

BAYESIAN SAE WITH AN EMPHASIS ON LOW- AND MIDDLE-INCOME COUNTRIES: LECTURE 6: SAE4HEALTH SOFTWARE

Jon Wakefield

Departments of Statistics and Biostatistics
University of Washington

SOFTWARE OVERVIEW

FUTURE WORK

SOFTWARE OVERVIEW

THE MOTHERSHIP

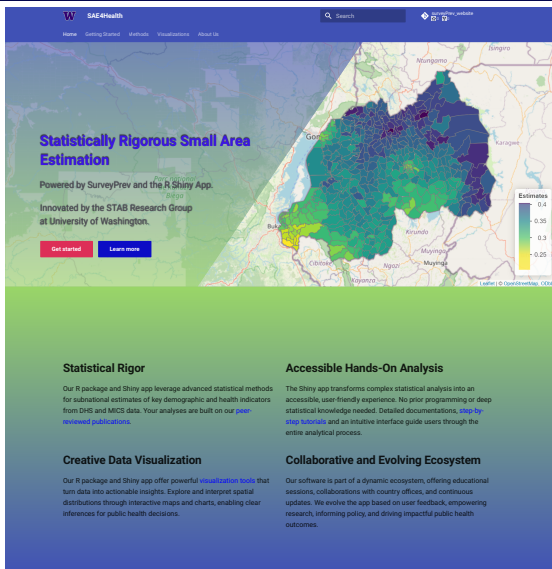


FIGURE 1: <https://sae4health.stat.uw.edu/>

MOTIVATION AND OVERVIEW

- ▶ While there are many R packages that can be used to perform SAE, they do not have mapping capabilities or graphical interfaces.
- ▶ The Space Time Analysis Bayes (STAB) group was always focused on SAE in low- and middle-income countries (LMICs) and collaborations with National Statistical Offices (NSOs), government officials and international agencies (UNICEF, WHO) based in those countries.
- ▶ Hence, **computational efficiency** and **user-friendly software** are key.
- ▶ Integrated Nested Laplace Approximation (INLA) provides the former, and we have been working on the latter for over 10 years.

- ▶ **SUMMER**: original package primarily written by Zehang Richard Li for space-time modeling of the under-5 mortality rate.
 - ▶ Initially using area-level and unit-level models with BYM2 and RW smoothing in space and time.
 - ▶ Range of models not extensively extended.
 - ▶ A large array of plotting facilities.
- ▶ **surveyPrev**: Primarily designed for DHS data — Qianyu Dong was the driver.
- ▶ **surveyMICS**: Zihang Yu and Josh Yang.
- ▶ In Thomas Lumley's **survey** package and written by Peter Gao: `smoothArea()` and `smoothUnit()` functions.

Web-based apps with all computation done in the cloud:

- ▶ R Shiny App for DHS data: All countries with DHS data since 2000 and 165 indicators – brainchild of Yunhan Wu and Victoria Knutson.
- ▶ R Shiny App for MICS data: 9 countries with MICS data and 3 indicators – Jieyi Xu.

Pre-modeled Estimates App:

- ▶ 12 countries and 20 indicators – Jieyi Xu and Qianyu Dong.

Umbrella term for all methods and software is [sae4health](#)

- ▶ Designed for non-expert users.
- ▶ 165 indicators currently, and more being added all the time – soon several hundred to be added, with the help of Claude Code.
- ▶ At the moment we can analyze DHS and MICS data, more types of survey to come (harmonization!).
- ▶ We have all the DHS data that users can access (not download!).
- ▶ Successfully used at WHO and UNICEF in-person and online workshops – courses given to NSOs, government officials, UNICEF and WHO officials from around 20 countries in sub-Saharan Africa and Latin America.

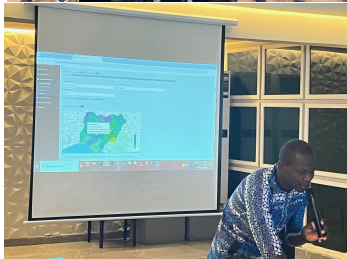
INGREDIENTS OF ALL SOFTWARE

- ▶ Statistically peer-reviewed methodology.
- ▶ Direct estimates along with **area-level** and **unit-level models** – computation is extremely fast.
- ▶ Extensive **graphical summaries** and report generation, in different languages – more languages to be added.
- ▶ Advanced features included **nested models** and **covariate modeling**.
- ▶ **Data sparsity checks** to start.

Over the years:

- ▶ NIH R01 grants.
- ▶ UNICEF.
- ▶ UN IGME.
- ▶ WHO.
- ▶ The Gates Foundation.

COUNTRY ENGAGEMENT



- ▶ The point of this work is to put SAE software into the hands of researchers and NSOs in LMICs.
- ▶ Countries exposed to methods and software, either through collaborations or workshops, in-person or remote:
- ▶ Benin
- ▶ Burkina Faso
- ▶ Central African Republic
- ▶ Côte d'Ivoire
- ▶ Democratic Republic of the Congo
- ▶ Guinea-Bissau
- ▶ Kenya
- ▶ Nigeria
- ▶ Rwanda
- ▶ Senegal
- ▶ Sierra Leone
- ▶ United Republic of Tanzania
- ▶ Zambia

FUTURE WORK

NEXT STEPS FOR SOFTWARE

- ▶ Expand the number of MICS indicators.
- ▶ Aggregation from Admin2 $\rightarrow$ Admin1 $\rightarrow$ National, using WorldPop.
- ▶ SPDE for continuous models.
- ▶ Bivariate plots of different indicators.
- ▶ Temporal plotting and modeling of indicators from multiple surveys.
- ▶ Model checking/diagnostics.
- ▶ New data types.
- ▶ Using AI for various endeavors.

NEXT STEPS FOR SOFTWARE

- ▶ Nationally benchmarked child mortality survival curves over time (UN-IGME).
- ▶ Fertility modeling.
- ▶ Covariate modeling:
 - ▶ Suite of available covariates.
 - ▶ Errors-in-variables.
 - ▶ Flexible machine learning models, with theoretically valid uncertainty estimates.
- ▶ And most important: **Country engagement!!!**
- ▶ Through collaborations including joint software development and workshops

REFERENCES