Methods for Subnational Estimation of Child Mortality

Jon Wakefield[1,2] and Katie Wilson[1]

[1] Department of Biostatistics, University of Washington
[2] Department of Statistics, University of Washington

Regional Workshop on Child Mortality and Stillbirth Estimation,
Johannesburg, South Africa

April 23–26, 2019

Subnational Estimation

Sustainable Development Goal 3.2: “By 2030, end preventable deaths of newborns and children under 5 years of age, with all countries aiming to reduce ... under-5 mortality to at least as low as 25 per 1,000 live births”.

Additionally:

- Paragraph 74.g, with reference to review processes: “They will be rigorous and based on evidence, informed by country-led evaluations and data which is high-quality, accessible, timely, reliable and disaggregated by income, sex, age, race, ethnicity, migration status, disability and geographic location and other characteristics relevant in national contexts.”
- Paragraph 17.18, under data, monitoring and accountability: “By 2020, enhance capacity-building support to developing countries, including for least developed countries and small island developing States, to increase significantly the availability of high-quality, timely and reliable data disaggregated by income, gender, age, race, ethnicity, migratory status, disability, geographic location and other characteristics relevant in national contexts.”

Outline

Logistics

Website:

http://faculty.washington.edu/jonno/UNICEF-WORKSHOPS.html

Content:

- These notes.
- Datasets and R code.
- Additional materials, including a link to a course on small-area estimation (SAE) that JW taught with Richard Zehang Li.

This website will stay live, and feel free to share with colleagues.

Subnational Estimation

Data and Methodology:

- With a civil registration system, one can obtain accurate estimates of child mortality directly.
- Without such a system, one must combine available data, from surveys and censuses, for example, to produce the best possible estimates (with uncertainty).
- To obtain estimates at a useful geographical and temporal scale, smoothing across space and time is beneficial.
- I will focus on U5MR estimation based on full birth history (FBH) data, in which birth and death times are known for each child of interviewed mothers.
- As an example, we will use data from the 2014 Kenya DHS (KDHS).

2014 Kenya DHS

Study Design:

- In the 2014 KDHS, the stratification was county (47) and urban/rural (2).
- Nairobi and Mombasa are entirely urban, so there are 92 strata in total.
- In each strata, Enumeration Areas (EAs) are selected with probability proportional to size using a sampling frame developed from the 2009 Census. In each of these clusters, households are ultimately selected. Within each household, women between the ages of 15 and 49 are interviewed.

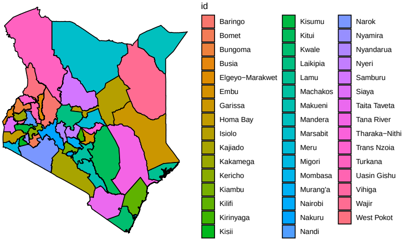


Figure 1: Counties of Kenya.

- An example of a stratified cluster design.
- We have data from a total of 1584 EAs across the 92 strata. In the second stage, 40,300 households are sampled.

Motivation for Smoothing: Temporal Case

- When looking at estimates over space or time, we want to know if the differences we see are "real", or simply reflecting sampling variability.
- Temporal setting: Even if the underlying prevalence is the same over time, we will see estimates in the empirical estimates.
- Figure 2 demonstrates: We sampled binomial data with $n = 10, 20, 200$ and $p = 0.2$ (shown in blue) in all cases.
- In the top plot in particular, we might conclude large temporal variation, but all we are seeing is sampling variation.
- Figure 3 summarizes estimates from a second simulation in which there is a real temporal pattern – here we would not want to oversmooth and remove the trend.
- Later we will apply temporal smoothing models to these two sets of data.
- The same principles apply to spatial data, it's just more difficult to gain insight, because two dimensions are harder than one!

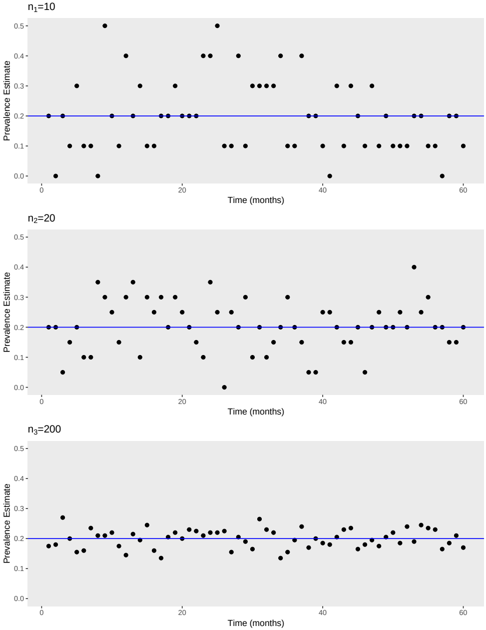


Figure 2: Prevalence estimates over time from simulated data with true prevalence of $p = 0.2$ (blue solid lines).

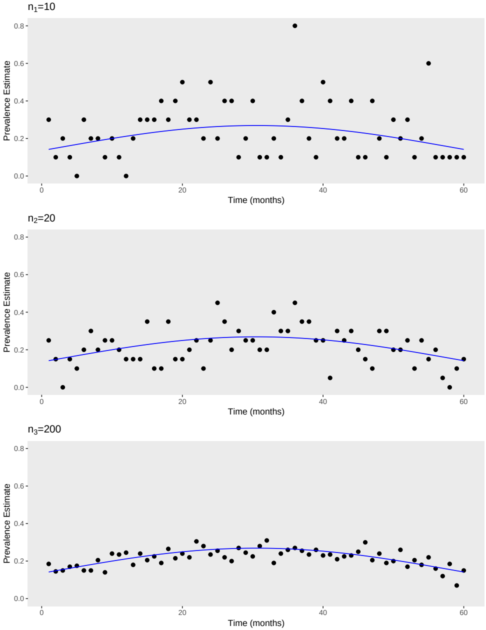


Figure 3: Prevalence estimates over time from simulated data, true prevalence corresponds to curved blue solid line.

Motivation for Smoothing: Spatial Case

- We repeat the previous simulation example, but now for spatial data.
- Counts Y_i are simulated for each area i from a binomial distribution with prevalence p_i and sample size n_i:

$$Y_i \mid p_i \sim \text{Binomial}(n_i, p_i).$$

- We look varying sample sizes $n_i = 50, 100, 500, 1000$, so that the influence of sampling variability can be examined.
- We examine two sets of simulated data:
  - Figure 4: Constant prevalence.
  - Figure 5: Spatially varying prevalence.

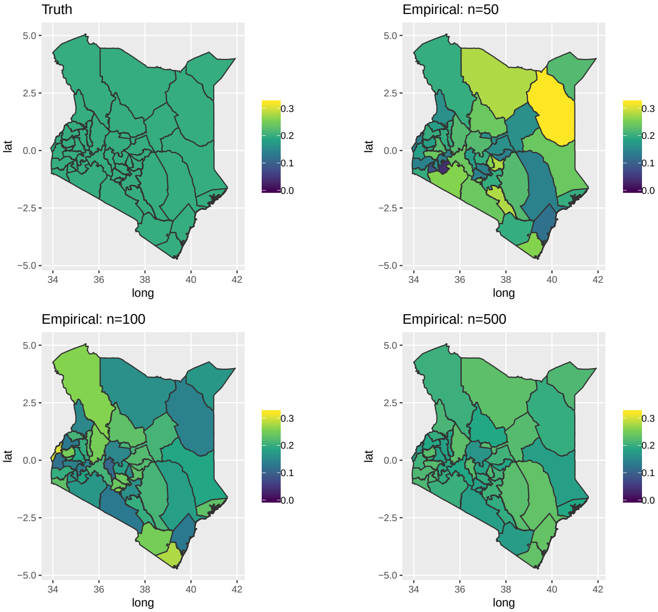


Figure 4: Prevalence estimates over space for simulated data with sample sizes of $n = 50, 100, 500$. True prevalence is 0.2 in all areas.

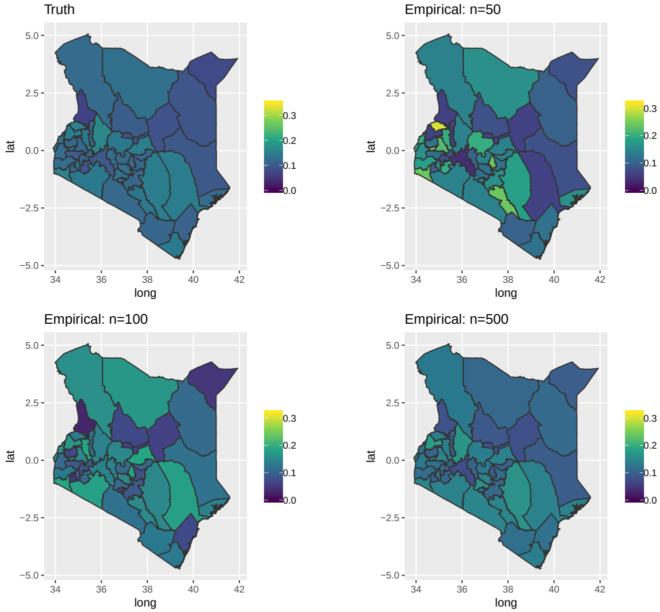


Figure 5: Prevalence estimates over space for simulated data with sample sizes of $n = 50, 100, 500$. True prevalence is spatially varying.

Complex Survey Data

Stratified Cluster Sampling

Most national surveys have a **stratified cluster sampling design** in which:

- The country is partitioned into a set of strata (e.g., province by urban/rural).
- Within each strata, clusters are sampled.
- Within each strata, households are sampled.
- Within each household, individuals are selected for interview.

The responses of the individuals provide the sample data with which we try to infer population characteristics.

When inference on the sample is performed, the design must be acknowledged:

- Ignoring the **stratified sampling** gives an estimate susceptible to **bias**, and an incorrect **variance** estimate.
- Ignoring the **cluster sampling** gives an incorrect **variance** estimate.

Weighted Estimation

Suppose we wish to estimate the **prevalence** P of some condition in an area, e.g., smoking, attaining an educational level, dying within the first month of life.

Let $y_1, \dots, y_n$ be 0/1 variables which indicate absence/presence of the condition of interest, with w_k the accompanying **design weight**.

The design weight is the reciprocal of the probability of being sampled, i.e.,

$$w_k = \frac{1}{\pi_k},$$

where π_k is the sampling probability and depends on the strata of the person.

Weighted Estimation

The weight w_k can be thought of as the number of people in the population represented by sampled person k.

Example 1: Simple Random Sampling

Suppose an area contains 1000 people:

- Using simple random sampling (SRS), 100 people are sampled.
- Sampled individuals have weight $w_k = 1/\pi_k = 1000/100 = 10$.

Example 2: Stratified Simple Random Sampling

Suppose an area contains 1000 people, 200 urban and 800 rural.

- Using stratified SRS, 50 urban and 50 rural individuals are sampled.
- Urban sampled individuals have weight $w_k = 1/\pi_k = 200/50 = 4$.
- Rural sampled individuals have weight $w_k = 1/\pi_k = 800/50 = 16$.

Weighted Estimation

To account for the design we use a **weighted estimate** of the prevalence:

$$\widehat{P} = \frac{\sum_k w_k y_k}{\sum_k w_k} = \frac{\text{Estimate of Total with Condition}}{\text{Population Size}}$$

A variance estimate V can be obtained, which takes into account the design.

A **95% uncertainty interval** for the prevalence is:

$$\widehat{P} \pm 1.96 \times \sqrt{V}$$

For small samples sizes, this interval will be wide.

Example 2 Revidited: Stratified Simple Random Sampling

Suppose an area contains 1000 people, 200 urban and 800 rural.

- Urban risk = 0.1.
- Rural risk = 0.2.
- **True risk = 0.18.**

Take a stratified SRS, 50 urban and 50 rural individuals sampled:

- Urban sampled individuals have weight 4; 5 cases out of 50.
- Rural sampled individuals have weight 16; 10 cases out of 50.
- **Simple mean is** $15/100 = 0.15 \neq 0.18$**.**
- **Weighted mean is**

$$\frac{4 \times 5 + 16 \times 10}{4 \times 50 + 16 \times 50} = \frac{180}{1000} = 0.18.$$

ENSANUT-ECU Yearly Estimates of U5MR

- Figure 6 displays a yearly time series of U5MR weighted estimates with 95% confidence intervals, also shown, for comparison, are IGME and IHME estimates, each of which draw on more data than the KDHS alone.
- IGME estimates obtained using the B3 model.

We would like to sift the **signal** (true differences) from the **noise** (sampling variation) — **hierarchical models** are suited to this purpose:

- **Stage One:** Sampling Model for the Data.
- **Stage Two:** Smoothing Model for the Parameters of Stage One.
- **Stage Three:** Prior Model for the Parameters of Stage Two.

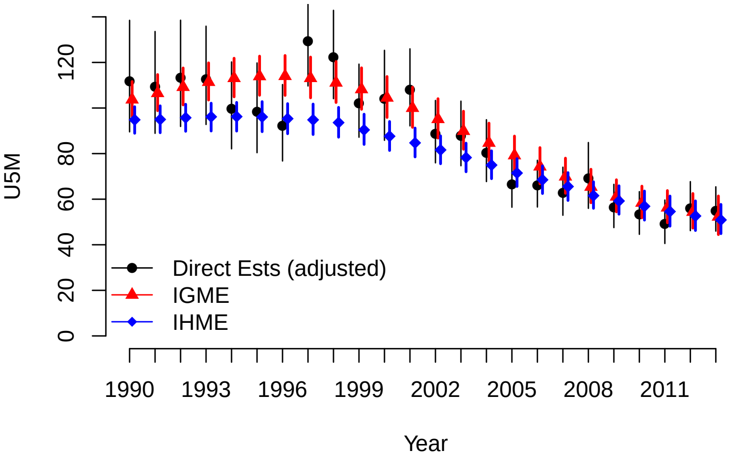


Figure 6: Yearly weighted estimates of under-5 mortality in Kenya, along with IGME and IHME estimates; with 95% uncertainty intervals for each.

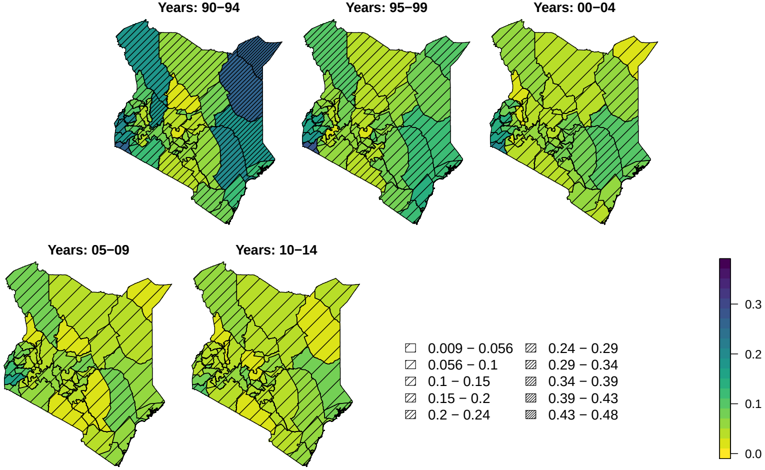


Figure 7: Yearly weighted estimates (from a discrete survival model) of under-5 mortality in Kenya, with uncertainty indicated by density of hatching; more hatching → more uncertainty, with the latter measured though width of 95% uncertainty interval.

Space-Time Smoothing

Smoothing

When faced with n different estimates of the **mortality rate** (say over time), there are two model choices:

- The true underlying mortality rates are **ALL THE SAME**.
- The true underlying mortality rates are **SIMILAR IN SOME SENSE**.

The latter seems more reasonable, but how do we model "similarity"?

There are a number of possibilities:

- The mortality rates are drawn from some **COMMON** probability distribution, but are not ordered in any way. We refer this as the independent and identically distributed, or **IID** model. We could think of this as saying we think the rates are likely to be of the same order of magnitude.
- The mortality rates are **CORRELATED** over time. We refer to this as the **SMOOTHING** model.

These are both examples of **HIERARCHICAL** or **RANDOM EFFECTS MODELS** — a key element is estimating the **SMOOTHING PARAMETER**.

Smoothing over Time

Rationale and overview of models for temporal smoothing:

- We often expect that the true underlying mortality in an area will exhibit some degree of smoothness over time.
- A linear trend in time is unlikely to be suitable for more than a small number of years, and higher degree polynomials can produce erratic fits.
- Hence, local smoothing is preferred.
- Splines (as used in B3) and random walk models have proved successful as local smoothers.
- And to emphasize again, in either approach, the choice of smoothing parameter is crucial.

Random Walk Models

We use random walk models which encourage the mean responses (e.g., prevalences) across time to not deviate too greatly from their neighbors.

The true underlying mean of the mortality at time t is modeled as a function of its neighbors:

$$\mu_t \mid \mu_{\text{NE}(t)} \sim \text{Normal}(m_t, v_t),$$

where

- μ_t is the mean mortality (or some function of it such as the logit) at time t.
- $\mu_{\text{NE}(t)}$ is the set of neighboring means – with the number of neighbors chosen depending on the model used – typically 2 or 4.
- m_t is the mean of some set of neighbors – for a **first order random walk** or **RW1** it is simply $\frac{1}{2}(\mu_{t-1} + \mu_{t+1})$.
- v_t is the variance, and depends on the number of neighbors – for the RW1 model it is $\sigma^2/2$, where σ^2 is a smoothing parameter – small values give large smoothing.

The smoothing parameter σ^2 is estimated from the data, and determines the extent deviations from the mean are penalized.

The penalty term for the RW1 model is:

$$p(\mu_t \mid \mu_{t-1}, \mu_{t+1}, \sigma^2) \propto \exp\left\{-\frac{1}{2\sigma^2}\left[\mu_t - \tfrac{1}{2}\left(\mu_{t-1} + \mu_{t+1}\right)\right]^2\right\}.$$

Hence:

- Values of μ_t that are close to $\frac{1}{2}(\mu_{t-1} + \mu_{t+1})$ are favored (higher density).
- The relative favorability is governed by σ^2 – if this variance is small, then μ_t can't stray too far from its neighbors.

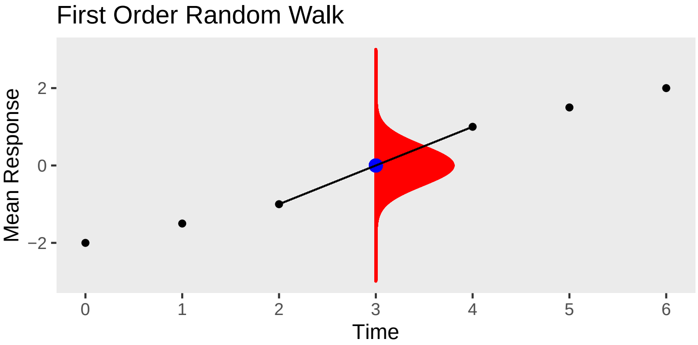


Figure 8: Illustration of the RW1 model for smoothing at time 3. The mean of the smoother is the average of the two adjacent points (and is highlighted as ●), and deviations from this mean are penalized via the normal distribution shown in **red**.

RW2 Model

- The second order RW (RW2) model produces smoother trajectories than the RW1, and has more reasonable short term predictions, which is desirable for modeling child mortality.
- In terms of second differences:

$$(\mu_t - \mu_{t-1}) - (\mu_{t-1} - \mu_{t-2}) \sim \text{Normal}(\ 0, \sigma^2\ ),$$

  showing that deviations from linearity are discouraged.
- Forecasts S steps ahead have a normal distribution with mean:

$$\text{E}[\mu_{T+S} \mid \mu_1, \ldots, \mu_T] = \mu_T + S(\mu_T - \mu_{T-1})$$

  which is a linear function of the values at the last two time points.
- The variance is

$$\text{var}(\mu_{T+S} \mid \mu_1, \ldots, \mu_T) = \frac{\sigma^2}{6} \times S(S+1)(2S+1)$$

  which is cubic in the number of periods S, so blows up very quickly.

Temporal Smoothing Model Summary

We have three models:

IID MODEL:

$$\mu_t \sim \text{Normal}(0, \sigma^2),$$

smooth towards zero.

RW1 MODEL:

$$\mu_t - \mu_{t-1} \sim \text{Normal}(0, \sigma^2),$$

smooth towards the previous value.

RW2 MODEL:

$$(\mu_t - \mu_{t-1}) - (\mu_{t-1} - \mu_{t-2}) \sim \text{Normal}(0, \sigma^2),$$

smooth towards the previous slope.

Bayesian Inference

Bayesian inference:

- A Data Model (Likelihood) is probabilistically combined with
- A Penalization (Prior) that expresses beliefs about the parameters θ encoding the model.
- Combination occurs via **Bayes Theorem:**

$$\underbrace{p(\theta|y)}_{\textbf{Posterior}} \propto \underbrace{L(\theta)}_{\textbf{Likelihood}} \times \underbrace{\pi(\theta)}_{\textbf{Prior}}.$$

- On the log scale:

$$\underbrace{\log p(\theta|y)}_{\textbf{Updated Beliefs}} = \underbrace{\log L(\theta)}_{\textbf{Data Model}} + \underbrace{\log \pi(\theta)}_{\textbf{Penalization}}.$$

Bayesian Inference

- In a Bayesian analysis the complete set of unknowns (parameters) is summarized via the **multivariate posterior distribution**.
- The marginal distribution for each parameter may be summarized via its mean, standard deviation, or quantiles.
- It is common to report the **posterior median** and a **90% or 95% posterior range** for parameters of interest.
- The range that is reported is known as a credible interval.
- The computations required for Bayesian inference (integrals) is often not trivial and many be carried out using a variety of analytic, numeric and simulation based techniques.
- We use the integrated nested Laplace approximation (INLA).

Bayes Example

Imagine the data model is normal with an unknown mean μ:

$$\bar{y} \mid \mu \sim \text{Normal}(\mu, \sigma^2/n),$$

where σ^2/n is assumed known ($\sigma/\sqrt{n}$ is the standard error).

We also imagine the prior is normal:

$$\mu \sim \text{Normal}(m, v),$$

so that values of the mean μ that are (relatively) far from m are penalized.

The log posterior is:

$$\underbrace{\log p(\mu \mid y)}_{\text{Updated Beliefs}} = -\underbrace{\frac{n}{2\sigma^2}(\bar{y} - \mu)^2}_{\text{Data Model}} - \underbrace{\frac{1}{2v}(\mu - m)^2}_{\text{Penalization}}.$$

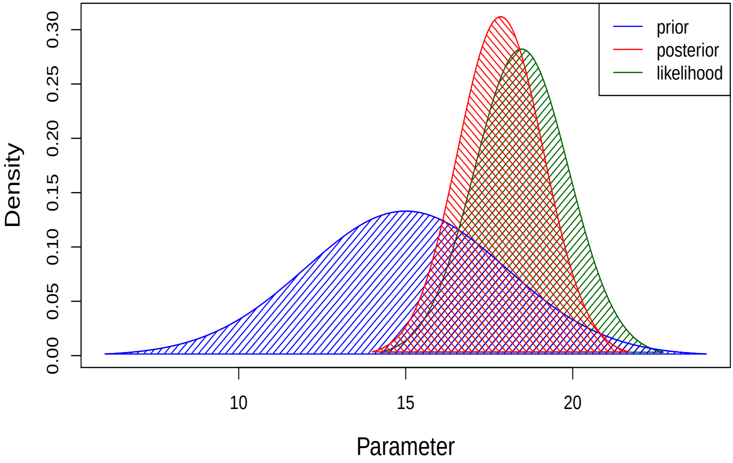


Figure 9: Normal data model with $n = 10$, $\overline{y} = 19.3$ and standard error 1.41. The prior for μ has mean $m = 15$ and $v = 3^2$. The posterior for the parameter μ is a compromise between the two sources of information: the posterior mean is 18.5 and the posterior standard deviation is 1.28.

RW Fitting to Simulated Data

- We illustrate fitting with the RW2 model, using the simulated data seen earlier.
- The model is:

$$
\begin{aligned}
Y_t | p_t &\sim \text{Binomial}(n_t, p_t) \\
\frac{p_t}{1 - p_t} &= \exp(\alpha + \phi_t) \\
(\phi_1, \dots, \phi_T) &\sim \text{RW2}(\sigma^2) \\
\sigma^2 &\sim \text{Prior on Smoothing Parameter} \\
\alpha &\sim \text{Prior on Intercept}
\end{aligned}
$$

- On Figures 10 and 11 the fitted values are shown in red – in both the constant prevalence and curved prevalence cases, the reconstruction is reasonable.

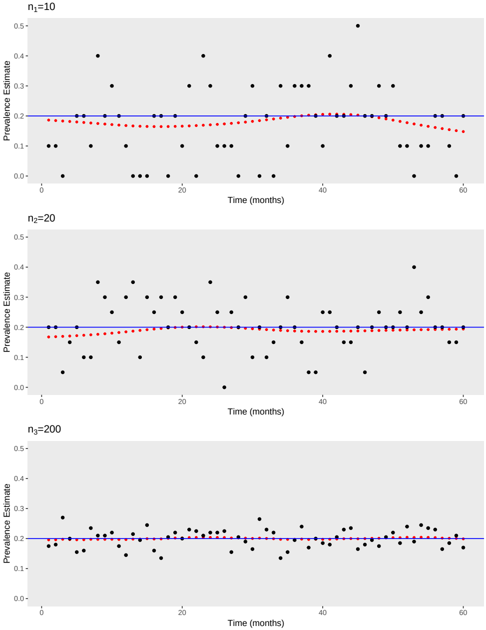


Figure 10: Prevalence estimates over time from simulated data, true prevalence $p = 0.2$ (blue solid lines). Smoothed random walk estimates in **red**.

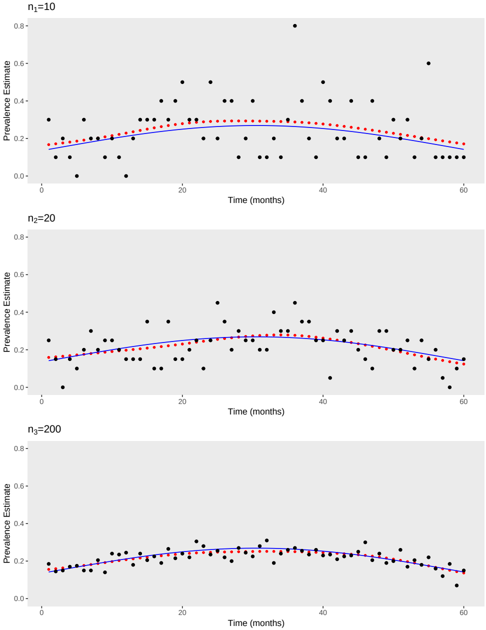


Figure 11: Prevalence estimates over time from simulated data, true prevalence corresponds to curved blue solid line. Smoothed random walk estimates in **red**.

Spatial Smoothing of Simulated Data

Data Model: For area i:

$$\underbrace{Y_i}_{\text{Count}} \mid \underbrace{p_i}_{\text{Prevalance}} \sim \underbrace{\text{Binomial}(n_i, p_i)}_{\text{Data Model}}.$$

Smoothing Model: For the odds in area i:

$$\frac{p_i}{1-p_i} = \exp(\alpha + \phi_i).$$

We consider two choices for the smoothing model:

- IID model: Smooth to the overall mean with no spatial structure $\phi_i \sim \text{Normal}(0, \sigma^2)$ where σ^2 controls the amount of smoothing — small/large corresponds to strong/weak smoothing.
- BYM2[1] model: Add a spatial component that encourages local similarity analogously to the random walk model with a suitable choice of neighbors, sharing a common boundary being the commonest choice.

[1] named after the paper that introdced the model, Besag, York and Mollié (1991)

Spatial Modeling of Simulated Data for $n = 50$ Constant Risk Case

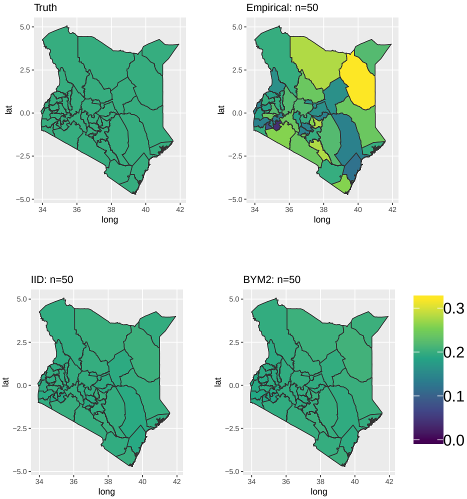


Figure 12: Results with $n = 50$ when true prevalence is 0.2. Top Left: Truth. Top Right: raw proportions. Bottom Left: Estimates with IID model. Bottom Right: smoothing with BYM2.

Spatial Modeling of Simulated Data for $n = 100$ Constant Risk Case

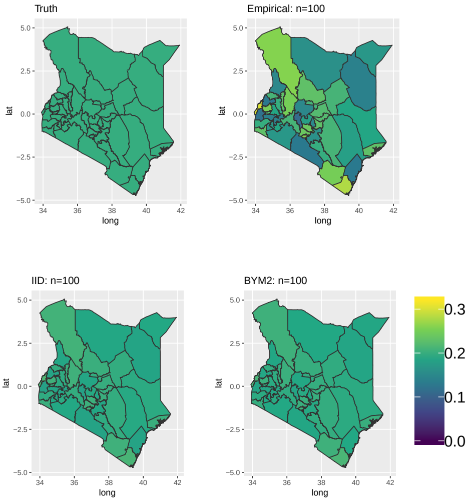


Figure 13: Results with $n = 100$ when true prevalence is 0.2. Top Left: Truth. Top Right: raw proportions. Bottom Left: Estimates with IID model. Bottom Right: smoothing with BYM2.

Spatial Modeling of Simulated Data for $n = 500$ Constant Risk Case

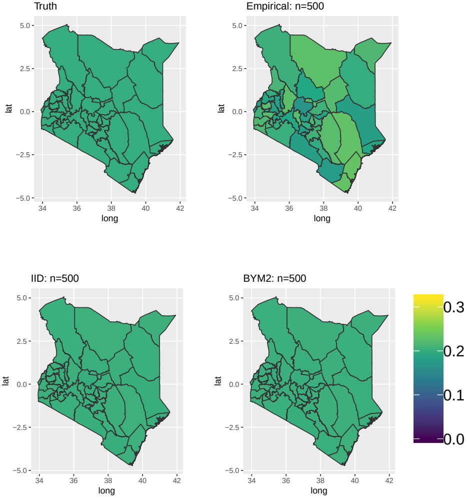


Figure 14: Results with $n = 500$ when true prevalence is 0.2. Top Left: Truth. Top Right: raw proportions. Bottom Left: Estimates with IID model. Bottom Right: smoothing with BYM2.

Spatial Modeling of Simulated Data for $n = 50$ Varying Risk Case

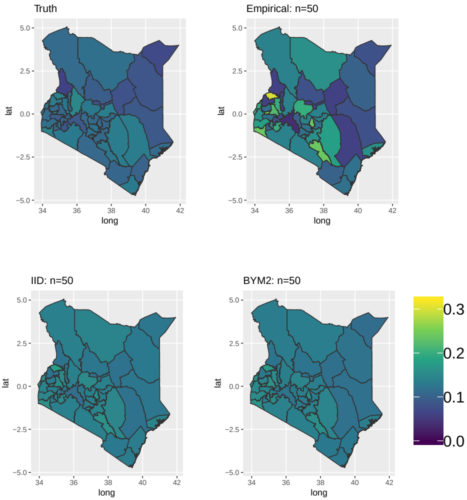


Figure 15: Results with $n = 100$ when true prevalence is varying. Top left: Truth. Top right: Raw proportions, Bottom left: smoothing with IID model. Bottom right: smoothing with BYM2.

Spatial Modeling of Simulated Data for $n = 100$ Varying Risk Case

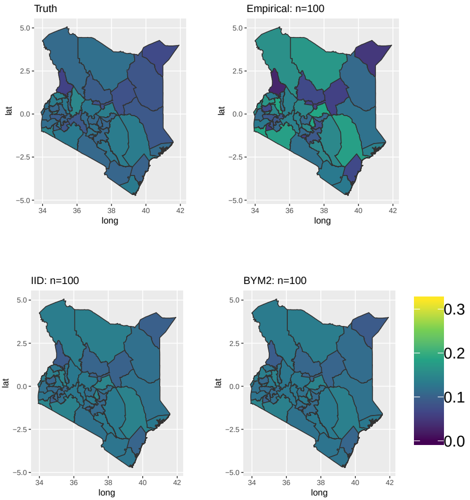


Figure 16: Results with $n = 100$ when true prevalence is varying. Top left: Truth. Top right: Raw proportions, Bottom left: smoothing with IID model. Bottom right: smoothing with BYM2.

Spatial Smoothing of Simulated Data for $n = 500$ Case

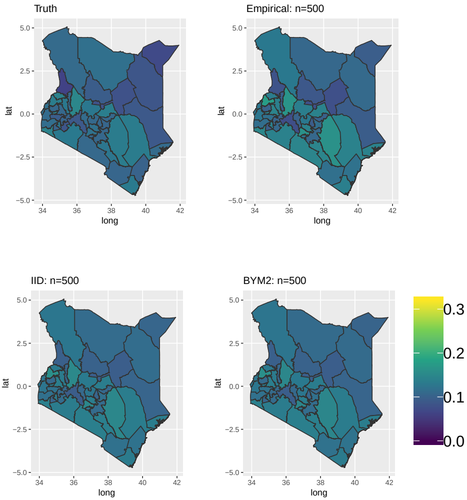


Figure 17: Results with $n = 500$ when true prevalence is varying. Top left: Truth. Top right: Raw proportions. Bottom left: smoothing with IID model. Bottom right: smoothing with BYM2.

Child Survival Modeling

Discrete Survival Model

We know that infant mortality varies greatly over the first 5 years of life and two possible approaches to modeling how mortality varies with age:

- A continuous function of age, via a parametric model (e.g., weibull, gamma,...).
- A discrete function of age, which involves splitting age into intervals.

For flexibility, we follow the latter route and assume a **discrete survival model**, with six **discrete hazards** (probabilities of dying in a particular interval, given survival to the start of the interval) for each of the age bands:

1. $[0, 1)$,
2. $[1, 12)$,
3. $[12, 24)$,
4. $[24, 36)$,
5. $[36, 48)$,
6. $[48, 60]$.

The first category corresponds to neonatal, the first two, infant mortality, and all six, under-5 mortality.

Discrete Survival Model

- Each child contributes up to 60 months of observation time, and can contribute less if censoring.
- For a generic calendar period:

$$\begin{aligned}
\text{Survival to 60 months} &= \text{Survival in month 1} \\
&\times \text{Survival in month 2} \mid \text{survived to end of month 1} \\
&\cdots \\
&\times \text{Survival in month 60} \mid \text{survived to end of month 59}
\end{aligned}$$

- Hence, we are following a **synthetic cohort** approach.
- The hazards are estimated using a logistic regression model, with weighting to account for the survey design.
- At the end of this process we have an estimate $\widehat{\text{U5MR}}$ in each area and to period, along with its variance.
- We also apply an HIV adjustment.

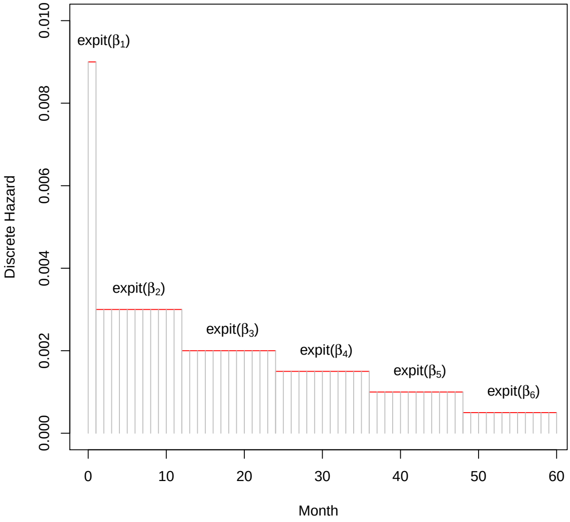


Figure 18: Representation of the conditional probabilities (hazards) of death by month.

The SUMMER Package

Spatial Smoothing Model

Key Idea: Take as data the **weighted estimator** – in large samples this follows a normal distribution.

Hierarchical Model:

1. **The Data Model:** Specifically, we take as data in area **area** i, the logit of the weighted estimates: $y_i = \text{logit}(\widehat{P}_i)$
$$\underbrace{y_i \mid \lambda_i \sim \text{Normal}\left(\lambda_i, V_i\right)}_{\text{Survey design acknowledged here}}$$
where V_i is the design variance.
2. **The Smoothing Model:**
$$\lambda_i \quad = \quad \underbrace{\mu}_{\text{Intercept}} + \underbrace{\epsilon_i}_{\text{Independent}} + \underbrace{S_i}_{\text{Spatial}}$$

The model is implemented in the `R` package `SUMMER`:

- A design object being created in the `survey` package.
- The `INLA` package is used for Bayesian computation.
- It is computationally inexpensive – country-specific estimates in seconds.

Space-Time Smoothing Model

Hierarchical Model:

1. **The Data Model**:

$$\underbrace{y_{it} \mid \lambda_{it} \sim \text{Normal}\left(\lambda_{it}, \widehat{V}_{it}\right),}_{\text{Survey design acknowledged here}}$$

   where

   - y_{it} is the logit of the direct estimator in area i and period t,
   - λ_{it} is the logit of the true U5MR in county i and period t, with $\widehat{V}_{it}$ known.

2. **The Smoothing Model**: We decompose λ_{it} into temporal, spatial and space-time components:

$$\lambda_{it} = \underbrace{\mu}_{\text{Intercept}}$$

$$+ \underbrace{\alpha_t}_{\text{Independent}} + \underbrace{\gamma_t}_{\text{Random Walk}} \qquad \text{Temporal Model}$$

$$+ \underbrace{\epsilon_i}_{\text{Independent}} + \underbrace{S_i}_{\text{Spatial}} \qquad \text{Spatial Model}$$

$$+ \underbrace{\delta_{it}}_{\text{Interaction}} \qquad \text{Space-Time Model}$$

Smoothing of ENSANUT Data

- We calculate 5-year weighted estimates of U5MR using a discrete survival model for the periods 85–89, 90–94, 95–99, 00–04, 05–09, 10–14.
- We smooth these estimates in time only using the model in the `SUMMER` package.
- Figure 19 compares smoothed estimates with IGME and IHME estimates.
- Figures 20 and 21 give the estimates with projections for U5MR and NMR, respectively.

RW2 Model Applied to Kenya Survey Data

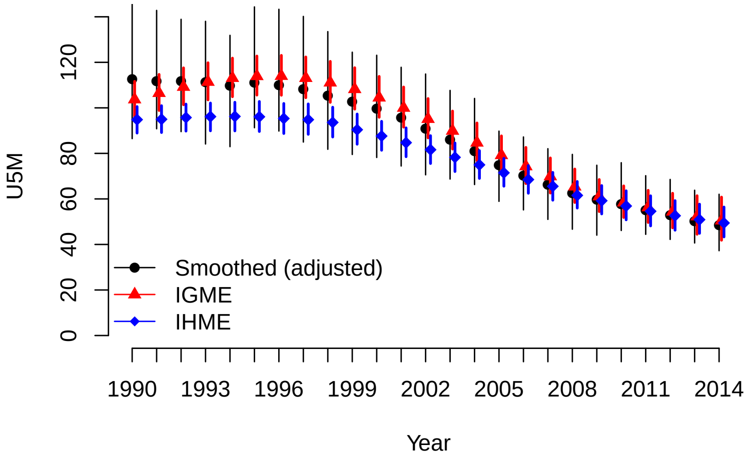


Figure 19: Yearly smoothed (with RW2 and adjusted for HIV bias) estimates of under-5 mortality in Kenya, along with IGME (B3)and IHME yearly estimates, with 95% uncertainty intervals.

Yearly U5MR Smoothed Estimates for Kenya

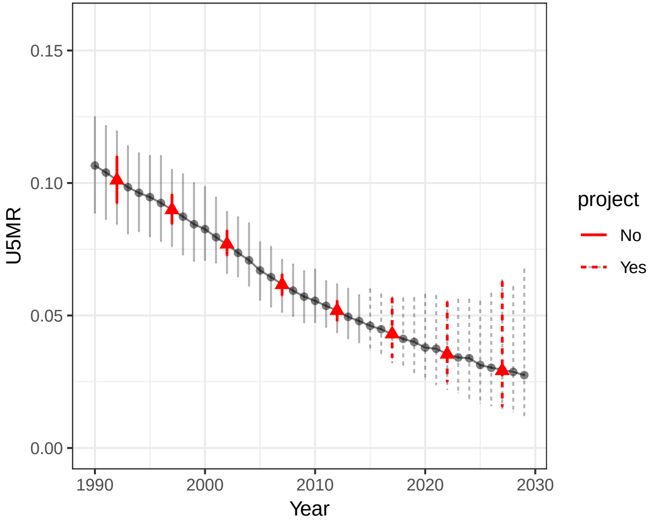


Figure 20: Yearly RW2 smoothing of weighted estimates of under-5 mortality in Kenya, with 95% uncertainty intervals. The dashed lines on the right are projections.

Yearly NMR Smoothed Estimates for Kenya

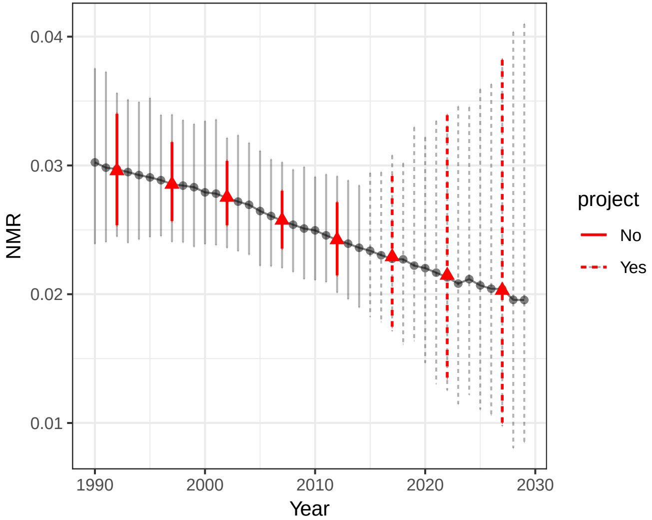


Figure 21: Yearly RW2 smoothing of weighted estimates of nenonatal mortality in Kenya, with 95% uncertainty intervals. The dashed lines on the right are projections.

Space-Time Smoothing Model Applied to Kenya Survey Data

We now turn to **space-time smoothing** using `SUMMER`:

- Figure 22 gives the **weighted estimates** with hatching representing uncertainty.
- Figure 23 gives the **smoothed estimates** with hatching representing uncertainty – these estimates show less spatial variability and reduced uncertainty.
- Figure 40 clearly shows the drop in U5MR over time, and reduced **between-province variability**. The uncertainty in estimates is also apparent.

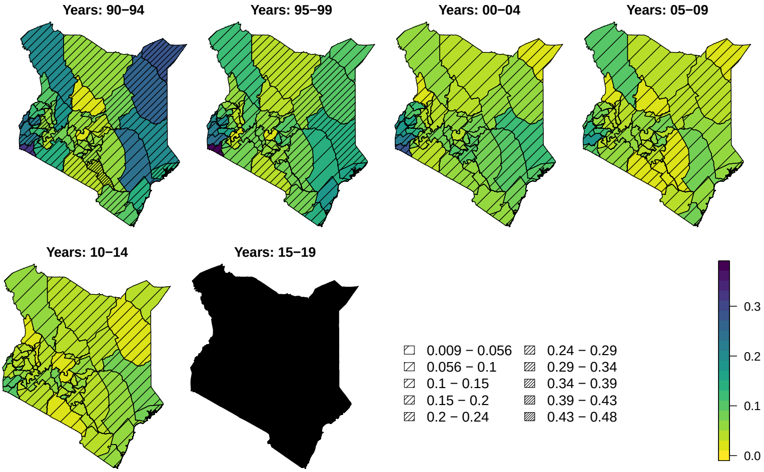


Figure 22: Yearly **weighted estimates** (from a discrete survival model) of under-5 mortality in Kenya, with uncertainty indicated by density of hatching; more hatching $\rightarrow$ more uncertainty, with the latter measured though width of 95% uncertainty interval.

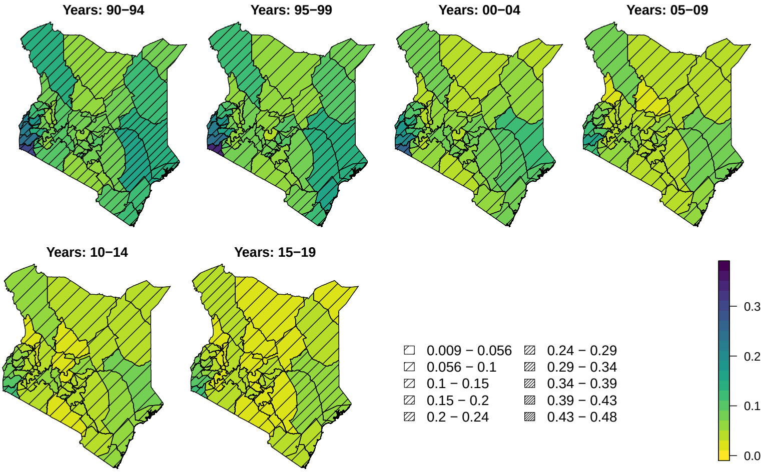


Figure 23: Yearly **smoothed estimates** (from a discrete survival model) of under-5 mortality in Kenya, with uncertainty indicated by density of hatching; more hatching → more uncertainty, with the latter measured though width of 95% uncertainty interval.

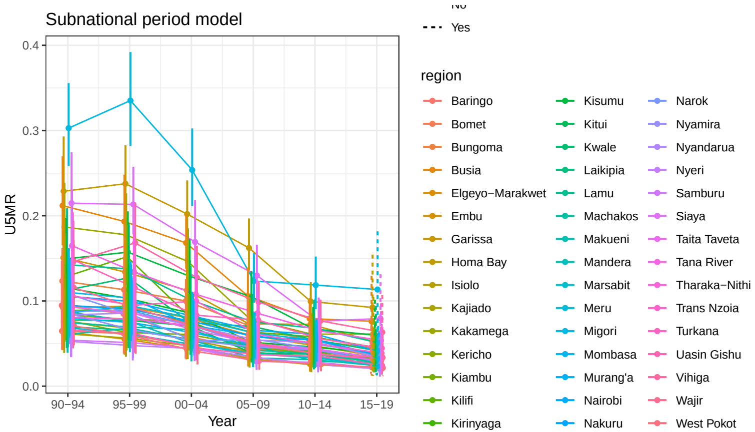
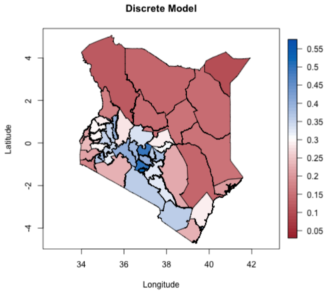


Figure 24: Five-yearly smoothed estimates (from a discrete survival model) of under-5 mortality in Kenya, by province, with 95% uncertainty intervals

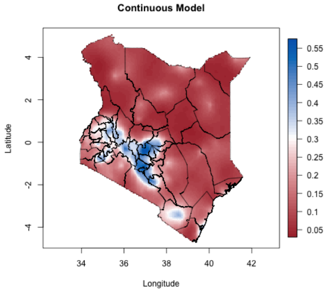
Final Thoughts

Two Approaches to Spatial Smoothing

- Model at the area level using a discrete spatial model. These are the models that are implemented in the SUMMER package.



- Model at the point level using a continuous spatial model. Gaussian Process (GP) models abound and have many different implementations.



Continuous Modeling

We are also pursuing the use of **continuous spatial models**:

- These are routinely used by both WorldPop and IHME, but continuous modeling is a more hazardous approach to estimation.
- However, it is the way forward to allow **multiple data sources** at different spatial resolutions to be combined.
- And reporting can be on a relevant discrete scale.

Surface Reconstructions for U5MR in Kenya

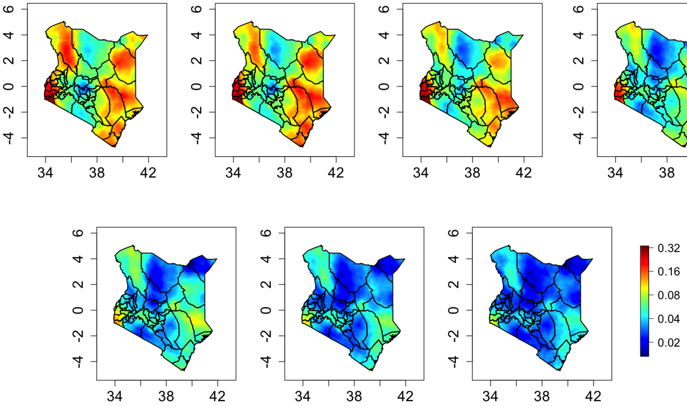


Figure 25: Posterior medians of U5MR for 1990, 1995, 2000, 2005, 2010, 2015, 2020.
Important Point: These are point estimates and the uncertainty at each pixel is in general very large.

Estimates for U5MR in Malawi

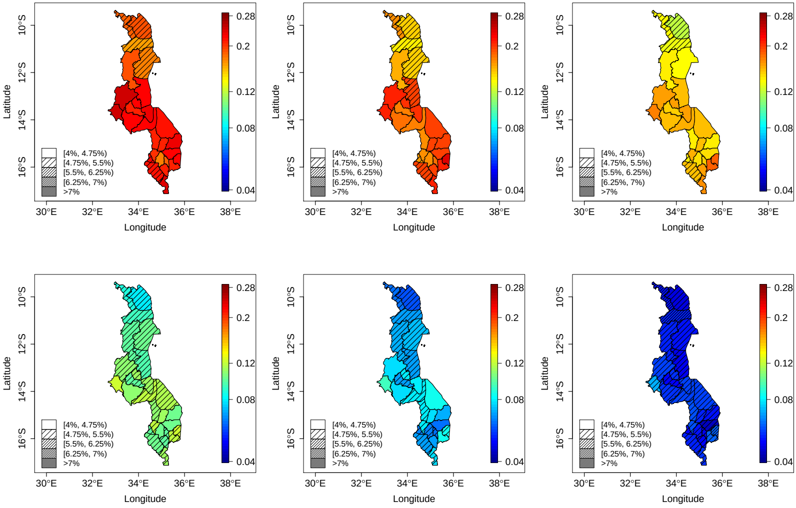
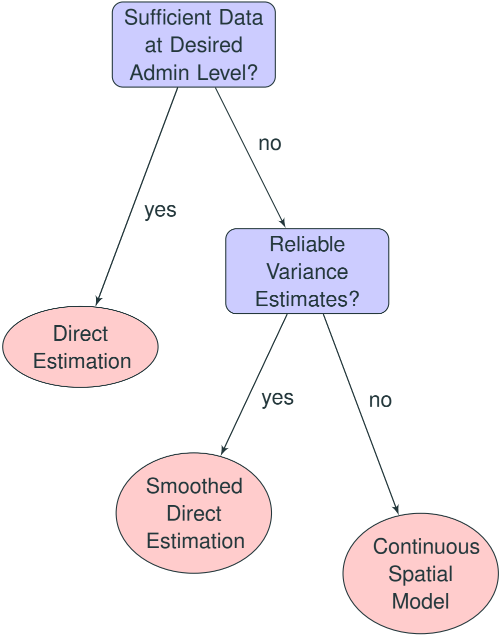


Figure 26: Estimates of U5MR for Malawi for 1990, 1995, 2000, 2005, 2010, 2015.

Recommended Methods for Routine Work



Concluding Comments

Final thoughts:

- `SUMMER` allows mortality to be examined for **different age groups (e.g., NMR, infant,...)** and also by **gender**.
- **Multiple surveys** can also be combined.
- **Summary Birth History (SBH)** data from census may be added using the same approach – soon to appear in `SUMMER`.
- Beyond that: Estimate mortality for **ages 5–14**.
- Work in progress on **cause of death**.

Concluding Comments

Feel free to contact Jon (**jonno@uw.edu**) or Katie (**wilsonkl@uw.edu**) with:

- Follow up questions on methods or use of the `SUMMER` package.
- On the website is a link to a paper with Admin-1 results for 35 African countries.
- If you would like to collaborate on subnational estimation please let us know!

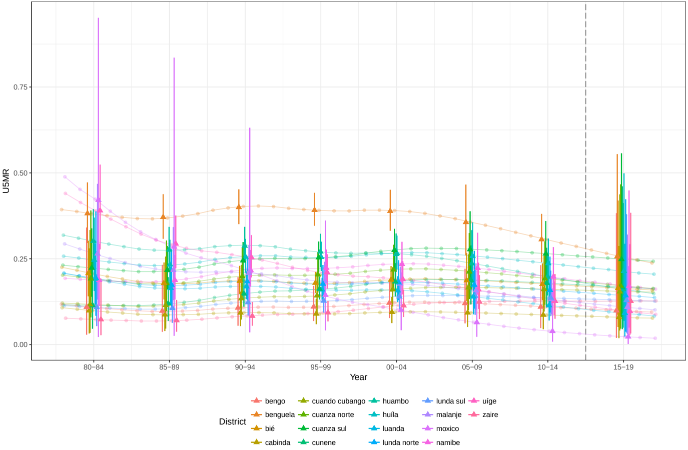


Figure 27: Admin-1 Estimates for Angola.

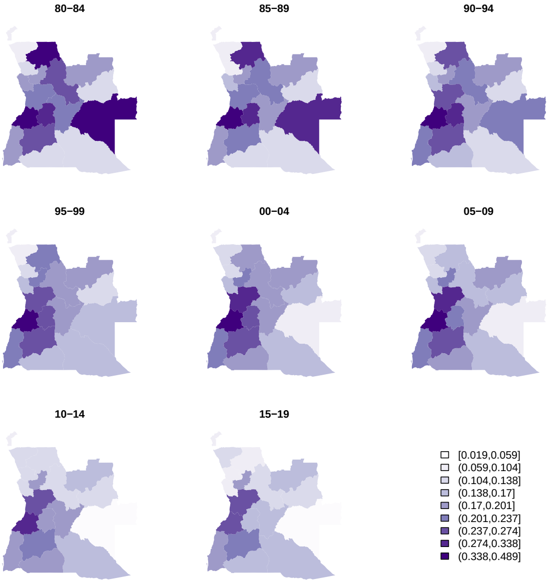


Figure 28: Admin-1 Estimates for Angola.

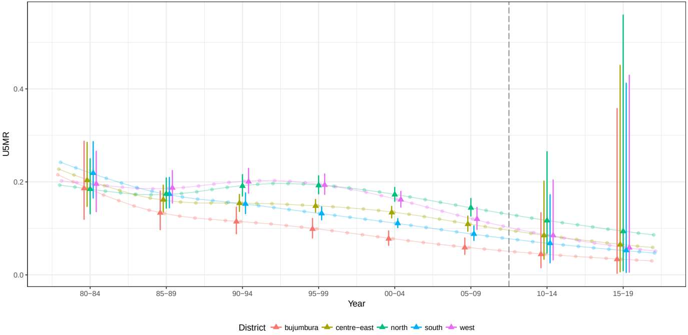


Figure 29: Admin-1 Estimates for Burundi.

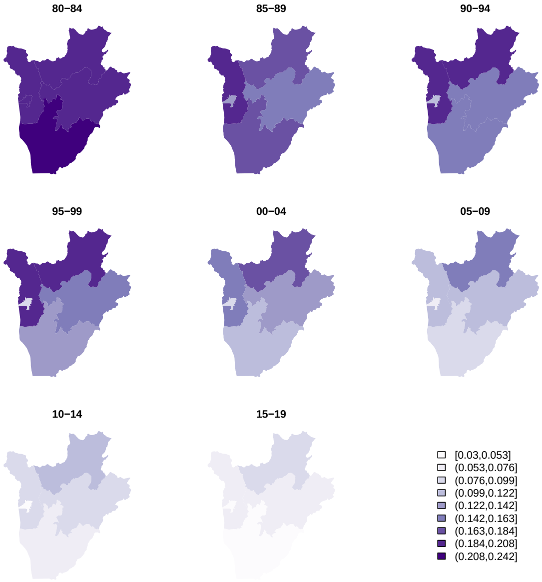


Figure 30: Admin-1 Estimates for Burundi.

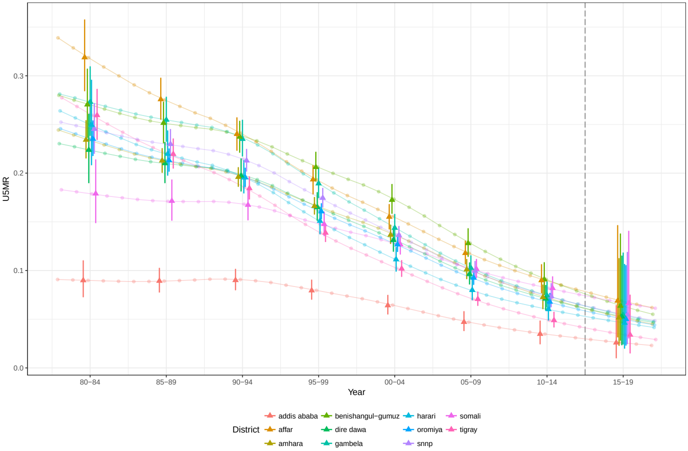


Figure 31: Admin-1 Estimates for Ethiopia.

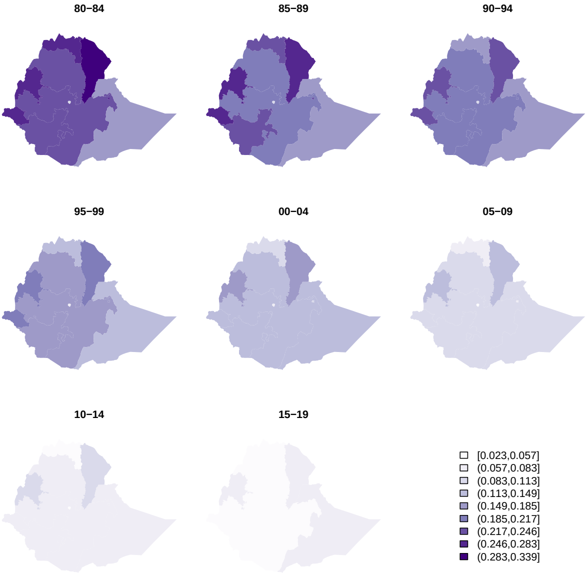


Figure 32: Admin-1 Estimates for Ethiopia.

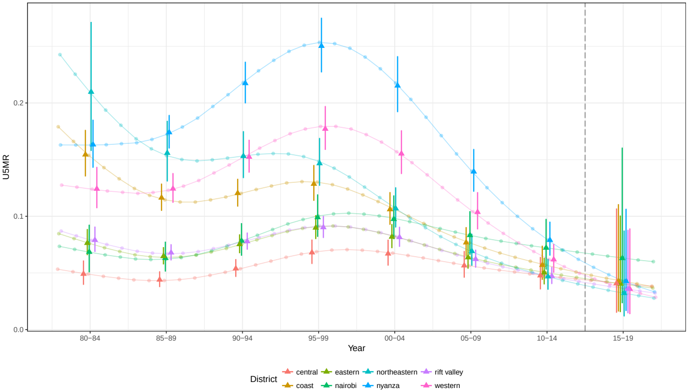


Figure 33: Admin-1 Estimates for Kenya.

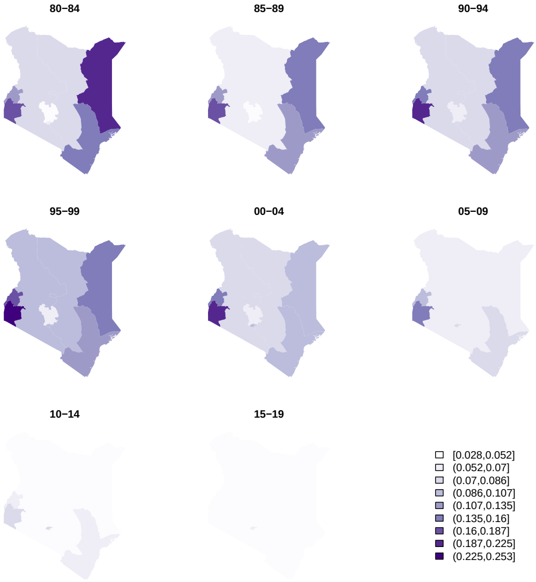


Figure 34: Admin-1 Estimates for Kenya.

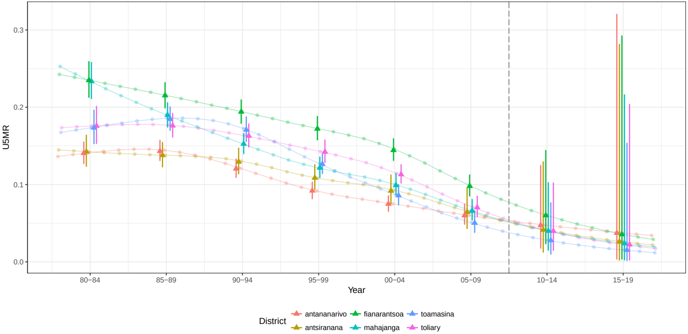


Figure 35: Admin-1 Estimates for Madagascar.

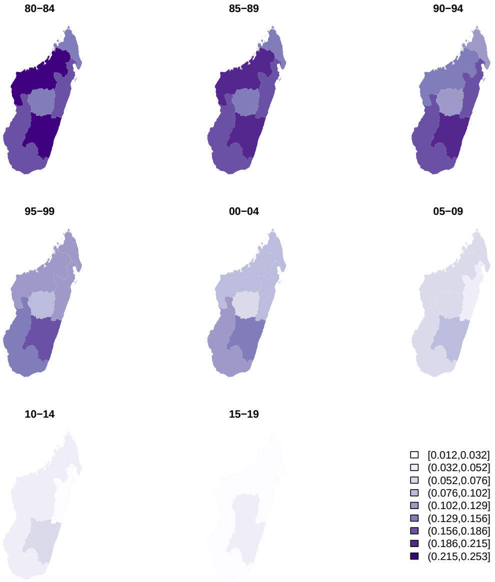


Figure 36: Admin-1 Estimates for Madagascar.

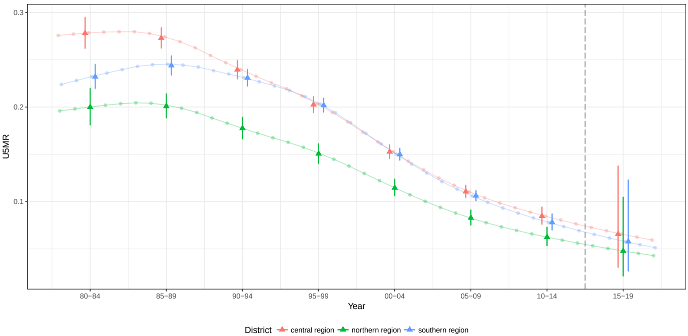


Figure 37: Admin-1 Estimates for Malawi.

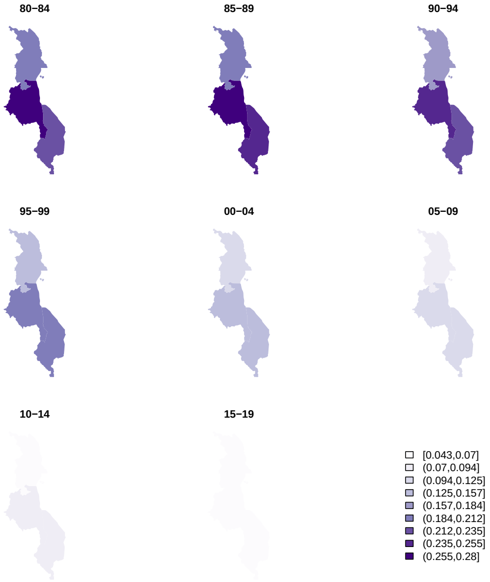


Figure 38: Admin-1 Estimates for Malawi.

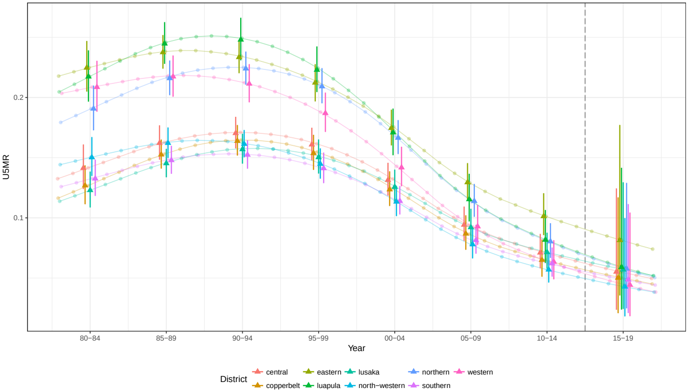


Figure 39: Admin-1 Estimates for Zambia.

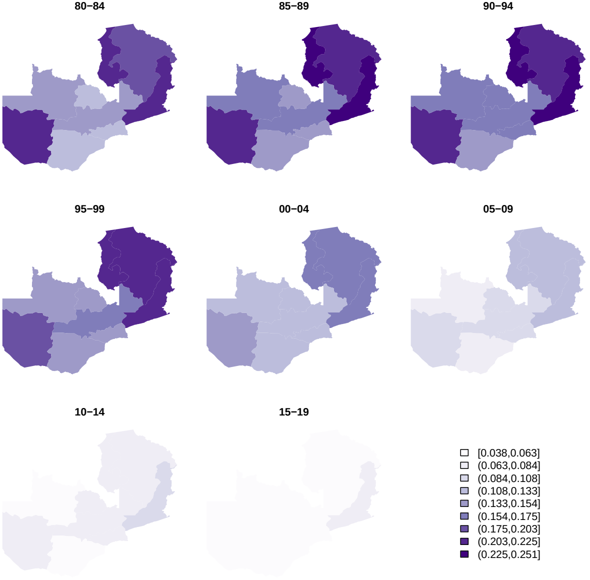


Figure 40: Admin-1 Estimates for Zambia.

Background Literature

- Smoothing of direct estimates (Fay and Herriot, 1979; Chen et al., 2014; Mercer et al., 2015).
- Comparison of discrete and continuous models (Wakefield et al., 2018).
- Application of space-time smoothing model to 40 African countries (Li et al., 2019).
- Modeling of SBH data (Brass, 1964; Sullivan, 1972; Brass, 1975; Trussell, 1975; Feeney, 1976; Coale and Trussell, 1977; Hill et al., 1983; Rajaratnam et al., 2010; Wilson and Wakefield, 2018a).
- Combining point and area data (Wilson and Wakefield, 2018b).
- INLA (Rue et al., 2009; Lindgren et al., 2011; Blangiardo and Cameletti, 2015; Wang et al., 2018; Krainski et al., 2018).

References

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*, 1–59.

Blangiardo, M. and M. Cameletti (2015). *Spatial and Spatio-Temporal Bayesian Models with R-INLA*. John Wiley and Sons.

Brass, W. (1964). *Uses of census or survey data for the estimation of vital rates*. United Nations. Paper prepared for the African Seminar on Vital Statistics, Addis Ababa, 14–19 December, 1964.

Brass, W. (1975). *Methods for Estimating Fertility and Mortality from Limited and Defective Data*. Chapel Hill, North Carolina: Population Center, Laboratories for Population Statistics.

Chen, C., J. Wakefield, and T. Lumley (2014). The use of sample weights in Bayesian hierarchical models for small area estimation. *Spatial and Spatio-Temporal Epidemiology 11*, 33–43.

Coale, A. J. and J. Trussell (1977). Annex i: estimating the time to which Brass estimates apply. *Population Bulletin of the United Nations* (10), 87–89.

Fay, R. and R. Herriot (1979). Estimates of income for small places: an application of James–Stein procedure to census data. *Journal of the American Statistical Association 74*, 269–277.

Feeney, G. (1976). Estimating infant mortality rates from child survivorship data by age of mother. *Asian and Pacific Census Newsletter 3*, 12–16.

Freire, W. B., P. Belmont, D. F. López-Cevallos, and W. F. Waters (2015). Ecuador's national health and nutrition survey: objectives, design, and methods. *Annals of Epidemiology 25*, 877–878.

Godwin, J. and J. Wakefield (2019). Space-time modeling of child and young adult mortality at the admin 2 level in a low and medium income countries context. *In Preparation*.

Hill, K., H. Zlotnik, and J. Trussell (1983). *Demographic estimation: a manual on indirect techniques. Manual X*.

Krainski, E. T., V. Gómez-Rubio, H. Bakka, A. Lenzi, D. Castro-Camilo, D. Simpson, F. Lindgren, and H. Rue (2018). *Advanced Spatial Modeling with Stochastic Partial Differential Equations Using R and INLA*. Chapman and Hall/CRC.

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*. Published January 22, 2019.

Lindgren, F., H. Rue, and J. Lindström (2011). An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic differential equation approach (with discussion). *Journal of the Royal Statistical Society, Series B 73*, 423–498.

Mercer, L., J. Wakefield, A. Pantazis, A. Lutambi, H. Mosanja, and S. Clark (2015). Small area estimation of childhood mortality in the absence of vital registration. *Annals of Applied Statistics 9*, 1889–1905.

Rajaratnam, J. K., L. N. Tran, A. D. Lopez, and C. J. Murray (2010). Measuring under-five mortality: validation of new low-cost methods. *PLoS Medicine 7*, e1000253.

Rue, H., S. Martino, and N. Chopin (2009). Approximate Bayesian inference for latent Gaussian models using integrated nested Laplace approximations (with discussion). *Journal of the Royal Statistical Society, Series B 71*, 319–392.

Sullivan, J. M. (1972). Models for the estimation of the probability of dying between birth and exact ages of early childhood. *Population Studies 26*, 79–97.

Trussell, T. J. (1975). A re-estimation of the multiplying factors for the Brass technique for determining childhood survivorship rates. *Population Studies 29*, 97–107.

Wakefield, J., G.-A. Fuglstad, A. Riebler, J. Godwin, K. Wilson, and S. Clark (2018). Estimating under five mortality in space and time in a developing world context. *Statistical Methods in Medical Research*. Published online April 19.

Wang, X., Y. Yue, and J. J. Faraway (2018). *Bayesian Regression Modeling with INLA*. Chapman and Hall/CRC.

Wilson, K. and J. Wakefield (2018a). Child mortality estimation incorporating summary birth history data. *Under revision, available at arXiv*.

Wilson, K. and J. Wakefield (2018b). Pointless spatial modeling. *Biostatistics*. Published online 6 September, 2018.

Scaling Up the Smoothed Direct Model

The smoothed direct model has been used for 35 African countries to estimate U5MR in Admin-1 regions by year.

Includes space-time interactions that cross random walk models in time with ICAR models in space.

Data:

- 121 DHS in 35 countries
- 1.2 million children
- 192 million child-months

UN have supported this research and these estimates.

Takes around 2.5 hours to obtain estimates for all countries – separate models for each country.

Spatial and space-time smoothed direct estimates models are available in `R`, via the `SUMMER` package.

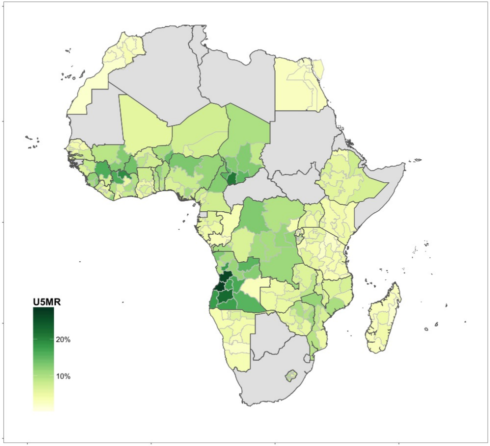


Figure 41: Predictions of U5MR for 2015, in 35 countries of Africa.

Smoothed Direct Estimates

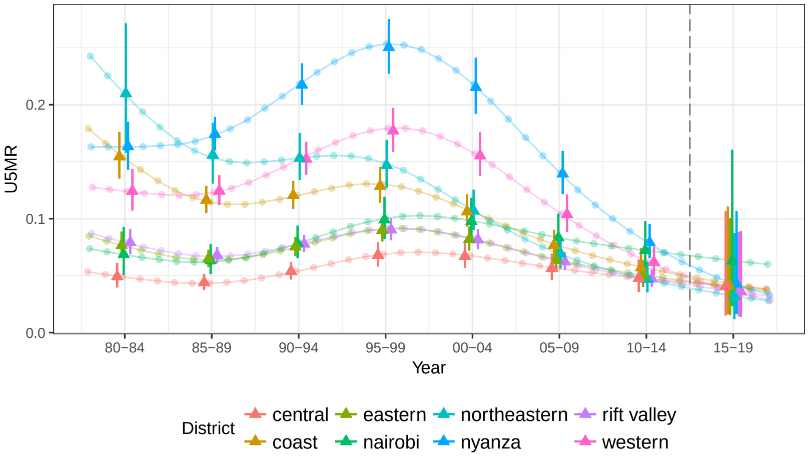


Figure 42: Posterior median estimates for Kenya districts.

Modeling Summary Birth History Data

The Brass Method

- SBH data consist of mother's age m and the number of child born and who have died, call this ratio r_m.
- The proportion of children who have died is:

$$\mathrm{E}[r_m] = \int_0^{A_m} \underbrace{c_m(a)}_{\text{Fertility}} \times \underbrace{q(a)}_{\text{Mortality}} \; da$$

  where
  - A_m is the relevant reproductive period for the mother's age group.
  - $c_m(a)$ is the proportion of births to women who are m at the time of the survey, a years prior to the survey,
  - $q(a)$ is the probability that a child born a years before the survey dies.
- By the mean value theorem this is equal to $q(a^\star)$ for some $0 < a^\star < A_m$.

Age Group	Mortality $q(a^\star)$
15–19	$q(1)$
20–24	$q(2)$
25–29	$q(3)$
30–34	$q(5)$
35–39	$q(10)$
40–44	$q(15)$
45–49	$q(20)$

- The idea of the Brass Method is to find $a^\star$ and then adjust $q(a^\star)$ to $q(5)$.

Combining Direct and Brass Estimates

- We can using the Mercer et al. (2015) smoothing model to combine direct and Brass estimates:

$$
\begin{aligned}
\text{logit}(\widehat{p}_{it}^{\text{DIR}}) &\sim \text{Normal}(\lambda_{it}, \widehat{V}_{it}^{\text{DIR}}) \\
\lambda_{it} &= \mu + \alpha_t + \gamma_t + S_i + \epsilon_i + \delta_{it} \\
\text{logit}(\widehat{p}_{it'}^{\text{BR}}) &\sim \text{Normal}(\lambda'_{it'}, \widehat{V}_{it'}^{\text{BR}}) \\
\lambda'_{it'} &= \mu' + \alpha_{t'} + \gamma_{t'} + S_i + \epsilon_i + \delta_{it'}
\end{aligned}
$$

  where $\widehat{V}_{it'}^{\text{BR}}$ is obtained via the jackknife.
- Direct estimate enters as a 5-year summary, and Brass in a single year, hence the time scales are different, but they both depend on the same underlying temporal process.
- Brass uncertainty estimates from the jackknife.
- Sources of bias are modeled in μ'.
- Full study reported in Godwin and Wakefield (2019).

Results for Bungoma County in Kenya

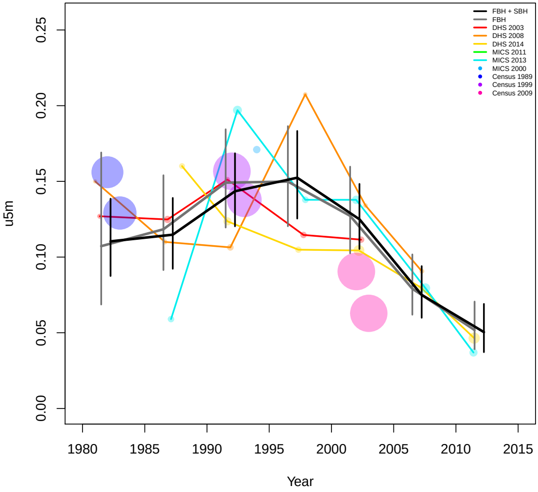


Figure 43: Five yearly estimates for Bungoma county, with different data sources. Circles are proportional to precision of estimates.

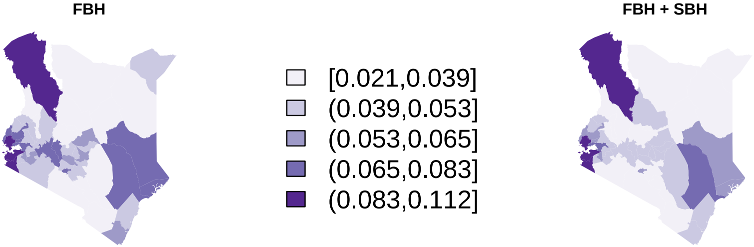


Figure 44: Subnational posterior medians for U5MR for 2010–2014.

Results for Kenya

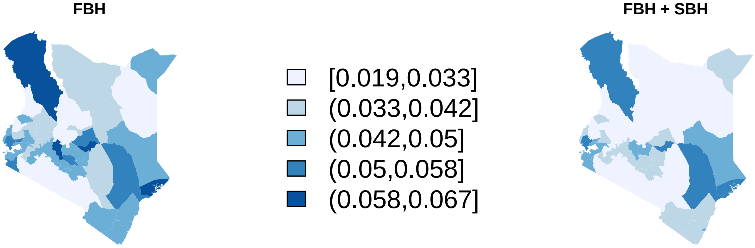


Figure 45: Subnational widths of 95% credible intervals for U5MR for 2010–2014.

IHME U5MR Estimates

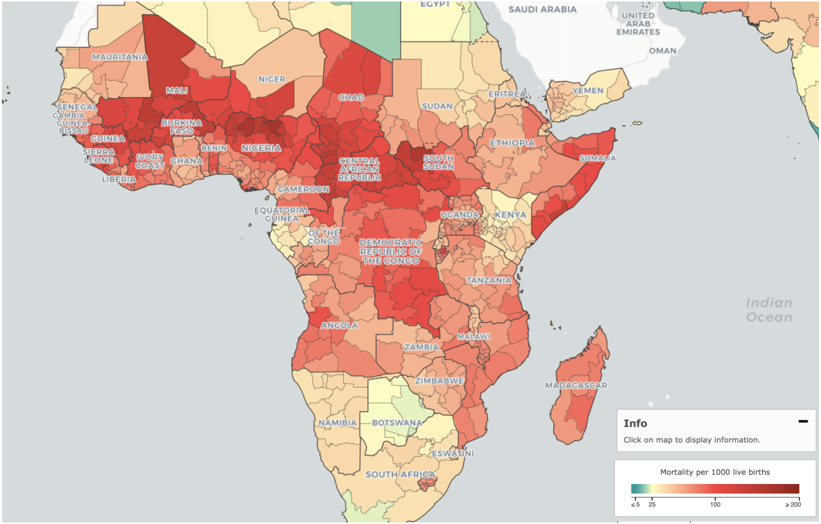


Figure 46: IHME estimates from a continuous space model; summarized at Admin1 level.