Gmacs

A generalized size-structured stock assessment model

The Gmacs Development Team

June 30, 2015

Outline

Notation

Generally

- a bold capital symbol $\mathbf{A}$ refers to a matrix
- a bold lowercase symbol $\mathbf{a}$ refers to a vector
- an unbolded italic symbol a refers to a scalar
- $\{a_i\}_{i=1}^n$ is an ordered n-tuple
- the terms $p(\cdot)$ or $\pi(\cdot)$ represent probability distributions
- $a|b$ means event a conditional on event b having occurred
- the symbol $\forall$ means for all values, usually referring to all of the values within an ordered tuple.
- we use red to indicate an estimable parameter
- we use blue to represent covariates and fixed parameters
- we use green to represent data

Indices

Symbol	Description
g	group
h	sex
i	year
j	time step (years)
k	gear or fleet
ℓ	index for size class
m	index for maturity state
o	index for shell condition

Notice no area index.

Leading model parameters

Symbol	Support	Description
M_0	$0 < M_0 < \infty$	Initial instantaneous natural mortality rate
R_0	$0 < R_0 < \infty$	Unfished average recruitment
$\ddot{R}$	$0 < \ddot{R} < \infty$	Initial recruitment
$\bar{R}$	$0 < \bar{R} < \infty$	Average recruitment
α_r	$\alpha_r > 0$	Mode of size-at-recruitment
β_r	$\beta_r > 0$	Shape parameter for size-at-recruitment
κ	$\kappa > 1$	Recruitment compensation ratio
ρ	$\rho > 1$	Recruitment autocorrelation

We group the leading model parameters into the vector

$$\boldsymbol{\theta} = \{M_0, R_0, \ddot{R}, \bar{R}, \alpha_r, \beta_r, \kappa\}.$$

Cubic splines

A spline is a numeric function that is piecewise-defined by polynomial functions, and which possesses a sufficiently high degree of smoothness at the places where the polynomial pieces connect (which are known as knots). A cubic spline is constructed of piecewise third-order polynomials. The second derivative of each polynomial is commonly set to zero at the endpoints, since this provides a boundary condition that completes the system of $m - 2$ equations. This produces a so-called "natural" cubic spline and leads to a simple tridiagonal system which can be solved easily to give the coefficients of the polynomials. However, this choice is not the only one possible, and other boundary conditions can be used instead.

Growth parameters

Symbol	Support	Description
α_h	$\alpha_h > 0$	Mode of size-at-recruitment
β_h	$\beta_h > 0$	Shape parameter for size-at-recruitment
φ_h	$\varphi_h > 0$	Instantaneous natural mortality rate
μ_h	$\mu_h > 0$	Length at 50% molting probability
c_h	$c_h > 0$	Coefficient of variation of molting probability

We group the growth parameters into the vector

$$\boldsymbol{\psi} = \{\alpha_h, \beta_h, \varphi_h, \mu_h, c_h\}.$$

Growth matrix

The average molt increment from size class ℓ to ℓ' is assumed to be sex-specific and is defined by the linear function

$$a_{h,\ell} = \frac{\alpha_h + \beta_h \ell}{\varphi_h}.$$

The probability of transitioning from size class ℓ to ℓ' assumes that variation in molt increments follows a gamma distribution
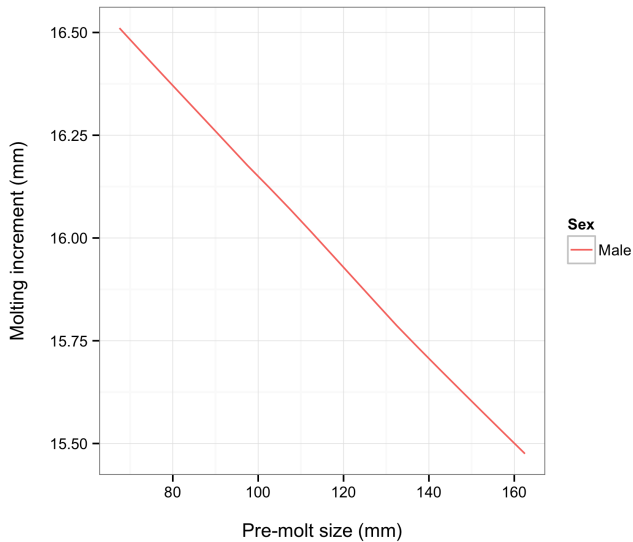
$$p(\ell'|\ell)_h = \boldsymbol{G}_h = \int_{\ell}^{\ell+\Delta\ell} \frac{\ell^{a_{h,\ell-1}} \exp\left(\frac{\ell}{\varphi_h}\right)}{\Gamma(a_{h,\ell})\ell^{a_{h,\ell}}} \quad \text{where} \quad \Delta\ell = \ell' - \ell.$$
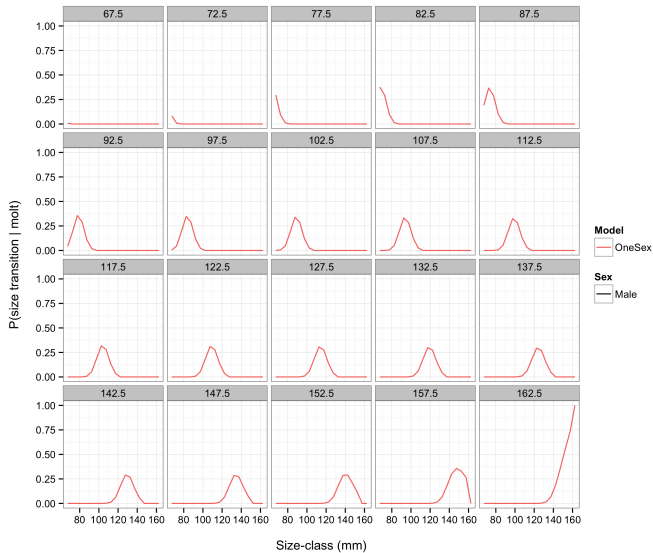
Specifically

$$\boldsymbol{G} = G_{\ell,\ell'} = \begin{pmatrix} G_{1,1} & G_{1,2} & \dots & G_{1,n} \\ G_{2,1} & G_{2,2} & \dots & G_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ G_{n,1} & G_{n,2} & \dots & G_{n,n} \end{pmatrix} \quad \text{where} \quad \sum_{\ell'} G_{\ell,\ell'} = 1 \quad \forall \ell.$$

Growth increments $(a_{h,\ell})$

Growth transitions ($\boldsymbol{G}_h$)

Molting probability ($\boldsymbol{P}_h$)

The standard deviation of molting probability (σ_h) is calculated from the length at 50% molting probability (μ_h) coefficient of variation of molting probability (c_h) as
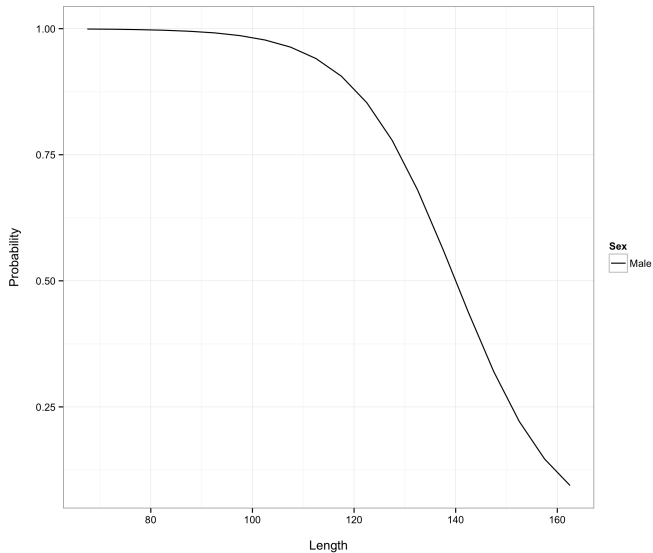
$$\sigma_h = \mu_h c_h.$$

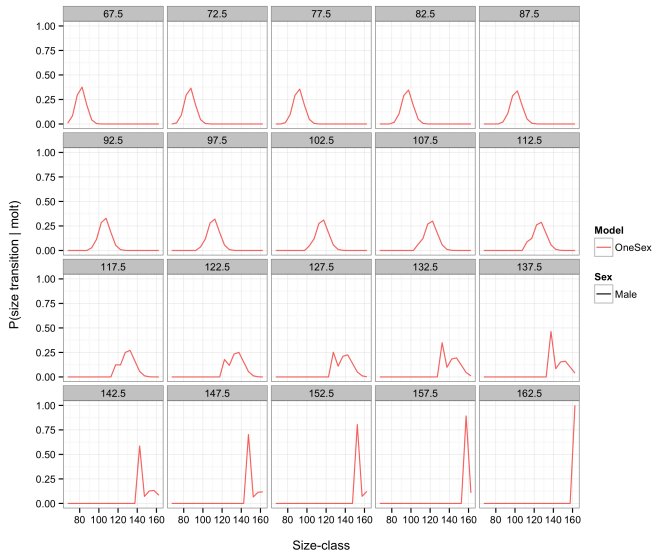The molting probability ($\boldsymbol{P}_h$) is calculated as

$$\boldsymbol{P}_h = 1 + (-1 - \exp(\mu_h - \ell)/\sigma_h)^{-1}.$$

The molting probability ($\boldsymbol{P}_h$) and the growth probability ($\boldsymbol{G}_h$) are combined to yield the size transition matrix ($\boldsymbol{P}_h \boldsymbol{G}_h$).

Molting probability ($\boldsymbol{P}_h$)

Size transitions ($\boldsymbol{P}_h\boldsymbol{G}_h$)

Natural mortality variables

Symbol	Description
$M_{0,h}$	Initial instantaneous natural mortality rate
σ_M	Standard deviation of natural mortality
δ_i	Natural mortality deviate
$M_{h,i}$	Natural mortality by sex h and year i

Natural mortality options

Natural mortality (M) is assumed to be sex-specific (h), size-independent (ℓ), and may or may not be constant over time (i). The options currently available in Gmacs include:

1. Constant natural mortality ($M_{h,i} = M_{0,h}$)
2. Random walk (deviates constrained by variance σ_M^2)
3. Cubic Spline (deviates constrained by nodes and node placement)
4. Blocked changes (deviates constrained by variance in specified blocks $\iota \in i$)

Natural mortality: option 2

If time-varying natural mortality is specified using the **random walk** option, the model constrains $M_{h,i}$ to be a random-walk process with variance σ_M^2

$$M_{h,i+1} = \begin{cases} M_{0,h} & \text{for} \quad i = 1 \\ M_{h,i}e^{\delta_i} & \text{for} \quad i > 1 \end{cases} \quad ,$$
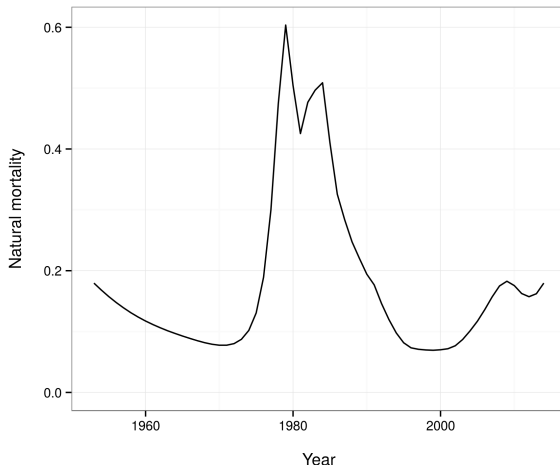
where

$$\delta_i \sim \mathcal{N}\left(0, \sigma_M^2\right).$$

A time-varying natural mortality can be estimated for all years (i), or for specified blocks of years ($\iota \in i$).
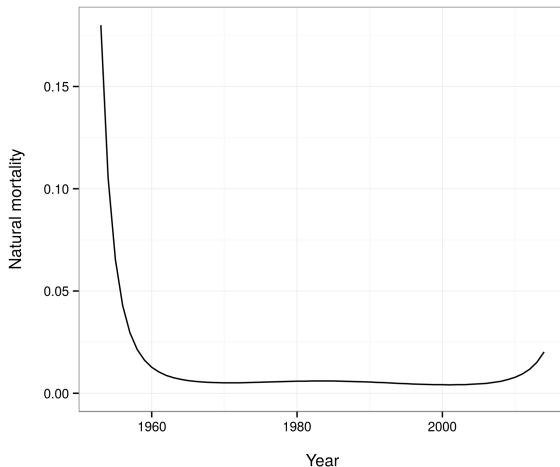
Natural mortality: option 2

Below we present an example in which time-varying natural mortality is estimated as a **random walk** process for all years (i)

Natural mortality: option 3

If time-varying natural mortality is specified using the **cubic spline** option, the model constrains $M_{h,i}$ to be a cubic spline process at specified knots. For example, setting

Natural mortality: option 4

If time-varying natural mortality is specified using the **blocked changes** option, the model constrains $M_{h,i}$ by the variance (σ_M^2). For example, setting $\sigma_M^2 = 0.04$ and four specific years (1976, 1980, 1985, 1994) we get
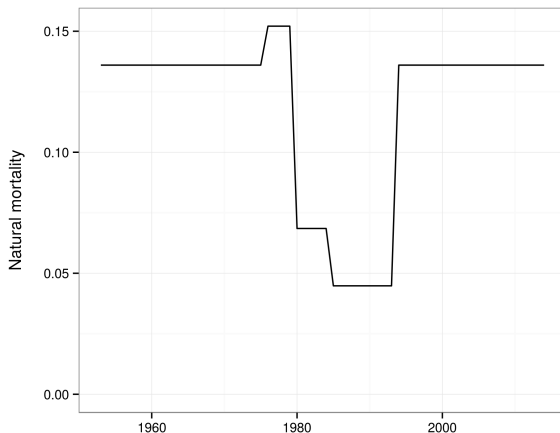
Selectivity, retention and fishing mortality

Symbol	Dimensions	Description
$a_{h,i,k}$	1	Length at 50% selectivity
$\sigma^s_{h,i,k}$	1	Standard deviation in length at selectivity
$\boldsymbol{s}_{h,i,k}$	$\ell \times 1$	Length at 50% selectivity in length interval ℓ
$r_{h,i,k}$	1	Length at 50%
$\sigma^y_{h,i,k}$	1	Standard deviation in length at retention
$\boldsymbol{y}_{h,i,k}$	$\ell \times 1$	Length at 50% retention in length interval ℓ
$\xi_{i,k}$	1	Discard mortality rate for gear k in year i
$\boldsymbol{\nu}_{h,i,k}$	$\ell \times 1$	Vulnerability due to fishing mortality for sex h
$\bar{\boldsymbol{f}}_k$	$i \times 1$	Average fishing mortality rate for gear k
$\boldsymbol{\Psi}_{i,k}$		Fishing mortality deviate for gear k in year i
$\boldsymbol{F}_{i,k}$		Annual fishing mortality rate for gear k in year i

Selectivity and retention

The probability of catching an animal of sex h, in year i, in fishery k, of length ℓ (i.e. selectivity) is

$$s_{h,i,k} = \left(1 + \exp\left(-(\ell - a_{h,i,k})/\sigma^s_{h,i,k}\right)\right)^{-1}.$$

The probability of an animal of sex h, in year i, in fishery k, of length ℓ being retained is

$$y_{h,i,k} = \left(1 + \exp\left(-(r_{h,i,k} - \ell)/\sigma^y_{h,i,k}\right)\right)^{-1}.$$

Selectivity, retention and fishing mortality

The joint probability of vulnerability due to fishing and discard mortality is

$$\boldsymbol{\nu}_{h,i,k} = \boldsymbol{s}_{h,i,k} \left[\boldsymbol{y}_{h,i,k} + (1 - \boldsymbol{y}_{h,i,k}) \xi_{i,k} \right],$$

where $\xi_{i,k}$ is the discard mortality rate for fishery k in year i. Finally the fishing mortality is calculated as

$$\boldsymbol{F}_{h,i} = \sum_k \exp \left(\bar{\boldsymbol{f}}_k + \boldsymbol{\Psi}_{i,k} \right) \boldsymbol{\nu}_{h,i,k},$$

The vector $\boldsymbol{F}_{h,i}$ represents all mortality associated with fishing, including discards in directed and non-directed fisheries.

Selectivity and retention

Assuming that selectivity for the NMFS trawl fishery is split into two blocks (1975-1981 and 1982-2014) and that retention is constant with time $\boldsymbol{y}_{h,i,k} = \boldsymbol{y}_{h,k}$
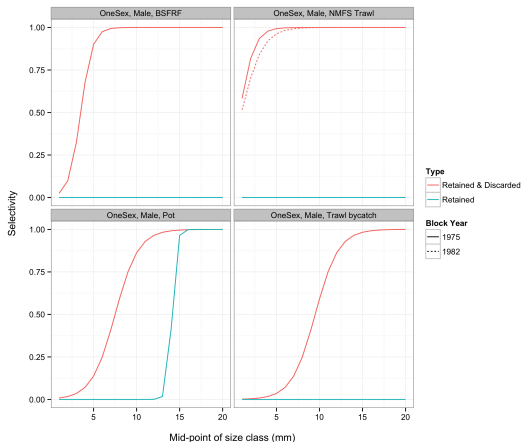
Recruitment

Recruitment size-distribution

$$\alpha = \frac{\alpha_r}{\beta_r},$$

$$p[x_\ell - 0.5\Delta x \le x \le x_\ell + 0.5\Delta x] = p[x] = \int_{x_\ell - 0.5\Delta x}^{x_\ell + 0.5\Delta x} \frac{x^{\alpha-1} \exp\left(\frac{x}{\beta_r}\right)}{\Gamma(\alpha) x^\alpha} dx.$$

Initial recruitment

$$\boldsymbol{r}_{h,i} = 0.5 p[x] \ddot{R} \quad \text{for} \quad i = 1.$$

Recruitment

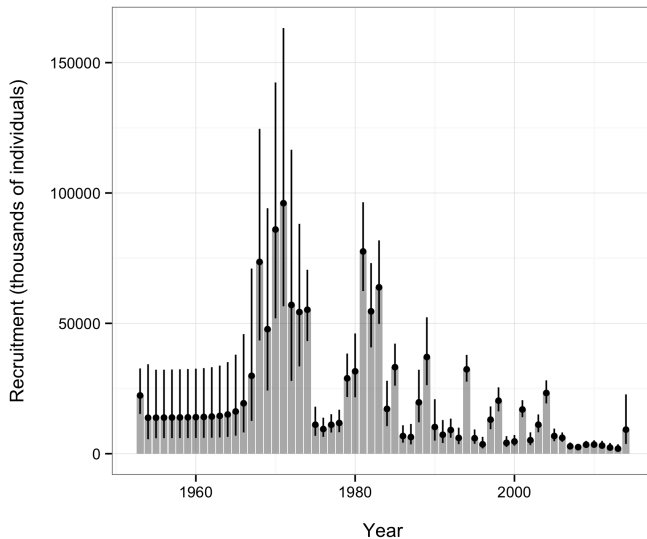$$\boldsymbol{r}_{h,i} = 0.5 p[x] \bar{R} e^{\delta_i} \quad \text{for} \quad i > 1,$$

where

$$\delta_i = \log(r_i) - (1 - \rho) \log(\bar{R}) - \rho \log(r_{i-1}) + 0.5 \sigma_R^2$$

and

$$r_i = \sum_h r_{h,i}.$$

Recruitment

Population dynamics

Growth and survival process are combined, represented by

$$\boldsymbol{A}_{h,i} = \begin{cases} \boldsymbol{G}_h \left[\exp(-M_{h,i})\boldsymbol{I}\right] & \text{for} \quad i = 1 \\ \boldsymbol{G}_h \left[\exp(-\boldsymbol{Z}_{h,i})\boldsymbol{I}\right] & \text{for} \quad i > 1 \end{cases} .$$

where

$$\boldsymbol{Z}_{h,i} = M_{h,i} + \boldsymbol{F}_{h,i}.$$

Assuming unit recruitment, then the growth and survivorship in unfished and fished conditions is given by the solutions to the equations

$$\boldsymbol{u}_{h,i} = -(\boldsymbol{A}_{h,i} - \boldsymbol{I})^{-1}(p[x]) \quad \forall i.$$

The vector $\boldsymbol{u}_{h,i}$ represent the unique equilibrium solution for the numbers per recruit in each size category.

Initial population

The mean weight at length (ℓ) by sex (h) is represented by the $\ell \times 1$ vector $\boldsymbol{w}_h$ and can take any form the user wishes

$$\boldsymbol{w}_h = f_w(\ell, \theta)$$

Similarly, the average proportion mature at length (ℓ) by sex (h) is represented by the $\ell \times 1$ vector $\boldsymbol{w}_h$ and can take any form the user wishes

$$\boldsymbol{m}_h = f_m(\ell, \theta)$$

Steady-state conditions

$$B_0 = R_0 \sum_h \lambda_h \sum_\ell \boldsymbol{u}_{h,i} \boldsymbol{w}_h \boldsymbol{m}_h \quad \text{for} \quad i = 1,$$

$$\tilde{B} = \tilde{R} \sum_h \lambda_h \sum_\ell \boldsymbol{u}_{h,i} \boldsymbol{w}_h \boldsymbol{m}_h \quad \text{for} \quad i > 1.$$

Population evolution

The total unfished numbers in each size category is defined as $R_0 \boldsymbol{u}_{h,i=1}$. Initial numbers at length

$$\boldsymbol{n}_{h,i} = \left[-\left(\boldsymbol{A}_h - \boldsymbol{I}\right)^{-1} \boldsymbol{r}_{h,i} \right] e^{\boldsymbol{\varepsilon}} \quad \text{for} \quad i = 1,$$

where $\boldsymbol{\varepsilon}$ is an $\ell \times 1$ vector of initial recruitment deviates. The numbers in each size-class in the following time-step ($\boldsymbol{n}_{h,i+1}$) is the product of the numbers in each size-class in the previous time-step ($\boldsymbol{n}_{h,i}$), size-specific growth and survival ($\boldsymbol{A}_{h,i}$), plus new recruits ($\boldsymbol{r}_{h,i}$)

$$\boldsymbol{n}_{h,i+1} = \boldsymbol{n}_{h,i} \boldsymbol{A}_{h,i} + \boldsymbol{r}_{h,i} \quad \text{where} \quad i \geq 1.$$

Likelihoods and penalties

Likelihoods

- likelihood of catch
- likelihood of relative abundance
- likelihood of size compositions
- likelihood of recruitment deviations
- likelihood of growth increment data

Penalties

- constrain $\log(\Psi_{i,k})$ to ensure they sum to zero
- constrain mean $\bar{\boldsymbol{f}}_k$ to regularize the solution
- constrain $M_{h,i}$ in random walk

Log-likelihood: catch

The standard deviation of the catch $(\sigma_{i,k})$ is calculated from the CV as

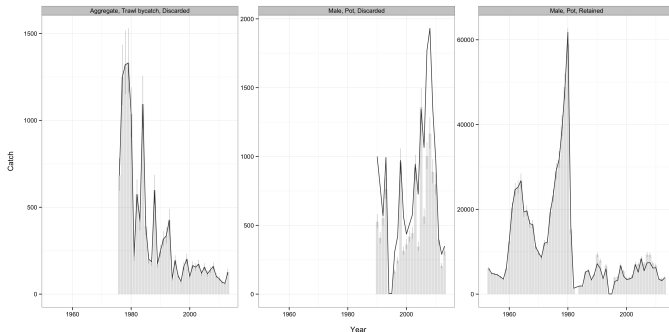$$\sigma_{i,k} = \sqrt{\log\left(1 + c_{i,k}{}^2\right)}.$$

The expected catch is calculated using the Baranov catch equation

$$\hat{C}_{i,k} = \sum_{\ell} \left[\boldsymbol{n}_{h,i}\boldsymbol{w}_h \frac{\boldsymbol{F}_{h,i}}{\boldsymbol{Z}_{h,i}} \left(1 - e^{-\boldsymbol{Z}_{h,i}}\right) \right] \quad \text{where} \quad \boldsymbol{Z}_{h,i} = M_{h,i} + \boldsymbol{F}_{h,i}.$$

The log-likelihood is

$$\ell(C_{i,k}) = 0.5\log(2\pi) + \log(\sigma_{i,k}) - \frac{1}{2\sigma_{i,k}{}^2}(C_{i,k} - \hat{C}_{i,k})^2.$$

Log-likelihood: catch

Log-likelihoods: relative abundance

The catchability coefficient q is treated as a nuisance parameter and integrated out of the model (Walters & Ludwig 1994).

$$q = \exp\left(\frac{1}{n}\sum_i \log\left(\frac{I_i}{V_i}\right)\right).$$

The standard deviation ($\sigma_{i,k}$) is calculated from the CV as

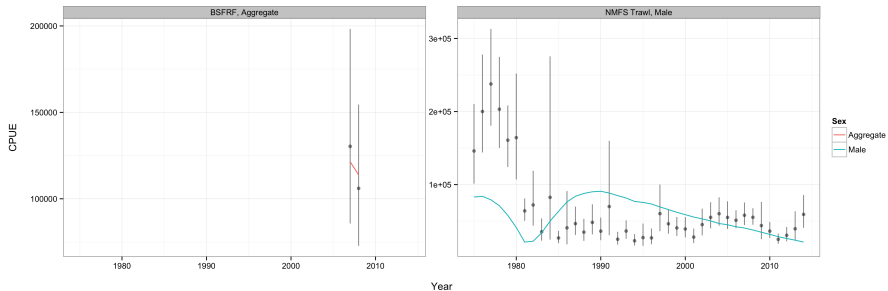$$\sigma_{i,k} = \sqrt{\log\left(1 + c_{i,k}{}^2\right)}.$$

The log-likelihood is

$$\ell(I_{i,k}) = \lambda\left(0.5\log(2\pi) + \log(\sigma_{i,k}) + \frac{1}{2\sigma_{i,k}^2}(I_{i,k} - qV_{i,k})^2\right).$$

Log-likelihood: relative abundance

Log-likelihoods: size composition

Size composition data is assumed to be multinomial distributed

$$\boldsymbol{P}_{h,i} = (P_\ell)_{h,i} = \mathcal{M}\text{ultinomial}\left(n_{h,i}, \boldsymbol{Q}_{h,i}\right)$$

Alternatively we could use

$$\boldsymbol{P}_{h,i} = (P_\ell)_{h,i} = \mathcal{D}\text{irichlet}\left(\lambda_0 n_{h,i} \boldsymbol{Q}_{h,i}\right).$$

In this context, λ_0 can be thought of as the data weight (which may be estimated in the model) and $n_{h,i}$ is the relative sample size between years.

Log-likelihoods

Natural mortality

$$\ell(M_{h,i}) = 0.5 \log(2\pi) + \log(\sigma_M) + 0.5 \frac{1}{\sigma_M^2} \sum_i \delta_i^2.$$

References

Walters, C. & Ludwig, D. (1994), 'Calculation of Bayes posterior probability distributions for key parameters', *Canadian Journal of Fisheries and Aquatic Science* **51**, 713–722.