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

Sean C. Anderson[1*], Author2, Author3, … (authorship order to be discussed)

1 Introduction

What is stock assessment simulation? Why is it increasingly critical?

- Stock-assessment simulation is a critical component to evaluating stock assessment methods and understanding their strengths and weaknesses. …
- need refs

What is SS3, why is it important, why simulate with it?

- SS software ref: Methot and Wetzel (2012)

- ADMB software ref: Fournier et al. (2012)

- Importance of integrated analysis with SS as an example: Maunder and Punt (2012)

- most widely used now

- facilitates rapid, reproducible analyses… focus on peer-review of the science not the modelling code

- integrated analysis - models population dynamics using a wide range of data (Maunder and Punt, 2012)

- 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)

- been used in XX stock assessments world wide (~60 as of 2012 - ask Rick) and involved in many more currently

- Piner et al. (2011) example of stock-assessment simulation research with SS3

- Methot and Taylor (2011) example of stock-assessment research with SS

Methot and Wetzel (2012):

> A comprehensive modeling framework such as SS enhances communication, efficiency, and education in the fishery assessment community (Methot, 2009). Communication is enhanced by creating a familiarity among users, reviewers, and clients regarding terminology and approach. Reviewers who are already familiar with SS can quickly focus on key issues for the assessment being reviewed, rather than spend time learning the features of a novel assessment model.

Therefore two benefits to simulating with SS: (1) much of the model has already been built (research can then progress rapidly and with less chance of errors) and checked and (2) the results are directly applicable to the tools used by stock assessment scientists.

However, there are many complications to conducting large-scale, rapid, and reproducible simulations.

- complications range from data, file, and folder management
- avoiding coding errors
- repeatedly manipulating operating models and estimation models to ask specific questions
- porting models and questions across stocks and species
- reproducible, understandable, and documented
- barrier to research
- R is the standard, but existing solutions are GUI and therefore not as flexible, scriptable, and repeatable.

<–Our goal is to provide a toolkit and general framework for fast, transparent, and reproducible stock assessment simulation.–>

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.

2 The ss3sim framework

2.1 Terminology

- operating model
- estimation model
- scenario
- case
- iteration
- a simulation = all scenarios and iterations
- SS the framework vs. `SS3` the binary of version 3 of SS

2.2 General philosophy

- Reproducible: documented in code or plain-text control files
- Flexible: specify your own OM and EM using all the possibilities of SS3
- Rapid: use SS3, which relies on ADMB; build on the existing SS and r4ss (Taylor *et al.*, 2013) framework
- Allow researchers to reuse control files as much as possible across scenarios to make the simulation easy to understand and less error prone
- Easy to deploy across multiple researchers, multiple computers, or multiple cores
- Leave a trace throughout of what has happened - log files etc.
- Make visualization easy so users are likely to use it to understand their models and detect errors quickly
- Minimize book keeping etc. so that researchers can concentrate on the science

2.3 General structure

- start with an OM and EM
- Specify how you want to affect the truth and your assessment of the truth (possibly in a time varying manner)
- ss3sim carries that out in SS3 by manipulating and running the models in R
- SS is controlled by the user through a series of plain-text configuration files
- ss3sim works, in general, by converting simulation arguments (e.g. a given natural mortality trajectory) into manipulations of these configuration files at the appropriate stage in conjunction with running appropriate models

2.4   Low-level ss3sim functions

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

2. Functions that conduct simulations. These functions generate a folder structure, call manipulation functions, run `SS3` as needed, and save the output. `run_ss3sim copy_models run_ss3model`

3. Functions for analyzing and plotting simulation output.

2.5   High-level ss3sim functions

- an example framework
- 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 …, … "Environmental correlates of recruitment fluctuations can be included in SS in two ways (Schirripa et al., 2009)" (Methot and Wetzel, 2012)

3   An example simulation with ss3sim

3.1   Setting up the SS models

- OM and EM models
- things to keep in mind
- running through SS to format as `.ss_new` files and renaming

3.2   File and folder setup

- required files
- nested vs. flat folders

3.3   Translating research questions into configuration files

- E.g. time-varying M

3.4   Deterministic model testing

- reduce recdevs, reduce sigma R, bias correction
- what to expect, what to plot

3.5 Output analysis and visualization

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

4 Discussion

4.1 How ss3sim complements other simulation software

4.2 The need for balance between generalizing and tailoring

maybe?

4.3 Research opportunities with ss3sim

5 Acknowledgements

6 Figure captions

Figure 1: Flow diagram of ss3sim

Figure X: Folder and file setup?

Figure X: Panels with example model and output

7 Tables

Table 1: User-facing ss3sim functions and a description of their purpose.

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?

8 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.

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/.