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, Quito, Ecuador

April 2–4, 2019

Outline

Motivation

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

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."

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 2012 Ecuador National Health and Nutrition Survey (ENSANUT-ECU).

ENSANUT Survey

Study Design:

- There are 26 strata (24 provinces + Quito + Guayaquil).
- In each strata, 64 clusters are selected with probability proportional to size. In each cluster, 12 households are ultimately selected. Within each household, 1 woman of fertile age was selected via simple random sampling (SRS).
- An example of a stratified cluster design.

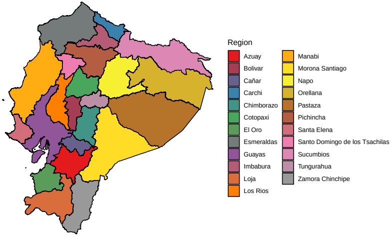

Figure 1: Provinces of Ecuador.

- To simplify the spatial modeling we remove the Galapagos Islands, and Quito and Guayaquil are placed in their provinces (Pichincha and Guayas, respectively).

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!

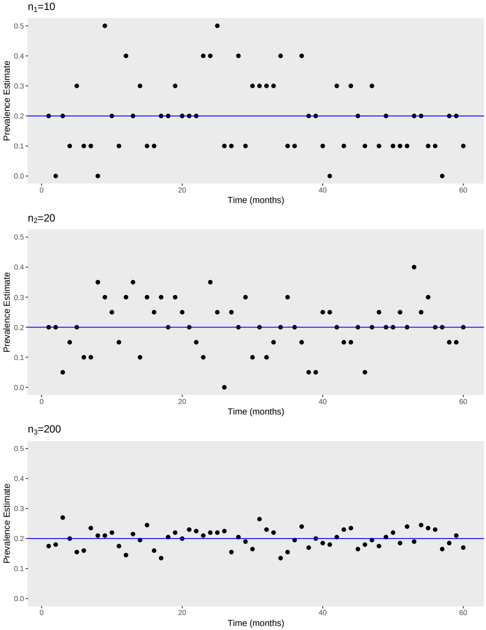

Figure 2: Prevalence estimates over time from simulated data with true prevalence of $p = 0.2$ (blue solid lines).

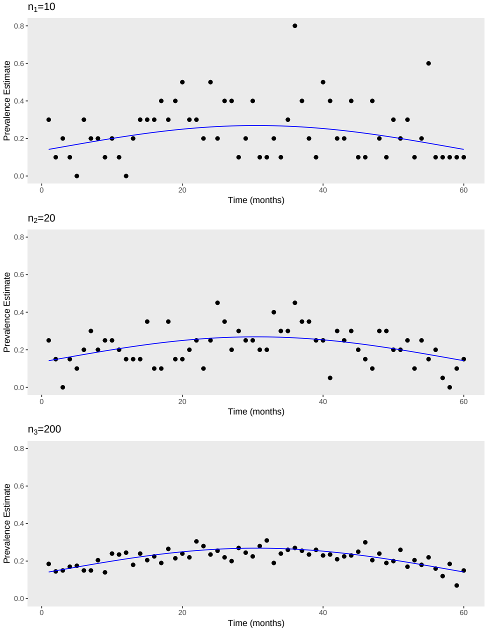

Figure 3: Prevalence estimates over time from simulated data, true prevalence corresponds to curved blue solid line.

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.

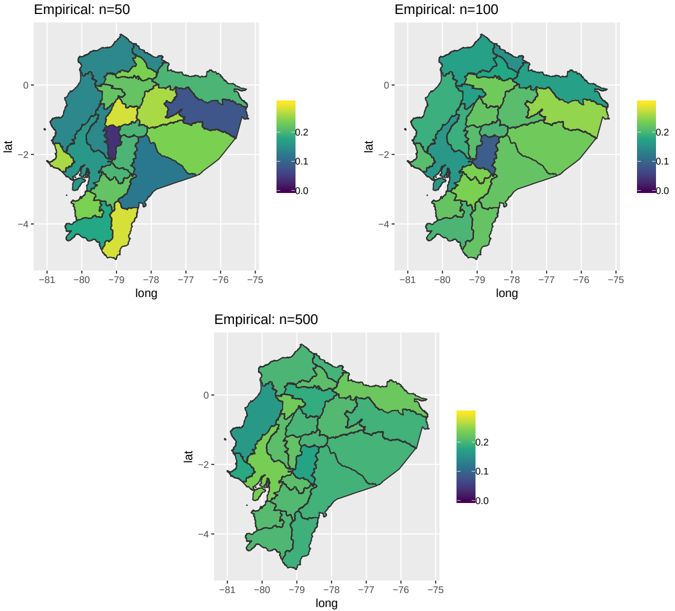

Figure 4: Prevalence estimates over space for simulated data with sample sizes of $n = 50, 100, 500$. True prevalence is 0.2 in all areas.

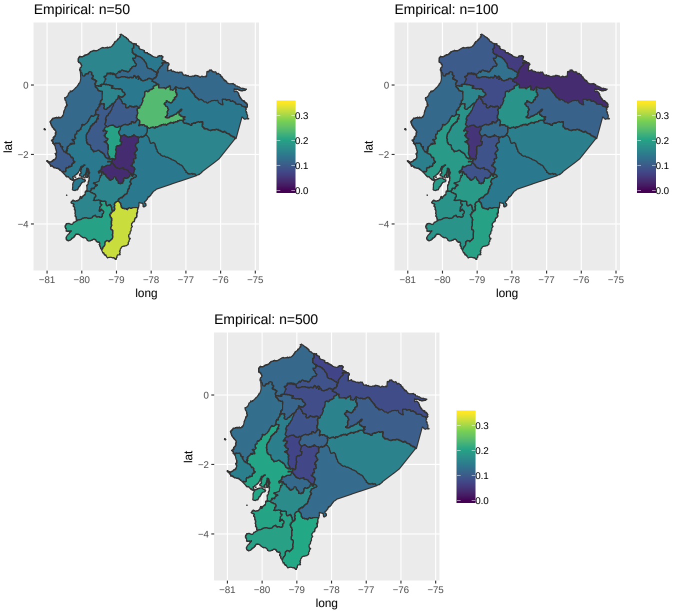

Figure 5: Prevalence estimates over space for simulated data with sample sizes of $n = 50, 100, 500$. True prevalence is spatially varying.

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 $y_1, \ldots, y_n$ 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 people 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.

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 ENSANUT-ECU survey 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.

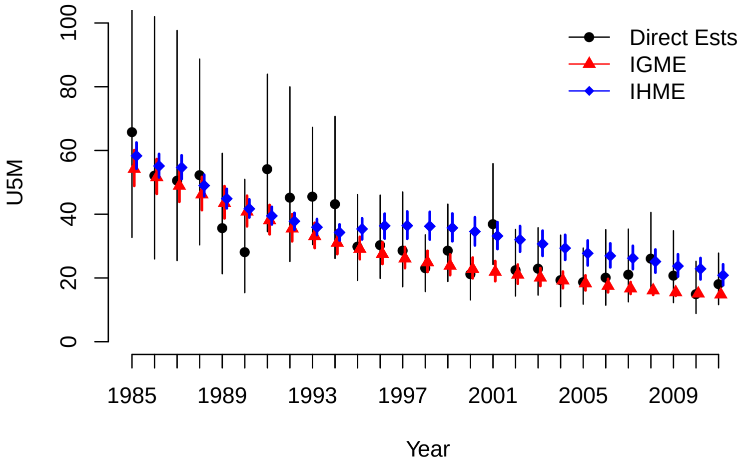

Figure 6: Yearly weighted estimates of under-5 mortality in Ecuador, with 95% uncertainty intervals for Direct and IHME, and 90% for IGME.

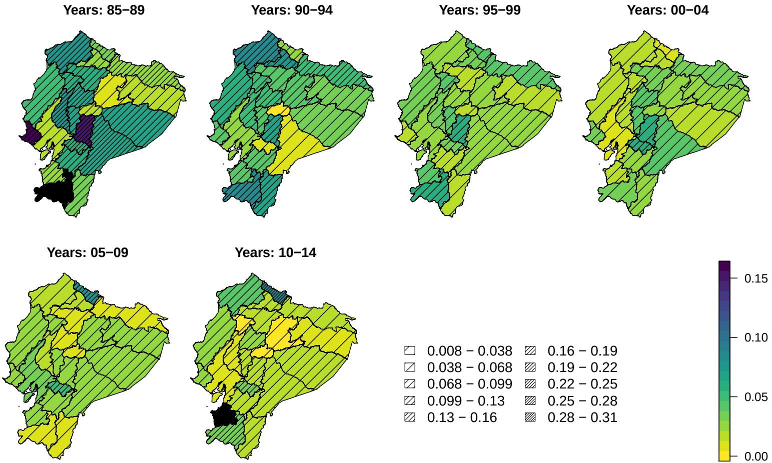

Figure 7: Yearly weighted estimates (from a discrete survival model) of under-5 mortality in Ecuador, with uncertainty indicated by density of hatching; more hatching → more uncertainty, with the latter measured though width of 95% uncertainty interval.

Space-Time Smoothing

Smoothing 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.
- 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.

Random Walk Models

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 neigbors.

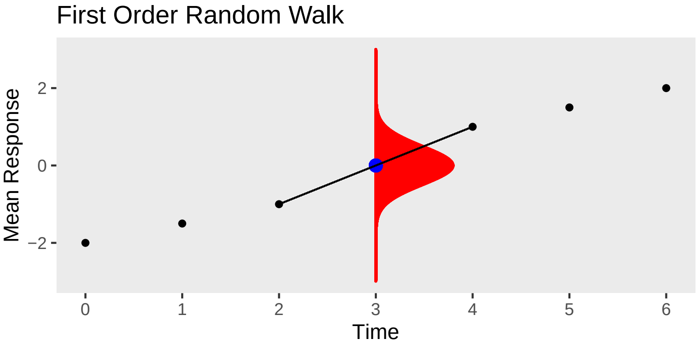

Figure 8: Illustration of the RW1 model for smoothing at time 3. The mean of the smoother is the average of the two adjacent points (and is highlighted as ●), and deviations from this mean are penalized via the normal distribution shown in **red**.

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:

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

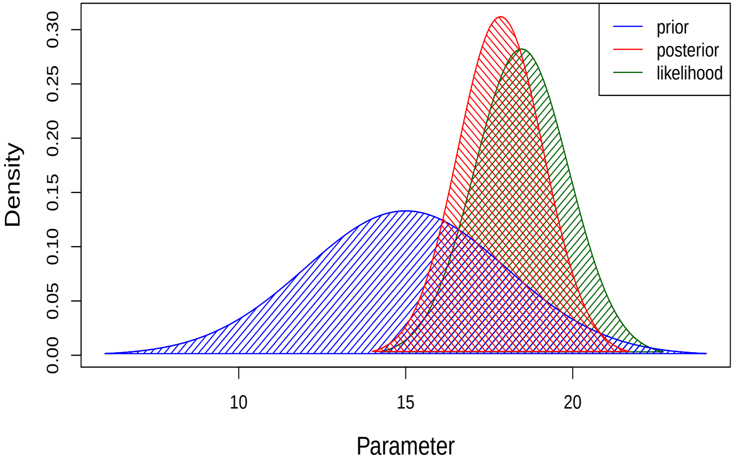

Figure 9: Normal data model with $n = 10$, $\overline{y} = 19.3$ and standard error 1.41. The prior for μ has mean $m = 15$ and $v = 3^2$. The posterior for the parameter μ is a compromise between the two sources of information: the posterior mean is 18.5 and the posterior standard deviation is 1.28.

RW Fitting to Simulated Data

- 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.
- 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.

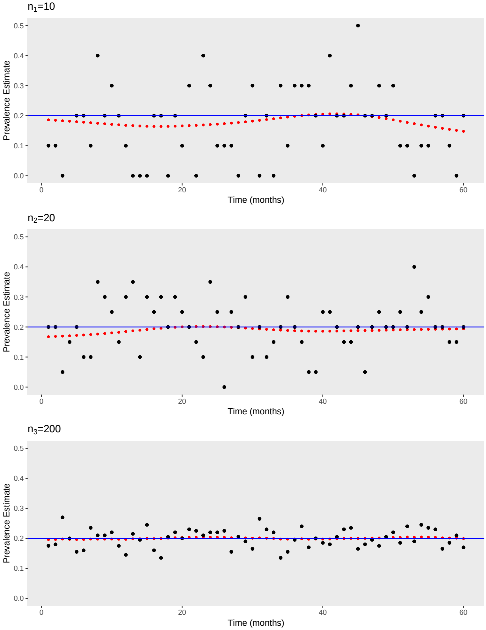

Figure 10: Prevalence estimates over time from simulated data, true prevalence $p = 0.2$ (blue solid lines). Smoothed random walk estimates in **red**.

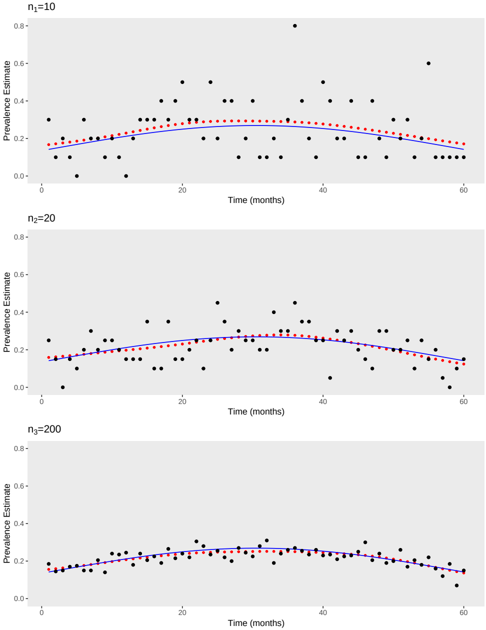

Figure 11: Prevalence estimates over time from simulated data, true prevalence corresponds to curved blue solid line. Smoothed random walk estimates in **red**.

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 Smoothing of Simulated Data for $n = 50$ Case

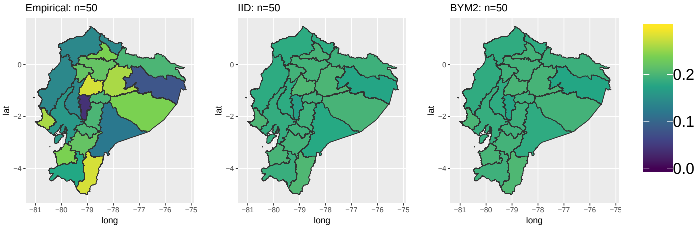

Figure 12: Left: raw proportions are on the left, middle: smoothing with IID model, right: smoothing with BYM2. True prevalence is 0.2.

Spatial Smoothing of Simulated Data for $n = 100$ Case

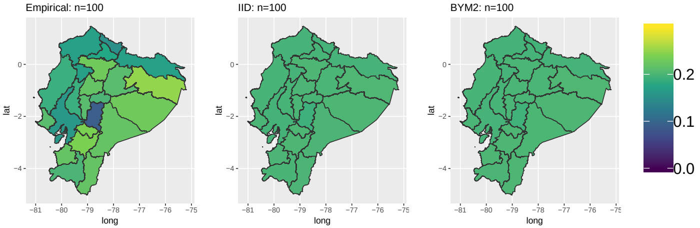

Figure 13: Left: raw proportions are on the left, middle: smoothing with IID model, right: smoothing with BYM2. True prevalence is 0.2.

Spatial Smoothing of Simulated Data for $n = 50$ Case

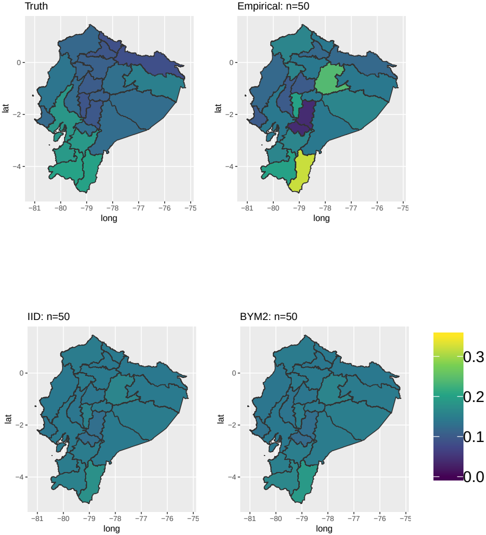

Figure 14: Top left: Truth, Top right: Raw proportions, Bottom left: smoothing with IID model, Bottom right: smoothing with BYM2.

Spatial Smoothing of Simulated Data for $n = 100$ Case

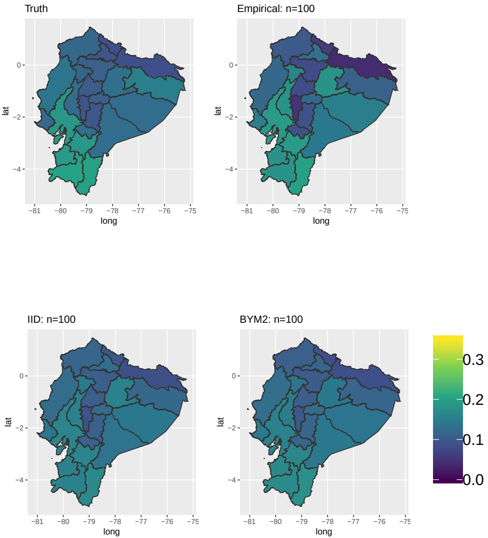

Figure 15: Top left: Truth, Top right: Raw proportions, Bottom left: smoothing with IID model, Bottom right: smoothing with BYM2.

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

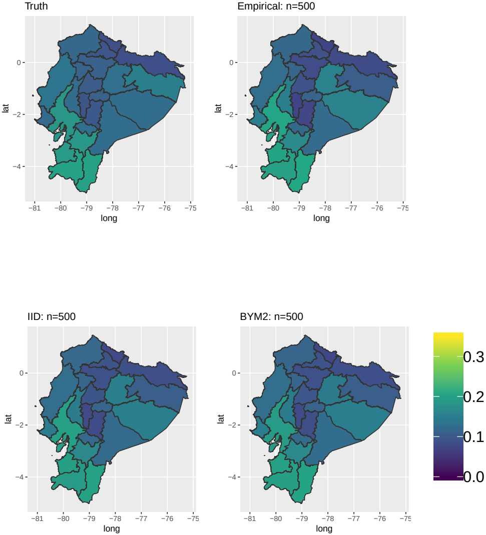

Figure 16: Top left: Truth, Top right: Raw proportions, Bottom left: smoothing with IID model, Bottom right: smoothing with BYM2.

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.

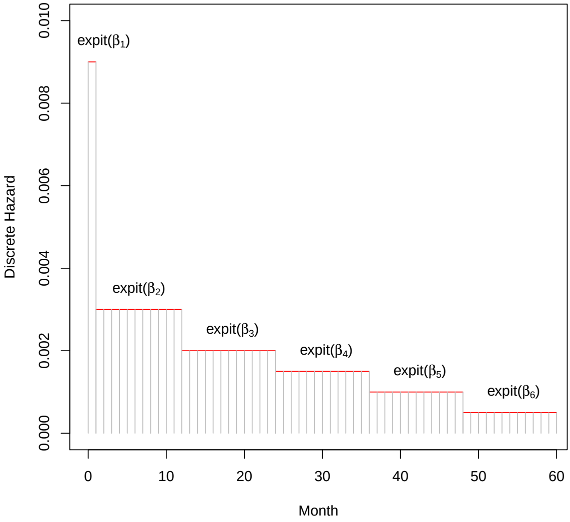

Figure 17: Representation of the conditional probabilities (hazards) of death by month.

The SUMMER Package

Spatial Smoothing Model

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

Data Model: Specifically, we take as data in area **area** i, the logit of the weighted estimates: $y_i = \text{logit}(\widehat{P}_i)$

$$\underbrace{y_i \mid \mu_i \sim \text{Normal}\,(\mu_i, V_i)}_{\text{Survey design acknowledged here}}$$

where V_i is the design variance.

Smoothing Model:

$$\mu_i \quad = \quad \underbrace{\alpha}_{\text{Intercept}} + \underbrace{S_i}_{\text{Spatial}} + \underbrace{\epsilon_i}_{\text{Independent}}$$

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.

Hierarchical Model:

1. The Data Model:

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

2. The Space-Time (Random Effects) Prior:

$$\underbrace{\lambda_{it} = f(\text{ space } i, \text{ time } t)}_{\text{Smoothing here}} .$$

Different space (e.g., area-based or pixel-based) and time (e.g. random walks, splines) smoothing models can be slotted into this framework.

Space-Time Smoothing Model

The data model is

$$y_{it}|\lambda_{it} \sim \mathrm{N}(\lambda_{it}, \widehat{V}_{it}),$$

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, and we emphasize that $\widehat{V}_{it}$ is known.

We decompose λ_{it} into temporal, spatial and space-time components:

$$\begin{aligned}
\lambda_{it} = \; & \underbrace{\mu}_{\text{Intercept}} \\
& + \underbrace{\alpha_t}_{\text{Independent}} + \underbrace{\gamma_t}_{\text{Random Walk}} && \text{Temporal Model} \\
& + \underbrace{\theta_i}_{\text{Independent}} + \underbrace{\phi_i}_{\text{Spatial}} && \text{Spatial Model} \\
& + \underbrace{\delta_{it}}_{\text{Interaction}} && \text{Space-Time Model}
\end{aligned}$$

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 using the model in the `SUMMER` package.
- Figures 18 and 19 give the estimates.

RW2 Model Applied to Ecuador Survey Data

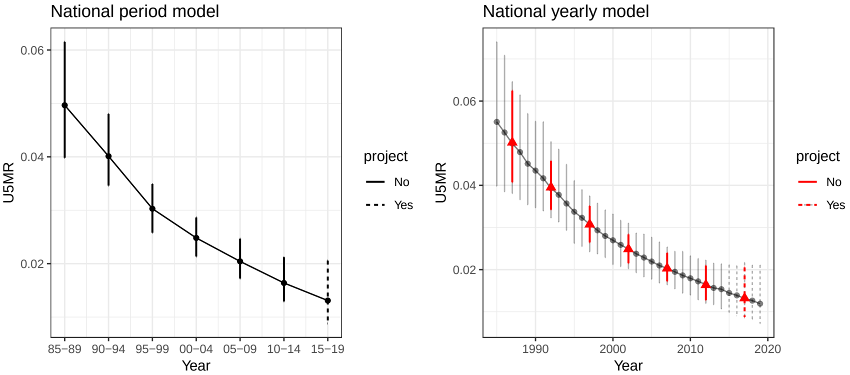

Figure 18: Yearly RW2 smoothing of weighted estimates of under-5 mortality in Ecuador, with 95% uncertainty intervals. On the left we apply to data aggregated over 5 years and on the right by 1 year. The dashed lines on the right of each plot are projections.

RW2 Model Applied to Ecuador Survey Data

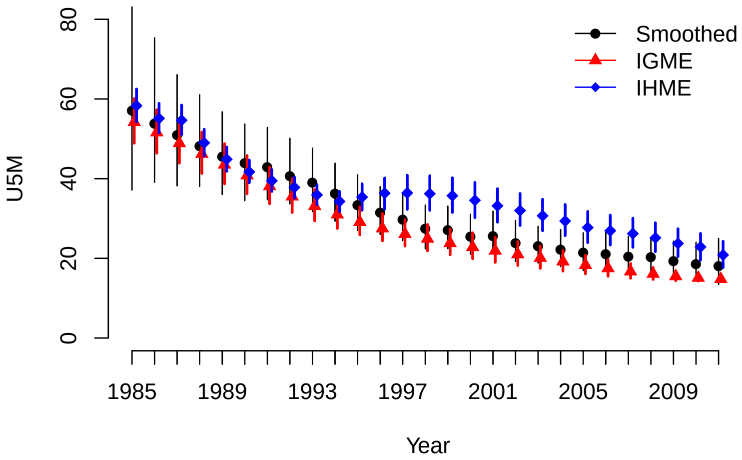

Figure 19: Yearly smoothed (with RW2) estimates of under-5 mortality in Ecuador, with 95% uncertainty intervals, along with IGME (B3) (with 90% intervals) and IHME (with 95% intervals) yearly estimates.

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

- Figure 20 gives the **weighted estimates** with hatching representing uncertainty.
- Figure 21 gives the **smoothed estimates** with hatching representing uncertainty – these estimates show less spatial variability and reduced uncertainty.
- Figure 23 clearly shows the drop in U5MR over time, and reduced **between-province variability**. The uncertainty in estimates is also apparent.

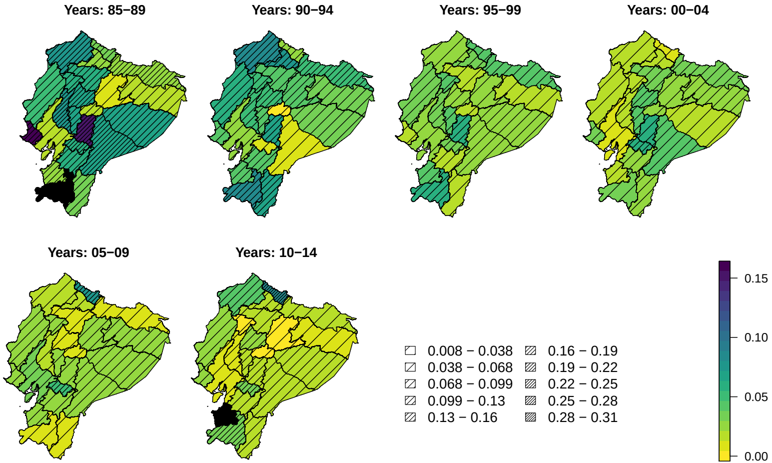

Figure 20: Yearly **weighted estimates** (from a discrete survival model) of under-5 mortality in Ecuador, with uncertainty indicated by density of hatching; more hatching → more uncertainty, with the latter measured though width of 95% uncertainty interval.

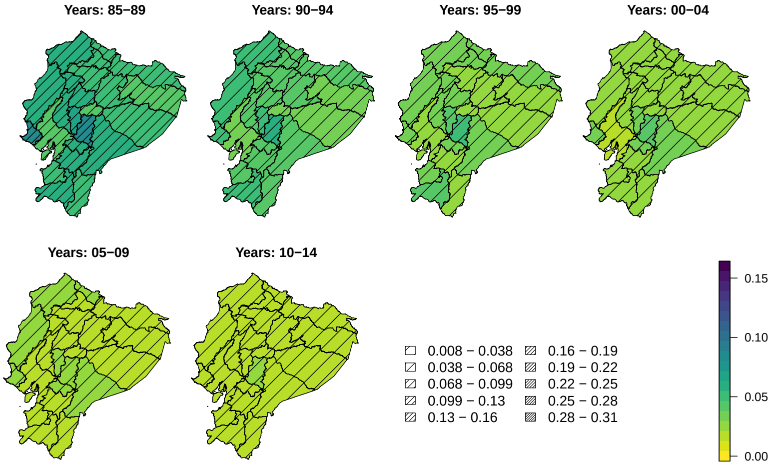

Figure 21: Yearly **smoothed estimates** (from a discrete survival model) of under-5 mortality in Ecuador, with uncertainty indicated by density of hatching; more hatching → more uncertainty, with the latter measured though width of 95% uncertainty interval.

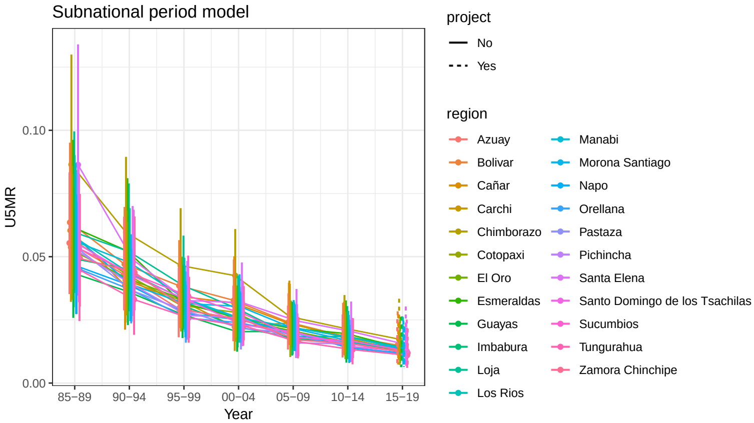

Figure 22: Five-yearly smoothed estimates (from a discrete survival model) of under-5 mortality in Ecuador, by province, with 95% uncertainty intervals

Final Thoughts

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.

An Example of Continuous Surface Modeling

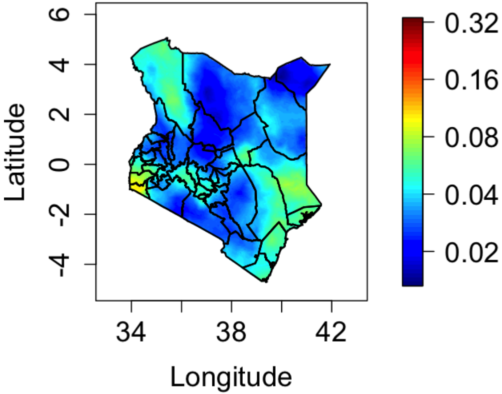

Figure 23: Continuous surface reconstruction of U5MR for Kenya in 2015

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**.
- Would like to include **geographical variables** in the model, to understand spatial inequality.
- Possible to make **HIV adjustments**, where needed.

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.

Background Literature

- Ecuador National Health and Nutrition Survey (ENSANUT-ECU) (Freire *et al.*, 2015).
- 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., York, J., and Mollié, A. (1991). Bayesian image restoration with two applications in spatial statistics. *Annals of the Institute of Statistics and Mathematics*, **43**, 1–59.

Blangiardo, M. and Cameletti, M. (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*. Population Center, Laboratories for Population Statistics, Chapel Hill, North Carolina.

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

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

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

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

Li, Z. R., Hsiao, Y., Godwin, J., Martin, B. D., Wakefield, J., and Clark, S. J. (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., Rue, H., and Lindström, J. (2011). An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic differential equation approach (with discussion). *Journal of the Royal Statistical Society, Series B*, **73**, 423–498.

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

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

Rue, H., Martino, S., and Chopin, N. (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., Fuglstad, G.-A., Riebler, A., Godwin, J., Wilson, K., and Clark, S. (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., Yue, Y., and Faraway, J. J. (2018). *Bayesian Regression Modeling with INLA*. Chapman and Hall/CRC.

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

Wilson, K. and Wakefield, J. (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.

Smoothed Direct Estimates

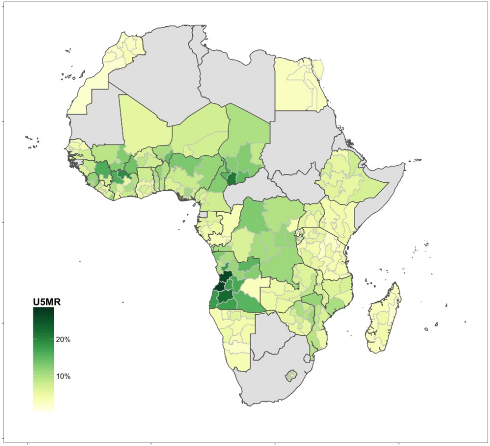

Figure 24: Predictions of U5MR for 2015, in 35 countries of Africa.

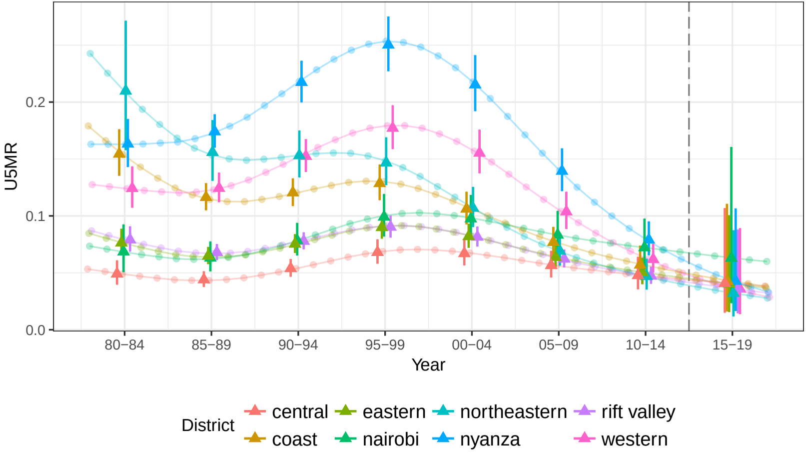

Figure 25: Posterior median estimates for Kenya districts.

IHME U5MR Estimates

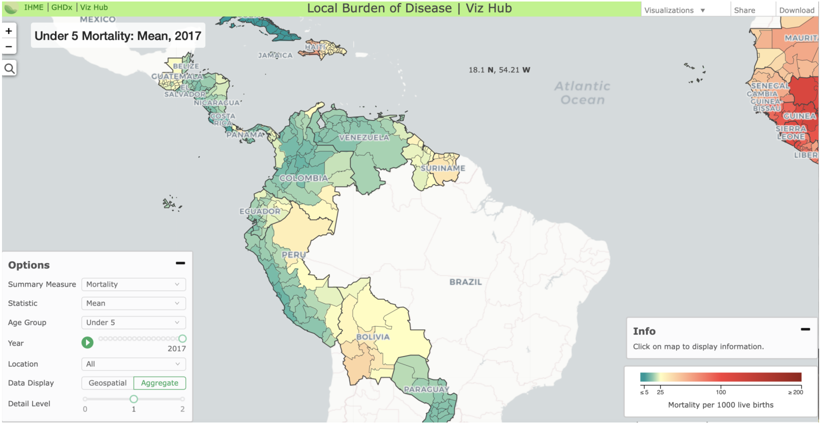

Figure 26: IHME estimates from a continuous space model; summarized at Admin1 level.