

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

Sean C. Anderson<sup>1</sup>, ... (authorship and order to be discussed)

<sup>1</sup>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

## 8 Introduction

Fisheries stock assessments provide decision makers with information to manage fish  
10 stocks (Hilborn and Walters, 1992). Stock-assessment simulation allows testing as-  
sessment methods in a controlled environment where the true state of a fishery is  
12 known (REF). Given the expanding complexity of assessment models (REF), sim-  
ulation is becoming an increasingly critical component of testing assessment meth-  
14 ods (REF), and the field of stock-assessment simulation is rapidly developing (e.g.  
Hillary, 2009; Jiao *et al.*, 2012; Lee *et al.*, 2011; 2012; Methot and Taylor, 2011;  
16 Piner *et al.*, 2011).

Stock Synthesis (SS; Methot and Wetzel, 2012), is a widely-used statistical catch-at-  
18 age (SCAA) stock-assessment framework. It implements Integrated Analysis popu-  
lation dynamics models using a wide range of minimally-processed data (Maunder  
20 and Punt, 2012). By using this already-developed assessment framework, stock-  
assessment scientists and peer reviewers can focus on the underlying science, instead  
22 of the model code (Methot and Wetzel, 2012). Owing to these advantages, SS3 (the  
third version of the software) is now one of the world’s most commonly-used stock-  
24 assessment tools, particularly on the west coast of the United States where it was  
used in 60%/XX assessments in 2012 (REF) and in Australia (REF).

26 While SS is increasingly the standard for fisheries stock assessment (Methot and  
Wetzel, 2012), and high-level programming languages such as R (R Core Team,  
28 2013) have become the standard for statistical computing and visualization, we lack  
a generalized framework to link these components in a simulation context. Here,  
30 we introduce ss3sim, an R software package that facilitates large-scale, rapid, and  
reproducible stock-assessment simulation with the widely-used SS framework. We  
32 begin by outlining the general philosophy of ss3sim and describing its functions. We

then demonstrate the software by developing a simple example where we consider  
the effect of increasing or decreasing survey effort and estimating or not estimating  
natural mortality  $M$ . 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

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 researchers  
to reuse control files as much as possible 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`) 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 (Fournier *et al.*, 2012). 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 you will 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 (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:

- 82 1. Functions that manipulate SS configuration files. These manipulations gener-  
84 ate an underlying “truth” (OM) and control our assessment of those models  
(EM).
- 86 2. Functions that conduct simulations. These functions generate a folder struc-  
ture, call manipulation functions, run **SS3** as needed, and save the output.
3. Functions for analyzing and plotting simulation output.

## 88 **High-level tailored ss3sim functions**

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

## **An example simulation with ss3sim**

96 (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  
98 elements to give a flavour for what can be done in the main paper)

## **Setting up the SS models**

- 100 • choosing a specific conditioning model or generic conditioning type
- setting up the OM and EM SS models

- 102 • things to keep in mind
- running through SS to format as `.ss_new` files and renaming

## 104 **File and folder setup**

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

## 108 **Translating research questions into configuration files**

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

## 112 **Deterministic model testing**

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

## **Output analysis and visualization**

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

## 118 **Discussion**

- Other sections?
- 120 • 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*

- 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”*

- 144 • <http://fisherysimulation.codeplex.com>, Windows only, GUI..., works on bootstrapped data only, therefore isn't as flexible as ss3sim. Used in:

- 146 1. Lee et al. (2012)
2. Piner et al. (2011)
- 148 3. Lee et al. (2011)

## Research opportunities with ss3sim

- 150 • there are lots, we should brainstorm some key ones

## Acknowledgements

- 152 • funding: Fulbright Canada, NSERC, Simon Fraser University, many others...
- discussions and advice: André Punt, Richard Methot, Ian Taylor, James Thor-
- 154 son, ...



## Figures

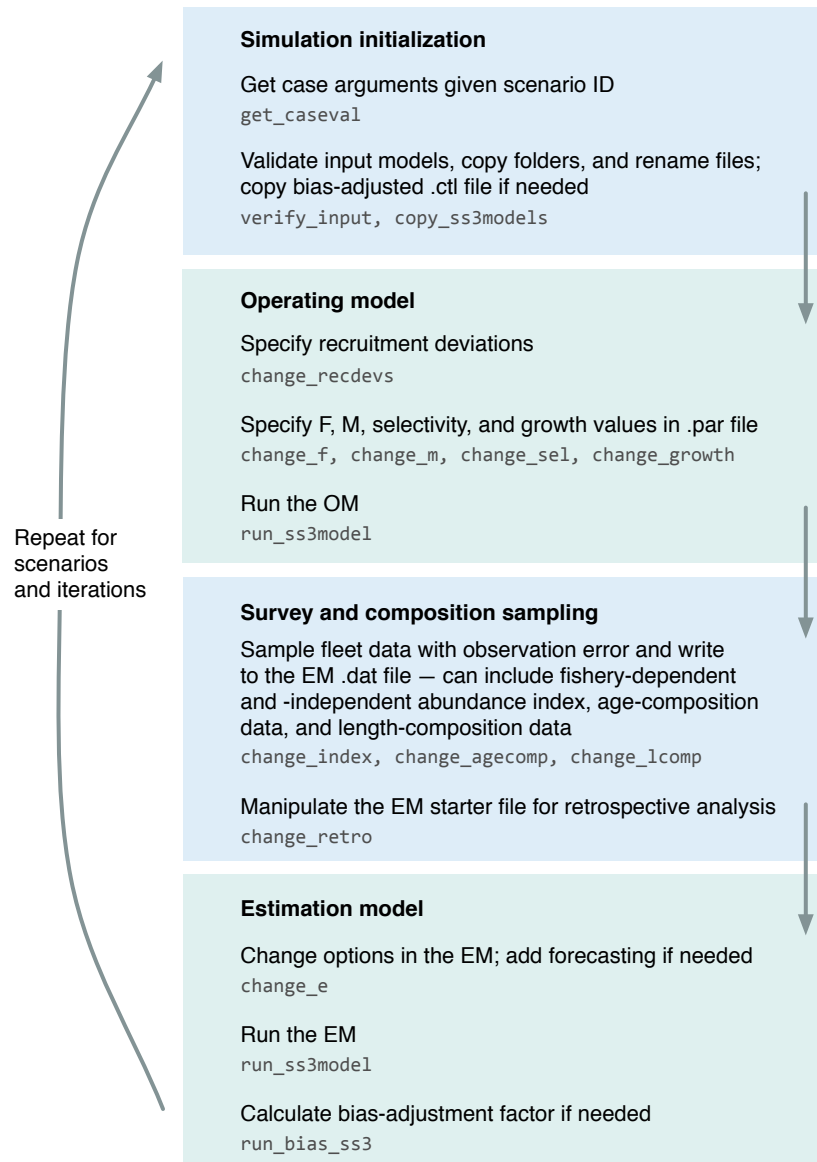


Figure 1: Flow diagram of `run_ss3sim()` stock-assessment simulation steps.

<sup>156</sup> Figure 2: Panels with output from the example

## Tables

158 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  
160 itself.

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

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

## References

Fournier, D. A., Skaug, H. J., Ancheta, J., Ianelli, J., Magnusson, A., Maunder, M. N., and Nielsen, 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.

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.

- 188 Maunder, M. N., and Punt, A. E. 2012. A review of integrated analysis in fisheries  
stock assessment. *Fisheries Research*, 142: 61–74.
- 190 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.
- 192 Methot, R. D., and Wetzel, C. R. 2012. Stock Synthesis: A biological and statistical  
framework for fish stock assessment and fishery management. *Fisheries Research*,  
194 142: 86–99.
- Piner, K. R., Lee, H.-H., Maunder, M. N., and Methot, R. D. 2011. A simulation-  
196 based method to determine model misspecification: examples using natural mortality  
and population dynamics models. *Marine and Coastal Fisheries*, 3: 336–343.
- 198 R Core Team. 2013. R: A Language and Environment for Statistical Computing.  
R Foundation for Statistical Computing, Vienna, Austria. [http://www.R-project.](http://www.R-project.org/)  
200 [org/](http://www.R-project.org/).
- Taylor, I., Stewart, I., Hicks, A., Garrison, T., Punt, A., Wallace, J., and Wetzel, C.  
202 2013. r4ss: R code for Stock Synthesis. <http://code.google.com/p/r4ss/>.