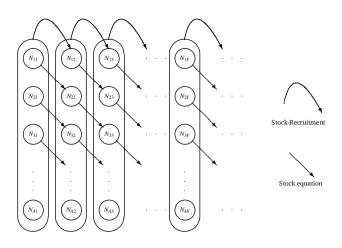
#### Basic configurations

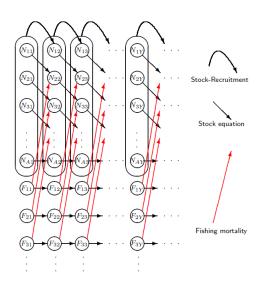
Olav N.R. Breivik



#### Stock assessment



#### Stock assessment



### State space stock assessment

From yesterday we remember that SAM assumes:

$$\begin{split} \log N_{1,y} &= \log R(\mathbf{N}_{y-1}) + \eta_{1,y} \\ \log N_{a,y} &= \log N_{a-1,y-1} - F_{a-1,y-1} - M_{a-1,y-1} + \eta_{a,y} \\ \log N_{A,y} &= \log (N_{A-1,y-1}e^{-F_{A-1,y-1}-M_{A-1,y-1}} + N_{A,y-1}e^{-F_{A,y-1}-M_{A,y-1}}) + \eta_{A,y} \end{split}$$

were

$$\log \mathbf{F}_y = \log \mathbf{F}_{y-1} + \boldsymbol{\xi}_y$$

### State space stock assessment

From yesterday we remember that SAM assumes:

$$\begin{split} \log N_{1,y} &= \log R(\mathbf{N}_{y-1}) + \eta_{1,y} \\ \log N_{a,y} &= \log N_{a-1,y-1} - F_{a-1,y-1} - M_{a-1,y-1} + \eta_{a,y} \\ \log N_{A,y} &= \log (N_{A-1,y-1} e^{-F_{A-1,y-1} - M_{A-1,y-1}} + N_{A,y-1} e^{-F_{A,y-1} - M_{A,y-1}}) + \eta_{A,y} \end{split}$$

were

$$\log \mathbf{F}_y = \log \mathbf{F}_{y-1} + \boldsymbol{\xi}_y$$

Observe:

$$\log C_{a,y} = \log \mu_{C_{a,y}} + \epsilon_{a,y}^{c}$$
$$\log I_{y}^{(s)} = \log \mu_{I_{a,y}}^{(s)} + \epsilon_{a,y}^{s}$$

Assumes  $\eta_y$ ,  $\xi_y$  and  $\epsilon_y^C$  and  $\epsilon_y^s$  all Gaussian distributed.

#### Recap of standard SAM run:

```
library(stockassessment)
  cn <- read.ices("cn.dat")
  cw <- read.ices("cw.dat")
  dw <- read.ices("dw.dat")</pre>
  lw <- read.ices("lw.dat")</pre>
6 mo <- read.ices("mo.dat")
  nm <- read.ices("nm.dat")
8 pf <- read.ices("pf.dat")</pre>
  pm <- read.ices("pm.dat")
10 sw <- read.ices("sw.dat")</pre>
  lf <- read.ices("lf.dat")</pre>
  surveys <- read.ices("survey.dat")</pre>
14
  dat <- setup.sam.data(surveys=surveys,
                        residual.fleet=cn,
                        prop.mature=mo.
                        stock.mean.weight=sw,
18
                        catch.mean.weight=cw,
                        dis.mean.weight=dw,
                        land.mean.weight=lw.
                        prop.f=pf,
                        prop.m=pm,
                        natural.mortalitv=nm.
                        land.frac=lf)
  conf<- defcon(dat)
  par <- defpar(dat,conf)
  fit <- sam.fit(dat, conf, par)
  saveConf(fit$conf,file = "conf.cfg")
31 #Modify conf.cfg and load it
  conf2 = loadConf(dat,file = "conf.cfg")
  par2 <- defpar(dat.conf2)
  fitNew <- sam.fit(dat, conf2, par2)
```

## **Basic configurations**

We will now elaborate basic configurations for:

- Recruitment function
- Survival process
- Fishing mortality process
- Survey catchability
- Observation uncertainty
- Prediction-variance relation
- Correlated observations
- Age plus groups in surveys and catch

### State space stock assessment

We assume the standard stock equations:

$$\begin{split} \log N_{1,y} &= \log R(\mathbf{N}_{y-1}) + \eta_{1,y} \\ \log N_{a,y} &= \log N_{a-1,y-1} - F_{a-1,y-1} - M_{a-1,y-1} + \eta_{a,y} \\ \log N_{A,y} &= \log (N_{A-1,y-1} e^{-F_{A-1,y-1} - M_{A-1,y-1}} + N_{A,y-1} e^{-F_{A,y-1} - M_{A,y-1}}) + \eta_{A,y} \end{split}$$

were

$$\log \mathbf{F}_y = \log \mathbf{F}_{y-1} + \boldsymbol{\xi}_y.$$

$$\log C_{a,y} \sim N(\mu_{C_{a,y}}, \sigma_{c,a}^2)$$
  
 $\log I_{a,y}^{(s)} \sim N(\mu_{I_{a,y}}^{(s)}, \sigma_{s,a}^2)$ 

#### Configurations automatically generated based on data:

minAge

```
$\pi\int \text{minAge}
# The minimium age class in the assessment
3
```

maxAge

```
1 $maxAge
2 # The maximum age class in the assessment
3 11
```

These provide the maximum and minimum age in the assessment data.

**Note**: These should not be changed without changing the input data accordingly

#### **Recruitment function**

We assume the standard stock equations:

$$\begin{split} \log N_{1,y} &= \log \frac{R(N_{y-1}) + \eta_{1,y}}{\log N_{a,y}} \\ &\log N_{a,y} = \log N_{a-1,y-1} - F_{a-1,y-1} - M_{a-1,y-1} + \eta_{a,y} \\ &\log N_{A,y} = \log (N_{A-1,y-1} e^{-F_{A-1,y-1} - M_{A-1,y-1}} + N_{A,y-1} e^{-F_{A,y-1} - M_{A,y-1}}) + \eta_{A,y} \end{split}$$

were

$$\log \mathbf{F}_y = \log \mathbf{F}_{y-1} + \boldsymbol{\xi}_y.$$

Observe:

$$\log C_{a,y} \sim N(\mu_{C_{a,y}}, \sigma_{c,a}^2) \ \log I_{a,y}^{(s)} \sim N(\mu_{I_{a,y}}^{(s)}, \sigma_{s,a}^2)$$

Example with random walk:

\$stockRecruitmentModelCode
0

#### Recruitment functions

0: plain random walk  $\log N_{1,v} \sim N(\log N_{1,v-1}, \sigma_r^2)$ 

 $\log N_{1,v} \sim N(f_r(SSB_{v-1}), \sigma_r^2)$ 1: Ricker

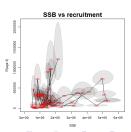
 $\log N_{1,v} \sim N(f_b(SSB_{v-1}), \sigma_r^2)$ 2: Beverton-Holt

 $\log N_{1,V} \sim N(\mu_V, \sigma_r^2)$ 3: Piece-wise constant

#### Example with random walk:

\$stockRecruitmentModelCode 2

- Recruitment is difficult to predict
- More options included that are typically used in MSY analysis.



## **Recruitment and survival process**

We assume the stock equations:

$$\begin{split} \log N_{1,y} &= \log R(\mathbf{N}_{y-1}) + \tfrac{\eta_{1,y}}{\eta_{1,y}} \\ \log N_{a,y} &= \log N_{a-1,y-1} - F_{a-1,y-1} - M_{a-1,y-1} + \tfrac{\eta_{a,y}}{\eta_{A,y}} \\ \log N_{A,y} &= \log (N_{A-1,y-1} e^{-F_{A-1,y-1} - M_{A-1,y-1}} + N_{A,y-1} e^{-F_{A,y-1} - M_{A,y-1}}) + \tfrac{\eta_{A,y}}{\eta_{A,y}} \end{split}$$

were

$$\log \mathbf{F}_{y} = \log \mathbf{F}_{y-1} + \boldsymbol{\xi}_{y}.$$

Observe:

$$\log C_{a,y} \sim \mathcal{N}(\mu_{C_{a,y}}, \sigma_{c,a}^2) \ \log I_{a,y}^{(s)} \sim \mathcal{N}(\mu_{I_{a,y}}^{(s)}, \sigma_{s,a}^2)$$

1 \$keyVarLogN

# Coupling of the recruitment and survival process variance parameters

0111111111:

# **Recruitment and survival process**

```
\begin{split} \log N_{1,y} &= \log R(\mathbf{N}_{y-1}) + \textcolor{red}{\eta_{1,y}} \\ \log N_{a,y} &= \log N_{a-1,y-1} - F_{a-1,y-1} - M_{a-1,y-1} + \textcolor{red}{\eta_{a,y}} \\ \log N_{A,y} &= \log (N_{A-1,y-1}e^{-F_{A-1,y-1}-M_{A-1,y-1}} + N_{A,y-1}e^{-F_{A,y-1}-M_{A,y-1}}) + \textcolor{red}{\eta_{A,y}} \end{split}
```

```
| SkeyVarLogN | # Coupling of the recruitment and survival process variance parameters | 0 1 1 1 1 1 1 1 1 1 1 1
```

- Typically a separate variance parameter for recruitment and survival
- Typically a common variance parameter for survival



## Fishing mortality process

We assume the stock equations:

$$\begin{split} \log N_{1,y} &= \log R(\textbf{N}_{y-1}) + \eta_{1,y} \\ \log N_{a,y} &= \log N_{a-1,y-1} - \textbf{\textit{F}}_{a-1,y-1} - M_{a-1,y-1} + \eta_{a,y} \\ \log N_{A,y} &= \log (N_{A-1,y-1} e^{-\textbf{\textit{F}}_{A-1,y-1} - M_{A-1,y-1}} + N_{A,y-1} e^{-\textbf{\textit{F}}_{A,y-1} - M_{A,y-1}}) + \eta_{A,y} \end{split}$$

were

$$\log \mathbf{F}_y = \log \mathbf{F}_{y-1} + \boldsymbol{\xi}_y.$$

$$\log C_{a,y} \sim N(\mu_{C_{a,y}}, \sigma_{c,a}^2)$$
  
 $\log I_{a,y}^{(s)} \sim N(\mu_{I_{a,y}}^{(s)}, \sigma_{s,a}^2)$ 

# Fishing mortality process

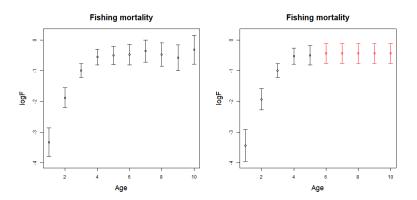
$$\begin{split} \log N_{1,y} &= \log R(\mathbf{N}_{y-1}) + \eta_{1,y} \\ \log N_{a,y} &= \log N_{a-1,y-1} - F_{a-1,y-1} - M_{a-1,y-1} + \eta_{a,y} \\ \log N_{A,y} &= \log (N_{A-1,y-1}e^{-F_{A-1,y-1}-M_{A-1,y-1}} + N_{A,y-1}e^{-F_{A,y-1}-M_{A,y-1}}) + \eta_{A,y} \end{split}$$

- The configuration keyLogFsta defines how we couple the fishing mortality states.
- Here  $(F_{1,y},...,F_{11,y}) = (F_{1,y},...,F_{5,y},F_{6,y},...,F_{6,y})$
- Often a flat selectivity pattern for older ages is selected



# Fishing mortality process

```
$keyLogFsta
# Coupling of the fishing mortality states processes for each age
0 1 2 3 4 5 5 5 5 5 5 5
-1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1
```



- Defines dimension of latent F
- Does not effect number of model parameters.

## Fishing mortality variance

We assume the stock equations:

$$\begin{split} \log N_{1,y} &= \log R(\mathbf{N}_{y-1}) + \eta_{1,y} \\ \log N_{a,y} &= \log N_{a-1,y-1} - F_{a-1,y-1} - M_{a-1,y-1} + \eta_{a,y} \\ \log N_{A,y} &= \log (N_{A-1,y-1} e^{-F_{A-1,y-1} - M_{A-1,y-1}} + N_{A,y-1} e^{-F_{A,y-1} - M_{A,y-1}}) + \eta_{A,y} \end{split}$$

were

$$\log \mathbf{F}_y = \log \mathbf{F}_{y-1} + \boldsymbol{\epsilon}_y^F.$$

$$\log C_{a,y} \sim N(\mu_{C_{a,y}}, \sigma_{c,a}^2)$$
  
 $\log I_{a,y}^{(s)} \sim N(\mu_{I_{a,y}}^{(s)}, \sigma_{s,a}^2)$ 

# Fishing mortality variance

$$\log \mathbf{F}_y = \log \mathbf{F}_{y-1} + \epsilon_y^F.$$

- The configuration *keyVarF* defines how we couple the variances of the fishing mortality increments.
- Here two variances parameters are included:  $\sigma_{1,F}$  and  $\sigma_{2,F} = \sigma_{3,F} = \cdots = \sigma_{11,F}$

### Fishing mortality correlation

We assume the standard stock equations:

$$\begin{split} \log N_{1,y} &= \log R(\mathbf{N}_{y-1}) + \eta_{1,y} \\ \log N_{a,y} &= \log N_{a-1,y-1} - F_{a-1,y-1} - M_{a-1,y-1} + \eta_{a,y} \\ \log N_{A,y} &= \log (N_{A-1,y-1} e^{-F_{A-1,y-1} - M_{A-1,y-1}} + N_{A,y-1} e^{-F_{A,y-1} - M_{A,y-1}}) + \eta_{A,y} \end{split}$$

were

$$\log \mathbf{F}_y = \log \mathbf{F}_{y-1} + \boldsymbol{\epsilon}_y^F.$$

$$\log \textit{C}_{\textit{a},\textit{y}} \sim \textit{N}(\mu_{\textit{C}_{\textit{a},\textit{y}}},\sigma_{\textit{c},\textit{a}}^2) \ \log \textit{I}_{\textit{a},\textit{y}}^{(\textit{s})} \sim \textit{N}(\mu_{\textit{I}_{\textit{a},\textit{y}}}^{(\textit{s})},\sigma_{\textit{s},\textit{a}}^2)$$

```
$corFlag
# Correlation of fishing mortality across ages (0 independent, 1 compound symmetry,
# 2 AR(1),
2
```

# **Fishing mortality correlation**

$$\log \mathbf{F}_y = \log \mathbf{F}_{y-1} + \boldsymbol{\xi}_y.$$

```
$corFlag
# Correlation of fishing mortality across ages (0 independent, 1 compound symmetry,
# 2 AR(1),
2
```

The configuration *corflag* defines the correltion structure of the fishing mortality increments.

- 0: Independent no correlation between F across age
- 1: Compound symmetry equal correlation across all ages
- 2: AR1 ages close to each other are more correlated

## **Survey catchability**

We assume the standard stock equations:

$$\begin{split} \log N_{1,y} &= \log R(\mathbf{N}_{y-1}) + \eta_{1,y} \\ \log N_{a,y} &= \log N_{a-1,y-1} - F_{a-1,y-1} - M_{a-1,y-1} + \eta_{a,y} \\ \log N_{A,y} &= \log (N_{A-1,y-1} e^{-F_{A-1,y-1} - M_{A-1,y-1}} + N_{A,y-1} e^{-F_{A,y-1} - M_{A,y-1}}) + \eta_{A,y} \end{split}$$

were

$$\log \mathbf{F}_y = \log \mathbf{F}_{y-1} + \boldsymbol{\epsilon}_y^F.$$

$$\begin{split} \log C_{a,y} &\sim \textit{N}(\mu_{\textit{C}_{a,y}}, \sigma_{\textit{c},a}^2) \\ \log \textit{J}_{a,y}^{(s)} &\sim \textit{N}(\log \textit{Q}_{a}^{(s)}\textit{N}_{a,y}, \sigma_{s,a,y}^2) \end{split}$$

# **Survey catchability**

- The catchability parameter is the proportionality constant between stock size and the observed index.
- keyLogFpar defines how the catchability parameters are coupled across age groups.

Survey observation equation:

$$\log I_{a,y}^{(s)} \sim N(\log Q_a^{(s)} N_{a,y}, \sigma_{s,a,y}^2)$$

Here it is included a unique catchability for the youngest age, a common catchability for the four next ages, and a common catchability for the older ages.

#### **Observation variance**

We assume the standard stock equations:

$$\begin{split} \log N_{1,y} &= \log R(\mathbf{N}_{y-1}) + \eta_{1,y} \\ \log N_{a,y} &= \log N_{a-1,y-1} - F_{a-1,y-1} - M_{a-1,y-1} + \eta_{a,y} \\ \log N_{A,y} &= \log (N_{A-1,y-1} e^{-F_{A-1,y-1} - M_{A-1,y-1}} + N_{A,y-1} e^{-F_{A,y-1} - M_{A,y-1}}) + \eta_{A,y} \end{split}$$

were

$$\log \mathbf{F}_y = \log \mathbf{F}_{y-1} + \boldsymbol{\epsilon}_y^F.$$

$$\log C_{a,y} \sim \mathcal{N}(\mu_{C_{a,y}}, \sigma_{c,a}^2) \ \log I_{a,y}^{(s)} \sim \mathcal{N}(\mu_{I_{a,y}}^{(s)}, \sigma_{s,a}^2)$$

#### **Observation variance**

- Observations include noise
- keyVarObs couples the observation variance parameters

#### Observation equations:

$$\log C_{a,y} \sim N(\mu_{C_{a,y}}, \sigma_{c,a}^2) \ \log I_{a,y}^{(s)} \sim N(\mu_{l_{a,y}}^{(s)}, \sigma_{s,a}^2)$$

#### Four variance parameters included

- Three for catch:  $\sigma_{c,1}$ ,  $\sigma_{c,2}$  and  $\sigma_{c,3} = \cdots \sigma_{c,11}$
- One for the survey:  $\sigma_{s,1} = \cdots \sigma_{s,8}$



#### Link between observation mean and variance

- When using the log-normal the relationship between mean  $\mu$  and variance v is  $v=\alpha\mu^2$
- This relationship may not be correct, so instead the power  $\beta$  can be estimated in  $\mathbf{v}=\alpha\mu^{\beta}$

- In the configuration above this is configured for the first two fleets
- Used in situations where the size of e.g. catches vary greatly in the time period
- Residuals can also be inspected (plot residual versus predicted)
- Go trough SAM-predVar.pdf





#### **Observation correlation**

We assume the standard stock equations:

$$\begin{split} \log N_{1,y} &= \log R(\textbf{N}_{y-1}) + \eta_{1,y} \\ \log N_{a,y} &= \log N_{a-1,y-1} - F_{a-1,y-1} - M_{a-1,y-1} + \eta_{a,y} \\ \log N_{A,y} &= \log (N_{A-1,y-1}e^{-F_{A-1,y-1}-M_{A-1,y-1}} + N_{A,y-1}e^{-F_{A,y-1}-M_{A,y-1}}) + \eta_{A,y} \end{split}$$

were

$$\log \mathbf{F}_y = \log \mathbf{F}_{y-1} + \boldsymbol{\epsilon}_y^F.$$

$$egin{aligned} \log \mathbf{C}_y &\sim \mathcal{N}(oldsymbol{\mu}_{C_y}, oldsymbol{\Sigma_c}) \ \log \mathbf{I}_y^{(s)} &\sim \mathcal{N}(oldsymbol{\mu}_y^{(s)}, oldsymbol{\Sigma_s}) \end{aligned}$$

#### **Correlated observations**

- Setting this option require two fields \$obsCorStruct and \$keyCorObs
- If only the first is set to "AR", then is does not work (interface design flaw?)
- Example:

- The bottom coupling matrix must only be filled in if the AR-structure is used.
- This matrix only has columns for successive paris of ages (one less than number of ages)
- Set to NA if unstructured or independent



### Independent observations

From the example on the previousl slide, we then assume:

$$\log \mathbf{C}_{y} \sim \mathcal{N}\Big(\boldsymbol{\mu}_{\mathcal{C}_{y}}, \boldsymbol{\Sigma}\Big) \tag{1}$$

where

$$oldsymbol{\Sigma} = egin{pmatrix} \sigma_{C,a_1}^2 & 0 & 0 & \cdots & 0 \ 0 & \sigma_{C,a_2}^2 & 0 & \cdots & 0 \ 0 & 0 & \sigma_{C,a_3}^2 & \cdots & 0 \ dots & dots & dots & dots \ 0 & 0 & 0 & \cdots & \sigma_{C,A}^2 \end{pmatrix}$$

# (Ir)regular grid AR

- The observation vector  $o_y^{(f)}$  for fleet f in year y is assumed  $o_y^{(f)} \sim \mathcal{N}(\mu_y^{(f)}, \Sigma)$
- In the regular AR structure the covariance is defined as:

$$\Sigma_{ij} = 
ho^{|i-j|} \sqrt{\Sigma_{ii} \Sigma_{jj}}$$

- So correlation only depende on distance between i and j, not which i and j.
- First realize that we can get the same covariance structure by:

$$\Sigma_{ij} = 0.5^{lpha|i-j|} \sqrt{\Sigma_{ii}\Sigma_{jj}}$$
 , where  $lpha > 0$ 

- Notice that this implies a regular grid.
- We can extend this structure by defining

$$\Sigma_{\textit{ij}} = 0.5^{|\theta_{\textit{i}} - \theta_{\textit{j}}|} \sqrt{\Sigma_{\textit{ii}} \Sigma_{\textit{jj}}} \quad , \quad \text{where} \quad \theta_1 = 0 \leq \theta_2 \leq \dots \leq \theta_{\textit{A}}$$

- This corresponds to having the points on an irregular grid.
- If all deltas are the same, then it is a regular AR structure



Olav N.R. Breivik

#### **Unstructured covariance**

The fully unstructured covariance can be constructured in the following way.

$$\Sigma_{ij} = (D^{-\frac{1}{2}}LL^tD^{-\frac{1}{2}})_{ij}\sqrt{\Sigma_{ii}\Sigma_{jj}}$$

 Here L is a lower triangle matrix (Cholesky of the correlation) and D is the diagonal matrix of (LL<sup>t</sup>)

$$L = \begin{pmatrix} 1 & 0 & 0 \\ \theta_1 & 1 & 0 \\ \theta_2 & \theta_3 & 1 \end{pmatrix}$$

- The model parameters are the elements in L and the log-standard deviations
- This is very flexible, but also requires a lot of parameters to be estimated
- Now we have a lot of options (ID, AR, IGAR, US)
- How can we go about choosing an optimal configuration?
- Useful diagnostics: res <- residuals(fit), plot(res), empirobscorrplot(res)

## Plus groups

#### Plus group age can very between fleets

```
1 $maxAgePlusGroup
2 # Is last age group considered a plus group for each fleet (1 yes, or 0 no)
3 1 1
```

Here both catch and survey consist of observations where the oldest age group contains that age and above.

The plus group implies the following update to the stock equation:

$$N_{A,y} = N_{A-1,y-1}e^{-Z_{A-1,y-1}} + N_{A,y-1}e^{-Z_{A,y-1}}$$

 If A<sub>f</sub> is less than A, then the stock number used to predict observations at age A<sub>f</sub> for the fleet is the sum of the stock numbers from A<sub>f</sub> to A.

# F-bar range

#### Age range in $\bar{F}$ is controlled by *fbarRange*

```
$fbarRange
# lowest and higest age included in Fbar
4 7
```

Here ages 4 to 7 are included in  $\bar{F}$ , i.e.,:

$$\bar{F}_y = \frac{1}{4} \sum_{a=4}^{7} F_{a,y}$$