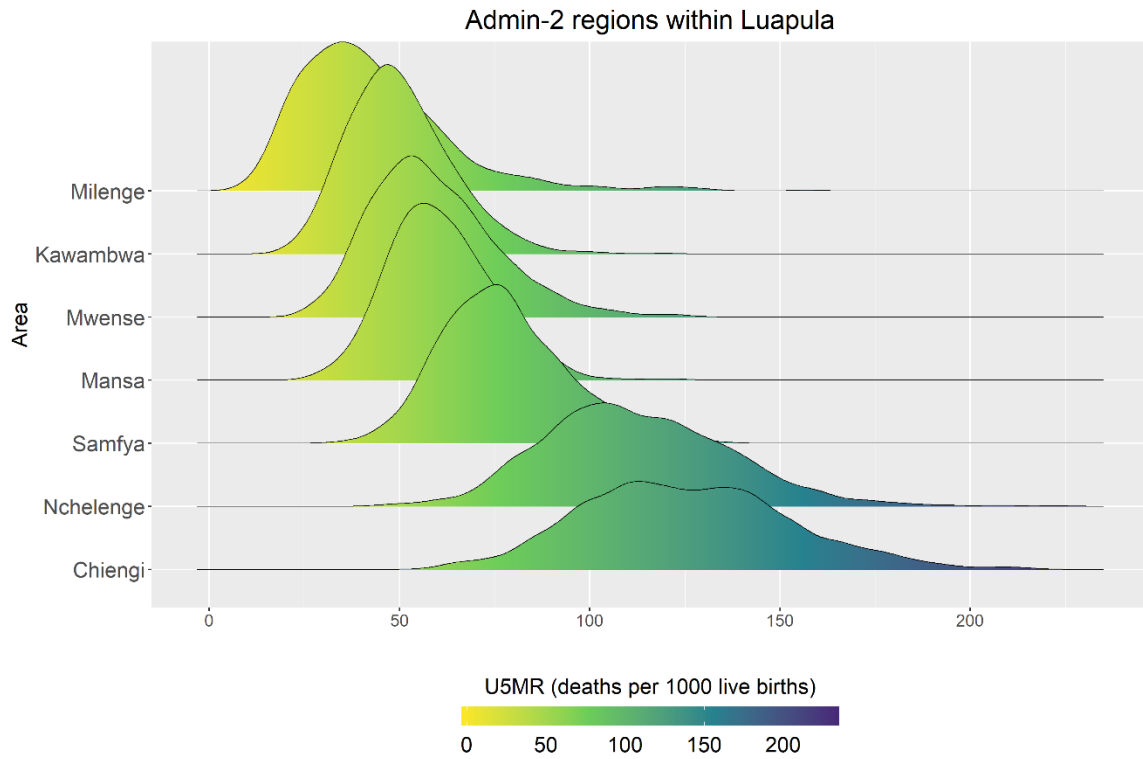
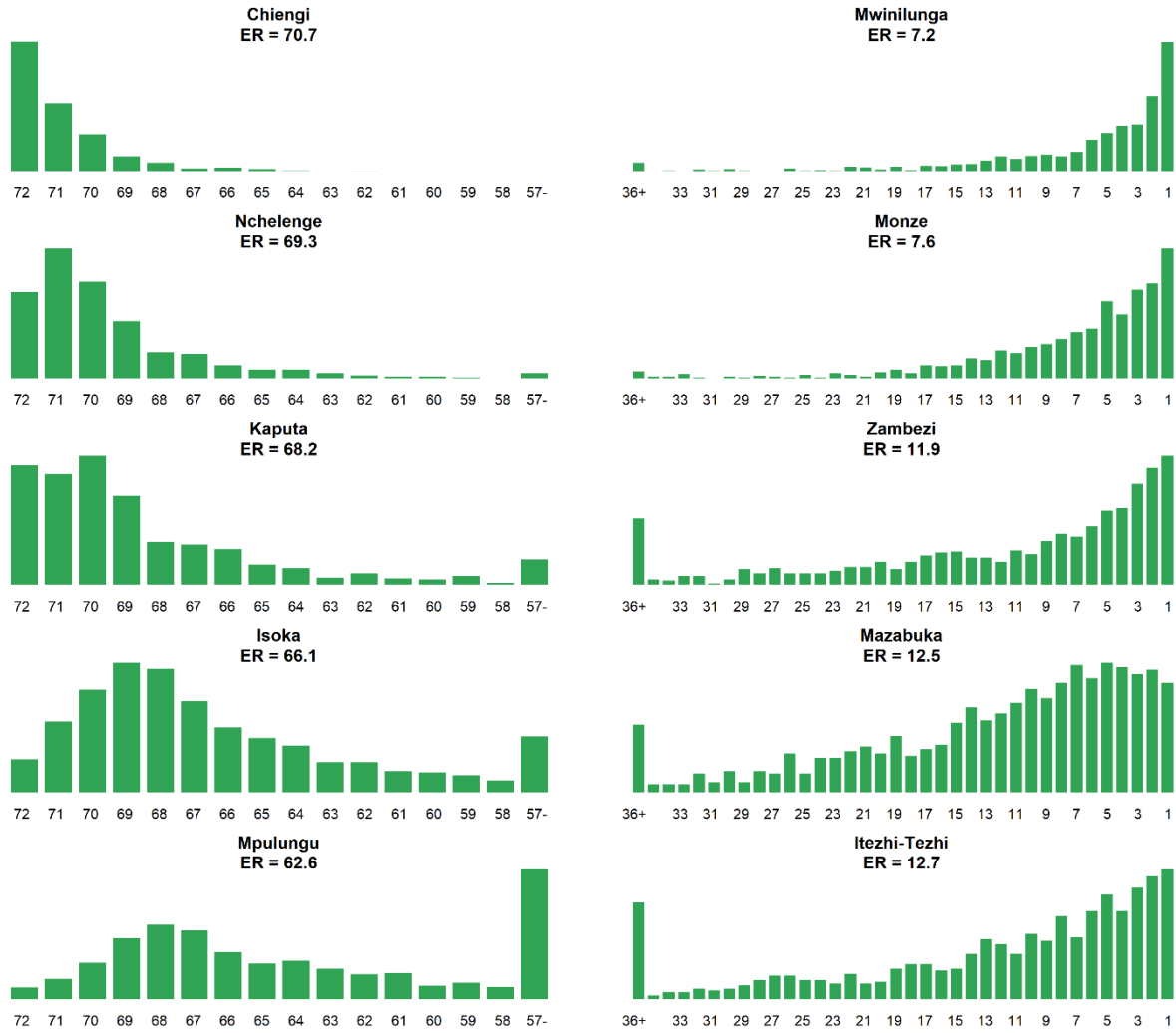


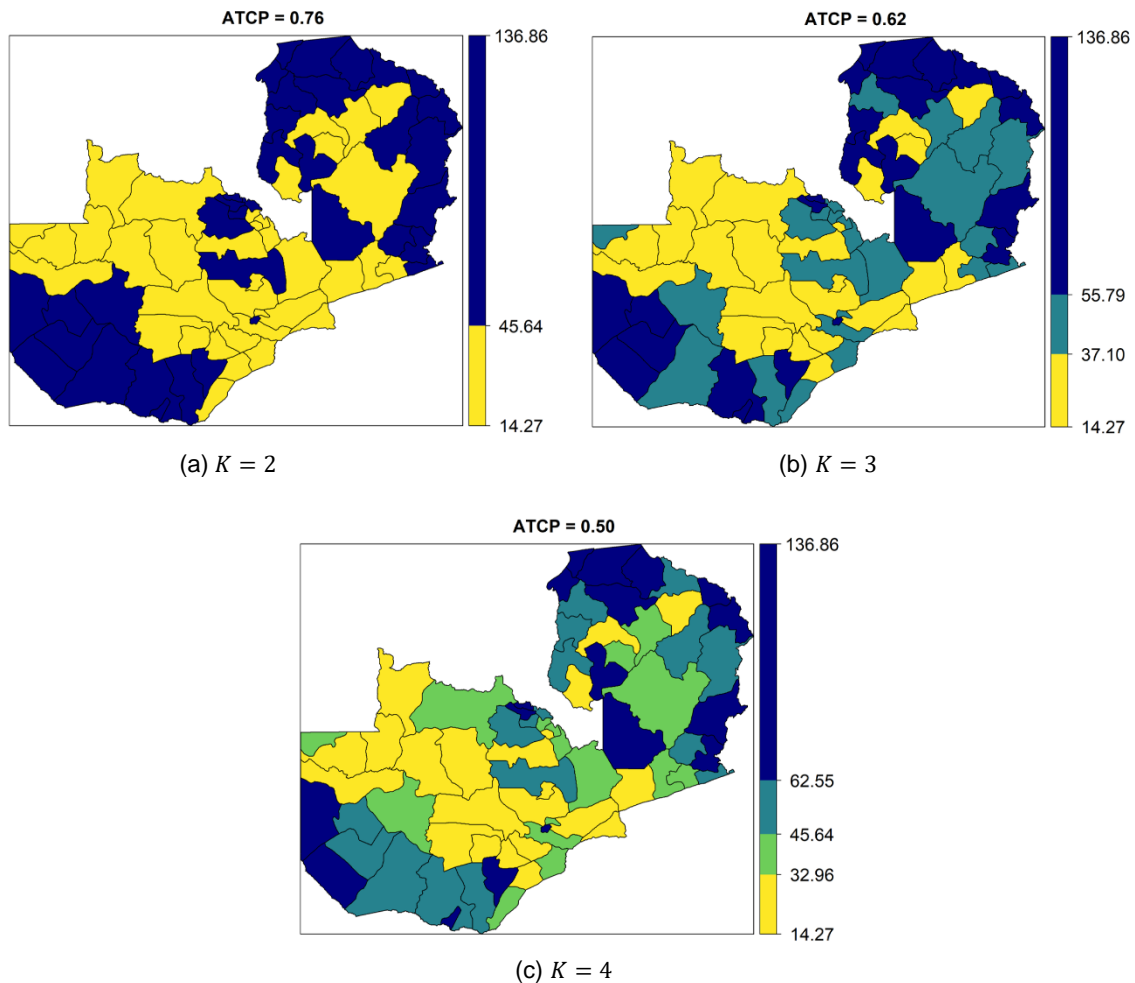
Figure 17 shows the posterior distribution of rankings of a subset of Admin 2 area in Zambia, where 1 indicates the area with the lowest U5MR among all areas, and 72 indicates the area with the highest U5MR. The uncertainty in the rankings is apparent.

Figure 17 Admin 2 U5MR ranking in Zambia, 2018. On the left we show the regions with the highest U5MR, and on the right are those with the lowest U5MR.



We end the discussion of different visual summaries by showing True Classification Probability (TCP) maps. The concept, originally described in Dong and Wakefield (2021), is to allow only a smaller number of colors on maps. Our previous maps effectively use a continuous color legend to represent the posterior summary estimate. Figure 18 (a) illustrates with $K = 2$ colors. Areas in blue have a greater than 0.5 chance that the posterior distribution is greater than 45.62, while areas in yellow have less than 0.5. For any area, there is a corresponding complementary probability of less than 0.5 of incorrect classification. Across areas, we can report the average TCP, which here is 0.76. One way of interpreting this number is to randomly pick an area, and then ask about the probability that it is correctly colored. As we increase the number of colors, this probability will decrease. We illustrate with $K = 3$ and $K = 4$ in Figures 18 (b) and (c), and have ATCPs of 0.62 and 0.50, respectively. Thus, as we increase the number of colors, the correct classification of areas decreases. The thresholds are set according to the quantiles of the posterior samples for U5MR in all Admin 2 areas in 2018. See Dong and Wakefield (2021) for full details.

Figure 18 Average True Classification Probabilities (ATCP) for (a) $K = 2$, (b) $K = 3$, and (c) $K = 4$ for Zambia in 2018



6.4 Covariate Modeling

We summarize some simple covariate modeling exercises and, specifically, we explore the use of area-level covariate information in our models. We do not fit models with covariates that carry at the cluster-level, because of the aggregation difficulties discussed in Section 4.1. The SUMMER package is equipped to use area-level covariates in the cluster-level model. Using covariate maps obtained from WorldPop, we fitted the cluster-level model to the Zambia DHS data with fixed effects for five area-level covariates: night time lights (National Oceanic and Atmospheric Administration), elevation (National Oceanic and Atmospheric Administration), vegetation (National Aeronautics and Space Administration), travel time to nearest settlement (Malaria Atlas Project), and precipitation (Climatic Research Unit gridded Time Series). We entered the covariates at both Admin 1 and Admin 2, but we only discuss the latter here. Results for the Admin 1 level covariate model are discussed in the Appendix.

Table 5 summarizes the posterior distributions on the fixed effects (β in equation (18)). The 95% credible interval for vegetation does not include zero, which indicates an association between U5MR risk and this covariate, and higher vegetation levels are associated with lower risk. Elevation shows some association with higher elevation being associated with higher risk.

Table 5 Fixed effects posterior summaries for cluster-level model with Admin 2 level covariates

Covariates	2.5 th Percentile	Posterior Median	97.5 th Percentile
Travel Time	-0.081	0.112	0.304
Elevation	-0.015	0.162	0.347
Vegetation	-0.463	-0.279	-0.096
Night Time Lights	-0.194	-0.021	0.154
Precipitation	-0.125	-0.021	0.081

Figure 19 shows the yearly small area estimates from both the base model, with no covariates, and the covariate model, along with 95% interval estimates. We see virtually no difference between the point or interval estimates, although the covariate model estimates are somewhat less smooth.

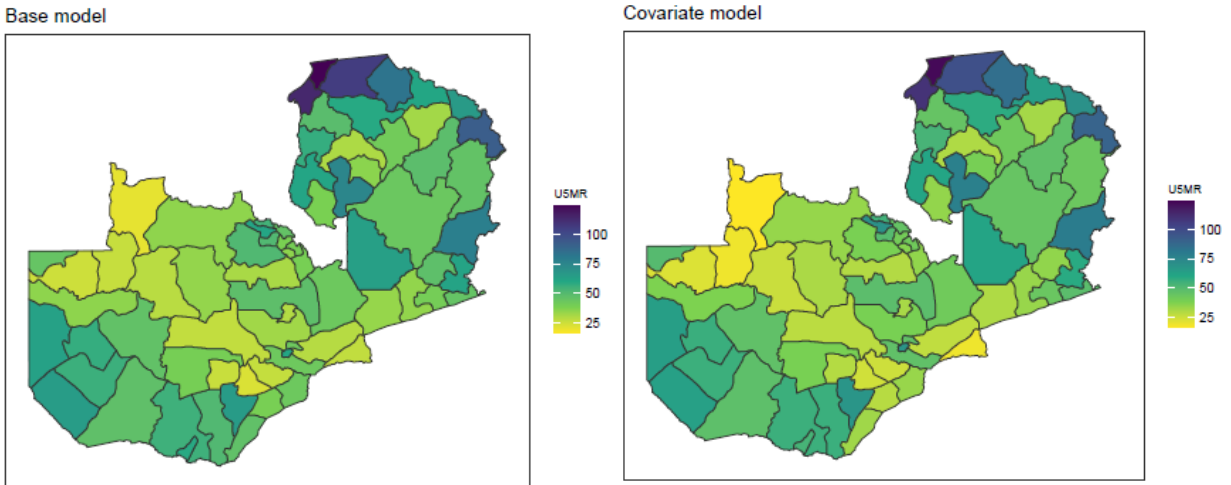
In conclusion, the area-level estimates at both the Admin 1 and Admin 2 levels were not substantially different from those obtained via the base (no covariate) model. This suggests that the area-level covariates are not able to capture much of the between-area variation in risk. Although our initial analysis found no improvement from including area-level covariate information in our cluster-level model, cluster-level covariates could conceivably provide more useful information. More work is needed to develop effective methods for incorporating covariate information for estimation of child mortality rates at the subnational level. In particular, the aggregation aspect, which is far more difficult for cluster-level covariates, requires further examination.

Figure 19 Comparison of yearly Admin 2 U5MR estimates for the beta-binomial cluster level model with Admin 2 level covariates and base (no covariate) model

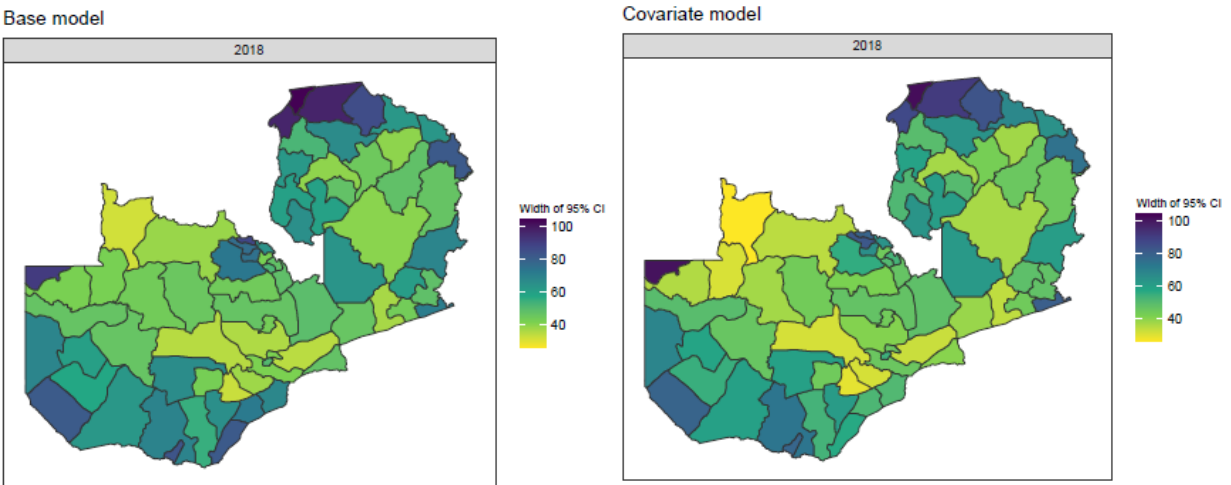


Figure 20 confirms the lack of impact of the covariate modeling, and shows maps of point estimates and uncertainty estimates under the base and covariate models. It is difficult to see any differences in either map.

Figure 20 Comparison of 2018 admin 2 U5MR (a) point and (b) uncertainty estimates for Zambia using cluster-level model with Admin 2 level covariates and base model



(a) Point estimates



(b) Width of the 95% interval estimates

7 DISCUSSION

In this report, we have described models for conducting small area estimation of U5MR across Admin 2 regions. We recommend including discrete spatial random effects at the inferential level, Admin 2, and exploring the use of an urban/rural fixed effect to acknowledge stratification in the design.

The models we propose are smoothing models in time and space, and apply shrinkage. Such models improve inference for the complete collection of areas, but introduce bias in order to reduce variance. This is beneficial overall, although it may produce poorer estimates than the direct estimates in some areas by over-shrinking. While this shrinkage decreases with increasing sample size, it may take attention away from highs in particular areas that do not have abundant data.

Area-level models, such as those proposed by Fay and Herriot (1979), do not require an aggregation step, since they use direct modeling at the level of inference. Unfortunately, it is generally not feasible to use such models at the Admin 2 level because of data sparsity. Cluster-level models, such as the one we recommend, can overcome this difficulty. Aggregation is required to move from the cluster to the area level, and this requires information on the under-5 population at risk, perhaps by urban/rural status. This step introduces an extra level of uncertainty. If the population estimates are inaccurate, inaccurate U5MR estimates may result. The Fay-Herriot models avoid this, since they directly model the weighted estimates and population information is encoded in these weights. Such an approach also accounts for other aspects of the design and implementation such as the PPS sampling and the nonresponse adjustment, neither of which are accounted for in the cluster-level models. These shortcomings mean that it is important to compare Admin 1 aggregated estimates with direct estimates. We would expect to see shrinkage, which manifests itself in a reduced range of the cluster-level estimates. Estimates that overall tend to be lower or higher may indicate that the modeled estimates are biased. We are working on methods for benchmarking (adjusting estimates to known totals), which may be useful at the national level if reliable estimates are available.

If the number of Admin 2 areas is large relative to the number of clusters, then smoothing methods are likely to fail because of the paucity of information. In LMICs, it is difficult to find accurate covariates at the cluster level, the level at which covariates will be the most beneficial. Aggregation with cluster-level covariates is also difficult.

Finally, the models we have described will not correct for systematic problems with the raw survey data.

REFERENCES

Banerjee, S., B. P. Carlin, and A. E. Gelfand. 2015. *Hierarchical Modeling and Analysis for Spatial Data*. Second Edition. New York, USA: Chapman and Hall/CRC Press.
<https://doi.org/10.1201/b17115>.

Battese, G. E., R. M. Harter, and W. A. Fuller. 1988. “An Error-Components Model for Prediction of County Crop Areas Using Survey and Satellite Data.” *Journal of the American Statistical Association* 83 (401): 28–36.
<https://doi.org/10.2307/2288915>.

Besag, J., J. York, and A. Mollié. 1991. “Bayesian Image Restoration, with Two Applications in Spatial Statistics.” *Annals of the Institute of Statistics and Mathematics* 43 (1): 1–59.
<https://doi.org/10.1007/BF00116466>.

Binder, D. 1983. “On the Variances of Asymptotically Normal Estimators from Complex Surveys.” *International Statistical Review* 51 (3): 279–292.
<https://doi.org/10.2307/1402588>.

Blangiardo, M. and M. Cameletti. 2015. *Spatial and Spatio-Temporal Bayesian Models with R-INLA*. Chichester, UK: John Wiley and Sons, Ltd.
<https://www.wiley.com/en-us/Spatial+and+Spatio+temporal+Bayesian+Models+with+R+INLA-p-9781118326558>.

Burstein, R., N. J. Henry, M. L. Collison, L. B. Marczak, A. Sligar, S. Watson, N. Marquez, M. Abbasalizad-Farhangi, M. Abbasi, F. Abd-Allah, et al. 2019. “Mapping 123 Million Neonatal, Infant and Child Deaths Between 2000 and 2017.” *Nature* 574 (7778): 353–358.
<https://doi.org/10.1038/s41586-019-1545-0>.

Diggle, P. J. and E. Giorgi. 2019. *Model-based Geostatistics for Global Public Health: Methods and Applications*. Boca Raton, FL, USA: Chapman and Hall/CRC.
<https://doi.org/10.1201/9781315188492>.

Dong, T. and J. Wakefield. 2021. “Modeling and Presentation of Health and Demographic Indicators in a Low- and Middle-Income Countries Context.” *Vaccine* 39 (18): 2584–2594.
[doi:10.1016/j.vaccine.2021.03.007](https://doi.org/10.1016/j.vaccine.2021.03.007).

Dwyer-Lindgren, L., F. Kakungu, P. Hangoma, M. Ng, H. Wang, A. D. Flaxman, F. Masiye, and E. Gakidou. 2014. “Estimation of District-Level Under-5 Mortality in Zambia using Birth History Data, 1980–2010.” *Spatial and Spatio-Temporal Epidemiology* 11: 89–107.
<https://doi.org/10.1016/j.sste.2014.09.002>.

Dwyer-Lindgren, L., M. A. Cork, A. Sligar, K. M. Steuben, K. F. Wilson, N. R. Provost, B. K. Mayala, J. D. VanderHeide, M. L. Collison, J. B. Hall, et al. 2019. “Mapping HIV Prevalence in Sub-Saharan Africa between 2000 and 2017.” *Nature* 570 (7760): 189–193.
<https://doi.org/10.1038/s41586-019-1200-9>.

Fay, R. E. and R. A. Herriot. 1979. "Estimates of Income for Small Places: An Application of James–Stein Procedure to Census Data." *Journal of the American Statistical Association* 74 (366a): 269–277.

<https://doi.org/10.1080/01621459.1979.10482505>.

Gething, P. W., D. C. Casey, D. J. Weiss, D. Bisanzio, S. Bhatt, E. Cameron, K. E. Battle, U. Dalrymple, J. Rozier, P. C. Rao, et al. 2016. "Mapping Plasmodium Falciparum Mortality in Africa between 1990 and 2015." *New England Journal of Medicine* 375 (25): 2435–2445.

<https://doi.org/10.1056/NEJMoa1606701>.

Godwin, J. and J. Wakefield. 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): 1593–1638.

<https://doi.org/10.1002/sim.8854>.

Golding, N., R. Burstein, J. Longbottom, A. J. Browne, N. Fullman, A. Osgood-Zimmerman, L. Earl, S. Bhatt, E. Cameron, D. C. Casey, 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.

[https://doi.org/10.1016/S0140-6736\(17\)31758-0](https://doi.org/10.1016/S0140-6736(17)31758-0).

Gómez-Rubio, V. 2020. *Bayesian Inference with INLA*. Boca Raton, FL, USA: Chapman and Hall/CRC Press.

<https://doi.org/10.1201/9781315175584>.

Gutreuter, S., E. Igumbor, N. Wabiri, M. Desai, and L. Durand. 2019. "Improving Estimates of District HIV Prevalence and Burden in South Africa using Small Area Estimation Techniques." *PLoS One* 14 (2): e0212445.

<https://doi.org/10.1371/journal.pone.0212445>.

Hájek, J. 1971. "Comment on 'An Essay on the Logical Foundations of Survey Sampling, Part I'." by D. Basu. In *Foundations of Statistical Inference*, edited by V. Godambe and D. Spratt. Toronto, Canada: Holt, Rinehart, and Winston.

Heaton, M. J., A. Datta, A. O. Finley, R. Furrer, J. Guinness, R. Guhaniyogi, F. Gerber, R. B. Gramacy, D. Hammerling, M. Katzfuss, et al. 2018. "A Case Study Competition among Methods for Analyzing Large Spatial Data." *Journal of Agricultural, Biological, and Environmental Statistics* 24 (3): 398–425.

<https://doi.org/10.1007/s13253-018-00348-w>.

Horvitz, D. G. and D. J. Thompson. 1952. "A Generalization of Sampling Without Replacement from a Finite Universe." *Journal of the American Statistical Association* 47 (260): 663–685.

<http://dx.doi.org/10.1080/01621459.1952.10483446>.

Hosseinpour, A. R., N. Bergen, A. Schlottheuber, and J. Grove. 2015. "Measuring Health Inequalities in the Context of Sustainable Development Goals." *Bulletin of the World Health Organization* 96 (9): 654–659.

<http://dx.doi.org/10.2471/BLT.18.210401>.

- Knorr-Held, L. 2000. “Bayesian Modelling of Inseparable Space-Time Variation in Disease Risk.” *Statistics in Medicine* 19 (17-19): 2555–2567.
- Krainski, E. T., V. Gómez-Rubio, H. Bakka, A. Lenzi, D. Castro-Camilo, D. Simpson, F. Lindgren, F., and H. Rue. 2018. *Advanced Spatial Modeling with Stochastic Partial Differential Equations Using R. and INLA*. New York, USA: Chapman and Hall/CRC Press.
- Li, Z. R., B. D. Martin, T. Dong, G.-A. Fuglstad, J. Godwin, J. Paige, A. Riebler, S. Clark, and J. Wakefield. 2020. “Space-time Smoothing of Demographic and Health Indicators Using the R Package SUMMER.”
<https://arxiv.org/pdf/2007.05117.pdf>.
- Li, Z. R., Y. Hsiao, J. Godwin, B. D. Martin, J. Wakefield, and S. J. Clark. 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): e0210645.
<https://doi.org/10.1371/journal.pone.0210645>.
- Li, Z. R., B. D. Martin, Y. Hsiao, J. Godwin, J. Paige, J. Wakefield, S. J. Clark, G.-A. Fuglstad, and A. Riebler. 2021. SUMMER: Small-Area-Estimation Unit/Area Models and Methods for Estimation in R. R package version 1.2.1.
<https://www.r-pkg.org/pkg/SUMMER>.
- Mayala, B., T. Dontamsetti, T. D. Fish, and T. N. Croft. 2019. *Interpolation of DHS Survey Data at Subnational Administrative Level 2*. DHS Spatial Analysis Reports No. 17. Rockville, Maryland, USA.
<https://dhsprogram.com/pubs/pdf/SAR17/SAR17.pdf>.
- Mercer, L., J. Wakefield, A. Pantazis, A. Lutambi, H. Mosanja, and S.J. Clark. 2015. “Small Area Estimation of Childhood Mortality in the Absence of Vital Registration.” *Annals of Applied Statistics* 9 (4): 1889–1905.
<http://www.jstor.org/stable/43826448>.
- Moraga, P. 2019. *Geospatial Health Data: Modeling and Visualization with R-INLA and Shiny*. Boca Raton, FL, USA: CRC Press.
<https://doi.org/10.1201/9780429341823>.
- Mosser, J. F., W. Gagne-Maynard, P. C. Rao, A. Osgood-Zimmerman, N. Fullman, N. Graetz, R. Burstein, R. L. Updike, P. Y. Liu, S. E. Ray, et al. 2019. “Mapping Diphtheria-Pertussis-Tetanus Vaccine Coverage in Africa, 2000–2016: A Spatial and Temporal Modelling Study.” *The Lancet* 393 (10183): 1843–1855.
[https://doi.org/10.1016/S0140-6736\(19\)30226-0](https://doi.org/10.1016/S0140-6736(19)30226-0).
- Osgood-Zimmerman, A. and J. Wakefield. 2020. Template Model Builder (TMB), a Flexible Alternative to the Integrated Nested Laplace Approximation. Pending publication.
<https://arxiv.org/pdf/2103.09929.pdf>.
- Paige, J., G.-A. Fuglstad, A. Riebler, and J. Wakefield. 2020. “Model-Based Approaches to Analysing Spatial Data from Complex Surveys.” *Journal of Survey Statistics and Methodology*.

- Pedersen, J. and J. Liu. 2012. “Child Mortality Estimation: Appropriate Time Periods for Child Mortality Estimates from Full Birth Histories.” *PLoS Medicine* 9 (8): e1001289.
<https://doi.org/10.1371/journal.pmed.1001289>.
- Rao, J. N. K. and I. Molina. 2015. *Small Area Estimation, Second Edition*. New York, USA: John Wiley.
<https://www.wiley.com/en-us/Small+Area+Estimation%2C+2nd+Edition-p-9781118735787>.
- Riebler, A., S. H. Sørbye, D. Simpson, and H. Rue. 2016. “An Intuitive Bayesian Spatial Model for Disease Mapping that Accounts for Scaling.” *Statistical Methods in Medical Research* 25 (4): 1145–1165.
<https://doi.org/10.1177%2F0962280216660421>.
- Schlüter, B-S. and B. Masquelier. 2021. “Space-Time Smoothing of Mortality Estimates in Children Aged 5-14 in Sub-Saharan Africa.” *PLoS One* 16 (1): e0245596.
<https://doi.org/10.1371/journal.pone.0245596>.
- Simpson, D., H. Rue, A. Riebler, T. Martins, and S. Sørbye. 2017. “Penalising Model Component Complexity: A Principled, Practical Approach to Constructing Priors.” *Statistical Science* 32 (1): 1–28.
<https://doi.org/10.1214/16-STS576>.
- Skinner, C. and J. Wakefield. 2017. “Introduction to the Design and Analysis of Complex Survey Data.” *Statistical Science* 32 (2): 165–175.
<https://doi.org/10.1214/17-STS614>.
- Stein, M. L. 1999. *Interpolation of Spatial Data: Some Theory for Kriging*. New York, USA: SpringerVerlag.
<https://doi.org/10.1007/978-1-4612-1494-6>.
- Stevens, F. R., A. E. Gaughan, C. Linard, and A. J. Tatem. 2015. “Disaggregating Census Data for Population Mapping using Random Forests with Remotely-Sensed and Ancillary Data.” *PloS One* 10 (2): e0107042.
<https://doi.org/10.1371/journal.pone.0107042>.
- United Nations Inter-Agency Group for Child Mortality Estimation (UN IGME). 2021. *Subnational Under-five Mortality Estimates, 1990–2019: Estimates Developed by the United Nations Inter-agency Group for Child Mortality Estimation*. New York, USA: United Nations Children’s Fund.
<https://childmortality.org/>.
- United Nations General Assembly. 2015. *Transforming our World: The 2030 Agenda for Sustainable Development*. New York, USA: United Nations Population Fund.
<https://sdgs.un.org/sites/default/files/publications/21252030%20Agenda%20for%20Sustainable%20Development%20web.pdf>.
- Utazi, C. E., J. Thorley, V. A. Alegana, M. J. Ferrari, S. Takahashi, C. J. E. Metcalf, J. Lessler, and A. J. Tatem. 2018. “High Resolution Age-Structured Mapping of Childhood Vaccination Coverage in Low and Middle Income Countries.” *Vaccine* 36 (12): 1583–1591.
<https://doi.org/10.1016/j.vaccine.2018.02.020>.

Utazi, C. E., K. Nilsen, O. Pannell, W. Dotse-Gborgbortsi, and A. J. Tatem. 2021. "District-Level Estimation of Vaccination Coverage: Discrete vs. Continuous Spatial Models." *Statistics in Medicine* 40 (9): 2197-2211.
<https://doi.org/10.1002/sim.8897>.

Wakefield, J. 2008. "Ecologic Studies Revisited." *Annual Review of Public Health* 29: 75–90.
<https://doi.org/10.1146/annurev.publhealth.29.020907.090821>.

Wakefield, J. 2020. "Prevalence Mapping." In *Wiley StatsRef: Statistics Reference Online*, 1–7, Chichester, UK: John Wiley & Sons, Ltd.
<https://doi.org/10.1002/9781118445112.stat08238>.

Wakefield, J., G.-A. Fuglstad, A. Riebler, J. Godwin, K. Wilson, and S. Clark. 2019. "Estimating Under Five Mortality in Space and Time in a Developing World Context." *Statistical Methods in Medical Research* 28 (9): 2614–2634.
<https://doi.org/10.1177%2F0962280218767988>.

Wakefield, J., T. Okonek, and J. Pedersen. 2020. "Small Area Estimation for Disease Prevalence Mapping." *International Statistical Review* 88 (2): 398–418.
<https://doi.org/10.1111/insr.12400>.

Wickremasinghe, D., I. E. Hashmi, J. Schellenberg, and B. I. Avan. 2016. "District Decision-Making for Health in Low-Income Settings: A Systematic Literature Review." *Health Policy and Planning* 31 (Supplement 2): ii12–ii24.
<https://doi.org/10.1093/heapol/czv124>.

Wolter, K. 2007. *Introduction to Variance Estimation, Second Edition*. New York, USA: Springer-Verlag.
<https://link.springer.com/content/pdf/10.1007%2F978-0-387-35099-8.pdf>.

Yaya, S., O. A. Uthman, F. Okonofua, and G. Bishwajit. 2019. "Decomposing the Rural-Urban Gap in the Factors of Under-Five Mortality in Sub-Saharan Africa? Evidence from 35 Countries." *BMC Public Health* 19 (1): 1-10.
<https://doi.org/10.1186/s12889-019-6940-9>.

You, Y. and J. Rao. 2002. "Small Area Estimation Using Unmatched Sampling and Linking Models." *Canadian Journal of Statistics* 30 (1): 3–15.
<https://doi.org/10.2307/3315862>.