

Self-Tuned Rejection Sampling within Gibbs and a Case Study in Small Area Estimation

Guide to Replication Materials

Andrew M. Raim, Kyle M. Irimata, and James A. Livsey

1 Introduction

This document describes replication materials for the article “Self-Tuned Rejection Sampling within Gibbs and a Case Study in Small Area Estimation”. It includes an R package, code for two simulations, and a data analysis. The following folders are included in the materials.

- **01-sim-cond**: simulation generating from a single conditional of the SAE model under consideration.
- **02-sim-post**: simulation generating from the posterior of the SAE model under consideration.
- **03-sae**: data analysis with the SAIPE dataset.
- **data**: the SAIPE dataset and codes /inputs to assemble it.
- **vws**: An R package that supports construction of VWS samplers.
- **saevws**: an R package for the Gibbs sampler and supporting functions. It can be configured with basic vertical weighted strips (VWS), self-tuned VWS, and independent Metropolis-Hastings (IMH) for the $\sigma_1^2, \dots, \sigma_m^2$ steps.

A working installation of R ([R Core Team 2025](#)) is required with the following packages (and their dependencies). These packages are available on CRAN at the time of this writing.

- **devtools** ([Wickham et al. 2022](#)),
- **fnt1** ([Raim 2024](#))
- **jsonlite** ([Ooms 2014](#))
- **knitr** ([Xie 2014](#))
- **mcmcse** ([Flegal et al. 2021](#))
- **Rcpp** ([Eddelbuettel and Balamuta 2018](#)),
- **sae** ([Molina and Marhuenda 2015](#))
- **statmod** ([Smyth 2005](#))
- **tidyverse** ([Wickham et al. 2019](#)),
- **xtable** ([Dahl et al. 2019](#))

Additionally, the included **vws** package must be installed.

```
R> devtools::install_local("/path/to/vws")
```

Section 2 gives an overview of the **saevws** package, including installation and basic usage. Section 3 lists the scripts used for the conditional simulation, posterior simulation, and data analysis. Section 4 gives a comprehensive list of packages and versions used in the computational studies, for reproducibility.

2 saevws R Package

The `saevws` package can be installed using an R command like the following.

```
R> devtools::install_local("/path/to/saevws")
```

The package exports several functions for use in R:

- `gibbs` invokes the Gibbs sampler.
- `get_gibbs_control` constructs a control object with additional arguments for the Gibbs sampler.
- `get_vws_control` constructs a control object with additional arguments for VWS within the Gibbs sampler.
- `get_init` constructs an object with initial values for the Gibbs sampler.
- `get_fixed` constructs an object to indicate that certain elements of the MCMC chain are to remain fixed in the Gibbs sampler. This is primarily used for testing the code.
- `traceplots` is a utility to make trace plots from a matrix with draws.

These functions' interfaces are shown below; see the associated manual pages for details about arguments, outputs, etc..

```
gibbs =  
function (y, s2, X, Z, df, init = get_init(m = length(y), d1 = ncol(X),  
      d2 = ncol(Z)), control = get_gibbs_control(), fixed = get_fixed())  
  
get_gibbs_control =  
function (R = 1000, burn = 0, thin = 1, report = R + 1, save_latent = integer(0),  
      vws = get_vws_control())  
  
get_vws_control =  
function (tol1 = 0.01, tol2 = exp(-100), max_rejects = 1e+06, method = c("vws-tune",  
      "vws-basic", "imh"), N = 50)  
  
get_init =  
function (m, d1, d2, beta = NULL, gamma = NULL, phi2 = NULL, tau2 = NULL,  
      sigma2 = NULL, theta = NULL)  
  
get_fixed =  
function (beta = FALSE, gamma = FALSE, phi2 = FALSE, tau2 = FALSE, sigma2 = FALSE,  
      theta = FALSE)  
  
traceplots =  
function (x, xlab = "")
```

3 Scripts

The following codes in `01-sim-cond` produce the conditional simulation study in Section 5.1 of the article.

- `metro.R`: Simulations with IMH.
- `vws.R`: Simulations with VWS.
- `samplers.cpp`: C++ functions for samplers consisting of one IMH step and one VWS step.

The following codes in `02-sim-post` produce the posterior simulation study in Section 5.2 of the article.

- `gen.R`: Generate folder structure for the study. Each level of the study is placed into a separate folder.

- `sim.R`: Run one level of the study.
- `analyze.R`: Analyze the results from the completed study and produce summaries.

The `03-sae` folder contains a single script `saipe.R` that produces the data analysis in Section 6 of the article.

The `data` folder contains scripts and inputs to create the example SAIPE dataset.

- `saipe.csv`: the prepared dataset.
- `cntysnap.csv`: Supplemental Nutrition Assistance Program (SNAP) data for 2022, to be used as a covariate with 2023 SAIPE data.
- `co-est2023-alldata.csv`: 2023 population estimates from the U.S. Census Bureau’s Population Estimates Program. The data may be obtained from the Census Bureau [website](#).
- `build-data.R`: Code to construct the dataset `saipe.csv`.

4 Session Information

Results in the paper were produced in an R environment with the following packages and versions.

```
library(devtools)
library(tidyverse)
library(Rcpp)
library(knitr)
library(mcmcse)
library(jsonlite)
library(xtable)
library(vws)
library(statmod)
```

```
sessionInfo()
```

R version 4.4.0 Patched (2024-05-13 r86547 ucrt)

Platform: x86_64-w64-mingw32/x64

Running under: Windows 11 x64 (build 26100)

Matrix products: default

locale:

```
[1] LC_COLLATE=English_United States.utf8
[2] LC_CTYPE=English_United States.utf8
[3] LC_MONETARY=English_United States.utf8
[4] LC_NUMERIC=C
[5] LC_TIME=English_United States.utf8
```

time zone: America/New_York

tzcode source: internal

attached base packages:

```
[1] stats      graphics  grDevices  utils      datasets  methods    base
```

other attached packages:

```
[1] saevws_0.1.0
```

loaded via a namespace (and not attached):

[1] vctrs_0.6.5	cli_3.6.5	knitr_1.50	rlang_1.1.6
[5] xfun_0.53	generics_0.1.4	S7_0.2.0	jsonlite_2.0.0
[9] glue_1.8.0	htmltools_0.5.8.1	brio_1.1.5	scales_1.4.0
[13] rmarkdown_2.29	grid_4.4.0	vws_0.2.0	tibble_3.3.0
[17] evaluate_1.0.5	fastmap_1.2.0	yaml_2.3.10	lifecycle_1.0.4
[21] compiler_4.4.0	dplyr_1.1.4	RColorBrewer_1.1-3	testthat_3.2.3
[25] fftwtools_0.9-11	Rcpp_1.1.0	pkgconfig_2.0.3	rstudioapi_0.17.1
[29] farver_2.1.2	digest_0.6.37	R6_2.6.1	tidyselect_1.2.1
[33] pillar_1.11.0	magrittr_2.0.3	mcmcse_1.5-0	ellipse_0.5.0
[37] tools_4.4.0	gtable_0.3.6	ggplot2_4.0.0	

References

- Dahl, David B., David Scott, Charles Roosen, Arni Magnusson, and Jonathan Swinton. 2019. *xtable: Export Tables to LaTeX or HTML*. <https://CRAN.R-project.org/package=xtable>.
- Eddelbuettel, Dirk, and James Joseph Balamuta. 2018. “Extending R with C++: A Brief Introduction to Rcpp.” *The American Statistician* 72 (1): 28–36. <https://doi.org/10.1080/00031305.2017.1375990>.
- Flegal, James M., John Hughes, Dootika Vats, Ning Dai, Kushagra Gupta, and Uttiya Maji. 2021. *mcmcse: Monte Carlo Standard Errors for MCMC*. Riverside, CA,; Kanpur, India.
- Molina, Isabel, and Yolanda Marhuenda. 2015. “sae: An R Package for Small Area Estimation.” *The R Journal* 7 (1): 81–98. <https://journal.r-project.org/archive/2015/RJ-2015-007/RJ-2015-007.pdf>.
- Ooms, Jeroen. 2014. “The jsonlite Package: A Practical and Consistent Mapping Between JSON Data and R Objects.” *arXiv:1403.2805 [Stat.CO]*. <https://arxiv.org/abs/1403.2805>.
- R Core Team. 2025. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org>.
- Raim, Andrew M. 2024. *Fntl: Numerical Tools for 'Rcpp' and Lambda Functions*. <https://doi.org/10.32614/CRAN.package.fntl>.
- Smyth, Gordon K. 2005. “Numerical Integration.” *Encyclopedia of Biostatistics*, 3088–95.
- Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D’Agostino McGowan, Romain François, Garrett Grolemond, et al. 2019. “Welcome to the tidyverse.” *Journal of Open Source Software* 4 (43): 1686. <https://doi.org/10.21105/joss.01686>.
- Wickham, Hadley, Jim Hester, Winston Chang, and Jennifer Bryan. 2022. *devtools: Tools to Make Developing R Packages Easier*. <https://CRAN.R-project.org/package=devtools>.
- Xie, Yihui. 2014. “Knitr: A Comprehensive Tool for Reproducible Research in R.” In *Implementing Reproducible Computational Research*, edited by Victoria Stodden, Friedrich Leisch, and Roger D. Peng. Chapman; Hall/CRC.