Bayesian computation in R

By Hedibert F. Lopes, Insper

May 29th 2026

This document contains excerpt from two books and a paper on software for Bayesian Analyis, as well as a short description of several of these packages including the ones I have used in an undergraduate in economics ellective course this spring named *Bayesian Dynamic Modeling in Macro & Finance*.

Gamerman, D., Lopes, H.F., & Gonçalves, F.B. (2026). *MCMC: Stochastic Simulation for Bayesian Inference (3rd edition)*. Taylor & Francis – Section 5.6 Software for Bayesian Anaysis

"This section starts with the nicely organised and relatively old review chapter by Goel (1988) that appeared in the *Bayesian Statistics 3*, the brick-like Proceedings of the Third Valencia International Meeting (June 1–5, 1987). The chapter is on "Software for Bayesian Analysis: Current Status and Additional Needs":

> *The starting point for this article was the workshop on Bayesian Statistical Computing, which was held at The Ohio State University during May 1986 and attended by approximately 40 people involved in statistical computing, Bayesian decision analysis and AI. We organised this workshop to Bayesian analysis software. We believe that Bayesian methodology will be used routinely by a widespread group of data analysts and scientists if a user-friendly, general purpose Bayesian Statistical Analysis Package was available which could be used for classroom teaching as well as for experimental data analysis.*

Goel continues by saying that

> *An overwhelming majority agreed with the notion that it is too early to push for a new statistical package. Instead, we should strive for new Bayesian software to be compatible with a package like "S", in order to use its data handling and graphics capabilities.*

Despite being a 40-year-old paper, Goel's chapter looks quite contemporaneous regarding the constant, pressing need to make Bayesian data analysis, inference, model criticism, and decision analysis more readily available to a broader community of scientists, data analysts, and decision makers. Goel says that, at that time,

> *Only a few people have devoted their energy to developing general purpose Bayesian Analysis software. We believe that this must change quickly if we want to see the Bayesian the 21st Century.*

Here is a short list taken from Goel (1988) cataloguing the first attempts to make Bayesian statistical software more accessible, contributed by prestigious Bayesian scholars, including Arnold Zellner, Enrique de Alba, Edward Leamer, Adrian Smith, Herman van Dijk, Luke Tierney and Mark Schervish, among many others: CADA (general purpose computer assisted data analysis monitor), BRAP (Bayesian regression analysis program), SHAZAM (General econometrics program), BTS (Bayesian time series), BAYES FOUR (Interactive numerical computation of multiple integrals), BAYES 3/3D, and SBAYES.

See also Press (1980) for an earlier attempt to list and discuss the available software for Bayesian computation, about a decade before Goel (1988), and the more recent review published in the journal *Statistical Science* by Štrumbelj et al. (2024) that discusses the past, the present, and the future of software for Bayesian inference. They address Goel's main concern regarding the lack of devoted energy for software development by saying that

> *In the past three decades, Bayesian inference has established itself as a viable alternative to more classical approaches to statistical inference and is now a must-have tool for every statistician's toolbox. Many theoretical and methodological developments have contributed to the success of Bayesian statistics. However, no development has been as important for mass adoption as was the emergence of accessible and robust software.*

S, S-Plus, and R packages

During the last three and a half decades, since the appearance of the seminal papers of Tanner and Wong (1987) and Gelfand and Smith (1990), software development, at various levels of sophistication, has been part of the agenda of any applied Bayesian statistician, biostatistician, econometrician, and other Bayesian practitioners both in academia and in the industry.

The explosion of Bayesian computation is partially due to statistical packages like S (mentioned above), S-Plus (its commercial version) and R (open-source version). S grew up in the statistics research departments at AT&T Bell Labs, starting in 1976, a department which was heavily influenced by John Tukey's approach to Exploratory Data Analysis. Its main developers were Richard Becker and John Chambers, among others. See Becker (1994) for a brief history of the S package.

The S package inspired the appearance of the commercially available S-plus language. Soon after, an open-source alternative became available, the R package, which almost immediately led to the development of a large, active, and vibrant community of contributors. Chronologically, between 1991 and 1996, Ross Ihaka and Robert Gentleman created the R package in the Department of Statistics at the University of Auckland, then the first announcement of R was made to the public, and a paper appeared in the prestigious Journal of Computational and Graphical Statistics: Ihaka and Gentleman (1996). In 1997 the R Core Group was formed and in 2000 R version 1.0.0 was released to the public. From a historical perspective, the second author arrived at Duke in August 1996 for his PhD, already proficient in S-Plus, C, C++, and Fortran 90. In 1997, Peter Müller introduced him and his fellow students to a beta version of R; we have been dedicated users and developers ever since.

It is an understatement to say that R has become the most important and used statistical environment in academia and industry for the pressing needs stated above with respect to data analysis and decision making. Nevertheless, below, we briefly review the two most influential probabilistic programming languages that appeared over the last 30 plus years: BUGS and Stan.

Probabilistic programming languages

Practical Bayesian data analysis and decision making have benefited from probabilistic programming languages, such as BUGS and Stan, both of which interacted with the R package, and can be thought of as the representatives of two generations of probabilistic programming languages for MCMC-based Bayesian inference, with BUGS based on the Gibbs sampler (and extensions) and Stan mainly based on HMC algorithms. Štrumbelj et al. (2024) says that

> *General-purpose Bayesian computation has had two distinct periods, each dominated by a certain type of Bayesian computation and software. From the early 1990s to the 2010s, it was Gibbs sampling and the quintessential representative of software is BUGS. From the 2010s up to now, it is HMC and the quintessential representative is Stan.*

The Bayesian Analysis Using Gibbs Sampler (BUGS) project was developed in 1989 at the Medical Research Council Biostatistics Unit in Cambridge, UK. The BUGS project has evolved through several versions, including Classic BUGS, WinBUGS and OpenBUGS, and the R interface package R2WinBugs. WinBUGS updates the BUGS language and the MCMC algorithms, while OpenBUGS runs on Linux. BUGS, WinBUGS, and OpenBUGS are no longer being developed. See Lunn et al. (2000), Lunn et al. (2009) and Spiegelhalter et al. (2014) for further details.

Similarly, Stan is named after Stanislaw Ulam, one of the fathers of the Monte Carlo revolution, starting with Metropolis and Ulam (1949) who introduced the Monte Carlo method. See Carpenter et al. (2017), for further details about Stan. For more historical accounts of Stanislaw Ulam's contributions to the Monte Carlo revolution, see Ulam (1987), Metropolis (1987) and Eckhardt (1987), all part of the special issue *Los Alamos Science: Special Issue, Stanislaw Ulam 1909-1984*."

Agresti, A., Kateri, M., Grove, R., & Mira, A. (2026). *Foundations of Bayesian Statistics for Data Scientists*. Taylor & Francis - Appendix A.1.1 - R Software Packages for Bayesian Data Analysis

"Many Bayesian statistical methods require computationally intensive algorithms, which has limited its widespread use and acceptance by data scientists and practitioners in the past. In the previous 30 years, Bayesian methods have gained increasing popularity due to the development of software for implementing them that uses simulation methods presented in Chapter 6 of this book. The earliest software package used *Gibbs sampling*: **`BUGS`** (Bayesian inference Using Gibbs Sampling) was introduced in 1989. More recent Gibbs-sampling based software is **`JAGS`** (Just

Another Gibbs Sampler), which is written in C++. **Nimble** is an extension of BUGS that also implements other Bayesian computation methods, such as Metropolis-Hastings and Hamiltonian Monte Carlo. All of these packages have interfaces for implementation in R.

One of the most widely used software platforms for Bayesian inference that uses modern computational methods is **Stan** (https://mc-stan.org), a powerful and flexible platform for Bayesian inference. Its R interface is provided through the **rstan** package. The **brms** (Bayesian Regression Models using "Stan") package that we used in Chapters 3–7 of this book to fit a wide variety of models uses Stan, thus giving data scientists a simple way to conduct many analyses without learning about Stan itself. It provides a simple user interface for rstan, with which the user can define the formula, data, prior distributions, and family of distributions for the response variable. It then transforms these elements to the form required in Stan modeling code and then executes Stan. Besides brms, another package that can implement Bayesian modeling is **rstanarm**, which contains pre-compiled modules to simplify the specification and implementation of Stan models, offering a consistent and very fast fit. rstanarm is easier to use, having a simple syntax, similar to familiar basic R modeling functions (such as lm and glm). However, it offers a fixed set of possible models and a restricted choice of prior distributions. By contrast, brms is more flexible and can handle a wider range of models, including more complex models such as non-linear and multivariate models. Furthermore, brms allows customized priors. With brms, the user can control nearly all aspects of model specification, and it has more extensive options for post-processing and model diagnostics.

Other R packages are popular for certain Bayesian data analyses. For example, **coda** provides diagnostics and analysis of MCMC output. The **BayesFactor** package computes Bayes factors for contingency tables, ANOVA designs, and linear regression. The **bayestestR** package provides a greater variety of indices describing the posterior distribution. The **bayesplot** packages provides visualization of posterior draws, MCMC diagnostics, and prediction checking. The **MCMCpack** package performs MCMC computations for a variety of statistical models. The **LearnBayes** package is a collection of functions helpful in learning the Bayesian approach to statistical inference."

Erik Štrumbelj, Alexandre Bouchard-Côté, Jukka Corander, Andrew Gelman, Håvard Rue, Lawrence Murray, Henri Pesonen, Martyn Plummer, and Aki Vehtari (2024) Past, Present and Future of Software for Bayesian Inference, *Statistical Science*, Vol. 39, No. 1, 46-61. https://doi.org/10.1214/23-STS907

Abstract. Software tools for Bayesian inference have undergone rapid evolution in the past three decades, following popularisation of the first generation MCMC-sampler implementations. More recently, exponential growth in the number of users has been stimulated both by the active development of new packages by the machine learning community and popularity of specialist

software for particular applications. This review aims to summarize the most popular software and provide a useful map for a reader to navigate the world of Bayesian computation. We anticipate a vigorous continued development of algorithms and corresponding software in multiple research fields, such as probabilistic programming, likelihood-free inference and Bayesian neural networks, which will further broaden the possibilities for employing the Bayesian paradigm in exciting applications.

Key words and phrases: Statistics, data analysis, MCMC, computation, probabilistic programming.

TABLE 2

Total RStudio [104] CRAN mirror downloads for Bayesian inference-related R packages referenced in this paper for the period between January 1, 2022 and December 31, 2022. We used the cranlogs package [26]. We include ggplot2 [135], the most popular R package for statistical graphics, as a baseline for comparison. While these counts should in most cases be a good proxy for relative popularity, we have to keep in mind that users can also download these packages from other CRAN mirrors or directly from code repositories. Also, some popular R packages are not available on CRAN, for example, R-INLA, cmdstanr, the R interface to Stan or R2MultiBUGS, the R interface to MultiBUGS

Package	Download count	Description
ggplot2	31,457,872	Create Elegant Data Visualisations Using the Grammar of Graphics
mgcv	1,523,237	Mixed GAM Computation Vehicle with Automatic Smoothness Estimation
coda	1,190,640	Output Analysis and Diagnostics for MCMC
rstan	993,086	R Interface to Stan
loo	738,325	Efficient Leave-One-Out Cross-Validation and WAIC for Bayesian Models
bayestestR	599,283	Understand and Describe Bayesian Models and Posterior Distributions
prophet	338,276	Automatic Forecasting Procedure
posterior	314,669	Tools for Working with Posterior Distributions
bayesplot	308,747	Plotting for Bayesian Models
bnlearn	286,003	Bayesian Network Structure Learning, Parameter Learning and Inference
shinystan	272,855	Interactive Visual and Numerical Diagnostics and Posterior Analysis for Bayesian Models
BayesFactor	239,538	Computation of Bayes Factors for Common Designs
rjags	228,433	Bayesian Graphical Models using MCMC
brms	215,302	Bayesian Regression Models using Stan
MCMCpack	186,124	Markov Chain Monte Carlo (MCMC) Package
rstanarm	164,469	Bayesian Applied Regression Modeling via Stan
bridgesampling	155,278	Bridge Sampling for Marginal Likelihoods and Bayes Factors
R2WinBUGS	61,926	Running WinBUGS and OpenBUGS from R SPLUS
nimble	36,471	MCMC, Particle Filtering and Programmable Hierarchical Modeling
abc	36,251	Tools for Approximate Bayesian Computation (ABC)
R2OpenBUGS	27,284	Running OpenBUGS from R
greta	8453	Simple and Scalable Statistical Modeling in R
abctools	6404	Tools for ABC Analyses
EasyABC	5344	Efficient Approximate Bayesian Computation Sampling Schemes

Some of the packages at a glance (by Hedibert F. Lopes)

The following notes summarize each of the software tools mentioned above, with the year each was first proposed and its principal author(s).

S. Developed from 1976 at AT&T Bell Laboratories, Murray Hill, NJ, by John Chambers, Richard A. Becker, and Allan Wilks. An interactive language and environment for data analysis and graphics -- the "S" stands for *statistics* -- strongly influenced by John Tukey's approach to Exploratory Data Analysis. It established the "data analysis as programming" paradigm that every later system inherited; John Chambers received the 1998 ACM Software System Award for it.

S-Plus. The commercial implementation of S, first released in 1988 by Statistical Sciences, Inc. (founded by R. Douglas Martin, University of Washington). It layered a graphical user interface, commercial support, and additional libraries on top of the S language. Ownership passed through MathSoft and Insightful Corporation before being acquired by TIBCO Software, where it was marketed as TIBCO Spotfire S+.

R. A free, open-source dialect of S created between 1991 and 1996 by Ross Ihaka and Robert Gentleman at the University of Auckland, New Zealand — the name plays both on "S" and on the shared first initial of its two authors. First announced publicly in 1993 and described in Ihaka & Gentleman (1996); the R Core Team formed in 1997, and version 1.0.0 was released in 2000. Distributed under the GNU General Public License and now maintained by the R Core Team and the R Foundation, it has become the dominant environment for statistical computing.

RStudio. An integrated development environment (IDE) for R, first released in 2011 by RStudio, Inc., founded by J. J. Allaire. It combines a code editor, console, plotting, debugging, and package and version-control tools in a single interface, available in open-source and commercial editions and in both desktop and server versions. The company was renamed Posit in 2022 to reflect its broadening support for other languages, notably Python.

BUGS (Bayesian inference Using Gibbs Sampling). Proposed in 1989 (first released as free software in 1991) by David J. Spiegelhalter and colleagues at the MRC Biostatistics Unit, Cambridge. It was the earliest general-purpose package for Bayesian inference via Gibbs sampling; later branches include WinBUGS (Lunn, Thomas, Best & Spiegelhalter, 2000), OpenBUGS, and MultiBUGS.

JAGS (Just Another Gibbs Sampler). Developed by Martyn Plummer and introduced in 2003 (first public release December 2007). Written in C++, it is a cross-platform reimplementation of the BUGS language designed as a flexible, open-source engine for MCMC analysis of Bayesian hierarchical models.

NIMBLE. An extension of the BUGS language proposed around 2014–2016 by Perry de Valpine, Daniel Turek, Christopher Paciorek, and collaborators. Beyond Gibbs sampling it supports other algorithms such as Metropolis–Hastings and Hamiltonian Monte Carlo, and lets users program their own statistical algorithms in R.

Stan. A probabilistic programming language released around 2012 and described in print by Bob Carpenter, Andrew Gelman, Matthew Hoffman, and co-authors (2017). It uses modern gradient-based methods, principally the No-U-Turn variant of Hamiltonian Monte Carlo. Its R interface is provided through the rstan package.

brms (Bayesian Regression Models using “Stan”). Created by Paul-Christian Bürkner and introduced in 2017 (Journal of Statistical Software). It offers a high-level formula interface that translates model specifications into Stan code, supporting a wide range of (non-)linear, multivariate, and multilevel models with customizable priors.

rstanarm. Developed by the Stan Development Team (Goodrich, Gabry, Ali, Brilleman) from around 2016. It supplies pre-compiled Stan models behind a syntax mirroring familiar R functions such as lm and glm, giving fast, consistent fits at the cost of a fixed model and prior menu.

coda (Convergence Diagnosis and Output Analysis). Authored by Martyn Plummer, Nicky Best, Kate Cowles, and Karen Vines (2006). It provides convergence diagnostics and summary/analysis tools for MCMC output.

BayesFactor. Developed by Richard D. Morey and Jeffrey N. Rouder (first released around 2012). It computes Bayes factors for common designs, including contingency tables, ANOVA, and linear regression.

bayestestR. Part of the easystats ecosystem, introduced in 2019 by Dominique Makowski, Mattan S. Ben-Shachar, and Daniel Lüdecke. It provides a broad set of indices for describing and testing posterior distributions.

bayesplot. Created by Jonah Gabry and the Stan team (around 2017). It offers plotting functions for posterior draws, MCMC diagnostics, and posterior-predictive checking, built on ggplot2.

MCMCpack (Markov Chain Monte Carlo in R). Introduced by Andrew D. Martin, Kevin M. Quinn, and Jong Hee Park (2011, Journal of Statistical Software). It performs posterior simulation for many standard statistical models and includes assorted MCMC utility functions.

LearnBayes. Written by Jim Albert to accompany his textbook on Bayesian computation with R (package available from the mid-2000s). It is a teaching-oriented collection of functions for learning the Bayesian approach to inference.

List of R packages presented to the junior/senior undergrad econ students for the elective course *Bayesian Dynamic Modeling in Macro & Finance*.

Bayesian regression

Package	Summary	Typical applications
`Brms (already listed above)`	High-level interface to Stan for flexible Bayesian regression and hierarchical models.	General regression, mixed-effects models, psychometrics, ecology, health sciences.
`Rstanarm (already listed above)`	Simplified Stan interface for standard regression models using R formula syntax.	Applied regression modeling, logistic/probit regression, teaching Bayesian methods.

`bayesreg`	Bayesian linear/logistic regression with shrinkage priors (ridge, lasso, horseshoe).	High-dimensional regression, variable selection, predictive modeling.
`sparsevb`	Sparse regression via variational Bayes inference.	Fast approximate inference for large-scale, sparse linear models.

Time series & dynamic models

Package	Summary	Typical applications
`bsts`	Bayesian structural time series with trend, seasonality, and spike-and-slab regression.	Forecasting, causal impact analysis, marketing analytics, anomaly detection.
`bayesSSM`	Particle MCMC inference for nonlinear/non-Gaussian state-space models.	Complex dynamic systems, ecological/epidemiological time series, signal processing.
`dlbayes`	Bayesian dynamic linear models for time-varying regression and forecasting.	Forecasting with state-space models, real-time updating, econometrics.
`kDGLM`	Dynamic generalized linear models with Kalman filter inference for non-Gaussian data.	Count/time series modeling (e.g., Poisson, binomial), control systems.
`BayesARIMAX`	Bayesian ARIMA/ARIMAX modeling for forecasting with exogenous regressors.	Economic and financial forecasting with explanatory variables.
`bayesforecast`	Stan-based Bayesian framework for ARIMA, VAR, and related time-series models.	General-purpose Bayesian forecasting, macroeconomic and financial time series.

Time-varying parameter, VAR & macroeconomic models

Package	Summary	Typical applications
`shrinkTVP`	Bayesian regression with time-varying coefficients and hierarchical shrinkage.	Dynamic regressions, evolving relationships in economics and finance.
`shrinkTVPVAR`	Time-varying parameter VAR models with global-local shrinkage priors.	Macroeconomic forecasting, monetary policy analysis, structural VARs.
`bsvars`	Bayesian structural VAR estimation with flexible identification and priors.	Macroeconomic impulse response analysis, causal inference in time series.

Volatility & financial econometrics

Package	Summary	Typical applications
`stochvol`	Univariate stochastic volatility modeling via efficient MCMC.	Asset return volatility, risk modeling, financial econometrics.
`factorstochvol`	Multivariate factor stochastic volatility model estimation.	Portfolio risk modeling, multivariate volatility estimation.
`bsarsv`	Stochastic autoregressive volatility modeling with Bayesian inference.	Financial time series with persistence in volatility dynamics.
`bayesGARCH`	Bayesian estimation for GARCH-type volatility models.	Stock and exchange rate volatility, risk forecasting.
`bayesDccGarch`	Bayesian dynamic conditional correlation (DCC-GARCH) for multivariate series.	Correlated financial markets, portfolio risk and contagion studies.
`bmgarch`	Stan-based Bayesian multivariate GARCH modeling with flexible priors.	Multivariate financial modeling, volatility spillovers.

Dynamic factor models & dimension reduction

Package	Summary	Typical applications
`bayesdfa`	Bayesian dynamic factor analysis for multivariate time series using Stan.	Environmental or biological monitoring, latent trend detection, dimensionality reduction.

Bibliography

References cited in Section 5.6 (“Software for Bayesian analysis”) of Gamerman, Lopes & Gonçalves (2026) *MCMC: Stochastic Simulation for Bayesian Inference, third edition.*

Becker, R. A. (1994). *A Brief History of S*. AT&T Bell Laboratories, Murray Hill, NJ.

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P., and Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of Statistical Software*, 76.

Eckhardt, R. (1987). Stan Ulam, John von Neumann, and the Monte Carlo method. *Los Alamos Science: Special Issue—Stanislaw Ulam 1909–1984*, 15:131–141.

Gelfand, A. E. and Smith, A. F. M. (1990). Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association*, 85:398–409.

Goel, P. K. (1988). Software for Bayesian analysis: Current status and additional needs. In Bernardo, J. M., DeGroot, M. H., Lindley, D. V., and Smith, A. F. M., editors, *Bayesian Statistics 3*, pages 611–622. Oxford University Press, Oxford.

Ihaka, R. and Gentleman, R. (1996). R: A language for data analysis and graphics. *Journal of Computational and Graphical Statistics*, 5(3):299–314.

Lunn, D., Spiegelhalter, D., Thomas, A., and Best, N. (2009). The BUGS project: Evolution, critique and future directions. *Statistics in Medicine*, 28:3049–3067.

Lunn, D. J., Thomas, A., Best, N., and Spiegelhalter, D. (2000). WinBUGS—a Bayesian modelling framework: Concepts, structure, and extensibility. *Statistics and Computing*, 10:325–337.

Metropolis, N. (1987). The beginning of the Monte Carlo method. *Los Alamos Science: Special Issue—Stanislaw Ulam 1909–1984*, 15:125–130.

Metropolis, N. and Ulam, S. (1949). The Monte Carlo method. *Journal of the American Statistical Association*, 44(247):335–341.

Press, S. J. (1980). Bayesian computer programs. In Zellner, A., editors, *Bayesian Analysis in Econometrics and Statistics: Essays in Honor of Harold Jeffreys*, pages 429–442. North-Holland, Amsterdam.

Spiegelhalter, D., Thomas, A., Best, N., and Lunn, D. (2014). *OpenBUGS User Manual*. Version 3.2.3.

Tanner, M. A. and Wong, W. (1987). The calculation of posterior distributions by data augmentation (with discussion). *Journal of the American Statistical Association*, 82:528–550.

Ulam, S. (1987). Stan Ulam: Vita's excerpts from *Adventures of a Mathematician*. Los Alamos Science: Special Issue, Stanislaw Ulam 1909–1984, 15:8–22.

Štrumbelj, E., Bouchard-Côté, A., Corander, J., Gelman, A., Rue, H., Murray, L., Pesonen, H., Plummer, M., and Vehtari, A. (2024). Past, present and future of software for Bayesian inference. *Statistical Science*, 39(1):46–61.