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

Abstract

Introduction

Simulation is a critical component to testing fishery stock-assessment methods [1]. With simulation, we can evaluate the precision and bias of increasingly complex assessment methods (REF) in a controlled environment where the true state of a fishery is known. Recently, simulation studies have been key to improving methods to assess, for example, natural mortality (REF), XX (REF), and XX (REF). (Possible refs: [2–6])

Stock Synthesis [7], is a widely-used stock-assessment framework. It implements a statistical catch-at-age Integrated Analysis population dynamics models using a wide range of minimally-processed data [7,8]. By using this already-developed assessment framework, those conducting stock assessments and peer reviewers can focus on the underlying science, instead of the model code [7]. Owing to these advantages, SS3 (the third version of the software) is one of the world's most commonly-used stock-assessment tools, particularly on the west coast of the United States and Australia, where it was used in 60?/XX (REF) and XX/XX (REF) assessments in 2012.

While SS is increasingly the standard for fisheries stock assessment, and the programming language R [9] has become the standard for statistical computing and visualization (REF?), 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 philosophy 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] Throughout this paper we refer to a number of terms, which we define here. We use the term *operating model* (OM) to refer to the model that represents the underlying true dynamics of the system. We use the term *estimation model* (EM) to refer to the model used to estimate quantities of interest. 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.

General philosophy

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 [10]. 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.

General structure

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 by converting case file arguments (e.g. a given natural mortality trajectory) into manipulations of SS3 configuration files, running the OM, sampling pseudo data, and running the EM, and storing the output (Figure 1).

RE-WORK THIS:

Low-level generic ss3sim functions

See Table 1 for a description of the main functions. We show how the functions fit into the general structure of a stock assessment simulation in Figure 1. ss3sim functions are divided into three types of functions:

1. Functions that manipulate SS configuration files. These manipulations generate an underlying "truth" (OM) and control our assessment of those models (EM).

2. Functions that conduct simulations. These functions generate a folder structure, call manipulation functions, run `SS3` as needed, and save the output.

3. Functions for analyzing and plotting simulation output.

High-level tailored ss3sim functions

- `run_ss3sim` also see `run_fish600` for an example custom wrapper function for a specific set of projects
- because it relies on manipulation of these configuration files, it's important the config files match a specific format
- general framework, because you start with your own OM and EM, and a wide variety of questions are then available through manipulations of …, …

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

File and folder setup

- required files
- Why we chose a flat-file structure
- see vignette

Translating research questions into 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

Discussion

- Other sections?
- how we validated it
- benefit of using one well tested and well-understood modeling framework (SS) (but disadvantages too) — 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 higher level functions
- but researchers are free to develop their own higher level functions
- because in an open-source MIT(?) licensed R package, users are free to modify functions as needed

How ss3sim complements other generic stock-assessment simulation software

- focus on "generic" software, e.g. not software the just works for salmon simulation

r4ss

- Reference 11
- 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 12 for FLR and Reference 13 for simulation in FLR
- statistical catch-at-age only?
- not integrated analysis, not SS
- but particularly relevant to Europe

"Hooilator"

- http://fisherysimulation.codeplex.com, Windows only, GUI…, works on bootstrapped data only, therefore isn't as flexible as ss3sim. Used in:
  1. Reference 6
  2. Reference 3
  3. Reference 2

Research opportunities with ss3sim

- there are lots, we should brainstorm some key ones

Acknowledgements

- funding: Fulbright Canada, NSERC, Simon Fraser University, many others…
- discussions and advice: André Punt, Richard Methot, Ian Taylor, James Thorson, …

Figures

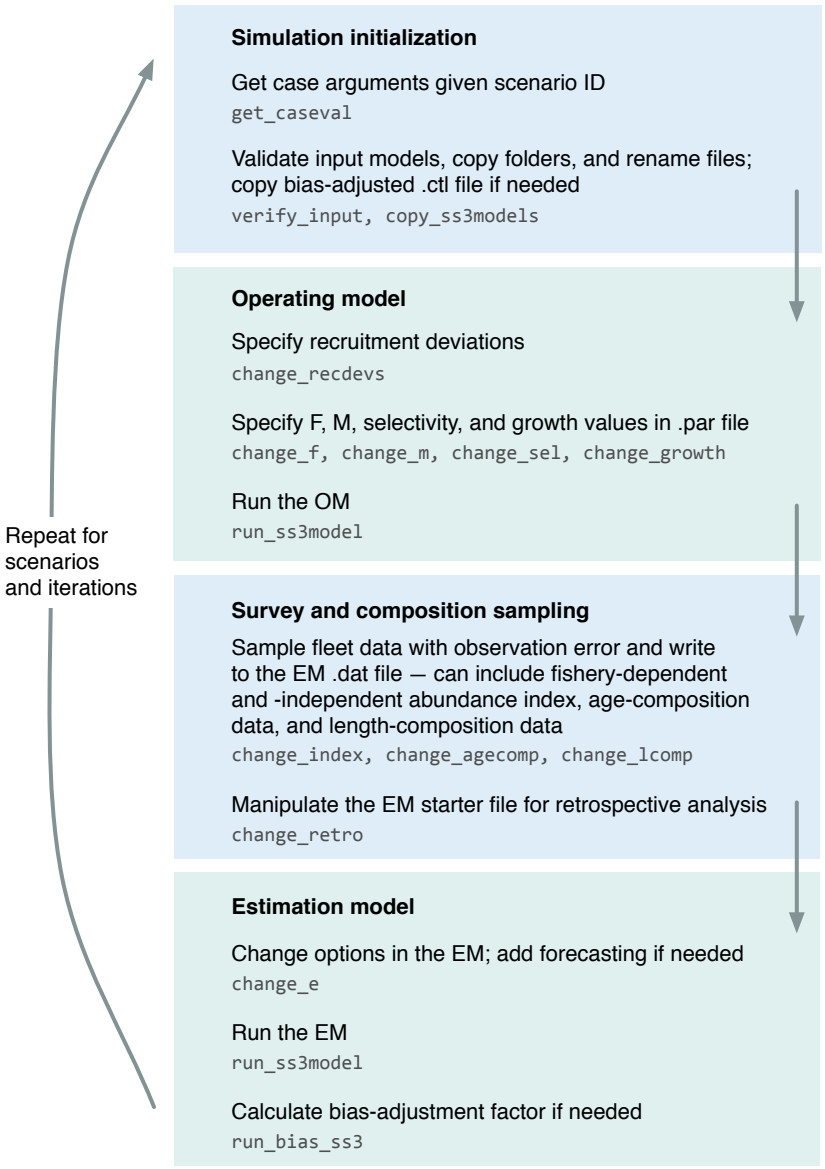

Figure 1: Flow diagram of `run_ss3sim()` stock-assessment simulation steps.

Figure 2: Panels with output from the example

Tables

Table 1: User-facing ss3sim functions and a description of their purpose. This is now a bit redundant with Fig. 1, the main body text, and the package documentation itself.

Function name	Description
`change_f`	Changes the fishing mortality
`change_m`	Adds time-varying natural mortality features
`change_growth`	Adds time-varying growth features
`change_sel`	Adds time-varying selectivity
`change_e`	Controls what and how parameters are estimated
`change_lcomp`	Controls how length composition data are sampled
`change_agecomp`	Controls how age composition data are sampled
`change_index`	Controls how the fishery and survey indices operate
`change_rec_devs`	Substitutes recruitment deviations
`change_retro`	Controls the number of years to discard for a retrospective analysis
`run_ss3sim`	Master function that runs an ss3sim simulation
`run_fish600`	Wrapper function that facilitates one particular simulation setup
`get_results_all`	Extract results from a series of scenarios
`get_results_scenario`	Extract the results for a single scenario
`plotting functions!!`	Plot the output…

Table X: Comparison with related software? - maybe a table with the possible columns: software, reference, platform (e.g. R, GUI…), Short description/comparison, examples of papers using it

References

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

2. 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.

3. 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.

4. Methot RD, Taylor IG (2011) Adjusting for bias due to variability of estimated recruitments in fishery assessment models. Canadian Journal of Fisheries and Aquatic Sciences 68: 1744–1760.

5. 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.

6. 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.

7. Methot RD, Wetzel CR (2012) Stock Synthesis: A biological and statistical framework for fish stock assessment and fishery management. Fisheries Research 142: 86–99.

8. Maunder MN, Punt AE (2012) A review of integrated analysis in fisheries stock assessment. Fisheries Research 142: 61–74.

9. R Core Team (2013) R: A Language and Environment for Statistical Computing.

182 Vienna, Austria: R Foundation for Statistical Computing.. Available: http://www.R-project.org/.

184 10. Fournier DA, Skaug HJ, Ancheta J, Ianelli J, Magnusson A, et al. (2012) AD Model Builder: using automatic differentiation for statistical inference of highly
186 parameterized complex nonlinear models. Optimization Methods and Software 27: 233–249.

188 11. 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/.

190 12. 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
192 strategies. ICES Journal of Marine Science 64: 640–646.

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