

SFS 2024 Short Course – Bayesian Applications in Environmental and Ecological Studies with R and Stan

Song S. Qian and Mark R. DuFour

6/1/2024

The short course will provide a concise introduction to Bayesian applications in environmental and ecological studies, featuring examples and code from [Qian et al \(2022\)](#). To facilitate practice, we will supply multiple Rmarkdown files. To ensure the best learning experience, it is recommended that you bring your own laptop to the workshop, with R and RStudio already installed. The workshop will have an informal atmosphere, allowing time for questions to be addressed as they arise. Instead of using a traditional PowerPoint presentation, we will guide you through the Rmarkdown files and provide explanations of the materials. Therefore, we kindly request that you make the necessary preparations. You can access all course materials, including data, code, and references, on [Github](#), follow the link to [SFS](#) for the Rmd files used in the short course.

Participants of the workshop can receive a 25% discount of the book by ordering the book at [the publisher's home page](#), using code `SFS24`.

Preparing R and RStan

To ensure smooth compatibility and avoid version-specific issues, it is strongly recommended that you update both R and RStudio. The materials provided in the course were thoroughly tested with recent versions: R 4.4.0, nicknamed “Puppy Cup,” and RStudio version 2024.04.0+735, called “Chocolate Cosmos.”

Throughout the course, we will utilize Stan for all Bayesian computations, which can be accessed through the R package ‘rstan’. It’s important to note that using ‘rstan’ requires access to a C compiling toolchain. If you are a Windows user, please ensure that you also install ‘Rtools’. The current recommended version is [Rtools44](#)).

For proper installation of ‘rstan’, please follow the instructions provided by the Stan developers:

- Windows users using recent versions of R (4.3 and above) should be aware that the current version of ‘RStan’ available from CRAN is often not compatible (because of misaligned updating schedules between R and Stan) and may result in error messages. To resolve this, you need to install the preview of ‘rstan’ 2.26 by running the following commands:

```
remove.packages(c("StanHeaders", "rstan"))
install.packages("StanHeaders",
  repos = c("https://mc-stan.org/r-packages/", getOption("repos")))
install.packages("rstan",
  repos = c("https://mc-stan.org/r-packages/", getOption("repos")))
```

- For Mac users, please refer to [Stan Github](#) for detailed instructions on configuring the C++ toolchain.
- For Linux users, you are most likely fine. Check [here](#) just in case.

Preparing your laptop

To prepare for the short course, it is recommended that you create an empty folder on your laptop dedicated to this course. Then, download all the .Rmd and .RData files from the [course repository](#) into this folder. During

the class, the data files will be accessed through the internet. However, if you anticipate not having internet access on-site, it is crucial that you download and clone the [GitHub page](#) onto your laptop **beforehand**. This ensures that you have the necessary files readily available. In case of any unforeseen circumstances, I will also bring a copy of all essential files with me.

Installing package `rv`

We use `rv` extensively for processing MCMC samples. Unfortunately, the author of `rv` does not actively maintain the package. As a result, it is often removed from R archive. You can still install it directly from [GitHub](#):

```
packages('devtools') # package devtools needed
devtools::install_github("jsta/rv")
```

Your Data Examples

To enhance the workshop's usefulness, I have structured the day into four sessions. Here's a breakdown of each session:

- **Introductory Session:** We will start with an introductory session that focuses on a simple problem that doesn't necessarily require Bayesian statistics or the computational power of Stan. This session serves as a foundation for understanding key concepts.
- **Bayesian Benefits and Model Evaluation:** After the morning tea-break, we will delve into the benefits of Bayesian analysis, particularly emphasizing the advantages of using Stan for model evaluation. We will explore two examples where the probability distribution of the response variable is complex, showcasing the power of Bayesian approaches.
- **Bayesian Hierarchical Models:** In the afternoon, I have prepared three examples centered around my favorite topic: Bayesian hierarchical models. These examples will provide an in-depth understanding of how hierarchical models can be effectively utilized in Bayesian analyses.
- **Participant Data Examples:** During the last couple of hours of the day, we will open the floor for discussion and analysis of data examples brought by participants. If you have an example you'd like to share, please send it (song.qian@utoledo.edu) to me ahead of time (before May 31, US time). I will do my best to include it in the final session and provide insights based on the shared data.

By structuring the day in this manner, we aim to cover various aspects of Bayesian analysis, progressively building knowledge and offering opportunities for practical application.