

# Gmacs

A generalized size-structured stock assessment model

The Gmacs Development Team

June 27, 2015

# Outline

## 1 Notation and some definitions

- Notation
- Indices
- Leading model parameters
- Growth parameters
- Latent states
- Other variables
- Cubic splines

## 2 Growth

- Growth
- Molting

## 3 Natural mortality and survival

- Natural mortality

## 4 Selectivity, retention, fishing

- Selectivity and retention

## 5 Recruitment

## 6 Initialization

- Initialization

# 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.

# Indices

Symbol	Description
$g$	group
$h$	sex
$i$	year
$j$	time step (years)
$k$	gear or fleet
$\ell$	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
$\alpha_r$	$\alpha_r > 0$	Mode of size-at-recruitment
$\beta_r$	$\beta_r > 0$	Shape parameter for size-at-recruitment
$\kappa$	$\kappa > 1$	Recruitment compensation ratio

We group the leading model parameters into the vector

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

# Growth parameters

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

We group the growth parameters into the vector

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

# Latent states

Symbol	Support	Description
$\nu$	$\ell \times 1$	Initial recruitment deviates
$\xi$		Discard mortality rate

We group the latent states into the vector

$$\omega = \{\nu, \xi\}.$$

## Other variables

Symbol	Dimensions	Description
$\boldsymbol{w}_h$	$\ell \times 1$	Mean weight at length ( $\ell$ ) by sex ( $h$ )
$\boldsymbol{m}_h$	$\ell \times 1$	Average proportion mature at length ( $\ell$ ) by sex

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

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



# 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

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

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

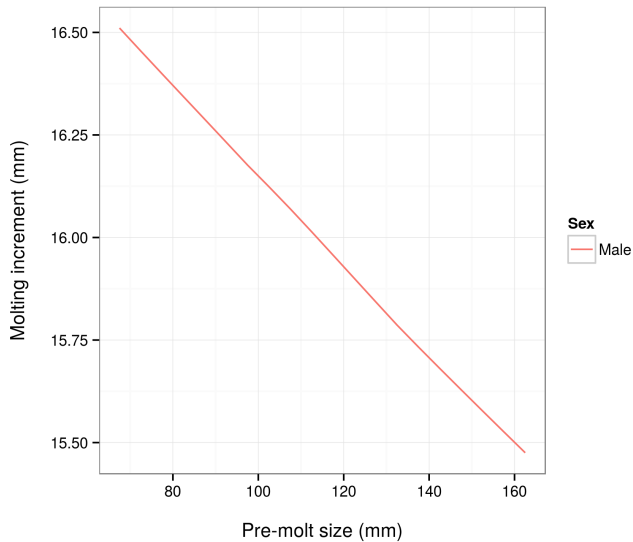
The probability of transitioning from size class  $\ell$  to  $\ell'$  assumes that variation in molt increments follows a gamma distribution

$$p(\ell'|\ell)_h = \mathbf{G}_h = \int_{\ell}^{\ell+\Delta\ell} \frac{\ell^{a_{h,\ell}-1} \exp\left(-\frac{\ell}{\varphi_h}\right)}{\Gamma(a_{h,\ell}) \ell^{a_{h,\ell}}} \quad \text{where} \quad \Delta\ell = \ell' - \ell.$$

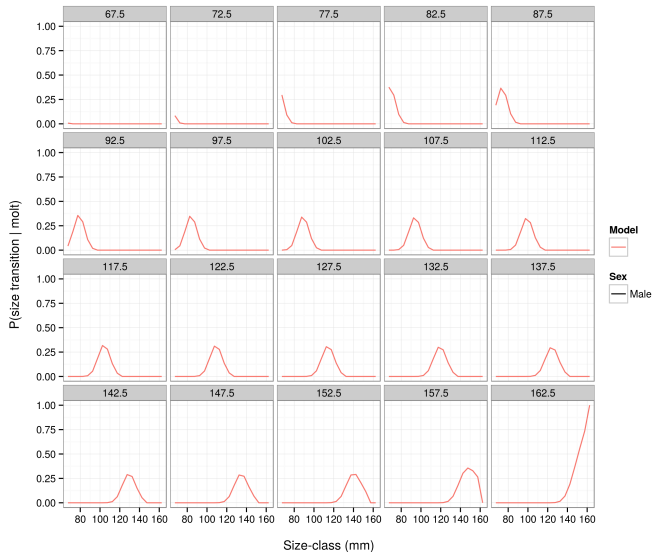
Specifically

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

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



# Growth transitions ( $G_h$ )



## Molt probability ( $\mathbf{P}_h$ )

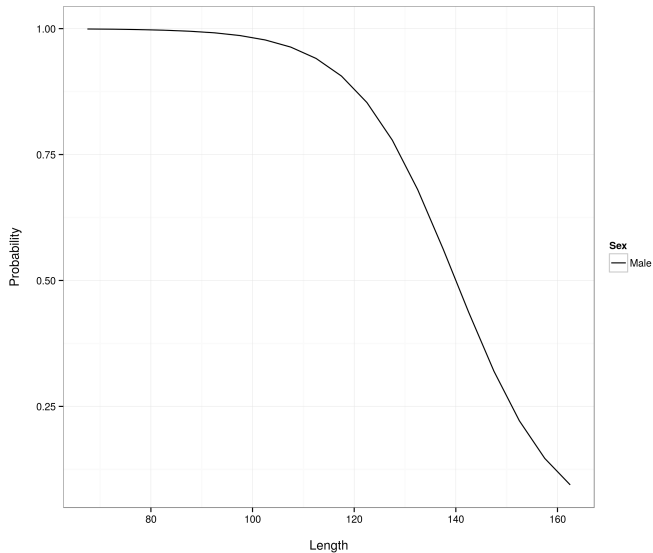
The standard deviation of molting probability ( $\sigma_h$ ) is calculated from the coefficient of variation of molting probability ( $c_h$ ) as

$$\sigma_h = \mu_h c_h.$$

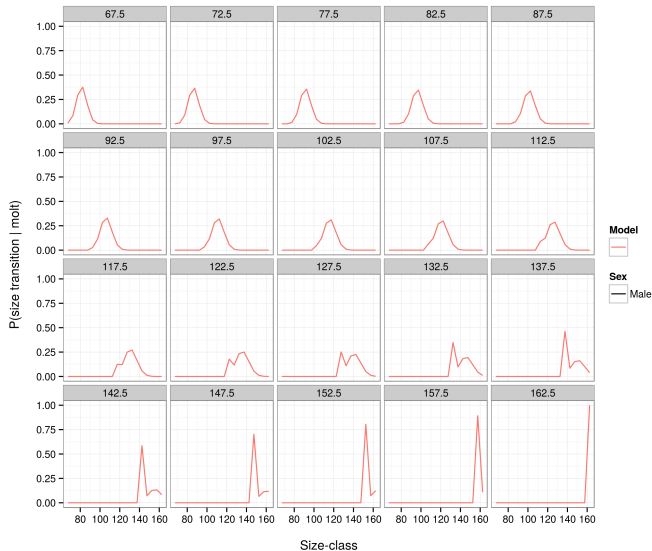
The molting probability ( $\mathbf{P}_h$ ) is calculated as

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

# Molt probability ( $P_h$ )



# Size transitions ( $P_h G_h$ )



# Natural mortality

Symbol	Description
$M_{0,h}$	Initial instantaneous natural mortality rate
$\sigma_M$	Standard deviation of natural mortality
$\delta_i$	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 ( $\ell$ ), 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  $\sigma_M^2$ )
- 3 Cubic Spline (deviates constrained by nodes and node placement)
- 4 Blocked changes (deviates constrained by variance at specific knots)

## 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  $\sigma_M^2$

$$M_{h,i+1} = \begin{cases} M_{0,h} \\ M_{h,i} e^{\delta_i} \end{cases},$$

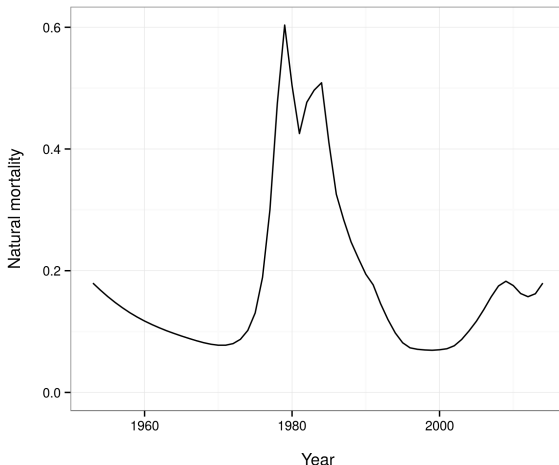
where

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

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

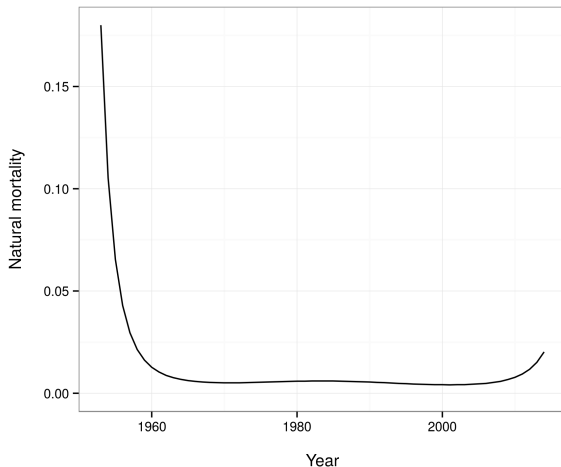
## Natural mortality: option 2

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



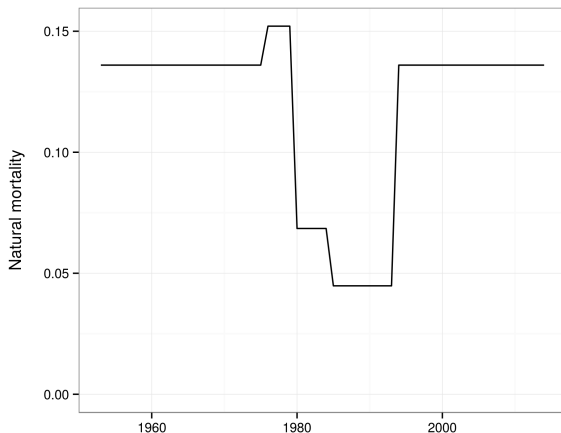
## Natural mortality: option 3

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 ( $\sigma_M^2$ ). 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
$\sigma_{h,i,k}^s$	1	Standard deviation in length at selectivity
$s_{h,i,k}$	$\ell \times 1$	Length at 50% selectivity in length interval $\ell$
$r_{h,i,k}$	1	Length at 50%
$\sigma_{h,i,k}^y$	1	Standard deviation in length at retention
$y_{h,i,k}$	$\ell \times 1$	Length at 50% retention in length interval $\ell$
$\xi_{h,k}$	1	Discard mortality rate for sex $h$ and gear $k$
$\nu_{h,i,k}$	$\ell \times 1$	Vulnerability due to fishing mortality for sex $h$
$\bar{f}_k$	$i \times 1$	Average fishing mortality rate for gear $k$
$\Psi_{k,i}$		Fishing mortality deviate for gear $k$ in year $i$
$F_{k,i}$		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  $\ell$  (i.e. selectivity) is

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

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

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

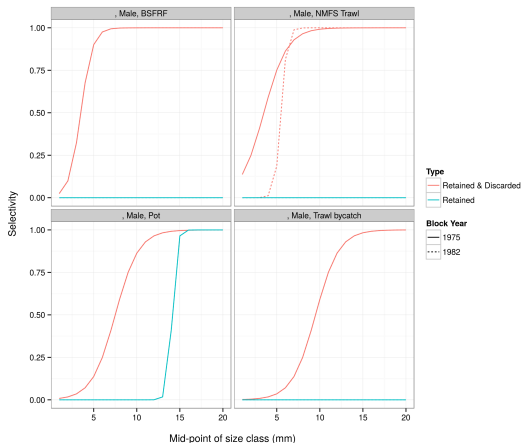
The joint probability of vulnerability due to fishing and discard mortality is

$$\nu_{h,i,k} = s_{h,i,k} \left[ y_{h,i,k} + (1 - y_{h,i,k}) \xi_{h,k} \right],$$

where  $\xi_{h,k}$  is the discard mortality rate for sex  $h$  in fishery  $k$ .

# 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  $y_{h,i,k} = y_{h,k}$





# Fishing mortality

$$\mathbf{F}_{h,i} = \sum_k \exp(\bar{\mathbf{f}}_k + \mathbf{\Psi}_{k,i}) \nu_{h,i,k},$$

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

# Recruitment

Recruitment size-distribution

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

$$p(\mathbf{r}_i) = \int_{x_\ell - 0.5\Delta x}^{x_\ell + 0.5\Delta x} \frac{x^{\alpha-1} \exp\left(\frac{x}{\beta_r}\right)}{\Gamma(\alpha)x^\alpha} dx,$$

$$\mathbf{r}_{h,i} = 0.5p(\mathbf{r}_i)\ddot{R}.$$

# Initialization

Growth and survival

$$\mathbf{A}_h = \mathbf{G}_h [\exp(-\textcolor{red}{M}_{0,h})\mathbf{I}] \quad \text{for } i = 1.$$

$$\mathbf{u}_h = -(\mathbf{A}_h - \mathbf{I})^{-1}(p[\mathbf{r}]),$$

Steady-state conditions

$$B_0 = \textcolor{red}{R}_0 \sum_h \lambda_h \sum_{\ell} \mathbf{u}_h \mathbf{w}_h \mathbf{m}_h,$$

The total unfished numbers in each size category is defined as  $R_0 \mathbf{u}_h$ .

Initial numbers at length

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

where

$$\boldsymbol{\nu} \sim \mathcal{N}(0, ).$$

## Survivorship to length

Assuming a non-zero steady-state fishing mortality rate vector  $\mathbf{f}_{h,i}$ , the equilibrium growth and survival process is represented by

$$\mathbf{B}_{h,i} = \mathbf{G}_h [\exp(-M_{h,i} - \mathbf{f}_{h,i})\mathbf{I}] .$$

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

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

$$\mathbf{v}_h = -(\mathbf{B}_h - \mathbf{I})^{-1}(p[\mathbf{r}]) .$$

The vector  $\mathbf{v}_h$  represent the unique equilibrium solution for the numbers per recruit in each size category.

$$\tilde{B} = \tilde{R} \sum_h \lambda_h \sum_{\ell} \mathbf{v}_h \mathbf{w}_h \mathbf{m}_h .$$

# Population dynamics

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 ( $\boldsymbol{B}_{h,i}$ ), plus new recruits ( $\boldsymbol{r}_{h,i}$ )

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

# 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

- $\log_{fdev}$  to ensure they sum to zero
- mean  $F$  to regularize the solution
- constrain  $M$  in random walk
- recruitment deviations

## Log-likelihoods

In general, if we have a random variable  $x$  that is normally distributed with mean  $\mu$  and variance  $\sigma^2$ , we write

$$x \sim \mathcal{N}(\mu, \sigma^2).$$

The log-likelihood is

$$\ell(\mu, \sigma^2; x_1, \dots, x_n) = -\frac{n}{2} \log(2\pi) - n \log(\sigma) - \frac{1}{2\sigma^2} \sum_i^n (x_i - \mu)^2.$$

If a coefficient of variation  $c$  is defined, rather than a standard deviation  $\sigma$  we write

$$\sigma = \sqrt{\log(1 + c^2)}.$$

If we have a random variable  $x$  that is log-normally distributed with location  $\mu$  and scale  $\sigma$ , we write

$$x \sim \log \mathcal{N}(\mu, \sigma^2).$$

The log-likelihood is

$$\ell(\mu, \sigma^2; x_1, \dots, x_n) = 0.5 \log(2\pi) + \log(\sigma) + \log(x) + \frac{1}{2\sigma^2} (x - \mu)^2.$$

## Log-likelihoods: catch

The standard deviation is calculated from the CV as

$$\sigma = \sqrt{\log(1 + c^2)}.$$

The log-likelihood is

$$\ell(\mu, \sigma^2; x_1, \dots, x_n) = \cancel{-\frac{n}{2} \log(2\pi)} - n \log(\sigma) - \frac{1}{2\sigma^2} \sum_i^n (x_i - \mu)^2.$$



## Log-likelihoods: relative abundance

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

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

$$I_i = qV_i$$

The standard deviation is calculated from the CV as

$$\sigma = \sqrt{\log(1 + c^2)}.$$

The log-likelihood is

$$\begin{aligned} \ell(\mu, \sigma^2; x_1, \dots, x_n) &= n\lambda - \cancel{\frac{n}{2} \log(2\pi)} - n \log(\sigma) - \frac{1}{2\sigma^2} \sum_i^n (x_i - \mu)^2 \\ &= n\lambda - n \log(\sigma) - \frac{1}{2\sigma^2} \sum_i^n (x_i - \mu)^2 \end{aligned}$$

# Log-likelihoods: size composition

Size composition data is assumed to be multinomial distributed

$$\mathbf{P}_{h,i} = (P_\ell)_{h,i} = \text{Multinomial}(n_{h,i}, \mathbf{Q}_{h,i})$$

Alternatively we could use

$$\mathbf{P}_{h,i} = (P_\ell)_{h,i} = \text{Dirichlet}(\lambda_0 n_{h,i} \mathbf{Q}_{h,i}).$$

In this context,  $\lambda_0$  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

Fishing mortality

$$\ell(F_k) = 0.5 \log(2\pi) + \log(\sigma_{F_k}) + 0.5 \frac{1}{\sigma_{F_k}^2} (\bar{F}_k - \hat{F})^2.$$

Natural mortality

$$\ell(M_h) = \begin{cases} 0.5 \log(2\pi) + \log(\sigma_{M_h}) + 0.5 \frac{1}{\sigma_{M_h}^2} \sum_i \delta_i^2 & \text{or} \\ 0.5 \log(2\pi) + \log(\sigma_{M_h}) + 0.5 \frac{1}{\sigma_{M_h}^2} \sum_i (\delta_{i+1} - \delta_i)^2 \end{cases}.$$

Abundance indices

$$\ell(I_k) = 0.5 \log(2\pi) + \log(\sigma_{F_k}) + 0.5 \frac{1}{\sigma_{F_k}^2} (\bar{F}_k - \hat{F})^2.$$