ss3sim: An R package for generalized 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

[*]sean_anderson@sfu.ca

Introduction

Fisheries stock-assessment simulation is a critical component to evaluating increasingly complex stock-assessment methods and understanding their strengths and weaknesses (Hilborn and Walters, 1992). Simulation lets us test our assessment methods in a controlled environment where we know the underlying state of a fishery system. Further, it lets us explore matches and mismatches of these known truths and are assumptions about those truths. Given the benefits of conducting stock-assessment simulations, increasing assessment-model complexity, and increasing computational power, stock-assessment simulations have become a rapidly expanding field (e.g. Hillary, 2009; Jiao *et al.*, 2012; Lee *et al.*, 2011; 2012; Methot and Taylor, 2011; Piner *et al.*, 2011).

Stock Synthesis (SS; Methot and Wetzel, 2012), is a widely used statistical catch-at-age (SCAA) stock-assessment framework. It implements Integrated Analysis (IA) style population dynamics models using a wide range of data (Maunder and Punt, 2012). SS3 is the third version of the software using this framework and is now one of the world's most widely-used stock-assessment tools, with researchers and modellers especially active on West Coast of United States and in Australia (REF). SS facilitates rapid, reproducible analyses, and thus allows users to focus on the underlying science and assessment of their stocks, not on the underlying model code. Therefore, there are two benefits to conducting stock-assessment simulation using SS. First, since the modelling framework has already been developed and tested, simulation research can progress rapidly and with less chance of errors. Second, since SS is so pervasive as a stock-assessment framework, the results are directly applicable to many stock assessments in practice.

There are, however, complications to conducting stock-assessment simulations that are on a large enough scale to render useful conclusions, fast enough to make conducting them feasible, and reproducible enough to make them credible and accurate. Complications include how to manage data and file structure, how to avoiding coding errors, how to repeatedly manipulate simulation models to ask specific questions, and how to translate models and questions across stocks and species. Further, while high-level scripting languages such as R and Python have become the standard for much of scientific analysis and visualization, and the stock-assessment framework Stock Synthesis is increasingly the standard for fisheries stock assessment, we lack a generalized framework to link these two components in a simulation context.

In this paper we introduce ss3sim, a software package for the popular statistical programming language R 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. Then, to demonstrate how a researcher might conduct a stock-assessment simulation with ss3sim, we work through an example starting at a research question and ending with plots and interpretation of the output. Our example includes considerations for setting up operating and estimation models, choosing a folder structure, model testing, and output manipulation and plotting. 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.

Extra notes:

- SS is increasingly used in simulation studies.
- Piner et al. (2011) example of stock-assessment simulation research with SS3
- Methot and Taylor (2011) example of stock-assessment research with SS
- been used in XX stock assessments world wide (˜60 as of 2012 - ask Rick) and

involved in many more currently

- allows a separation of research from stock assessment that informs management (Methot and Wetzel, 2012)
- been instrumental to investigating new stock assessment concepts: e.g. Piner et al. (2011), Methot and Taylor (2011)

The ss3sim framework

Terminology

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. Whereas the OM refers to the underlying truth, the EM generates our perception of that truth. We use the term *scenario* to refer to a combination of operating and estimation model *cases*. For example, an OM case might be natural mortality that follows a random walk, an EM case might be estimating a fixed parameter for natural mortality, and the combination of these two cases along with all other specified conditions creates a scenario. We refer to *iterations* or *replicates* as repeated simulations of a scenario with potentially 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 plaintext control files. Further, the plaintext control files refer to individual cases, which allows researchers to reuse control files as much as possible across scenarios. This reduces the chance for errors and makes the exploration of new scenarios rapid and simple. ss3sim than retains all SS3 output files as well as generating its own log files for documentation.

Flexible: ss3sim allows the user to specify their own OM and EM using all the possible configurations of SS3. ss3sim can take input in a number of forms (in R list format or through control files), and return output in a standard comma-separated-value (.csv) format allowing researchers to work with the output either using the package provided functions or their own tools.

Rapid: First, ss3sim relies on SS3, which uses ADMB as a backend optimization platform — the most rapid and robust optimization software available today (REF). Second, we built ss3sim so that it is easy to deploy across multiple computers or multiple researchers and re-combine the output. The scenarios are stored in a flat folder structure so they can be easily re-combined. Third, the package provides a number of functions to make visualization fast and easy. 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 simulation 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 you will assess that truth (an SS3 EM), (3) and a set of cases that deviate from these base models that you want to compare (configuration arguments provide as R list objects or plaintext control files). ss3sim works, in general, by converting simulation arguments (e.g. a given natural mortality trajectory) into manipulations of SS3 configuration files at an appropriate stage along with running the OM and EM as needed.

Low-level generic ss3sim functions

See Table 1 for 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)

How ss3sim complements other generic stock-assessment simulation software

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

r4ss

- Taylor et al. (2013)
- 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

- Kell et al. (2007) for FLR and Hillary (2009) 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. Lee et al. (2012)
    2. Piner et al. (2011)
    3. Lee et al. (2011)

The need for balance between generalizing and tailoring in simulation software

- maybe?
- 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

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: Andre Punt, Richard Methot, Ian Taylor, James Thorson, …

Figure captions

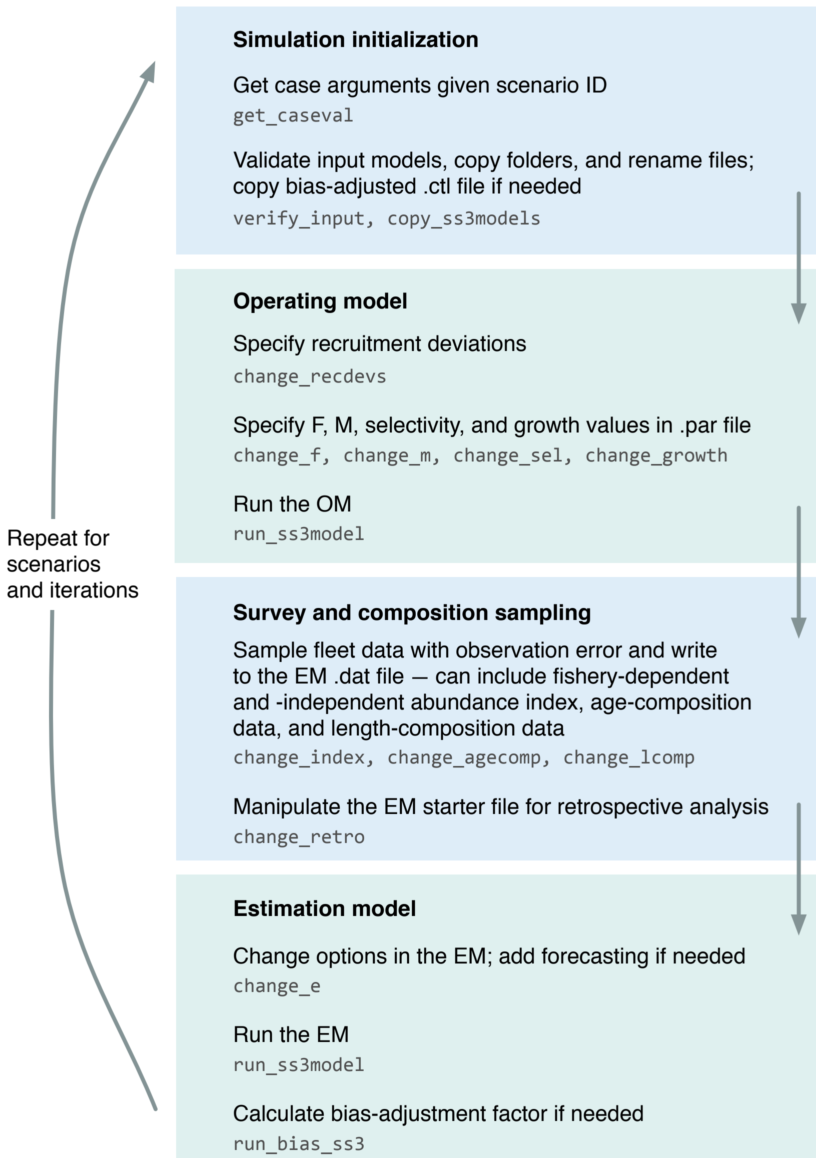

Figure 1: Flow diagram of `run_ss3sim` stock-assessment simulation steps.

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

Hilborn, R. W., and Walters, C. 1992. Quantitative Fisheries Stock Assessment: Choice, Dynamics, and Uncertainty. Chapman and Hall, London.

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

Jiao, Y., Smith, E. P., O'Reilly, R., and Orth, D. J. 2012. Modelling non-stationary natural mortality in catch-at-age models. ICES Journal of Marine Science, 69: 105–118.

Kell, L. T., Mosqueira, I., Grosjean, P., Fromentin, J.-M., Garcia, D., Hillary, R., and Jardim, E.*et al.* 2007. FLR: an open-source framework for the evaluation and development of management strategies. ICES Journal of Marine Science, 64: 640–646.

Lee, H.-H., Maunder, M. N., Piner, K. R., and Methot, R. D. 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.

Lee, H.-H., Maunder, M. N., Piner, K. R., and Methot, R. D. 2012. Can steepness of the stock-recruitment relationship be estimated in fishery stock assessment models?. Fisheries Research, 125–126: 254–261.

Maunder, M. N., and Punt, A. E. 2012. A review of integrated analysis in fisheries stock assessment. Fisheries Research, 142: 61–74.

Methot, R. D., and Taylor, I. G. 2011. Adjusting for bias due to variability of estimated recruitments in fishery assessment models. Canadian Journal of Fisheries and Aquatic Sciences, 68: 1744–1760.

Methot, R. D., and Wetzel, C. R. 2012. Stock Synthesis: A biological and statistical framework for fish stock assessment and fishery management. Fisheries Research, 142: 86–99.

Piner, K. R., Lee, H.-H., Maunder, M. N., and Methot, R. D. 2011. A simulation-based method to determine model misspecification: examples using natural mortality and population dynamics models. Marine and Coastal Fisheries, 3: 336–343.

Taylor, I., Stewart, I., Hicks, A., Garrison, T., Punt, A., Wallace, J., and Wetzel, C. 2013. r4ss: R code for Stock Synthesis. `http://code.google.com/p/r4ss/`.