

Five years of the bayesvl package: A journey through Bayesian statistical analysis

Hong-Hue Thi Nguyen

Shinhan Bank, Ho Chi Minh City, Vietnam

May 7, 2024

“No matter how many times it takes him to correct his plans, he does not mind, for he is immersed in these mathematical calculations.

– When night falls, he smiles with the utmost pride and satisfaction at the perfect plan to catch fish.”

–In “The Perfect Plan”; *The Kingfisher Story Collection* [1]

Five years ago, on May 24, 2019, the computer program ‘bayesvl’ was officially published on R under the name “bayesvl: Visually Learning the Graphical Structure of Bayesian Networks and Performing MCMC with ‘Stan’” [2]. This program (or package) was developed by two founders of the SM3D Portal, Vuong Quan Hoang and La Viet Phuong, to improve the productivity of conducting social research [3]. The package was designed with a pedagogical orientation, supporting users in familiarizing themselves with Bayesian statistical methods, MCMC simulation, and visualization of technical diagnoses and results.

As of now, the ‘bayesvl’ package has received nearly 22 thousand downloads from CRAN around the world. In an evaluation of packages for Bayesian analysis available on R, biostatistician and data scientist Darko Medin stated that “If you are looking into the Bayesian networks, visualizing them and using DAG (directed acyclic graphs) in a Bayesian framework, bayesvl is one of the best packages out there” [4].

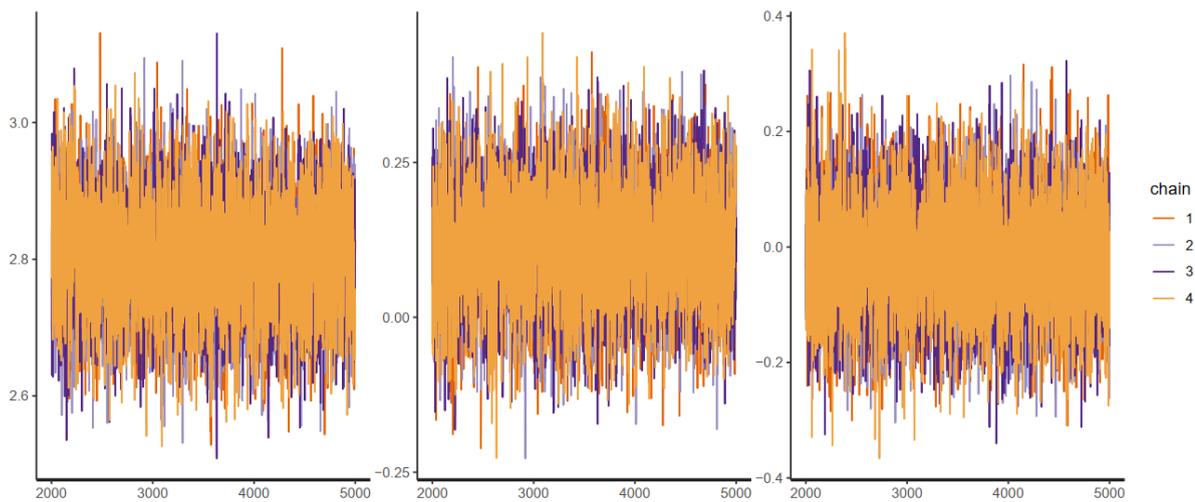


Figure 1. Trace plots generated by the 'bayesvl' package

Perhaps due to the good analytical and visualization capabilities of 'bayesvl', the number of social science research papers published using the software is noteworthy. Around 80 studies employing the 'bayesvl' package have been published. Among them, many were published by reputable academic journals such as Humanities and Social Sciences Communications (Nature Portfolio), Marine Policy (Elsevier), Research Evaluation (Oxford University Press), Pacific Conservation Biology (CSIRO and Australian Academy of Science), npj Climate Actions (Nature Portfolio), etc. Researchers using the package also have diverse backgrounds from developed and developing countries, such as Germany, Netherlands, South Korea, Indonesia, the United States, South Africa, New Zealand, Japan, Thailand, China, Australia, etc.

Additionally, thanks to the pedagogical support orientation, the package has continued to be developed and integrated with the Mindsponge Theory, forming the Bayesian Mindsponge Framework (BMF) analytics [5,6]. This method was introduced at the 2021 Seminar on Applied Statistics (organized by VIASM - Vietnam Institute for Advanced Study in Mathematics) and the 2023 Conference on "Innovating Mathematics Teaching Methods in Social Sciences" (organized by VIASM and Hanoi University). The Scientific Director of VIASM is Professor Ngo Bao Chau, a 2010 Fields Medalist.

During the 2023 VIASM-HANU Conference, 34 lecturers and researchers from 13 institutions collaborated to conduct a study using BMF analytics. After rounds of rigorous reviews, the research outcome was officially published in The VMOST Journal of Social Sciences and Humanities [7]. This is an academic journal established and prioritized investment by the Ministry of Science and Technology to provide a reputable open scientific publishing

platform for international knowledge dissemination.

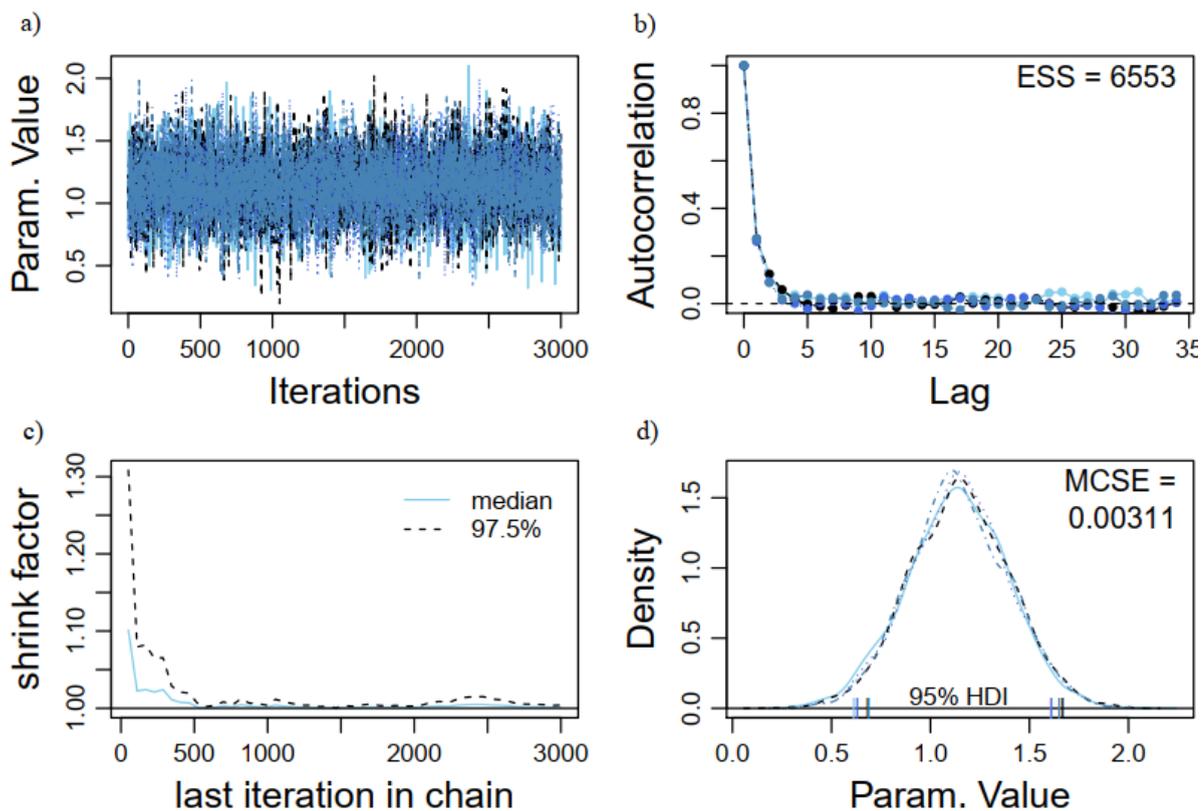


Figure 2. Analysis results using the 'bayesvl' package

Currently, the BMF analytics method has been used in training of more than 90 researchers from 54 organizations in 15 countries; notable names include Calcutta University (India), China University of Political Science and Law (China), Monash University (Australia), University of Pretoria (South Africa), Pepperdine University (USA), Sciences Po (France), University of Western Ontario (Canada), Saint Louis College (Thailand), Widya Mandala Catholic University (Indonesia), Ton Duc Thang University and Hanoi University of Science and Technology (Vietnam). Of these, 87.5% of the members are researchers from developing countries.

Because the 'bayesvl' package is built and running in R, an open platform, other researchers, especially early-career researchers and researchers from developing countries, can easily access and save costs for scientific research. Furthermore, to enable researchers to conduct analysis effectively, the development team has published two books (1 in English and 1 in Vietnamese) and five detailed guides. The detailed guides are all published openly and undergo rigorous review processes in reputable journals such as *SoftwareX*, *MethodsX*, *The VMOST Journal of Social Sciences and Humanities*, and *Software Impacts* [6,8-13]. Meanwhile, the guidebooks are being stored in many libraries around the world, such as the

libraries of Harvard University, the University of California system (Berkeley, Irvine, and San Diego), Bonn University, New York University, Pratt Institute, Ritsumeikan University, and Campinas State University, etc.

In addition to the above materials, R Basics, a community specializing in providing programming documentation for those new to R, has independently provided a guide for the 'bayesvl' package. The guide includes basic details, technical details, and analysis capabilities of the program [14].

References

[1] Vuong QH. (2022). *The Kingfisher Story Collection*. <https://www.amazon.com/dp/BOBG2NNHY6>

[2] La VP, Vuong QH. (2019). bayesvl: Visually Learning the Graphical Structure of Bayesian Networks and Performing MCMC with 'Stan'. *The Comprehensive R Archive Network*. <https://cran.r-project.org/package=bayesvl>

[3] Vuong QH. (Ed.) (2022). *A New Theory of Serendipity: Nature, Emergence and Mechanism*. Walter de Gruyter GmbH. <https://www.amazon.com/dp/8366675858>

[4] Medin D. (2022). Perspectives on 20 Bayesian libraries / R implementations. <https://www.linkedin.com/pulse/perspectives-16-bayesian-libraries-r-implementations-darko-medin/>

[5] Vuong QH. (2023). *Mindsponge Theory*. Walter de Gruyter GmbH. <https://www.amazon.com/dp/BOC3WHZ2B3>

[6] Vuong QH, Nguyen MH, La VP. (Eds.)(2022). *The mindsponge and BMF analytics for innovative thinking in social sciences and humanities*. Walter de Gruyter GmbH. <https://www.amazon.com/dp/BOC4ZK3M74>

[7] Nguyen MH, *et al.* (2024). Effects of water scarcity awareness and climate change belief on recycled water usage willingness: Evidence from New Mexico, United States. *The VMOST Journal of Social Sciences and Humanities*, 66(1), 62-75. https://d.vjst.vn/index.php/vmost_jossh/article/view/344

[8] Hoàng VQ, *et al.* (2021). *Ban hoa tau du lieu xa hoi*. Nxb Khoa hoc xa hoi.

- [9] Vuong QH, *et al.* (2020). Bayesian analysis for social data: A step-by-step protocol and interpretation. *MethodsX*, 7, 100924. <https://www.sciencedirect.com/science/article/pii/S2215016120301448>
- [10] Nguyen MH, *et al.* (2022). Introduction to Bayesian Mindsponge Framework analytics: An innovative method for social and psychological research. *MethodsX*, 9, 101808. <https://www.sciencedirect.com/science/article/pii/S2215016122001881>
- [11] La VP, *et al.* (2022). The bayesvl package: An R package for implementing and visualizing Bayesian statistics. *SoftwareX*, 20, 101245. <https://www.sciencedirect.com/science/article/pii/S2352711022001637>
- [12] Vuong QH, Nguyen MH, Ho MT. (2022). bayesvl: An R package for user-friendly Bayesian regression modelling. *The VMOST Journal of Social Sciences and Humanities*, 64(1), 85-96. https://d.vjst.vn/index.php/vmost_jossh/article/view/268
- [13] Vuong QH, *et al.* (2020). Improving Bayesian statistics understanding in the age of Big Data with the bayesvl R package. *Software Impacts*, 4, 100016. [https://www.softwareimpacts.com/article/S2665-9638\(20\)30003-8/fulltext](https://www.softwareimpacts.com/article/S2665-9638(20)30003-8/fulltext)
- [14] R Basics. (2024). The ultimate guide to the bayesvl package in R. <https://rbasics.org/packages/bayesvl-package-in-r/>

