

Peer review for Richie Erickson's software release of fishStan v2

<https://code.usgs.gov/rerickson/fishStan>

I want my identity withheld from the authors: *No*

Name: *Alison Appling*

Date review completed: *11/14/2020*

Areas of scientific expertise: *Bayesian modeling; R packages; biogeochemistry*

General Comments

This package is very readable and well documented, and it makes excellent use of available tools (especially `rstan`).

I used the UMESC code review checklist as my primary guide, aiming for consistency with <https://www.usgs.gov/products/software/software-management/types-software-review>.

I looked at the entire package but focused my attention on the new models in version 2 that were noted in the NEWS.md file.

Security

- Is the file path universal? *I found no issues with file paths when running the package locally.*
- Are usernames or passwords required to access data? *No.*
- Are computer names or network locations listed? *No.*

Functionality

The code seems to work as expected, with the exception of one test failure (see below).

- I installed successfully using `remotes::install_gitlab('rerickson/fishStan', host='code.usgs.gov')`, as recommended in the README (but with `rerickson` path for the review).*
- The package has an impressive 142 built-in tests. Of these, all but one passes. That one failure does seem to merit inspection.*

```
> devtools::test()
Loading fishStan
Testing fishStan
| OK F W S | Context
| 13      | test linear model [331.4 s]
| 2       | test VB without t0 functions [9.8 s]
| 3       | test VB functions [17.0 s]
| 2       | Test catch_curve
| 4       | catch_curve_stan test [4.5 s]
| 11      | test function to create projection inputs
| 2       | Gompertz functions test
x | 2 1     | test Gompertz stan functions [30.2 s]

Failure (test-gompertz_stan.R:72:5): Hierarchical von_b with t0 works
`group_t0` not equal to as.numeric(growth_coef[5:6, "mean"]).
```

```

1/2 mismatches
[1] 0.2 - 0.455 == -0.255

| 2      | Galluci Quinn functions test
| 78     | test binomial code and extract, rename functions [215.9 s]
| 3      | test gq function [30.9 s]
| 4      | helper functions test [19.3 s]
| 3      | test ling functions [36.9 s]
| 4      | Gompertz simulation function test
| 2      | Galluci Quinn functions test
| 3      | von B functions test
| 3      | von B t0 functions test using t0 = 0

Results
Duration: 696.1 s

[ FAIL 1 | WARN 0 | SKIP 0 | PASS 141 ]
>

```

- *I inspected the code files and found nothing of concern.*

Metadata

- What is the purpose of the code? *The paragraph in DESCRIPTION seems reasonable and up to date: “Methods for fitting hierarchical Bayesian models using R Stan, created with fisheries models in mind. Currently, this includes linear regresssion (such as length-weight models), growths curves, and logistic regression (such as length-maturity curves).”*
- Does the code appear to do what it was meant to do? *Yes; see **Code Functionality** section above. I have added a few typo-correction suggestions to the documentation here and there, and one meaningful suggestion to the vonBertalanffy.Rmd vignette (looks like a copy-paste issue relative to vonBertalanffy_t0.Rmd), but otherwise I see nothing of concern.*
- Is there contact information for authors? *Yes, the README.md contains contact information.*
- Is there a citation listed, with Digital Object Identifier? *Yes, the README.md contains a citation with placeholder DOIs.*
- Can all code files be found? *Yes.*
- Are all important sections available and up to date? (i.e. README, DISCLAIMER, LICENSE) *Yes. There’s one surprising file name in code.json: - downloadURL points to “.../pestpp_sensitivity_analysis-master.zip” - but that link does resolve to download a file named “fishStan-master-[hexcode].zip”, which seems fine (though of course outdated relative to fishStan v2).*
- Is there a statement provided estimating how long it takes the code to run? *Good enough: The README.md contains a statement of how long it takes to build the package (presumably with building vignettes). Also, the vignettes give notes on setting the number of iterations and the number of cores to optimize the runtime.*

Readability

- Can the code be understood? *Yes, the code is very readable.*
- Are there too many comments? Too little? *Just right.*
- Is there repeated code that could be done in a more concise way (functions)? *I saw no opportunities for this.*

- d. Is the code readable? *Yes, the code is very readable.*
- e. Is the format consistent? (i.e. spacing, indenting) *Yes, very consistent. `lintr` makes just one style recommendation, which I think is optional (but have added as a suggested change in the PR anyway):*

```
> lint_out <- lintr::lint_package()
```

In R/stanmodels.R, line 15: Place a space before left parenthesis, except in a function call.