Assessment For All (a4a)

The stock assessment model

Ernesto Jardim
Colin Millar
Iago Mosqueira
Chato Osio
Finlay Scott

European Commission
Joint Research Centre

Installing a4a

We have developed a4a under R-3.0

```
install.packages("FLCore",
repos="http://flr-project.org/Rdevel")

install.packages("devtools")

library(devtools)

install_github("a4a", "colinpmillar", subdir =
"packages/FLa4a")
```

see: https://github.com/colinpmillar/a4a

Context

Problem

- Lots of stocks
- Lots of data
- Limited resources

Solution

- Make modelling more accessible
- Automate some processes

Intuitive Modelling

Intuitive for who?

Fisheries Scientists

Fisheries scientist =

Biologists, Oceanographers, Gear technologists ...

Linear models are one of the most common modelling tools used in general scientific work.

Intuitive Modelling

It is not always obvious that stock assessments are often composed of linear models.
For example, the classical separable F assumption is simply that

$$F_{ay} = S_a \times F_y$$

which, in linear modelling parlance is

$$\log F \sim \text{age} + \text{year}$$

Intuitive Modelling

The "language" of linear models has been developing within the statistical community for many years:

- 1965 J. A. Nelder, notation for randomized block design
- 1973 Wilkinson and Rodgers, symbolic description for factorial designs
- 1990 Hastie and Tibshirani, introduced notation for smoothers
- 1991 Chambers and Hastie, further developed for use in S

Many modelling software use this language: Minitab, spss, genstat, SAS, R, S-plus.

Some examples

A separable model where the level of F is smooth through time

$$\log F \sim \text{age} + \text{s(year)}$$

Some examples

A separable model where F is smooth over age

$$\log F \sim \mathrm{s(age)} + \mathrm{year}$$

Some examples

F is smooth over age and year

$$\log F \sim s(\text{age}, \text{year})$$

Some examples

F is smooth over age and year, and there is limited interaction between age and year

$$\log F \sim \mathrm{s(age)} + \mathrm{s(year)} + \mathrm{s(age,\ year)}$$

Some examples

F is modelled by 2 separable periods, coded by **block**

$$\log F \sim \text{age:block} + \text{year}$$

Some examples

F is modelled by 2 separable periods, coded by **block**

$$\log F \sim \mathrm{s(age)}{:}\mathrm{w} + \mathrm{s(age)}{:}(1\text{-}\mathrm{w}) + \mathrm{year}$$

Some examples

A SAM or TSA mimic for F is would be

$$\log F \sim \mathrm{s(year)}{:}\mathrm{age} + \mathrm{s(age)}$$

Model Choices in a4a

For example in selectivity

$$\log Q \sim \overbrace{\text{log Contact Selectivity}}^{\text{offset}} + \underbrace{\text{log Availability}}_{\text{formula}}$$

Model Choices in a4a

With a linear model you can fit

- linear and smooth functions of age and year
- seperable models
- partially seperable
- non-seperable
- step changes (in level, in smoother form)
- covariates (smoothed and linear)

These can be applied to log **F**, log **catchability**, **stock recruit** parameters, observation **variance.**

Model detail

$$e^{\mathrm{E}[\log C]} = \frac{F}{F+M}\left(1 - e^{-F-M}\right) R e^{-\sum F+M}$$

and

$$e^{\mathrm{E}[\log I]} = QRe^{-\sum F+M}$$

and

$$\mathrm{Var}\left[\log C_{ay}\right] = \sigma^2_{ay} \qquad \mathrm{Var}\left[\log I_{ays}\right] = \tau^2_{ays}$$

Model detail

linear models for

- log F
- log Q
- log observation variances
- log initial age structure

Recruitment is modelled as a fixed variance random effect with linear models for

- log a
- log b

where relevant. Models available: Ricker, Beverton Holt, smooth hockeystick, geometric mean

What we can do, what we can't do

Can:

- missing values: missing at random
- multiple surveys
- variable Q, F, variance
- splines (fixed degreed of freedom)
- stock recruit relationship (fixed variance)
- stock recruit relationship (estimated variance) SLOW
- *fixed variance random effects: RW1, RW2, seasonal, user specified*

Can't:

- estimate random effect variance
- estimate smoothing parameters
- estimate growth parameters

What we can do

- simulate from the distribution of model params
  - normal approx
  - avoids the need for delta approx
  - can be biased, but we can also use MCMC if desired
- we can approximate the (joint) distribution of
  - terminal year Fs and Ns
  - terminal year Fbar and Fmsy
  - F / Fmsy

Current development

We are incorporating:

- length to age transformation
- a more intuitive interface
- model averaging
- use of biomass surveys

And thinking about:

- soft constraints on total catch weight?
- estimation of smoothness
- random effects variances

Length to age

Algorithm to generate 1 sample of F, Fbar, Fmsy etc.:

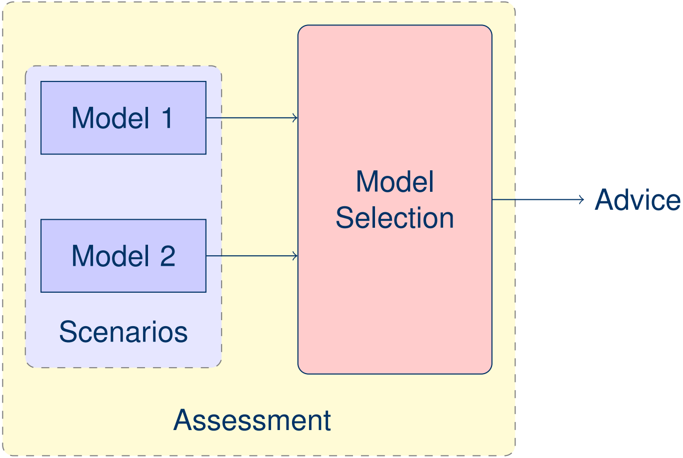
1. simulate some growth parameters
2. simulate age data, conditional on available length data
3. fit the a stock assessment model
4. simulate from the model
5. calculate summaries (F, Fmsy, ...) and store

A more intuitive interface...

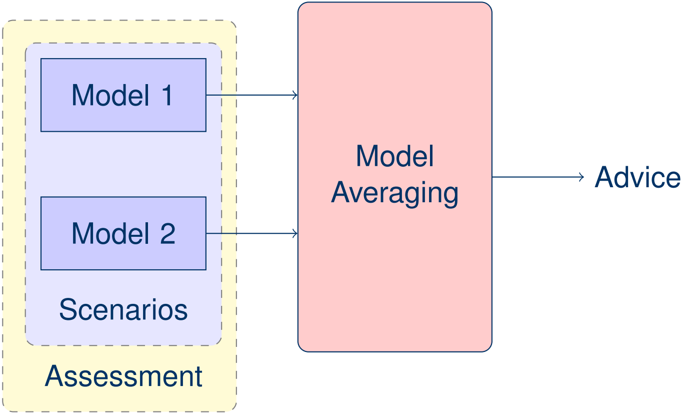
model builder functions:

- breakpoints(year, c(1990, 2000))
- trawl(plateau = 5)
- SAM(k = 10)
- acoustic(absolute = FALSE)
- ...

An Assessment Process



Model Averaging can help automation



Expert knowledge for model specification

Different plausible models for different levels

- Management area level (North Sea, Baltic Sea, ...)
- Species type (roundfish, flatfish, pelagic, Nephrops)
- specific groups (North Sea gadoids)

This provides a framework for setting up plausible models for new species.

Can lots of simple models averaged = a good model?
Kearns: Can a set of weak learners create a single strong learner

Thank you for listening!