The Journal of Open Source Software

fishStan: Hierarchical Bayesian models for fisheries

Richard A. Erickson[1], **Daniel S. Stich**[2], **and Jillian L. Hebert**[1]

1 U.S. Geological Survey, Upper Midwest Environmental Sciences Center, La Crosse, WI, United States of America **2** State University of New York at Oneonta, Oneonta, NY, United States of America

DOI: 10.21105/joss.03444

Software

- Review
- Repository
- Archive

Editor: Marcos Vital

Reviewers:

- @mikekaller
- @cole-monnahan-noaa

Submitted: 16 March 2021
Published: 21 March 2022

Summary

Fisheries managers and ecologists use statistical models to estimate population-level relations and demographic rates (e.g., length-maturity curves, growth curves, and mortality rates). These relations and rates provide insight into populations and inputs for other models. For example, growth curves may vary across lakes showing fish populations differ due to management actions or underlying environmental conditions. A fisheries manager could use this information to set lake-specific harvest limits or an ecologist could use this information to test scientific hypotheses about fish populations. The above example also demonstrates how populations exist within hierarchical structures where sub-populations may be nested within a meta-population. More generally, these hierarchical structures may be both biological (e.g., different lakes or river pools) and statistical (e.g., correlated error structures). Currently, limited options exist for fitting these hierarchical models and people seeking to use them often must program their own implementations. Furthermore, many fisheries managers and researchers may not have Bayesian programming skills, but many can use interactive languages such as R. Additionally, programs such as JAGS often require long run times (e.g., hours if not days) to fit hierarchical models and programs such as Stan can be more difficult to program because it is a compiled language. We created `fishStan` to share hierarchical models for fisheries and ecology in an easy-to-use R package.

Statement of need

We designed `fishStan` to be used by researchers, managers, and students wanting to apply hierarchical models to fisheries and ecological data (Erickson et al., 2020). An archived version of the package exists on Zenodo (10.5281/zenodo.6373008). The package enables users to take advantage of compiled languages (e.g., C++, Stan) for speed without losing the flexibility or ease-of-use of R. Specifically, `fishStan` uses RStan (Stan Development Team, 2020) to call Stan (Carpenter et al., 2017) for fisheries models including hierarchical growth models, hierarchical linear regression, hierarchical logistic regression, and a catch-curve model.

The initial model incorporated in `fishStan` was a hierarchical von Bertalanffy model presented in Midway et al. (2015) who used JAGS (Plummer, 2003) as their programming language. We included this model and other common fisheries models developed as part of ongoing research and management. These include hierarchical logistic models (e.g., maturity curves), hierarchical linear models (e.g., log-log length weight relations), and other growth curves.

We created the original package for Erickson et al. (2021b) that estimated bighead carp (*Hypophthalmichthys nobilis*) and silver carp (*H. molitrix)* demographic rates in the Illinois, Mississippi, and Ohio Rivers. This paper also includes a public, peer-reviewed dataset (Erickson et al., 2021a) and a public, peer-reviewed workflow (Erickson & Kallis, 2021). A subset of Erickson et al. (2021a) and Erickson & Kallis (2021) is included in Erickson et al. (2020) as a vignette to demonstrate an application of the package. Other applications of our Stan-based

models include the growth models in ASMFC -Atlantic States Marine Fisheries Commission (ASMFC) (2020) and Caves et al. (2020). Books such as Ogle (2016) provide an introduction and overview to the fisheries models included within the package. Books such as Gelman & Hill (2007) provide an introduction to and overview of hierarchical Bayesian modeling.

Acknowledgments

We thank Wayne Thogmartin for his feedback on a previous draft of this manuscript. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

References

Atlantic States Marine Fisheries Commission (ASMFC). (2020). *American shad benchmark stock assessment and peer review report*. ASMFC. https://doi.org/10.1002/nafm.10515

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P., & Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of Statistical Software*, *76*, 1–32. https://doi.org/10.18637/jss.v076.i01

Caves, S., Baumann, J. R., & Stich, D. S. (2020). Density-dependent changes in grass carp growth and mortality in long-term aquatic plant management. *North American Journal of Fisheries Management*, *41*, 355–365. https://doi.org/10.1002/nafm.10515

Erickson, R. A., Daniel S. Stich, & Hebert, J. L. (2020). *fishStan: Hierarchical bayesian models for fisheries*. U.S. Geological Survey software release. https://doi.org/10.5066/P9TT3ILO

Erickson, R. A., & Kallis, J. L. (2021). *Analysis of carp demographics data*. U.S. Geological Survey software release. https://doi.org/10.5066/P9Q6SUML

Erickson, R. A., Kallis, J. L., Coulter, A. A., Coulter, D., MacNamara, R., Lamer, J. T., Bouska, W., Irons, K. S., Solomon, L. E., Stump, A. J., Weber, M. J., Brey, M. K., Sullivan, C. J., Sass, G. G., Garvey, J. E., & Glover, D. C. (2021a). *Bighead and silver carp individual fish data from the Mississippi, Ohio, and Illinois rivers from 1997 to 2018*. U.S. Geological Survey data release. https://doi.org/10.5066/P9IAOZ8G

Erickson, R. A., Kallis, J. L., Coulter, A. A., Coulter, D., MacNamara, R., Lamer, J. T., Bouska, W., Irons, K. S., Solomon, L. E., Stump, A. J., Weber, M. J., Brey, M. K., Sullivan, C. J., Sass, G. G., Garvey, J. E., & Glover, D. C. (2021b). Demographic rate variability of bighead and silver carps along an invasion gradient. *Journal of Fish and Wildlife Management*, *12*, 338–353. https://doi.org/10.3996/JFWM-20-070

Gelman, A., & Hill, J. (2007). *Data analysis using regression and multilevelhierarchical models*. Cambridge University Press New York, NY, USA.

Midway, S. R., Wagner, T., Arnott, S. A., Biondo, P., Martinez-Andrade, F., & Wadsworth, T. F. (2015). Spatial and temporal variability in growth of southern flounder (paralichthys lethostigma). *Fisheries Research*, *167*, 323–332. https://doi.org/10.1016/j.fishres.2015.03.009

Ogle, D. H. (2016). *Introductory fisheries analyses with R*. CRC Press. https://doi.org/10.1201/9781315371986

Plummer, M. (2003). JAGS: A program for analysis of bayesian graphical models using gibbs sampling. *Proceedings of the 3rd International Workshop on Distributed Statistical Computing*, *124*, 1–10.

Stan Development Team. (2020). *RStan: The R interface to Stan*. http://mc-stan.org/