ss3sim: An R package for stock-assessment simulation with Stock Synthesis

Sean C. Anderson[1], … (authorship and order to be discussed)

[1]Department of Biological Sciences, Simon Fraser University, Burnaby BC, V5A 1S6, Canada

[*]Corresponding author: phone: 1-778-782-3989; email: sean_anderson@sfu.ca

Short title: ss3sim: Stock Synthesis simulation

Abstract

< 300 words

Introduction

Simulation is a critical component to testing fishery stock-assessment methods [1–4]. With simulation, we can evaluate the precision and bias of complex assessment methods in a controlled environment where we know the true state of nature [2]. Recently, simulation studies have been key to improving strategies for dealing with, for example, time-varying natural mortality [5–7], uncertainty in steepness of the stock-recruit relationship [8], and uncertainty in stock productivity [9]. [not necessarily the best examples]

Stock Synthesis [10], is a widely-used stock-assessment framework. It implements statistical age-structured population dynamics modeling using a wide range of minimally-processed data [10,11]. By using this framework, individuals conducting stock assessments and peer reviewers can focus on the underlying science, instead of the model code [10]. Owing to these advantages, SS3 (the third version of the software) is one of the world's most commonly-used stock-assessment tools, particularly in the United States and Australia, where it has been used in 35 and 12 stock assessments, respectively, as of 2012 [10].

Although SS is increasingly a standard for fisheries stock assessment, and the programming language R [12] has become the standard for statistical computing and visualization, we lack a generalized framework to link these components in a simulation context. Here, we introduce ss3sim, an R package that facilitates large-scale, rapid, and reproducible stock-assessment simulation with the widely-used SS framework. We begin by outlining the general structure of ss3sim and describing its functions. We then demonstrate the software by developing a simple example. We conclude by discussing how ss3sim complements other stock assessment simulation software and outlining research questions our accessible and general SS simulation framework could address.

The ss3sim framework

Terminology

[TODO abbreviate this paragraph substantially or cut it]

Throughout this paper we refer to a number of terms, which we define here. We use the term *operating model* (OM) [13] to refer to the model that represents the underlying true dynamics of the system (REF). We use the term *estimation method* (EM) to refer to the method used to estimate quantities of interest (REF). We use the term *scenario* to refer to a combination of operating and estimation model *cases*. For example, an OM case might specify that natural mortality follows a random walk, an EM case might estimate a single parameter for natural mortality, and the combination of these cases along with all other specified conditions creates a scenario. We refer to *iterations* or *replicates* as repeated simulations of a scenario, possibly with new process and observation error added each time. A simulation therefore refers to the combination of all scenarios and iterations.

Design goals of ss3sim

[This section is too long currently. I don't want to bore people.]

We designed ss3sim to be reproducible, flexible, and rapid. *Reproducible*: ss3sim allows for the simulation to be documented in code and plain-text control files. Further, the plain-text control files refer to individual cases, which allows for the reuse of control files across scenarios. This reduces the chance for errors and simplifies the exploration of new scenarios.

Flexible: ss3sim allows the user to specify their own OM and EM using all the possible configurations of SS3. ss3sim returns output in standard comma-separated-value (`.csv`). This means that the output can be easily processed with the package-provided functions or with other tools.

Rapid: First, ss3sim relies on SS3, which uses ADMB as a backend optimization platform — the most rapid and robust optimization software available [14]. Second, ss3sim allows simulations to be deployed across multiple computers or computer cores. Third, the package provides a number of functions to quickly visualize simulation output. Access to quick visualization tools means that users are more likely to graphically explore their models and are therefore more likely to detect errors and understand their simulation output as they introduce complexity. Finally, ss3sim minimizes the amount of bookkeeping code that researchers have to write so that they can concentrate on the science itself.

The general stucture of an ss3sim simulation

An ss3sim simulation requires three types of input: (1) a base model of the underlying truth (an SS3 OM), (2) a base model of how to assess that truth (an SS3 EM), (3) and a set of case files describing deviations from these base models. ss3sim works, in general, by converting case file arguments (e.g. a given natural mortality trajectory) into manipulations of SS3 configuration files (`change` functions), running the OM, sampling pseudo data, and running the EM (`run` functions), and facilitating the manipulation and visualization of output (`get` and `plot` functions) (Figure 1).

An example simulation with ss3sim

(unsure how much of this will go in the main paper and how much will just be in the appendix… probably many of these details should be appendix only with just enough elements to give a flavour for what can be done in the main paper)

Setting up the SS models:

- choosing a specific conditioning model or generic conditioning type
- setting up the OM and EM SS models
- things to keep in mind
- running through SS to format as `.ss_new` files and renaming
- required files

Setting up the configuration files:

- the (simple) research question (increasing or decreasing survey effort crossed with estimating M or fixing M)
- indicate which arguments to adjust

Deterministic model testing:

- reduce recdevs, reduce sigma R, bias correction
- what to plot, what to look for, how good is OK?

Output analysis and visualization:

- examples using the included functions
- brief take home of what we'd conclude

How ss3sim complements other simulation software

Probably turn this into a small table:

r4ss

- Reference 15
- r4ss has functions to facilitate aspects of simulations, mostly focused on reading and plotting output for stock assessment
- ss3sim uses r4ss functions for some reading, writing, and bias adjustment

FLR

- Reference 16 for FLR and Reference 17 for simulation in FLR
- statistical catch-at-age only?
- not integrated analysis, not SS
- but particularly relevant to Europe

"Hooalator"

- http://fisherysimulation.codeplex.com, Windows only, GUI…, works on bootstrapped data only, therefore isn't as flexible as ss3sim. Used in:
    1. Reference 8
    2. Reference 18
    3. Reference 5

Research opportunities with ss3sim

- there are lots, we should brainstorm some key ones

Conclusions

- benefit of using one well tested and well-understood modeling framework (SS) i.e. benefit to playing with all the switches and understanding one framework well versus having many tools that we superficially understand (based on Rick's comments at the conference)
- why we developed generic low-level functions and high-level functions
- researchers are free to develop their own low- and high-level functions because in an open-source MIT(?) licensed R package, users are free to modify functions as needed
- (these points are somewhat random at the moment)

Acknowledgements

- funding: Fulbright Canada, NSERC, Simon Fraser University, many others…

- discussions and advice: André Punt, Richard Methot, Ian Taylor, James Thorson, …
- Any FISH600 members not listed as authors

Tables

Table X: Comparison with related software? Possible columns: software, reference, platform (e.g. R, GUI…), short description/comparison, examples of papers using it

Figures legends

Figure 1: Flow diagram of `run_ss3sim()` stock-assessment simulation steps.

Figure 2: Panels with output from the example

References

1. Hilborn R, Walters CJ (1987) A general model for simulation of stock and fleet dynamics in spatially heterogeneous fisheries. Canadian Journal of Fisheries and Aquatic Sciences 44: 1366–1369.

2. Hilborn RW, Walters C (1992) Quantitative Fisheries Stock Assessment: Choice, Dynamics, and Uncertainty. London: Chapman and Hall.

3. Rosenberg AA, Restrepo VR (1994) Uncertainty and Risk Evaluation in Stock Assessment Advice for U.S. Marine Fisheries. Canadian Journal of Fisheries and Aquatic Sciences 51: 2715–2720.

4. Peterman R (2004) Possible solutions to some challenges facing fisheries scientists and managers. ICES Journal of Marine Science 61: 1331–1343.

5. Lee H-H, Maunder MN, Piner KR, Methot RD (2011) Estimating natural mortality within a fisheries stock assessment model: An evaluation using simulation analysis based on twelve stock assessments. Fisheries Research 109: 89–94.

6. Jiao Y, Smith EP, O'Reilly R, Orth DJ (2012) Modelling non-stationary natural mortality in catch-at-age models. ICES Journal of Marine Science 69: 105–118.

7. Deroba JJ, Schueller AM (2013) Performance of stock assessments with misspecified age- and time-varying natural mortality. Fisheries Research 146: 27–40.

8. Lee H-H, Maunder MN, Piner KR, Methot RD (2012) Can steepness of the stock-recruitment relationship be estimated in fishery stock assessment models?. Fisheries Research 125–126: 254–261.

9. Ianelli JN (2002) Simulation analyses testing the robustness of productivity determinations from West Coast Pacific ocean perch Stock Assessment Data. North American Journal of Fisheries Management 22: 301–310.

10. Methot RD, Wetzel CR (2013) Stock Synthesis: A biological and statistical framework for fish stock assessment and fishery management. Fisheries Research 142: 86–99.

11. Maunder MN, Punt AE (2013) A review of integrated analysis in fisheries stock assessment. Fisheries Research 142: 61–74.

12. R Core Team (2013) R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing.. Available: `http://www.R-project.org/`.

13. Linhart H, Zucchini W (1986) Model Selection. New York: John Wiley.

14. Fournier DA, Skaug HJ, Ancheta J, Ianelli J, Magnusson A, et al. (2012)

AD Model Builder: using automatic differentiation for statistical inference of highly parameterized complex nonlinear models. Optimization Methods and Software 27: 233–249.

15. Taylor I, Stewart I, Hicks A, Garrison T, Punt A, et al. (2013) r4ss: R code for Stock Synthesis. Available: http://code.google.com/p/r4ss/.

16. Kell LT, Mosqueira I, Grosjean P, Fromentin J-M, Garcia D, et al. (2007) FLR: an open-source framework for the evaluation and development of management strategies. ICES Journal of Marine Science 64: 640–646.

17. Hillary R (2009) An introduction to FLR fisheries simulation tools. Aquatic Living Resources 22: 225–232.

18. Piner KR, Lee H-H, Maunder MN, Methot RD (2011) A simulation-based method to determine model misspecification: examples using natural mortality and population dynamics models. Marine and Coastal Fisheries 3: 336–343.